Popular Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

UNIVERSITY OF MEDEA
Department of Process Engineering and Environment
Laboratory of Biomaterials and Transport Phenomena (LBMTP)

THESIS

Presented by:

Faiza OMARI

In order to obtain Doctor title in Process Engineering

Option: Chemical Engineering

Theme

MODELLING THE OPTICAL PROPERTIES OF BLOOD USING ARTIFICIAL INTELLIGENCE

The jury is composed of:

RAHAL Sofiane	MCA	UYFM	President
BEZARI Salah	MRA	CDER	Examiner
ATSAMNIA Djamel	MCA	UYFM	Examiner
KHAOUANE Latifa	Prof	UYFM	Thesis director
LAIDI Maamar	Prof	UYFM	Thesis co-director
HANINI Salah	Prof	UYFM	Guest

Dedications

To my parents,

I could never have done this without your faith, support, and encouragement. Thank you for teaching me to believe in myself and in my dreams.

I also dedicate this thesis to my sisters (Sadjia & Amel) and brothers (Mohamed & Fouad) who are my nearest surrounders and have provided me with a strong love shield that always surrounds me.

To my nephews Rashid, Rayan, Nassim, Djawad, Abd el Rahim, and Raid. They give me so many reasons to smile

To all OMARI and MELIKAOUI families

To my friends and all those who know and supported me closely or by far.

Faiza.

Acknowledgment and Gratitude

First of all, I want to thank ALLAH Almighty for giving me health, courage, and patience to do this work

I want to express my gratitude to my supervisor Pr. KHAOUANE Latifa for her help and guidance. I feel thankful to her for the insightful advice and suggestions.

I am also thankful to pr. LAIDI Maamar and Pr. HANINI Salah for giving me generous amount of time whenever I needed some help.

I am grateful to Dr. ROUBEHIE FISSA Mohamed & Dr. IBRIR Abdellah who I worked with during my research and learned a lot.

I also want to present my deep gratitude to: Dr. RAHAL Sofiane, Dr. BENZARI Salah, and Dr. ATSAMNIA Djamel for agreeing to judge this modest work.

Thank you to all those who helped me closely or by far during the years of studies and research, especially to all my teachers and colleagues.

I feel obliged to say "thanks" to so many people who have been doing so many good things for me, but I am afraid the list may go forever.

Summary

This thesis explored the crucial significance of the optical properties of blood in medical research in order to develop novel methodologies using artificial intelligence techniques, including artificial neural networks (ANN), support vector machines (SVM), and the dragonfly algorithm in combination with support vector machines (DA-SVM), to predict key optical properties of blood.

In the first contribution we focus on using feed-forward neural models to predict these properties based on easily measurable variables. To achieve this, three distinct feed-forward neural models were constructed to predict the refractive index, the absorption coefficient, and the scattering coefficient., the correlation coefficients for the test set were 0.9990, 0.9972, and 0.9987 for the neural networks describing the refractive index, absorption coefficient, and scattering coefficient, respectively. Additionally, the models exhibit low root mean square errors.

In the second contribution of this thesis, support vector machines (SVM) were employed to predict the optical properties of blood. The outcomes yielded an acceptable match between the experimental and predicted values. Notably, the correlation coefficients during the testing phase were found to be 0.9835, 0.8715, and 0.9788 for the refractive index, absorption coefficient, and scattering coefficient, respectively. The root mean square error values were 0.0043, 9.7411, and 6.3632.

In the third contribution, we introduced the DA-SVM approach to predict these properties using the same dataset. The results demonstrating that the established models exhibited exceptional predictive performance. Remarkably high correlation coefficients of 0.9971, 0.9994, and 0.9996 were achieved for the refractive index, absorption coefficient, and scattering coefficient, respectively. The RMSE values of 0.0028, 0.9792, and 2.9193 for the refractive index, absorption coefficient, and scattering coefficient, respectively.

The results comparison across the ANN, SVM, and DA-SVM models definitely demonstrates the exceptional performance of both the ANN and DA-SVM models in our research. Through a comprehensive comparative analysis, wherein we evaluated our models to those provided in the literature, it becomes evident that our ANN and DA-SVM models are superior and have a significant position in this domain.

Keywords: Optical Properties; Human Blood; Neural Networks; Support Vector Machines; Dragonfly Algorithm

الملخص

تناولت هذه لأطروحة الأهمية الحاسمة للخصائص البصرية للدم في البحوث الطبية من أجل تطوير منهجيات جديدة باستخدام تقنيات الذكاء الاصطناعي، بما في ذلك الشبكات العصبية الاصطناعية (ANN)، وآلات ناقلات الدعم (SVM)، وخوارزمية اليعسوب مقترنة مع ناقلات الدعم (DA-SVM) للتنبؤ بالخصائص البصرية للدم.

في المساهمة الأولى، نركز على استخدام نماذج عصبية متقدمة للتنبؤ بتلك الخصائص استنادًا إلى المتغيرات السهلة القياس. ولتحقيق ذلك، تم بناء ثلاثة نماذج عصبية متقدمة متميزة للتنبؤ بمعامل الانكسار ومعامل الامتصاص ومعامل الانتشار. وقد بلغت معاملات الارتباط للمجموعة التجريبية قيمًا بلغت 0.9990، 0.9972، و0.9987 على التوالي للشبكات العصبية التي تصف معامل الانكسار ومعامل الامتصاص ومعامل الانكسار. بالإضافة إلى ذلك، تتميز النماذج بأخطاء جذرية مربعة منخفضة.

في المساهمة الثانية لهذه الأطروحة، تم استخدام آلات ناقلات الدعم (SVM) للتنبؤ بالخصائص البصرية للدم .أسفرت النتائج عن تطابق مقبول بين القيم التجريبية والقيم المتوقعة .ومن الجدير بالذكر أن معاملات الارتباط خلال مرحلة الاختبار كانت 0.9835، 0.8715، و0.9788 لمعامل الانكسار، معامل الامتصاص، ومعامل الانتشار، على التوالي .وكانت قيم خطأ مربع الجذر المتوسط هي 0.0043 و9.7411 و6.3632.

في المساهمة الثالثة، قدمنا نهج DA-SVM للتنبؤ بهذه الخصائص باستخدام نفس مجموعة البيانات. أظهرت النتائج أن النماذج القائمة أظهرت أداء تنبؤيا استثنائيا .تم تحقيق معاملات ارتباط عالية بشكل ملحوظ بلغت 0.9971، 0.9994، و0.9996 لمعامل الانكسار، ومعامل الامتصاص، ومعامل الانتشار، على التوالي .قيم RMSE هي 0.0028 و0.9792 و2.9193 لمعامل الانكسار ومعامل الامتصاص ومعامل التشتت على التوالي.

توضح مقارنة النتائج عبر نماذج ANN و SVMو DA-SVMبالتأكيد الأداء الاستثنائي لكل من نماذج ANN و-DA SVMفي بحثنا. من خلال تحليل مقارن شامل، حيث قمنا بتقييم نماذجنا لتلك الواردة في الأدبيات، يصبح من الواضح أن نماذج ANNو DA-SVMالخاصة بنا متفوقة ولها مكانة مهمة في هذا المجال.

الكلمات المفتاحية: الخصائص البصرية؛ دم الإنسان؛ الشبكات العصبية؛ وآلات ناقلات الدعم؛ خوارزمية اليعسوب

Résumé

Cette thèse a exploré l'importance des propriétés optiques du sang dans la recherche médicale afin de développer de nouvelles méthodologies utilisant des techniques d'intelligence artificielle, notamment les réseaux de neurones artificiels (ANN), les machines à vecteurs de support (SVM) et l'algorithme de libellule en combinaison avec les machines à vecteurs de support (DA-SVM), pour prédire les propriétés optiques du sang.

Dans la première contribution, nous utilisons des modèles neuronaux de type feed-forward pour prédire les propriétés optiques en fonction de variables mesurables. Trois modèles ont été créés pour l'indice de réfraction, le coefficient d'absorption, et le coefficient de diffusion, montrant des coefficients de corrélation élevés (0,9990, 0,9972, 0,9987) et des faibles valeurs d'erreurs.

Dans la deuxième contribution de cette thèse, les machines à vecteurs de support (SVM) ont été utilisées pour prédire les propriétés optiques du sang. Les résultats ont montré une concordance acceptable entre les valeurs expérimentales et prédites. Notamment, les coefficients de corrélation lors de la phase de test étaient de 0,9835, 0,8715 et 0,9788 pour l'indice de réfraction, le coefficient d'absorption et le coefficient de diffusion, respectivement. Les valeurs de RMSE étaient de 0,0043, 9,7411 et 6,3632.

Dans la troisième contribution significative, nous avons introduit l'approche DA-SVM pour prédire ces propriétés. Les modèles établis ont montré d'excellentes performances avec des coefficients de corrélation élevés de 0,9971, 0,9994 et 0,9996 pour l'indice de réfraction, le coefficient d'absorption et le coefficient de diffusion, respectivement. Les valeurs de RMSE étaient de 0,0028, 0,9792 et 2,9193 pour ces propriétés dans la phase de test.

La comparaison des résultats entre les modèles ANN, SVM et DA-SVM démontre clairement la performance exceptionnelle des modèles ANN et DA-SVM dans notre recherche. Une analyse comparative complète, dans laquelle nous avons évalué nos modèles par rapport à ceux présentés dans la littérature, montre clairement que nos modèles ANN et DA-SVM sont supérieurs et occupent une place significative dans ce domaine.

Mots clés : Propriétés Optiques ; Sang Humain ; Réseaux Neuronaux ; Machines à Vecteurs de Support ; Algorithme de libellule.

Table of Contents

Abbreviations i
Nomenclatures and symbols iii
List of figures v
List of tables viii
List of work and research activities ix
General introduction 1
I. Chapter I: blood and its optical properties 4
I.1. Introduction 4
I.2. History of blood 4
I.3. The blood 5
I.4. Blood components 5
I.4.1. Red blood cells 7
I.4.2. White blood cells 8
I.4.2.1. Monocytes 9
I.4.2.2. Granulocytes 9
I.4.2.3. Lymphocytes 10
I.4.3. Platelets 11
I.4.4. Plasma 12
I.4.4.1. Criteria for choosing between serum and plasma in the medical analysis laboratory . 13
I.4.4.2 Plasma proteins 13
I.5. Blood function 16
I.5.1. Transportation 16
I.5.2. The regulation 16
I.5.3. Protection 16
I.6. The physical properties of blood 16
I.7. Therapeutic use of blood 17
I.8. The optical properties of blood 17
I.8.1. Refractive index 17
I.8.1.1. Definition 17
I.8.1.2. Measurement of a refractive index with the Abbe refractometer 18
I.8.1.3. Construction of the refractometer 18
I.8.1.4. Principle of operation 19

I.8.1.5. The use of refractometer 20
I.8.2. Absorption and diffusion 21
I.8.2.1. Definition of Absorption 21
I.8.2.2. Absorption measurement with spectrometry 21
I.8.2.3. Definition of scattering 23
I.8.2.3. Theoretical principle of the propagation of light in absorbing and scattering media 23
Conclusion 23
II. CHAPTER II: Trends in AI modeling approaches 24
II.1. Introduction 24
II.2. Historical 24
II.3. Basics of artificial neural networks (ANN) 24
II.3.1. Analogy between the biological neuron and the artificial neuron 24
II.3.2. Neural network architecture 27
II.3.3. Feed-forward static networks 28
II.3.5. Recurring Network "Feedback" 29
II.3.6. Activation functions 30
II.3.7 Learning phenomena of neural networks 31
II.3.7.1. Learning supervised learning 31
II.3.7.2. Unsupervised learning (unsupervised learning) 32
II.3.8. The most used neural networks 32
II. 3.8.1. The single-layer perceptron (simple perceptron) 33
II.3.8.2. The Multi-layer perceptron 33
II.3.8.3. Radial function networks 34
II.3.8.4. Kohonen Self-Organizing Maps (SOM) 34
II.3.8.5. Hopfield 's networks 35
II.3.8.6. ART networks 35
II.3.9. Designing a neural model 35
II.3.9.1. Determining the inputs/outputs of the neural network 35
II.3.9.2. Construction of the database and choice of samples 36
II.3.9.3. Definition of the network structure 36
II.3.9.3.1. Number of hidden layers 37
II.3.9.3.2. Number of hidden neurons 37
II.3.9.4. Neural network performance evaluation 38

II.3.10. Application areas 38
II.3.10.1. Predictive modeling 38
II.3.10.2. Pattern recognition 38
II.3.10.3 Optimization 38
II.3.10.4. Control 39
II.3.11. Advantages and disadvantages of ANNs 39
II.4. Support Victor Machine (SVM) 39
II.4.1. Introduction 39
II.4.2. Statistical learning 39
II.4.3. SVM for regression 40
II.4.3.1. Using Cores 42
II.4.3.2. The kernels 42
II.4.4. Model evaluation 43
II.5. Dragonfly algorithm 43
II.6. Conclusion 45
III. Chapter III: Theoretical background on the optical properties of blood 46
III.1. Critical reviews about optical properties 46
III.2. Experimental studies and empirical modeling of the optical properties of blood components 46
III.3. Optical properties modelling using Artificial intelligence approaches 53
IV. CHAPTER IV: MATERIALS AND METHODS 56
IV.1. Introduction 56
IV.2. Materials and methods 56
IV.2.1. Computer Tools used 56
IV.2.2. Data collection, pretreatment and analysis 57
IV.2.3. Artificial neural networks (ANN) modeling 62
IV.2.3.1. Division of data base 62
IV.2.3.2. Optimization of ANN architectures 62
IV.2.3.3. Selection of the best model 63
IV.2.3.4. Statistical analysis and errors calculation 64
IV.2.3.5. Sensitivity analysis using weight method 64
IV.2.4. Support Vector Machine (SVM) modeling 66
IV.2.5. Modeling by Support Vector Machine coupled with Dragonfly Algorithm (DA–SVM) 68

Table of Contents

V. Chapter V: Results and Discussion 71
V.1. Introduction 71
V.2. MLP-ANN models (*Contribution I*) 71
V.2.1. Comparison between the experimental refractive index and calculated by MLP-ANN ... 81
V.2.2. Mathematical expressions of the optimized neural models 86
V.2.3. Sensitivity analysis by "weight" method 87
V.3. SVM models (*Contribution II*) 93
V.3.2. Comparison between the experimental optical properties and calculated by SVM 100
V.4. DA-SVM models (*Contribution III*) 104
V.5. Comparison between MLP-ANN, SVM and DA-SVM models 115
V.6. Comparison between our models and the models available in the literature 120
V.7. Conclusion 121
General conclusion and perspectives 124
References 127
Annex 134

Abbreviations

Symbol	Signification
AANN	**A**rtificial **A**uto-**A**ssociative **N**eural **N**etwork
AI	**A**rtificial **I**ntelligence
ANN	**A**rtificial **N**eural **N**etworks
ART	**A**daptive **R**esonance **T**heory
BFGS	**B**royden **F**letcher **G**oldfarb **S**hanno
CCD	**C**harge-**C**oupled **D**evice
DA	**D**ragonfly **A**lgorithm
DB	**D**ata **B**ase
FFNN	**F**eed **F**orward **N**euronal **N**etwork
GA	**Genetic Algorithm**
HB	**H**emoglobin
HCT	**H**ematocrit
MAE	**M**ean **A**bsolute **E**rror
MAPE	**M**ean **A**bsolute **P**ercentage **E**rror
MATLAB	**M**atrix **L**aboratory
MLP	**M**ulti-**L**ayer **P**erceptron
MRE	**M**ean **R**elative **E**rror
NK	**N**atural **K**iller
NN	**N**eural **N**etworks
OCT	**O**ptical **C**oherence **T**omography
PC	**P**ersonal **C**omputer
PLTs	**P**latelets
PSO	**P**article **S**warm **O**ptimization
P-Value	**P**robability
RBF	**R**adial **B**asis **F**unction
RBCs	**R**ed **B**lood **C**ells

Abbreviations

RI	**R**efractive **I**ndex
RMSE	**R**oot **M**ean **S**quare **E**rror
SD	**S**tandard **D**eviation
SEP	**S**tandard **E**rror of **P**rediction
SOM	**S**elf-**O**rganizing **M**ap
SOS	**S**um **of S**quares
SVM	**S**upport **V**ector **M**achine
SVR	**S**upport **V**ector **R**egression
UV	**U**ltra -**V**iolet

Nomenclatures and symbols

Reading	Designation	Unit
A	**A**bsorbance	
C	**C**oncentration of the solution	(g/dL)
Eq	**E**quation	(-)
f	Activation **f**unction	(-)
T	**T**emperature	C°
K	**K**ernel function	(-)
N	Number of experimental points	(-)
R	Correlation coefficient	(-)
$\mathbf{R^2}$	Coefficient of determination	(-)
W	**W**eight	(-)
$\mathbf{b_i}$	**B**ias	(-)
RI	**R**elative **I**mportance	%
$\mathbf{Z_i}$	Output of the hidden layer	(-)

Greek Letters

Reading	Designation	Unit
α	Slope of the straight line of the best linear fit	(-)
B	Ordered at the origin of the straight line of the best linear adjustment	(-)
μ_a	Absorption coefficient	(1/mm)
μ_s	Scattering coefficient	(1/mm)
η	Real part of the refractive index	(-)
ε	Dissipation rate	(-)
σ	Bandwidth of the RBF Gaussian kernel	(-)
λ	Wavelength	(nm)

Exhibitors

Reading	Designation
cal	**Calculated**
exp	**Experimental**
O2 sat	**Oxygen saturation**
val	**Validation**

Indices

Reading	Designation
i	Item i of a vector
ij	Crossing line I with column j of a matrix
j	Item j of a vector
max	Maximum
min	Minimum

List of figures

Chapter I

Figure I. 1. An overview of the composition of human blood and the further subdivision of human blood cells 6
Figure I. 2. Composition of blood system (Version & Engineering, 2021). 6
Figure I. 3. Red blood cells 7
Figure I. 4. Monocyte cell 9
Figure I. 5. Type of granulocytes 10
Figure I. 6. Type of Lymphocytes 11
Figure I. 7. Blood platelets 12
Figure I. 8. Blood plasma 13
Figure I. 9. Structure of albumin 14
Figure I. 10. General structure of an immunoglobulin 15
Figure I. 11. The structure of fibrinogen 15
Figure I. 12. The compositions of the refractometer 18
Figure I. 13. The refractometer prisms 19
Figure I. 14. Sight adjustment 20
Figure I. 15. Reading the refractive index (Singh, 2002) 20
Figure I. 16. The spectrophotometer 22

Chapter II

Figure II. 1. Biological neuron 25
Figure II. 2. Artificial neuron 26
Figure II. 3. Comparison between a biological and an artificial neuron 27
Figure II. 4. Feed forward and recurrent ANN architecture (Pekel & Kara, 2017) 27
Figure II. 5. Structure of a non-looped neural network 28
Figure II. 6. A looped neural network with two inputs (Chamekh, 2007). 30
Figure II. 7. The most used neural networks (Talbi, 2011). 33
Figure II. 8. A visual architecture of multi-layer perceptron (Pekel & Kara, 2017). 34
Figure II. 9. Hyper-tube modeling the regression function 41
Figure II. 10. Use of the kernels for the resolution of the nonlinear regression 42

Chapter III

Figure III. 1. number of publications dealing with the optical properties of blood and its components 46
Figure III. 2. Number of publications dealing with modeling the optical properties of blood using AI methods 55

Chapter IV

Figure IV. 1. Heatmap of correlation matrix for the RI 60
Figure IV. 2. Correlation heatmap for absorption coefficient 61
Figure IV. 3. Correlation heatmap for scattering coefficient 61
Figure IV. 4. Flowchart describing the development process of ANN used in the current work. 62
Figure IV. 5. Screenshot of STATISTICA activation functions 63
Figure IV. 6. Diagram illustrating the sensitivity analysis by weight method 65
Figure IV. 7. Flowchart of SVM 67
Figure IV. 8. Flowchart of DA-SVM process 69

Chapter V

Figure V. 1. Representation of best percentages division of the data bases 72
Figure V. 2. Comparison of experimental and ANN calculated values of the refractive index: (a) total phase, (b) training phase, (c) testing phase, (d) validation phase 75
Figure V. 3. Comparison of experimental and ANN calculated values of the absorption coefficient µa: (a) training phase, (b) testing phase, (c) validation phase, (d) total phase 78
Figure V. 4. Comparison of experimental and ANN calculated values of the scattering coefficient µs: (a) training phase, (b) testing phase, (c) validation phase, (d) total phase 81
Figure V. 5. Comparison between experimental and predicted values of RI vs wavelength 82
Figure V. 6. Comparison between experimental and predicted values of RI vs concentration of Hb 82
Figure V. 7. Comparison between experimental and predicted values of RI vs T 83
Figure V. 8. Performance of MLP-ANN models for: (a) training phase, (b) testing phase, (c) validation phase, and (d) total phase 86
Figure V. 9. Schematic representation of optimized ANN models 87
Figure V. 10. Relative importance histogram for RI 89
Figure V. 11. Relative importance for µa 91
Figure V. 12. Relative importance for µs 93
Figure V. 13. Correlation plot between the experimental and SVM predicted values of the refractive index: (a) training phase, (b) testing phase, and (c) the total phase 95
Figure V. 14. Correlation plot between the experimental and SVM predicted values of absorption coefficient: (a) training phase, (b) testing phase, and (c) the total phase 98
Figure V. 15. Correlation plot between the experimental and SVM predicted values of scattering coefficient: (a) training phase, (b) testing phase, and (c) the total phase 100
Figure V. 16. Comparison between µs (exp) and µs calculated by SVM vs wavelength 101
Figure V. 17. Comparison between µs (exp) and µs calculated by SVM vs hematocrit 101
Figure V. 19. Comparison between µa (exp) and µa calculated by SVM vs wavelength 102
Figure V. 20. Comparison between µa (exp) and µa calculated by SVM vs hematocrit 103
Figure V. 21. Comparison between µa (exp) and µa calculated by SVM vs O2 sat 103
Figure V. 22. Scatter plots between experimental and DA-SVM predicted values of the refractive index: (a) training phase, (b) testing phase, and (c) the total phase 106
Figure V. 23. Scatter plots between experimental and DASVM predicted values of the absorption coefficient: (a) training phase, (b) testing phase, and (c) the total phase 109

List of figures

Figure V. 24. Scatter plots between experimental and DASVM predicted values of the scattering coefficient: (a) training phase, (b) testing phase, and (c) the total phase 112
Figure V. 25. Residual plots for: (a) training dataset, and (b) testing dataset for μa 113
Figure V. 26. Residual plots for: (a) training dataset, and (b) testing dataset for μs 114
Figure V. 27. William's diagram for detection outliers for DA-SVM model (for μa). 114
Figure V. 28. William's diagram for detection outliers for DA-SVM model (for μs) 115
Figure V. 29. comparison between the three modes developed for the refractive index 117
Figure V. 30. comparison between the three modes developed for the absorption coefficient 118
Figure V. 31. comparison between the three modes developed for the scattering coefficient 120

List of tables

Chapter I

Table I. 1. Human blood cells properties 8

Chapter II

Table II. 1. Usual transfer functions, a=f (n) (Si-Moussa,2009). 31
Table II. 2. The advantages and disadvantages of ANNs 39
Table II. 3. Examples of kernels 42

Chapter IV

Table IV. 1. the software used in this study 57
Table IV. 2. Statistical parameters of the refractive index were used in building the models 58
Table IV. 3. Statistical description of the Absorption data Adjusted in this work 58
Table IV. 4. Statistical description of the Scattering Data used in building the models 59
Table IV. 5. Ranges of initial settings for SVM 68

Chapter V

Table V. 1. Structure of optimized MLP-ANN models 72
Table V. 2. Linear regression vectors from the linear 83
Table V. 3. Weights and bias of the best MLP-ANN model for RI (4-20-1). 88
Table V. 4. Weights and bias of the best MLP-ANN model for µa (3-23-1). 90
Table V. 5. Weights and bias of the best MLP-ANN model for µs (3-25-1). 92
Table V. 6. Comparison with the existing models in the literature 121

List of work and research activities

Publications

Omari, F., Khaouane, L., Laidi, M., Ibrir, A., Roubehie Fissa, M., Hentabli, M., & Hanini, S. (2023). Dragonfly algorithm–support vector machine approach for prediction the optical properties of blood. *Computer Methods in Biomechanics and Biomedical Engineering*, *0*(0), 1–10. https://doi.org/10.1080/10255842.2023.2228957

National communications

Omari, F., Tared, S., Khaouane, L., & Laidi, M (2023). Modélisation de Coefficient d'absorption du sang par Réseau de Neurones Artificiel. 1er Séminaire National de Chimie Appliquée. Le 9-10 mai 2023. USD Blida. Algeria

International communications

Omari, F., Tared, S., Khaouane, L., Laidi, M., & Ibrir, A (2022). Artificial intelligence to predict the real part of the refractive index of hemoglobin. 4 th international conference on applied engineering and natural sciences. On 10-13 November in 2022 at Konya. Turkey

General introduction

General introduction

Blood is a remarkable and vital component of the human body, often referred to as the "river of life" due to its fundamental role in sustaining health and well-being. This complex biological fluid circulates throughout our bodies, transports essential nutrients, oxygen, and hormones while also removing waste products (Roggan et al., 1999).

Blood is composed of several components, including red blood cells that transport oxygen from the lungs to the tissues and remove carbon dioxide; white blood cells play a crucial role in the immune system by fighting infections and destroying pathogens; platelets are necessary to enable blood clotting and stop bleeding, and plasma, which carries blood cells (Roggan et al., 1999).

Beyond its role in oxygenating tissues and maintaining bodily functions, blood serves as an invaluable diagnostic tool in the realm of medicine. One of the interesting aspects of blood that has garnered significant attention in medical diagnostics is its optical properties. These properties encompass a range of characteristics, including the refractive index, absorption coefficient, and scattering coefficient, which collectively influence how blood interacts with light. The study and understanding of these optical features have emerged as a crucial facet of modern medical diagnostics (Jedrzejewska-Szczerska, 2013).

The significance of blood's optical properties in medical diagnostics cannot be overstated. These properties play a pivotal role in enhancing the ability to detect, diagnose, and monitor various diseases and health conditions. Conditions such as anemia, characterized by a deficiency in red blood cells or hemoglobin, can be readily identified through the assessment of the optical characteristics of blood. Similarly, in the realm of oncology, the optical properties of blood have proven invaluable in the early identification and monitoring of cancerous growths (Soh et al., 2020; Yıldırım et al., 2018).

Beyond anemia and cancer, the optical properties of blood hold promise for identifying a wide array of medical conditions, ranging from cardiovascular diseases to infectious diseases. By harnessing the interaction between light and blood, researchers and healthcare professionals can gain valuable insights into a patient's health status, providing a non-invasive and effective method of diagnosis and monitoring. However, despite the significant role these properties play in healthcare,

there exists a glaring gap in the literature. The scarcity of available data stems from the inherent complexity of blood, making accurate measurements extremely challenging (Jacques, 2013).

In light of this scarcity, there emerges a compelling necessity to pioneer innovative methodologies. Such techniques must implement advantages of the remarkable potential of artificial intelligence approaches, which have consistently demonstrated their capacity for prediction and modeling across diverse fields. Artificial intelligence has transcended its role as a mere technological achievement and evolved into an indispensable tool in multiple sectors, ushering in transformative changes in healthcare, finance, engineering, and beyond (Yahya et al., 2016).

Artificial intelligence, particularly through techniques like Artificial Neural Network (ANN), Support Vector Machine (SVM), and an array of cutting-edge algorithms, has risen to prominence as an unrivaled force in data analysis and prediction. ANN, SVM, and related algorithms have become the linchpin of innovation, enabling us to unlock the enigmatic world of blood's optical properties with superior precision and effectiveness (Soh et al., 2020; Van Dao et al., 2019).

The choice to employ Artificial Neural Networks (ANN), Support Vector Machines (SVM), and innovative algorithms such as the Dragonfly algorithm in our study stems from the principle advantages these approaches offer over conventional methods. These advantages have made them the preferred choice in our quest to unravel the intricacies of optical properties of blood. ANN is renowned for their exceptional ability to recognize complex patterns within large datasets, perform exceptionally well in capturing nonlinear relationships and adjusting to challenging, real-world situations. SVM, on the other hand, are very effective at classifying data points into distinct categories through the creation of optimal decision boundaries. The Dragonfly algorithm, known for its swarm intelligence-inspired optimization, is adept at exploring complex solution spaces efficiently. Their strength in generalization, adaptability, and resilience to noise makes them invaluable for predictive modeling, particularly in the context of medical research where data can be limited and noisy. Furthermore, these approaches have a proven track record in diverse fields and are well-suited to unraveling the complexities of blood's optical properties. We seek to develop medical diagnostics and contribute to the larger landscape of healthcare innovation by utilizing their strength.

The methodology employed in this thesis consists of five chapters, each tailored to systematically explore and contribute to our understanding of blood's optical properties and their predictive modeling using artificial intelligence techniques:

- Chapter 1 serves as the foundation, elucidating the comprehensive description and definition of blood and its constituent components. It delves into the intricate optical properties of blood, elucidating the methods employed to measure these properties and elucidating their critical role in the realm of medical diagnostics.
- Chapter 2 pivots towards the contemporary landscape of artificial intelligence modeling techniques, providing an in-depth examination of three important methodologies: Artificial Neural Networks (ANN), Support Vector Machines (SVM), and the Dragonfly Algorithm (DA). These modeling approaches are presented in detail to lay the groundwork for their application in subsequent chapters.
- Chapter 3 embarks on a critical review of existing studies pertaining to the optical properties of blood found in the literature.
- The practical application of our study is transitioned into in Chapter 4, which also outlines the materials used and clarifies the various modeling techniques applied. This includes data collection, pre-treatment, and data analysis, as well as the meticulous steps involved in model construction. The "weight method," which is a crucial introduction to the sensitivity analysis.
- Chapter 5 culminates in presenting the results obtained from our comprehensive research endeavor. It includes an insightful comparison between the performance of the models we've developed, shedding light on their efficacy in predicting blood's optical properties. Additionally, we undertake a rigorous comparison between our models and those existing in the literature, discerning the advancements and contributions our work brings to the field.

Chapter I: Blood and optical properties

I. Chapter I: blood and its optical properties

I.1. Introduction

The optical properties of blood play a crucial role in medical diagnosis and treatment, as well as in the design of new medical devices. Hemoglobin is a vital constituent of blood whose optical properties affect all optical properties of human blood. It has been reported that the optical properties of hemoglobin strongly depend on physiology and biochemical parameters such as its concentration, temperature, etc. (Yahya et al., 2016).

In this chapter we present the different theoretical notions concerning blood, then we talk about the main biological compositions of human blood, their characteristics and functions of the figured elements. Then we give generalities on the optical properties of blood and their measuring devices.

I.2. History of blood

Since the highest antiquity, man has considered blood as the very symbol of life "The river of life". It was attributed the possibility of reviving an organism exhausted by hemorrhage, of restoring, if this blood was "young and fresh", strength and vigor to the old man, or even common sense to the madman (Jacques, 2013). And even today, when hematology "the science of blood" provides objective knowledge, this ambiguity remains. It is in the liver that Galen in the 2nd century located the formation and regulation of blood. For 1500 years, no one dares to refute this idea. When William Harvey, an English doctor, asserted in 1623 that the blood circulates in a closed circuit, driven from the heart to the organs, then from the organs to the heart, he triggered a veritable revolution. However, he is right. After having been the subject of many anatomical drawings, the cutaways, the circulation of the blood is visualized today by all sorts of methods, such as angiography, and is repaired by drugs and surgery. The first descriptions of the composition of blood date back to the 17th century, with Leuwenhoek, the inventor of the microscope, Malpighi and Swammerdam.

Blood is both the mirror in which the good or bad functioning of all the organs is reflected, and the proof of the uniqueness of each human being. Each of our cells bears the mark of our individuality, thanks to a group of molecules called markers. First recognized on the membrane of red blood cells, these markers are now identifiable at the gene level. Each individual has a unique biological formula, an unforgeable identity card. These developments are well known for their applications,

among others, in the field of forensic science to confirm the identification of a criminal or a victim. Blood is also a raw material, which can be given, exchanged or sold (Jacques, 2013).

I.3. The blood

Blood is a living tissue, which supplies the cells of the human body with everything they need to sustain life: oxygen, nutrients, electrolytes, hormones, vitamins and substances necessary for protection against disease. It also rids the body of heat and its various waste products as it also serves to bring the cells and molecules of the immune system to the tissues, and to diffuse hormones throughout the body. It makes up about 8% of body weight. Blood volume in an adult male is 5–6 liters and 4–5 liters in females.

Generally, we imagine that the blood is red because it contains red blood cells. But on the contrary, it comes down to asking why red blood cells are red. Red blood cells are made up of hemoglobin, which is a red pigment when subjected to white light, which only sends back photons whose wavelength is red. Hemoglobin is made up of iron and oxygen: iron combined with oxygen creates oxyhemoglobin, which gives blood the red color**.**

The blood regulates the PH by means of the buffer system; it also regulates body temperature through the heat-absorbing and cooling properties of its aqueous content. In addition, excess body heat is transported by the blood to the skin from where it is eliminated in the environment (Zijlstra et al., 1991).

I.4. Blood components

Human blood is composed of approximately 45% blood cells (Figure I.1), that are suspended in a fluid called plasma. These blood cells, including red blood cells (erythrocytes), white blood cells (leukocytes), and platelets, all originate from hematopoietic stem cells found in the bone marrow. As the stem cells mature, they undergo differentiation, leading to the formation of distinct types of blood cells. Red blood cells are responsible for carrying oxygen to tissues and appear as disc-shaped cells. White blood cells play crucial roles in the immune system's defense mechanisms, combating infections and foreign substances. Platelets, on the other hand, aid in blood clotting by aggregating and forming a dense pellet at the bottom of a test tube. Together, these components contribute to the intricate balance and functionality of human blood (Wray et al., 1988).

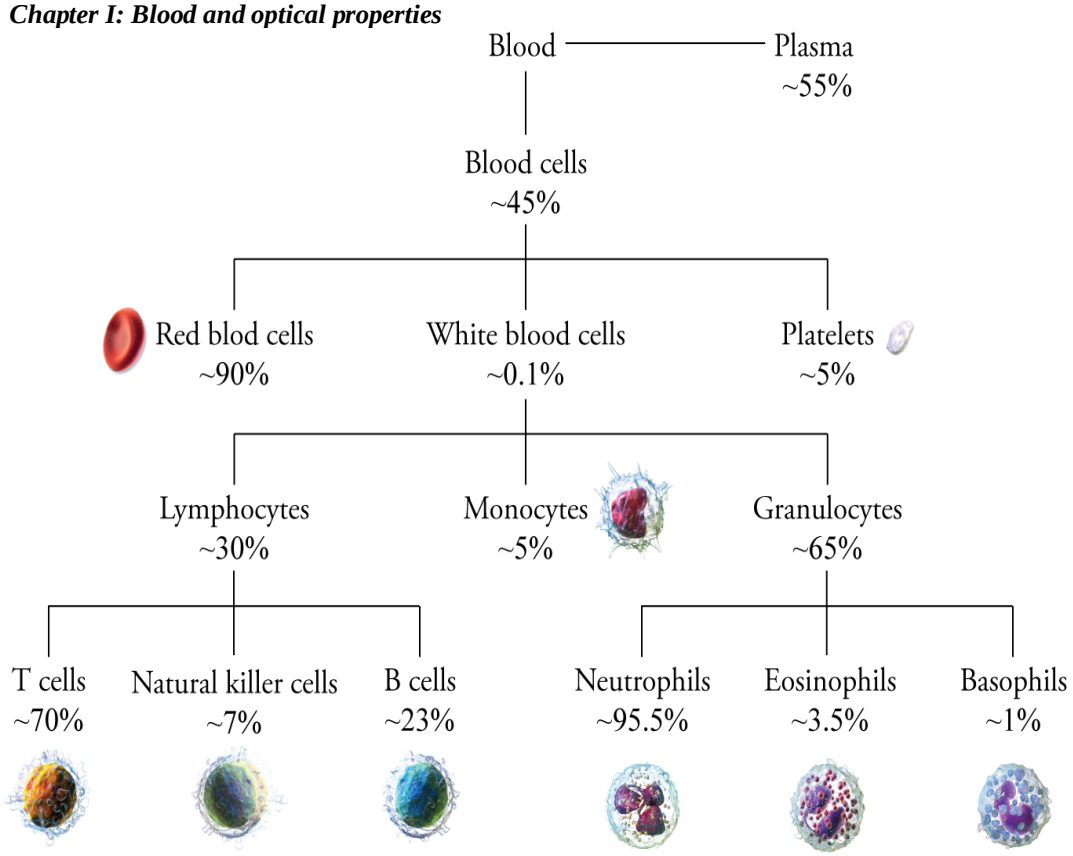


Figure I. 1. An overview of the composition of human blood and the further subdivision of human blood cells. (https://microbenotes.com/blood-components-formation-functions/)

Through the process of centrifugation, a blood sample can be separated into: red blood cells, the buffy coat, and plasma. The buffy coat, which lies between the red blood cells and plasma layers, is composed of platelets and white blood cells (Figure I.2).

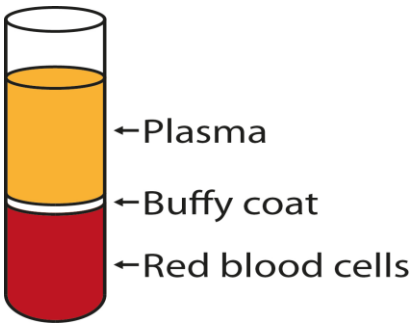


Figure I. 2. Composition of blood system (Version & Engineering, 2021).

I.4.1. Red blood cells

The red blood cells, also known as erythrocytes, are the most abundant type of cell in the blood, typically comprising 4 to 6 million units per mm3. They play a crucial role in maintaining blood viscosity. Erythrocytes are small biconcave discs with a diameter of 6 to 8 μm. Their unique shape, thinner in the center than at the ends, provides a large surface area to volume ratio, enabling efficient gas exchange. These cells lack a nucleus and have a lifespan of approximately 100 to 120 days. When they reach the end of their lifespan, erythrocytes fragment and are eliminated from the bloodstream by macrophages. Erythrocytes are produced in the red bone marrow, primarily in specific locations such as the ribs, vertebrae, sternum, and pelvis. The stem cells known as hemocytoblasts, capable of giving rise to any type of blood cell, are responsible for their production. In situations where tissues do not receive sufficient oxygen, the kidney secretes a hormone called erythropoietin, which stimulates erythrocyte production in the bone marrow. Conversely, excessive oxygen supply reduces erythropoietin secretion and slows down erythrocyte production. The remarkable characteristics of erythrocytes, including their biconcave shape, lack of nucleus, and deformability, enable them to effectively transport oxygen from the lungs to cells and facilitate the removal of carbon dioxide, ensuring proper gas exchange within the body's vascular system (Version & Engineering, 2021).



Figure I. 3. Red blood cells. (Version & Engineering, 2021).

I.4.2. White blood cells

White blood cells, also known as leukocytes, are a diverse group of cells that form an essential part of the immune system and are responsible for protecting the body against foreign substances and pathogens. Unlike red blood cells, the number of white blood cells in the blood is significantly lower, accounting for approximately 0.1% of all blood cells. They can be categorized into subclasses: monocytes, granulocytes, and lymphocytes, each with distinct tasks within the immune response. White blood cells possess nuclei and organelles, distinguishing them from red blood cells. They play a crucial role in combating diseases caused by viruses, bacteria, parasites, and tumor cells. When the body mobilizes leukocytes, their production is accelerated, and their numbers can double within a few hours. The number of white blood cells typically ranges from 4 to 11 thousand per mm^3 of blood. The ability of leukocytes to respond rapidly to immune challenges and increase in number, when necessary, contributes to the body's defense against various pathogens and infections. The characteristics and functions of blood are shown in (table I.1). system (Version & Engineering, 2021).

Table I. 1. Human blood cells properties

Cell	**Diameter (**μm)	**Function**	**Density (**g/cm3)
Erythrocyte	6 - 8	O_2 transportation	1.090 - 1.110
Platelets	2 – 4	Coagulation	1.054 - 1.062
Monocyte	15 - 14	Immune surveillance, macrophage precursor	1.055 – 1.062
Neutrophil	10 – 14	Defense of the body	1.080 – 1.084
Eosinophil	10 – 14	Defense against parasites and allergies	1.080 – 1.084
Basophil	10 – 14	Inflammation and allergy	1.080 – 1.084
T-lymphocyte	6 - 15	Immunity cellular	1.060 -1.072
B-lymphocyte	6 - 15	Antibody, precursor to plasma cells	1.060 – 1.072

I.4.2.1. Monocytes

Monocytes are produced in the bone marrow and circulate briefly in the bloodstream before migrating into various tissues. In the tissues, they differentiate into macrophages and dendritic cells. Macrophages engulf and remove dead cells and debris and also present antigens to T cells. Dendritic cells, on the other hand, activate T and B cells upon encountering pathogens. These cells play important roles in the immune response by clearing pathogens, activating immune cells, and initiating specific immune responses.

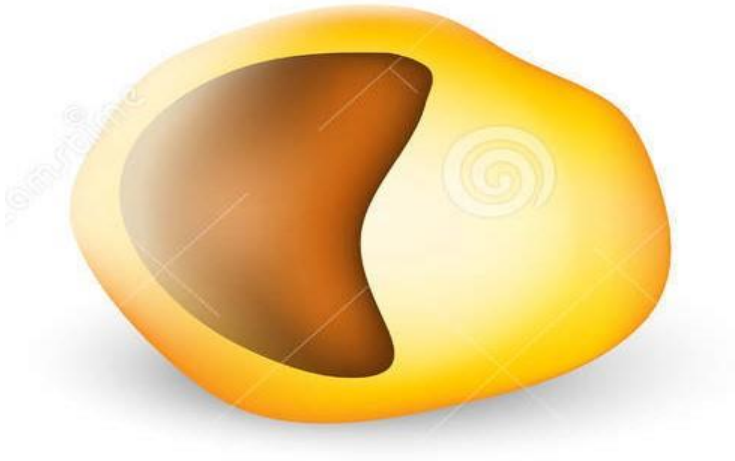


Figure I. 4. Monocyte cell. (http://www.clker.com/clipart-monocyte.html)

I.4.2.2. Granulocytes

Granulocytes are characterized by their distinctive nucleus shape and granules in the cytoplasm. They play a vital role in the immune response by producing substances that are toxic to microorganisms, helping to protect against infections. There are three types of granulocytes: neutrophils, eosinophils, and basophils. Neutrophils are involved in defending against bacterial and fungal infections, eosinophils primarily combat parasitic infections and participate in allergic reactions, and basophils release heparin to prevent blood clotting in areas of foreign invasion. Collectively, these granulocytes contribute to the body's defense against pathogens and participate in various immune responses.

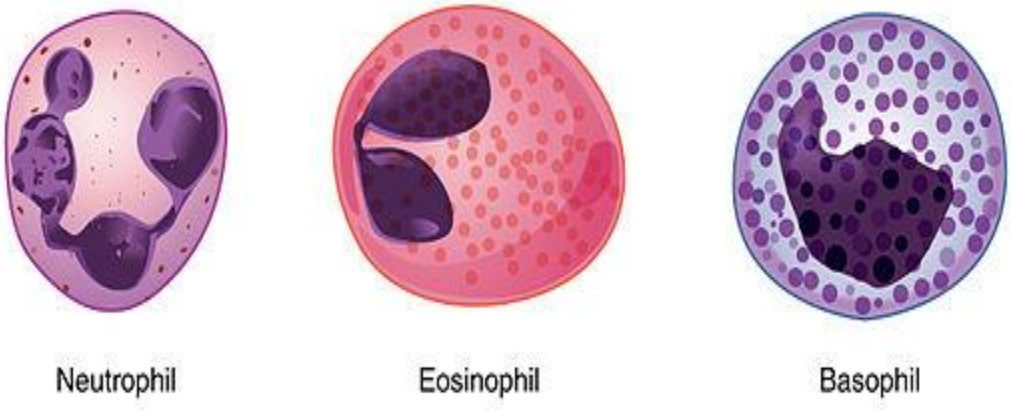


Figure I. 5. Type of granulocytes. (https://www.microscopemaster.com/granulocytes.html)

I.4.2.3. Lymphocytes

Lymphocytes play a vital role in the defense of the immune system against pathogens and foreign substances. They are primarily found in lymphoid tissues, including lymph nodes, the spleen, and the thymus. Lymphocytes undergo maturation and activation processes to develop specific receptors for antigens, allowing them to recognize and respond to diverse threats. The intricate interplay between different subsets of lymphocytes and their ability to mount both immediate and long-term immune responses make them indispensable in maintaining a strong and effective immune system. There are three main types of lymphocytes: B cells, T cells and NK cells (natural killer). system (Version & Engineering, 2021).

- ✓ **B cells**

Each of these cells has a special receptor on its surface that allows it to recognize a particular antigen. The B cells divide and develop into plasma cells after the identification stage ends, which release the particular receptor as an antibody (glycoprotein), a soluble version of the receptor. They are almost identical to lymphocyte membrane receptors and bind to the antigen that initially activated the B-cell.

- ✓ **T cells**

The activation of T lymphocytes by antigens induces a series of changes that eventually result in their differentiation. There are also several classes of mature T lymphocytes. Some assist the B lymphocyte for antigen recognition or interact with monocytes and help destroy pathogens: these are the helper T lymphocyte (TH for help). Others are specialized in the destruction of

parasite- or virus-infected cells. This activity is called cytotoxic, and these cytotoxic T cells (TC). These groups of T cells only recognize the antigen when they are associated with specific host cell markers, via reuptakes.

- ✓ **NK cells**

Natural Killer (NK) cells are part of the innate immune response and provide rapid responses to eliminate abnormal cells. NK cells are characterized by their ability to recognize and directly kill target cells without prior activation or specific antigen recognition. They can detect cells that display abnormal markers, such as infected or cancerous cells, and release cytotoxic substances to induce cell death. NK cells also contribute to the regulation of immune responses by producing cytokines and interacting with other immune cells. Their unique properties and functions make NK cells an integral component of the immune system's surveillance and defense mechanisms.

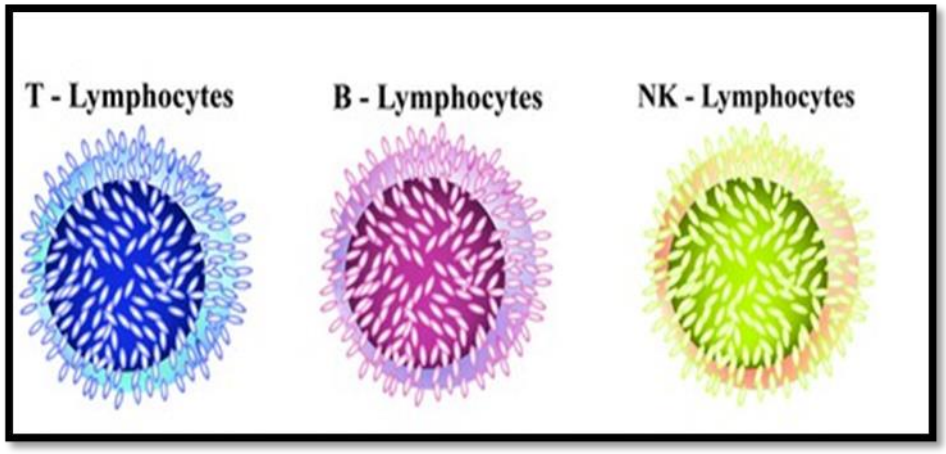


Figure I. 6. Type of Lymphocytes. (https://www.shutterstock.com/search/lymphocyte)

I.4.3. Platelets

Platelets, also known as thrombocytes, are small, nuclear-free, discoid-shaped cell fragments that play a crucial role in blood clotting. Their primary function is to form blood clots at the site of a vascular injury, preventing excessive blood loss. The surface of a platelet is covered with proteins that are essential for adhesion and aggregation. When a blood vessel is damaged, platelets are activated and change their shape, adhering to each other and to the injured area, creating a platelet plug. In addition, activated platelets can cross-link with the plasma protein fibrinogen, an important step in the blood clotting process. This cross-linking, along with the involvement of other clotting factors, initiates the coagulation cascade, ultimately leading to the formation of a stable blood clot,

sealing the break in the blood vessel. Platelets play a critical role in maintaining hemostasis, ensuring the balance between clot formation and dissolution, and contributing to the body's ability to repair damaged blood vessels.

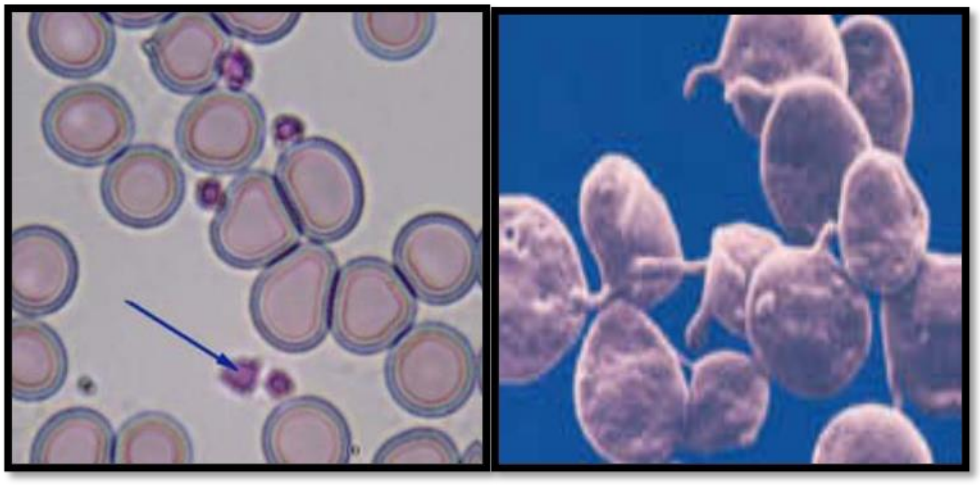


Figure I. 7. Blood platelets. (https://www.cedars-sinai.org/newsroom/research-reveals-blood-platelets-play-important-role-in-kawasaki-disease/)

I.4.4. Plasma

Plasma is the vital straw-colored liquid that serves as the medium for transporting all blood cells throughout the body. Comprising water and proteins, it plays a crucial role in maintaining the overall blood volume and osmotic pressure. The liver and lymphocytes produce the most common proteins found in plasma, including albumins, immunoglobulins, and fibrinogen. These proteins have diverse functions, such as regulating osmotic pressure, facilitating the transportation of non-water-soluble molecules like hormones, vitamins, and fatty acids, and contributing to the immune system's defense mechanisms through the production of antibodies. Additionally, plasma contains electrolytes, nutrients, waste products, RNA, DNA, and dissolved gases. It appears as a transparent, light-yellow liquid and constitutes approximately 55% of the total blood volume. With its high-water content (around 90%) and protein composition (approximately 7%), plasma also contains mineral salts, ions, carbohydrates, nucleotides, vitamins, hormones, lipids, and triglycerides. Without plasma, the solid blood cells would be unable to circulate effectively throughout the body. Furthermore, when the clotting factors, such as fibrinogen, are removed from plasma, the resulting liquid is known as serum (Andersson-engels, 2016; Version & Engineering, 2021).

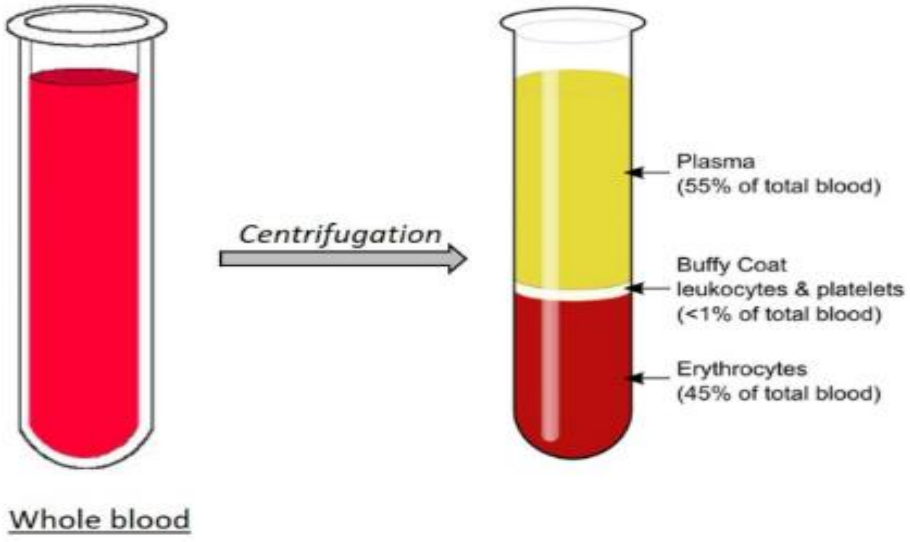


Figure I. 8. Blood plasma. (Version & Engineering, 2021)**.**

I.4.4.1. Criteria for choosing between serum and plasma in the medical analysis laboratory

Serum and plasma are blood products of different composition, that of serum having been modified by coagulation.

- Some compounds are consumed during coagulation (examples: glucose, fibrinogen). Their concentration is therefore lowered in the serum.
- Certain other compounds are produced or released into the serum during coagulation (examples: potassium, lactate, phosphates). Their concentration is therefore increased in the serum.
- To obtain plasma, the collection tube must contain an anticoagulant. Plasma is easier and faster to obtain. It is theoretically more representative of the physiological reality of non-coagulated blood.

I.4.4.2 Plasma proteins

With the exception of water, proteins make up the majority of blood plasma. Indeed, the latter contains more than a hundred proteins which can be classified into three main categories: albumin, globulins and fibrinogens.

- ✓ **Albumin**

Albumin, comprising approximately 60% of plasma proteins, plays a critical role in the body. Besides its function in transporting nutrients, albumin acts as a vital "water carrier." Its primary mission is to prevent excessive water loss from the blood as it circulates through narrow, water-permeable vessels and to maintain proper blood viscosity. In cases where an insufficient diet leads to

a deficiency of albumin, water can escape from the blood and accumulate in the tissues, resulting in a condition known as malnutrition edema. This condition arises due to the compromised ability of albumin to regulate fluid balance, leading to fluid retention in the tissues. By ensuring adequate levels of albumin through proper nutrition, the risk of malnutrition edema can be mitigated, allowing for the maintenance of healthy fluid distribution throughout the body (*Diagnostic Strategies for Blood Borne Infections in Sweden*, n.d.; Properties & City, n.d.).

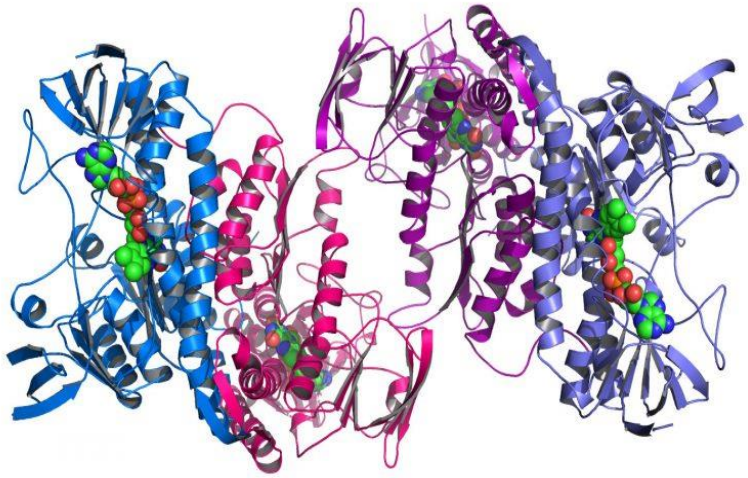


Figure I. 9. Structure of albumin. (https://www.researchgate.net/publication/296478720_Albumin-based_drug_delivery_Harnessing_nature_to_cure_disease)

- ✓ **Immunoglobulins**

Immunoglobulins, also known as antibodies, are vital components of the immune system and work in conjunction with leukocytes to play a crucial role in the body's specific defense against pathogens. They bind to foreign substances, such as bacteria and viruses, marking them for destruction and facilitating their removal from the body. Alongside immunoglobulins, the defense mechanism is further strengthened by the presence of over 20 proteins that constitute the complement system. While the complement system aids in the elimination of pathogens, excessive or inappropriate activation can lead to significant tissue damage. Therefore, the precise regulation of the complement system is crucial to ensure an effective immune response without causing harm to healthy tissues. The coordinated efforts of immunoglobulins, leukocytes, and the complement system provide a powerful defense mechanism against invading pathogens, safeguarding the body from infections and diseases.



Figure I. 10. General structure of an immunoglobulin. (https://images.app.goo.gl/rNQMX1XcS1PiMtAm9)

✓ **Fibrinogens**

Fibrinogen is a blood protein that plays a role in clotting. It participates in the formation of blood clots and also modulates the activity of blood platelets and vessel cells. Under the action of another protein, thrombin, it is transformed into fibrin. It is synthesized by liver cells. Its level in the blood normally varies from 2 to 4 g/L. However, the synthesis of this protein can increase following stress, during pregnancy or after the injection of certain drugs or growth hormone. An increase in the level of fibrinogen in the blood is also indicative of an inflammatory state.

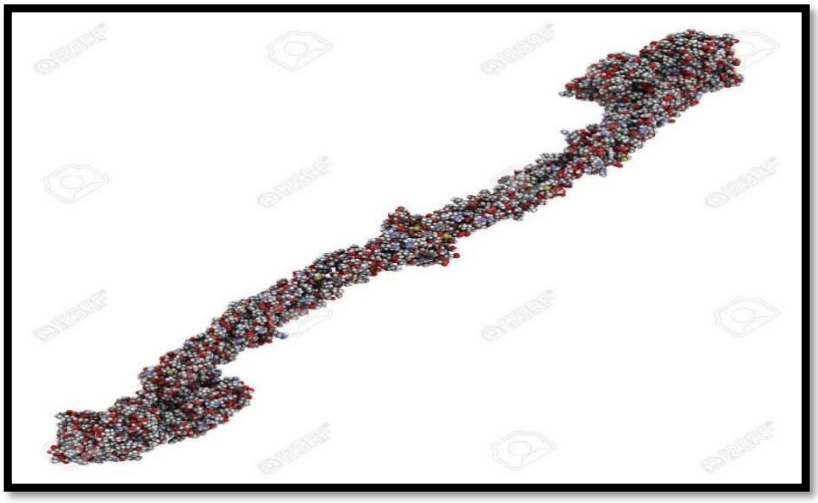


Figure I. 11. The structure of fibrinogen. (https://images.app.goo.gl/JbSbawN8TBXzUjAs7)

I.5. Blood function

I.5.1. Transportation

Blood serves a vital role in the body by facilitating the transportation of oxygen from the lungs to the cells and carrying carbon dioxide from the cells back to the lungs. It also delivers cells with nutrients from the digestive tract and hormones secreted by the endocrine glands. Furthermore, blood aids in the removal of waste products generated by cells, helping to maintain cellular health and overall bodily function (Petersson, 2018; Tuchin, n.d.).

I.5.2. The regulation

Blood performs several regulatory functions within the body, it regulates the pH through buffer systems, ensuring the stability of the internal environment, it also regulates body temperature through the heat-absorbing and cooling properties of its aqueous content. Excess body heat is efficiently transported by the blood to the skin, where it can be dissipated into the surrounding environment, helping to prevent overheating. Moreover, blood's osmotic pressure influences cellular water content, primarily through the presence of ions and proteins dissolved within it. This delicate balance ensures proper hydration and cellular function.

I.5.3. Protection

Blood serves as a protective shield for the body through various mechanisms. One such mechanism is coagulation, it also contains phagocytic white blood cells that engulf and destroy pathogens, along with specialized plasma proteins including antibodies, interferon, and complement. These components work together to defend the body against toxins and foreign microbes, bolstering the immune response and maintaining overall health (Petersson, 2018; Tuchin, n.d.).

I.6. The physical properties of blood

- **Viscosity:** Blood is a more viscous, thicker and denser liquid than water, it circulates more slowly than water and this is partly due to its viscosity.
- **Temperature:** it is around 38°C, slightly above normal body temperature.
- **PH:** The blood is slightly alkaline, its pH being between 7.35 and 7.45.
- **Resistivity:** the resistivity of a biologically balanced human is between 175 and 210 Ω.cm, and between 210 and 240 Ω.cm for sportsmen (very fluid blood). It decreases with age and accelerated aging phenomena such as heavy metal poisoning. There is a decrease in conductivity with increasing hematocrit.

- **Permittivity of blood:** At low frequency, the values of permittivity decrease with the drop in hematocrit. With increasing frequency, this drop becomes less and less significant. From this point, the opposite phenomenon is observed: a decrease in hematocrit induces an increase in permittivity (Petersson, 2018).

I.7. Therapeutic use of blood

Blood has many therapeutic properties and can be used in several forms: on the one hand labile blood products, which include leukodepleted red blood cell concentrate, leukodepleted platelet concentrate and plasma, and on the other hand blood products stable, prepared by plasma fractionation.

Transfusions of leukodepleted red blood cells are the most frequent transfusions. They are performed to correct a deficit of red blood cells responsible for tissue hypoxia or anoxia. They make it possible to treat anemia, whatever its origin, medical, surgical or obstetrical.

Leukodepleted platelet concentrate is mainly used to treat cases of thrombocytopenia below 20,000 platelets/mm ^{3}. These can occur, for example, after chemotherapy treatments (D. Gupta et al., 2007).

I.8. The optical properties of blood

The optical properties of blood play an important role for many diagnostic and therapeutic applications in laser medicine and routine medical diagnosis. There are three optical cleanliness of blood which are: refractive index, absorbance μ_a and diffusion μ_s (M. Meinke et al., 2007b; D. K. Sardar & Levy, 1998).

I.8.1. Refractive index

The refractive index of a medium material is an important optical parameter because it shows the optical properties of the material. Its values are often needed to interpret different types of spectroscopy data. Refractive index coefficients are important parameters in the design of a semiconductor laser. The problem of tampering is increasing day by day and therefore a simple automatic and accurate measurement of the refractive index of the material is of great importance these days. Rapid dice refraction measurements using simple techniques and refractometers can help control adulteration of common liquids to a greater extent (M. Yahya & Saghir, 2016).

I.8.1.1. Definition

We call refraction the change of direction that a light ray undergoes in passing from one given optical medium to another. This change is due to a change in the speed of propagation from

the point, called the point of incidence, where the incident light ray strikes the interface (Howard, 1920)

The refractive index is a quantity defined for any transparent material medium as the ratio of the speed of light in vacuum to the speed of light in this medium interface (Howard, 1920). The complex refractive index $n = n_{Rel} + in_{im}$ consists of a real part n_{Rel}, which represents the ratio of the speed of light in vacuum to the speed of light in material $n_{Rel} = c / v$, and an imaginary part n_{im}, which represents an absorption parameter characterizing the material, The imaginary part of the refractive index can be understood using the Bouguer-Lambert law (M. Yahya & Saghir, 2016)**.**

I.8.1.2. Measurement of a refractive index with the Abbe refractometer

The term refractometer is mainly used to name devices that allow the determination of the refractive index of a liquid, although there are also instruments that allow the determination of the refractive index of a solid. We will restrict ourselves here to the determination of the refractive index of liquids.

I.8.1.3. Construction of the refractometer

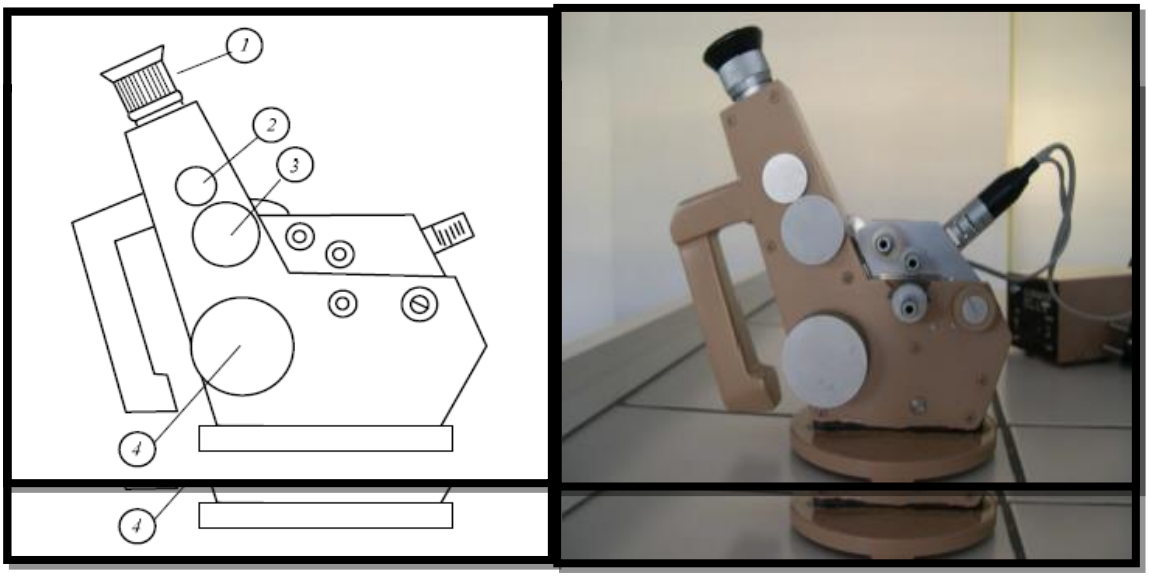


Figure I. 12. The compositions of the refractometer. (http://webphysique.fr/indice-de-refraction)

1. **Eyepiece:** Allows you to see the visual field which includes both the reticle and the scale. Equipped with a knurled screw for focusing;

2. **Cap for adjustment screw of the graduated scale:** divide the cap: then appears the adjustment screw. Adjustment is easily carried out by turning this screw using the accessory provided for this purpose;

3. **"Dispersion" button:** Achromatize to obtain a clear demarcation line between the bright and dark range of the visual field;

4. **Measurement knob:** turn the knob to make the dividing line between the light area and the dark area coincide with the crosshairs of the reticle in the visual field. This button is also connected to the graduated scale which is in the same visual field as the demarcation line (M. Yahya & Saghir, 2016)**.**

I.8.1.4. Principle of operation

The term refractometer is used to name devices which make it possible to determine the refractive index of a liquid, more rarely of a solid. Refraction is the change of direction that a light ray undergoes while passing from one optical medium to another. This change is due to a change in the propagation speed of light from the point of incidence where the light ray strikes the interface. The refractive index varies with the wavelength of light, temperature, and the concentration of constituents in the solution (Rawle, 2016).

The device consists of two prisms: the upper prism allows lighting; the lower prism allows measurement. A mirror returns the light beam through an eyepiece with a reticle graduated in units of the product to be dosed. The scale is made in the factory and cannot be modified. The reading is done at the intersection of the graduation and the dividing line between a light zone and a dark zone (Rawle, 2016).

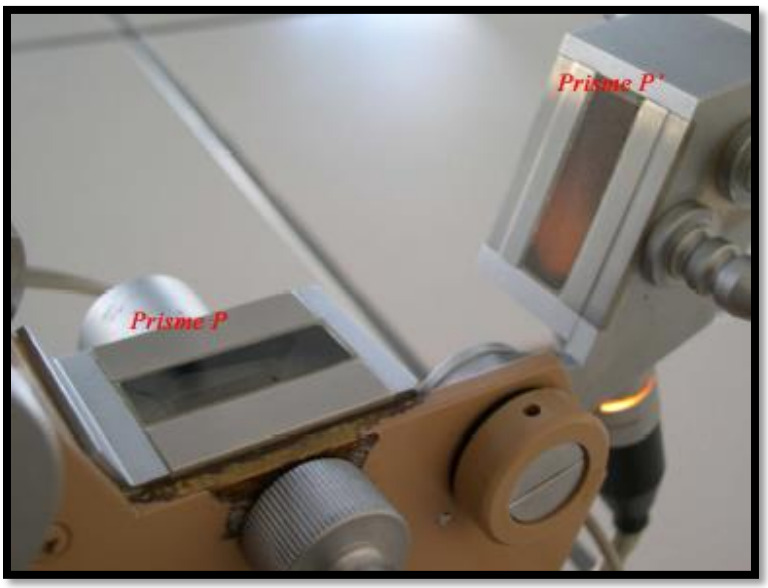
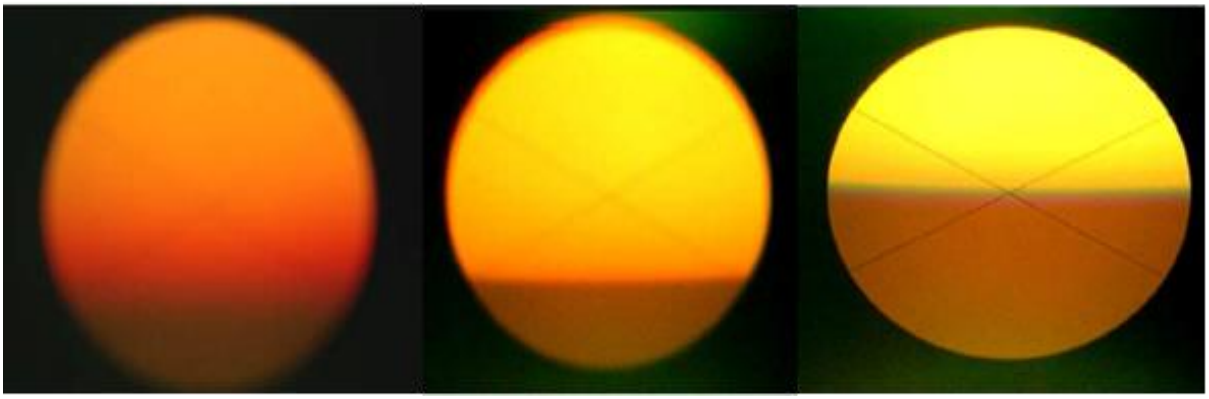


Figure I. 13. The refractometer prisms. (Rawle, 2016).

I.8.1.5. The use of refractometer

- ✓ Calibration of the instrument using distilled water.
- ✓ Clean prisms with acetone lab wipes after each calibration.
- ✓ Use a dropper to deposit a few drops of the liquid on the prism p and lower the prism p'.
- ✓ Turn knob 4 to be oriented near the limit angle; the demarcation line between the sober and light zones then appears in the eyepiece (see figure I.14. a).
- ✓ Turn knob 3 to remove the iridescence due to the use of polychromatic incident light: the demarcation line is then clear (see figure I.14. b).
- ✓ Use button 4 again to bring this line to the center of the target (see figure I.14. c)



(a) (b) (c)

Figure I. 14. Sight adjustment. (Singh, 2002)

- ✓ Reading of the refractive index: the reading window gives the value directly.

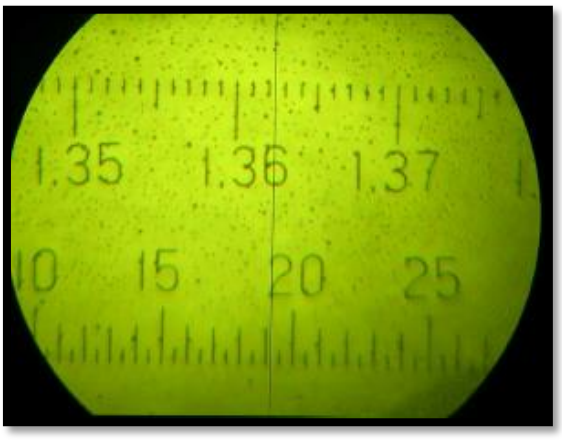


Figure I. 15. Reading the refractive index (Singh, 2002)

I.8.2. Absorption and diffusion

I.8.2.1. Definition of Absorption

For a body put in contact with a fluid or an energy, allows to pre-wash part of it. The absorbed body is included in the absorbed one, whereas in the case of absorption it is simply picked up on its surface, in physics absorption is the process by which electromagnetic radiation is captured by matter.

In biology, absorption also consists of a process corresponding to the crossing by a substance (water, nutrient, pollutant, etc.) of the biological barrier constituted by the plasma membrane of living cells. In summary, an absorption is the passage of a substance from the external environment to a living cell (García & Báez, 2012).

I.8.2.2. Absorption measurement with spectrometry

a) Definition The spectrophotometer

The spectrometer is a method of quantitative analysis which consists in measuring the absorbance or optical density of a given chemical substance in solution, the more this species is concentrated the more it absorbs light within the limits in proportionality stated by Beer - Lambert 's law.

$$A = \varepsilon\, b\, C \qquad (1)$$

A spectrophotometer measures the absorbance of a solution at a given wavelength, the spectrophotometer consists of:

- Visible light source
- A monochromator for generating visible light from a source. It is made up of a network that scatters white light.
- A tank containing the studied solution, the length of the tank is defined
- A photoelectric cell that provides an electric current proportional to the number of photons it receives
- A physical detector or receiver to measure incident flux
- A measuring device or recorder that receives, directly or through an amplifier. The current is delivered by the receiver (García & Báez, 2012).



Figure I. 16. The spectrophotometer. (https://chemistrytalk.org/spectrophotometer-chemistry-instrument)

b) The principle of the spectrophotometer

The device measures the intensity of the light it receives, which, passed through a transporting container (bowl whose material must be adopted at the wavelength), contains the solution to be studied. From the intensity of the light which is emitted (denoted I_o) and according to the measurement of the intensity of the transmitted light (I) the device calculates the absorbance (A). The algebraic formula for this operation is:

$$A = \log (I_o/I) \qquad (2)$$

These data can be made more complex by taking into account **Beer - Lambert 's law** which specifies that the absorbance depends proportionally to the concentration of a solution. In fact, it becomes possible to measure the absorbance and calculate the concentration

UV spectrophotometry, which can also measure the absorbance of gases or solids although this is rarer than solutions, however, UV-visible absorption is not a specific test for all compounds, because it can be disturbed by various elements (Zijlstra et al., 1991)**.**

c) ***Spectrophotometric* analysis :**

Absorption spectrophotometry is a technique of a respectable age still widely used in laboratories and in industry, it has benefited from recent technological progress;

- Technique at a reasonable cost.
- Verification and validation of well-documented data.
- It makes it possible to work on small quantities of substances, and it is not destructive with respect to the sample, it still remains one of the measurement methods (Zybin et al., 2005)

I.8.2.3. Definition of scattering

The term scattering is used in various fields including matter, it is an action of responding in such a way as to reach a uniform state, the result of this action in the case of light returning in all directions of the light which strikes the surface of a body, it is also called diffuse reflection, which differs from ordinary diffusion in that the direction of the scattered radiation is not related to that of the incident radiation (Sheffield et al. 2010).

I.8.2.3. Theoretical principle of the propagation of light in absorbing and scattering media

In order to perform spectrometry in a medium that is both diffusing and absorbing, such as biological tissues, it is necessary to establish an analytical relationship between the light intensities collected at a certain distance from a light source $I(r,\lambda)$, intensity of the source $I_o(0, \lambda)$ and the different optical parameters of the medium including its absorption coefficient $\mu_a(\lambda)$. One of the most complete models which describes the propagation of light in such a medium and which can give us such information of Boltzmann's transport equation. Although this takes into account all aspects concerning the propagation of optical radiation, this mathematical equation is difficult to solve and must be simplified (Brar & Verma, 2011). From this the absorption equation is

$$K = \mu_a \lambda / 4 \qquad (3)$$

Conclusion

In this chapter, we gave the basic concepts necessary for the continuation of our work on human blood. We presented its biological composition, functions of the figured elements and its optical properties.

Secondly, we presented in a general way notions on the optical properties of blood (named: the refractive index, the absorption coefficient, and the scattering coefficient), their definitions and their measurement. We ended this chapter with the device which allows the determination of the refractive index "Abb refractometer", and the absorption "spectrophotometry", we presented their construction, their principles and use.

Chapter II: Trends in AI modeling approaches

II. CHAPTER II: Trends in AI modeling approaches

II.1. Introduction

The use of the term "Artificial Intelligence" (AI) has become widespread among the public, as computer technology advances and spreads in human activities. AI is now part of our lives as evidenced by the existence of numerous articles and books relating to it more or less directly. Artificial intelligence techniques are numerical methods that exploit knowledge. The main techniques of AI are the presentation of knowledge, the search for a solution or reasoning and learning (or training) (Chaphalkar and Sandbhord 2013).

In this chapter we recall the basic definitions and notations relating to neural networks the stages of designing a neural network with a focus on the intricacies of variable selection, and the underlying structural framework. Additionally, we explore the fundamental principles of Support Vector Machines (SVM) and introduce the concept of Dragonfly Algorithm combined with Support Vector Machines (DA-SVM), shedding light on their distinctive advantages in various applications.

II.2. Historical

In 1943, Mac Culloch and Pitts proposed the first biological neuron model (Davalo and Naïm 1989). The latter, also called threshold neuron, was inspired by recent discoveries in biology. They are logical neurons (0 or 1). These two physicists showed that formal neurons can perform logical functions.

In 1949, the psychologist Donald Hebb introduced, in his work "The Organization of Behavior ", the term connectionism to speak of massively parallel and connected models. He proposed many rules for updating weights, the best known of which at that time was "Hebb's rule" (Hebb 1949; Davalo et Naïm 1989).

II.3. Basics of artificial neural networks (ANN)

II.3.1. Analogy between the biological neuron and the artificial neuron

The idea of artificial neural networks was inspired by the way biological neurons process information. This concept is used to implement software simulations of massively parallel processes that involve processing elements connected in a network architecture. Learning in the human brain

takes place through a network of neurons which are connected by the axon, synapses and dendrites, (figure II.1) (Tzanakou 2000).

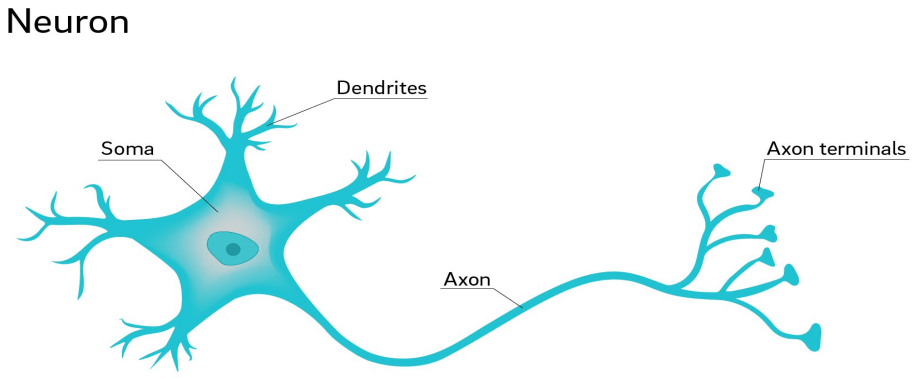


Figure II. 1. Biological neuron (Baillot, 2018).

The biological neuron is a cell composed of a cell body and a nucleus. The cell body branches to form dendrites. It is through the dendrites that information is conveyed from the outside to the soma, the body of the neuron. The information processed by the neuron then travels along the axon to be transmitted to the other neurons. The junction between two neurons is called the synapse. The dendrites, collect, via the synapses, the electrochemical signals or the action potentials coming from other cells or from outside and convey them towards the cellular body or the soma. The soma, in addition to its role in the metabolism of the cell, collects and concentrates information and sums it up. If the somatic potential exceeds a certain threshold, an action potential is emitted, which corresponds to a very brief electrical oscillation. The axon is a cellular extension responsible for transmitting nerve information to other cells through the synapses thanks to its terminal arborization.

Synapses are at the point of contact between neurons, they convert nerve impulses into chemical signals, that is to say into neurotransmitter molecules. Indeed, the arrival of an action potential at the end of an axon leads to the passage of a quantity of neurotransmitters in the synaptic interval. The number of synapses per neuron varies considerably from several hundred to ten thousand (Tzanakou, 2000).

The artificial neuron (Figure II.2) receives inputs that are analogous to the electrochemical impulses that the dendrites of biological neurons receive from other neurons. It is therefore a very simple mathematical operator which translates the mathematical representation of a biological

neuron. It has inputs which can be the outputs of other neurons, or external signal inputs, and an output. The value of the output results from the calculation of the sum of the inputs, weighted by coefficients (synaptic weights) and from the calculation of a bounded nonlinear function of this weighted sum. This function (activation function) is most often a sigmoid function or a step (Tzanakou, 2000).

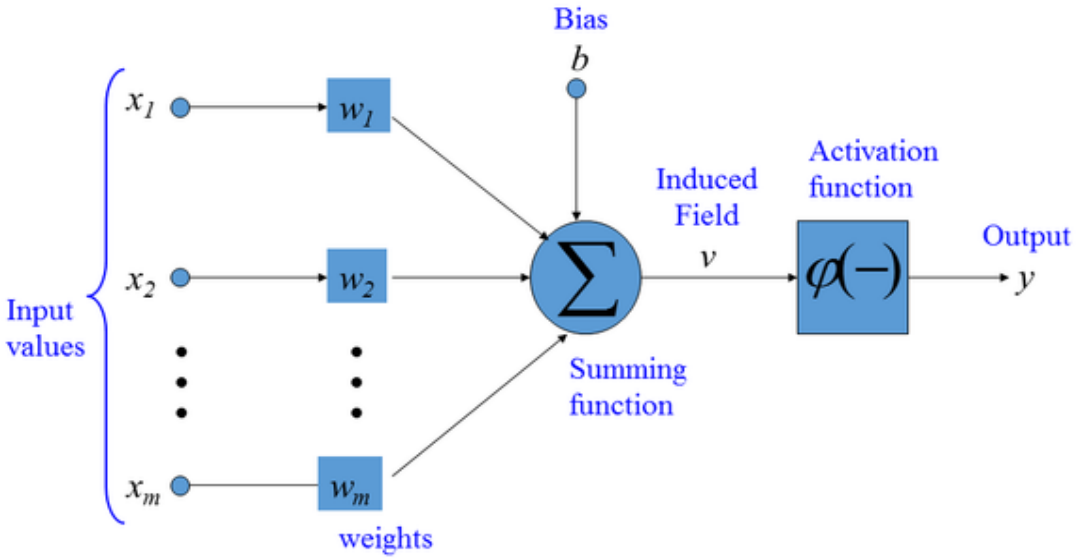


Figure II. 2. Artificial neuron. (Tzanakou, 2000)

Figure II.3 summarizes the correspondence between the artificial neuron and the biological neuron. We notice in this schematic description, that each component or constituent of the artificial neuron finds its biological inspiration

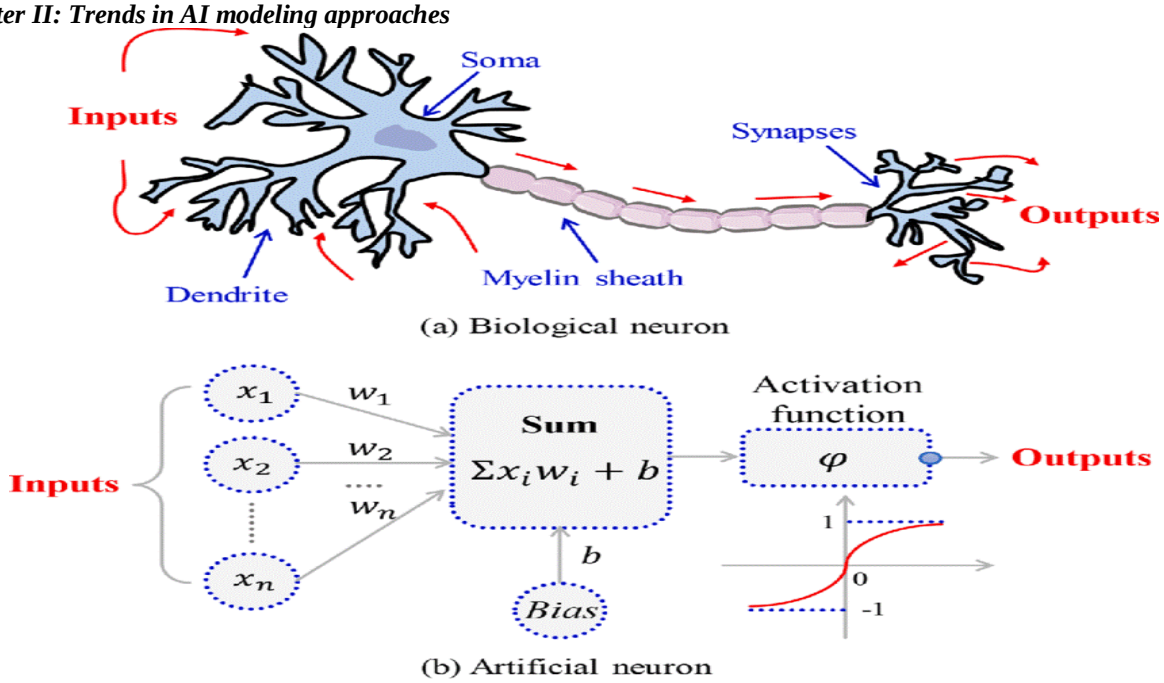


Figure II. 3. Comparison between a biological and an artificial neuron (X. Wang et al., 2021)

II.3.2. Neural network architecture

Artificial neural networks group together in networks a number of formal neurons connected to each other in various ways. A network is defined by its topology, which represents the type of connection existing between the various neurons of the network, by the function of activation which characterizes it and by the learning methods used to find an optimal nonlinear relation by approximation between the input and output variables (Talbi, 2011).

There are two main families of neural networks: looped networks (Recurrent) and non-looped (Feed-forward) networks (Figure II.4).

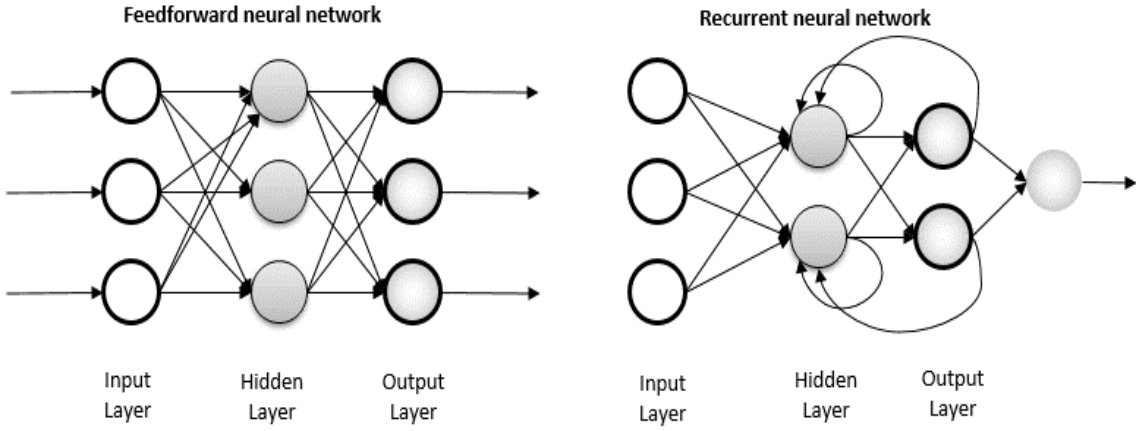


Figure II. 4. Feed forward and recurrent ANN architecture (Pekel & Kara, 2017).

II.3.3. Feed-forward static networks

A non-looping neural network, also known as a feedforward neural network (FFNN) or a feedforward artificial neural network, is graphically represented as a collection of interconnected neurons, with information flowing from inputs to outputs without any "backtracking." It's important to note that the graph of a non-looping network is acyclic. This means that when we traverse this type of network starting from any neuron and following the connections, we cannot return to the initial neuron (Bishop, 1995). The neurons responsible for the final computations in function composition are the output neurons. Those performing intermediate calculations are the hidden neurons (see Figure II.5).

Layered feedforward neural networks are structured in such a way that neurons within the same layer are not connected to each other. Instead, each layer receives signals from the preceding layer and transmits the results of its computations to the next layer. The two outermost layers correspond to the input layer, which receives inputs from the external environment, and the output layer, which provides the results of the performed computations. The intermediate layers are referred to as hidden layers, and their number can vary (Chamekh, 2007).

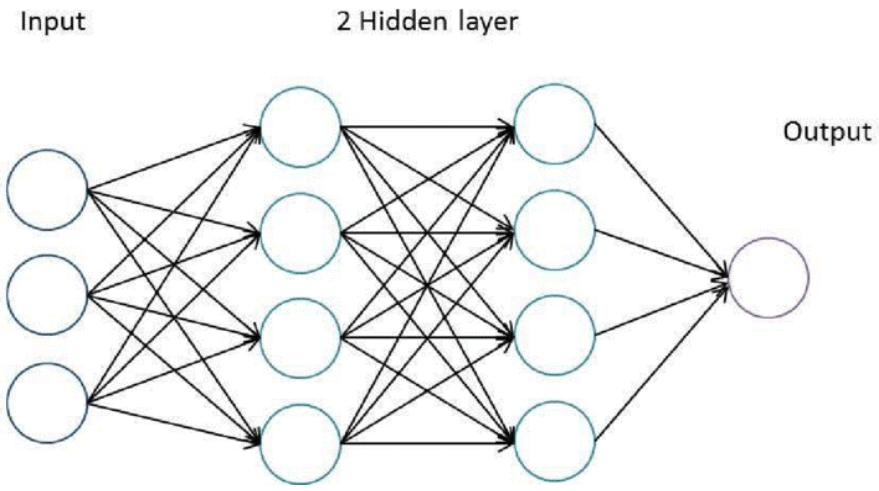


Figure II. 5. Structure of a non-looped neural network. (Chamekh, 2007)

In such a feedforward ANN a neuron, in a hidden layer or an output layer, has two tasks (Tzanakou, 2000). It adds the bias to the weighted inputs of several connections and then applies an activation function (also called transfer function or thresholding function) to this sum as given by equation 3.1 (for neuron *j* of the hidden layer), (Roubehie Fissa et al., 2019)

$$Z_j = f_n\left(\sum_{i=1}^{n} w_{ji}^{I}\, u_i + bhj\right);\ j=1,2,\ldots m \qquad (II.1)$$

It propagates the resulting value through outgoing connections to neurons in the next layer where it undergoes the same process as given by Equation II.1 (e.g. outputs from the hidden layer fed to neuron *k* in the output layer give the output v_k),

$$V_K = fo\left(\sum_{j=1}^{m} w_{kj}^{h}\, z_i + b_{okay}\right);\ k=1,2,\ldots 1 \qquad (II.2)$$

By combining equations II.1 and II.2 we obtain the relation between the output v_k and the inputs u_i from NN:

$$V_K = fo\left(\sum_{j=1}^{m} w_{kj}^{h} f_n\left(\sum_{i=1}^{n} w_{ji}^{I}\, u_i + b_{hj}\right) + b_{ok}\right);\ k=1,2,\ldots 1 \qquad (II.3)$$

The number of neurons in the input layer and the output layer is fixed by the number of independent and dependent variables respectively. The modeler defines the number of hidden layers and the number of neurons in each hidden layer. Model development is achieved through a learning process in which a set of experimental data of the independent variables is presented to the input layer of the network. Outputs from the output layer include a prediction of the dependent variables. The network learns the relations between the independent and dependent variables by the iterative comparison of the calculated outputs and the experimental outputs and the posterior adjustment of the matrix of weights and the vector of the biases of each layer by a learning algorithm of back propagation of the error gradient. Thus, the network develops a neural model capable of predicting with acceptable accuracy the output variables included in the domain of the training data. Therefore, the goal of neural modeling is to minimize the prediction errors of the validation data set presented to the network at the end of the training stage, (Roubehie Fissa et al., 2019).

II.3.5. Recurring Network "Feedback"

A neural network looped is schematized by a connection graph which is cyclic. When moving through the network following the direction of the connections, it is possible to find at least one path that returns to its starting point (such a path is referred to as a "cycle"). Thus, an integer delay multiple of the chosen time unit is attached to each connection of a looped neural network (or to each edge of its graph). A quantity at a given instant cannot be a function of its own value at the same instant. Any cycle of the network graph must have a non-zero delay (Chamekh, 2007).

Figure II.6 shows an example of a looped neural network. The numbers in the squares indicate the delay attached to each connection, expressed as a multiple of the time unit. This network

contains a cycle, which starts from neuron 3 and returns to it via neuron 4. The connection from 4 to 3 having a delay of one unit of time.

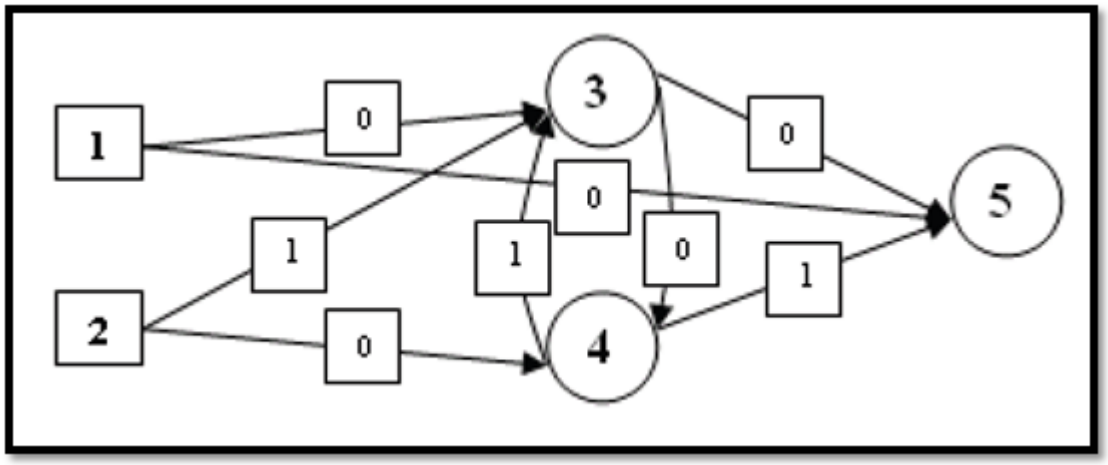
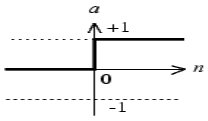
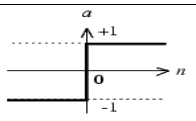
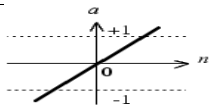
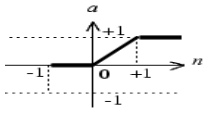


Figure II. 6. A looped neural network with two inputs (Chamekh, 2007).

II.3.6. Activation functions

The activation function calculates the value of the state of the neuron which is transmitted to the downstream neurons, it is also used to introduce a non-linearity in the functioning of the neuron, unlike biological neurons whose state is binary, most transfer functions are continued offering an infinity of possible values included in the interval [0, +1] or [-1, +1]. There are many possible forms for the transfer function, Table II.1 summarizes some transfer functions used in the technical calculation program MATLAB ® (Hagen et al., 1996)

Table II. 1. Usual transfer functions, a=f (n) (Si-Moussa,2009).

Function name	Input/output relationship	Function Graph	Function in MATLAB®
Threshold	$a = 0$ if $n < 0$ $a = 1$ if $n \geq 0$		hardlim
Symmetric threshold	$a = -1$ if $n < 0$ $a = 1$ if $n \geq 0$		hardlims
Linear	$a = n$		pureline
Saturated Linear	$a = 0$ if $n < 0$ $a = n$ if $0 \leq n \leq 1$ $a = 1$ if $n > 1$		satin

II.3.7 Learning phenomena of neural networks

Learning is probably the most interesting property of neural networks, it is considered a phase in the development of neural networks in which the weights of neurons are calculated in such a way that the outputs of the network are as close as possible to the outputs desired, the objective of the learning algorithms is to minimize the decision error made by the ANN by adjusting the weights at each presentation of a training vector. The types of learning can be divided into two broad categories; supervised learning and unsupervised learning (Hassoun, 1995).

II.3.7.1. Learning supervised learning

Supervised learning implies the existence of a teacher whose role is to evaluate the success (or failure) of the network when it is presented with a known stimulus (we say that this stimulus is an example belonging to the base learning). This supervision consists of sending information back to the network allowing it to change its connections (sometimes also its own architecture) in order to reduce its failure rate. The information can be explicit, in the form of a measurement of the error committed, for example, or globally on all the examples of the database. This is what happens in the case of the Perceptron. But it can be more implicit (reinforcement learning), in the form of a simple appreciation (good or bad, punishment or reward), without measurement of error and even be global,

on all the tasks that the network must perform. In the latter case, it is easy to understand that learning is the most difficult, the major difficulty being for the network to identify the stages of the process which are responsible for failure or success (Hassoun, 1995).

II.3.7.2. Unsupervised learning (unsupervised learning)

Unsupervised learning involves providing an autonomous network with a sufficient quantity of examples containing repetitions (in other words, redundancy), such that the latter automatically discovers the regularities. These networks are often called self-organizing, or competitive learning. In unsupervised learning, the data does not contain information about a desired output, there is no human supervisor or expert. This involves determining the parameters of the neural network according to a criterion to be defined. In this case, the examples presented at the input cause the network to self-adapt in order to produce output values that are close in response to similar input values. This type of learning often has less computational complexity compared to supervised learning. Of course, the network architecture, previously defined by its user, is a form of supervision (Hassoun, 1995).

II.3.8. The most used neural networks

Figure II.7 represents the most used neural networks in the literature: looped networks and non-looped networks. In the following we will make a description of each of them.

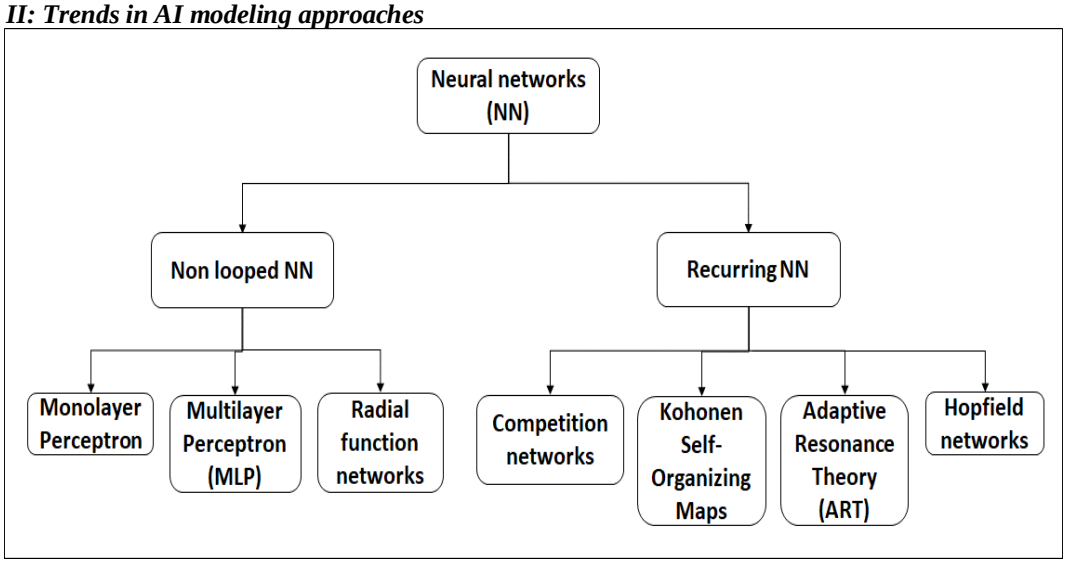


Figure II. 7. The most used neural networks (Talbi, 2011).

II. 3.8.1. The single-layer perceptron (simple perceptron)

The perceptron is the first model of neural networks invented in 1957 by Frank Rosenblatt (Rosenblatt, 1958). It is a simple network, since it only consists of an input layer and an output layer. It can also be used for classification and to solve simple logical operations, it generally follows supervised learning.

II.3.8.2. The Multi-layer perceptron

In our research, we employed the Multi-Layer Perceptron (MLP), a widely used neural network model that extends the previous concept by incorporating one or more hidden layers between the input and output layers (Talbi, 2011). This unique architecture facilitates connections between every neuron in a layer and those in the preceding and subsequent layers, enabling the network to address complex, nonlinear separable problems and intricate logic challenges. The MLP operates under a supervised learning framework, making it the preferred choice for our continued work. Its structure consists of successive layers where data flows unidirectionally from the input layer to the output layer, with no direct connections between neurons within the same layer, as illustrated in Figure II.8 (Keedwell & Narayanan, 2005). The network consists of various layers, named:

- Input layer: is a layer that associated with the system inputs, made up of input cells that are equivalent to the n input variables.
- Hidden layer(s): each layer is composed of k neurons. These neurons have no external connection and are called hidden neurons.

- Output layer: It is composed of the output value of the linear combination of the transfer functions of the hidden layer.

The neurons are interconnected, it is the weights of these connections that govern the operation of the network and program an application from the input space to the output space, using the nonlinear transformation (Khaouane et al., 2017; Hamadache et al., 2016; Laidi & Hanini, 2013).

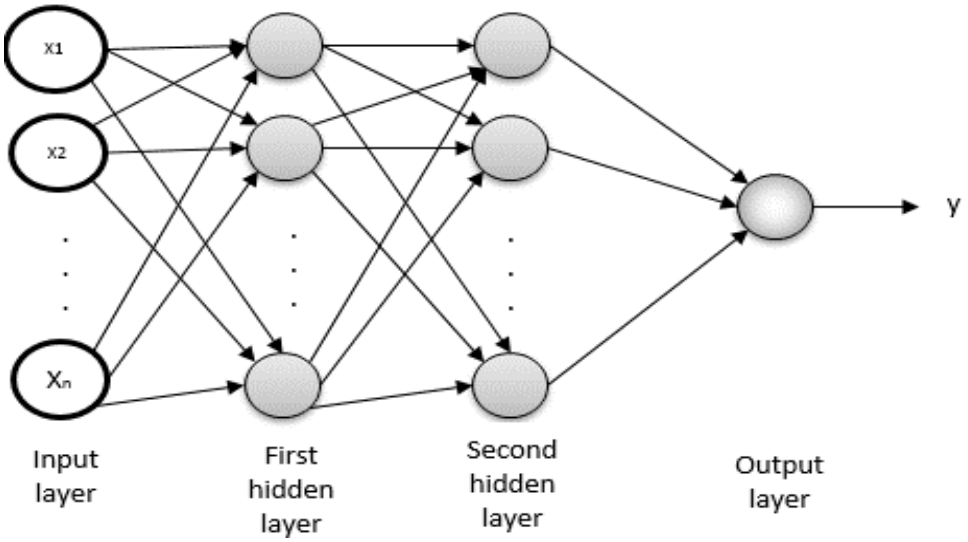


Figure II. 8. A visual architecture of multi-layer perceptron (Pekel & Kara, 2017).

II.3.8.3. Radial function networks

Radial basis function (RBF) networks or more simply radial basis networks have been proposed by Moody and Darken (Moody & Darken, 1989). RBFs will therefore be used in the same types of problems as multilayer perceptron, namely in classification and function approximation, particularly. The most commonly used learning for RBRs is hybrid mode.

II.3.8.4. Kohonen Self-Organizing Maps (SOM)

Topological maps or self-organizing maps were first introduced by Kohonen in 1982 (Kohonen, 1982). They are unsupervised learning networks that establish a discrete map, topologically ordered, according to input patterns. The network thus forms a sort of lattice in which each node is a neuron associated with a weight vector. The correspondence between each weight vector is calculated for each input. Subsequently, the weight vector having the best correlation, as well as some of its neighbors, will be modified in order to further increase this correlation.

II.3.8.5. Hopfield 's networks

Hopfield 's networks (Hopfield, 1982) are recurrent and fully connected networks. In this type of network, every neuron is connected to every other neuron and there is no difference between input and output neurons. They function as a non-linear associative memory and are able to find a stored object based on partial or noisy representations. The main application of Hopfield networks is the knowledge warehouse, but also the resolution of optimization problems. The learning mode used here is the unsupervised mode.

II.3.8.6. ART networks

ART (Adaptive Resonance Theory) networks are competitive learning networks (Carpenter & Grossberg, 1987; Carpenter & Grossberg, 1988). In this kind of network, the weight vectors will only be adapted if the input provided is sufficiently close to a prototype already known by the network. It should be noted that there are two main types of ART networks: ART-1 for binary inputs and ART-2 for continuous inputs. The learning mode of the ARTs can be supervised or unsupervised.

II.3.9. Designing a neural model

Neural networks perform parameterized nonlinear functions. Their implementations require:

- The determination of the relevant inputs and outputs, that is to say the quantities which have a significant influence on the phenomenon that one seeks to model;
- The construction of the database and the choice of samples;
- The determination of the number of hidden neurons necessary to obtain a satisfactory approximation;
- The achievement of learning;
- Evaluation of the performance of the neural network at the end of learning.

II.3.9.1. Determining the inputs/outputs of the neural network

For any model design, the selection of inputs must take into account two essential points:

- First, the intrinsic vector dimension of the inputs should be as small as possible, in other words, the representation of the inputs should be as compact as possible, while retaining essentially the same amount of information and keeping to keep in mind that the different inputs must be independent.

- Secondly, all the information presented in the inputs must be relevant for the quantity that one seeks to model: they must therefore have a real influence on the value of the output.

II.3.9.2. Construction of the database and choice of samples

The process of building a neural network begins with building a database and choosing samples. It is therefore necessary to determine the quantities which have a significant influence on the phenomenon that one seeks to model.

This step is crucial and helps to determine the most appropriate type of network to solve a given problem. In particular, the way in which the sample is presented conditions the type of network, the number of input and output cells and the way in which the training, the test and the validation will have to be carried out (Bishop, 1995).

In order to develop an application based on neural networks, it is necessary to have two databases, one to perform the learning and the other to test the obtained network and determine its performance.

To better control the learning phase, it is often preferable to have a third database called "cross-validation database" (Stone, 1977). The advantages linked to the use of this third database will be explained in the following paragraphs. Note that there is no rule to determine this sharing in a quantitative way, nevertheless each base must satisfy the constraints of representativeness of each class of data and must generally reflect the real distribution, that is to say the probability of occurrence of the various classes (Nascimento, 2000).

Generally, the databases undergo a pre-processing which consists in carrying out an appropriate normalization taking into account the amplitude of the values accepted by the network. Input and output values are normalized within a specific interval to give each parameter the same statistical influence. Moreover, an additional problem may arise due to a possible saturation of the output, if the weighted sum of the inputs of the multilayer perceptron ($\sum(w_i x_i)$) is in the asymptotic part of the activation function.

II.3.9.3. Definition of the network structure

Determining the optimal architecture has remained an open problem for a long time. There are currently various empirical methods involving in particular statistical tests for a large class of networks (Urbani et al., 1994).

As we mentioned, we retained the multilayer perceptron as the basis of the model. We structure this network by specifying the number of layers and hidden neurons so that the network is able to reproduce what is deterministic in the data.

II.3.9.3.1. Number of hidden layers

Apart from the input and output layers, one has to decide on the number of intermediate or hidden layers. Without a hidden layer, the network offers only weak possibilities of adaptation. Nevertheless, it has been demonstrated that a multilayer perceptron with a single hidden layer provided with a sufficient number of neurons, can approximate any function with the desired accuracy (Bermejo & Cabestany,1999).

II.3.9.3.2. Number of hidden neurons

Each neuron can take into account specific profiles of input neurons. A larger number therefore makes it possible to "stick" better to the data presented, but reduces the generalization capacity of the network. It is then necessary to find the adequate number of hidden neurons necessary to obtain a satisfactory approximation. The number of hidden neurons depends on the type of data available. In general, this number is determined by rules of thumb. According to the authors, the size of the hidden layer must be (Dames,2005):

- Is equal to that of the input layer;
- Or equal to 75% of it;
- Let be equal to the square root of the product of the number of neurons in the input and output layer.

For more sophisticated cases, the search would either consist of:

- Estimate a network comprising many neurons then simplify it by analyzing multi-collinearities or by a learning rule eliminating unnecessary neurons;
- Define an architecture taking into account the structure of the variables identified beforehand by a principal component analysis (Bermejo & Cabestany,1999; Bermejo & Cabestany,2001);
- Gradually increase the complexity of the model at each iteration, called " Cascading ", and reassess the generalization efficiency with each addition of a new neuron (Dames,2005).

II.3.9.4. Neural network performance evaluation

Finally, the last step must make it possible to estimate the quality of the network obtained by presenting it with examples that are not part of the training set. It allows both to assess the performance of the neural system and to detect the type of data that poses a problem. If the performances are not satisfactory, it will be necessary either to modify the architecture of the network, or to modify the learning base.

II.3.10. Application areas

The major fields of application of neural networks naturally follow from the properties stated above. We present in the following sections some examples to show the wide range of their applicability.

II.3.10.1. Predictive modeling

Several examples have been carried out in the literature on modeling for a predictive purpose using ANNs , such as the estimation of yields in biomass, in substrate such as glucose and in the product of ethanol by a recurrent network having as inputs: the measurements previous studies, process conditions and process control inputs, studied by Zhu and co- workers (Zhu et al., 1996), they used an ANN to describe the large-scale production of a product from on-line measurements and recursive simulation of sugar consumption. Galván and these co-authors (Galvan et al., 1996) used neural networks in the adjustment of complex kinetic data of chemical reactions, they demonstrated that this modeling is very useful when the kinetic mechanisms are not well understood and that enough process data is available. ANNs can replace semi-mechanistic models, because of its robustness, ease of use and speed of convergence (Fontain &Germain, 2001).

II.3.10.2. Pattern recognition

recognition is another application of *ANNs, it* involves detecting process irregularities, to predict process performance and to detect sensor faults (Glassey et al., 1997), in the study by Huang and coworkers, an artificial auto-associative neural network (AANN) was used online to detect deviations from normal fermentation in " *Streptomyces virginiae* " cells producing the antibiotic "virginiamycin" (Huang et al., 2002).

II.3.10.3 Optimization

ANNs have been applied for optimization of culture media in batch mode by Kennedy and co-authors (Kennedy et al., 1992; Kennedy & Spooner, 1996), this resulted in a significant decrease in experiments compared to the classical DOE-based approach. The combination of ANN with evolutionary algorithms, such as genetic algorithms, has given rise to new approaches that have

proven its effectiveness and robustness in optimizing *Lactobacillus plantarum culture media.* (Desai et al., 2006)

II.3.10.4. Control

Models developed by ANN are often applied within control schemes, state estimators or optimizer, Schenker and Agarwal (Schenker & Agarwal, 1998) studied the different types of neural structures for the predictive control of chemical reactors. Kovárová and co-workers (Kovarova, 2000) used a neural model that combines online optimization for feeding trajectory adaptation during riboflavin production.

II.3.11. Advantages and disadvantages of ANNs

The joint use of NRs makes it possible to draw the advantages and disadvantages of the ANN method. Table II.2 summarizes the advantages and disadvantages of artificial neural networks. (Benimam, 2019).

Table II. 2. The advantages and disadvantages of ANNs

Modeling by Artificial Neural Networks	
Advantages	**Disadvantages**
Mathematical model not required	Black Box.
No rule-based knowledge	Adapting to environments and relearning is often mandatory.
Several learning algorithms are available	

II.4. Support Victor Machine (SVM)

II.4.1. Introduction

The support vector machine algorithm was developed in the 90s by Vapnik (Vapnik, 2000). It was originally developed as a supervised binary classification algorithm. It turns out to be particularly efficient due to the fact that it can deal with problems involving large numbers of descriptors, that it provides a unique solution (no local minimum problems like neural networks) and that it has provided good results on real problems (Li, 2006).

II.4.2. Statistical learning

Statistical learning theory studies the mathematical properties of machine learning (Vapnik 2000). They represent the properties of the class of functions or models that can implement the machine. Statistical learning uses a limited number of inputs (called examples) of a system along

with the values of their outputs to learn a function that describes the existing, but not known, functional relationship between the inputs and the output of the system.

It is first assumed that the training examples, also called training examples, are generated according to a certain unknown (but fixed) probability, that is to say independent and identically distributed. This is a standard assumption in learning theory. The examples are of dimension m $\epsilon \Re^{m}$ and in the case of supervised learning, accompanied by labels characterizing their types or classes of membership. If these labels are countable, we speak of classification otherwise we speak of regression. In the case of a binary classification this label is either +1 or -1. The set of examples and their corresponding labels are called the training set.

In the case of support vector machines, the function sought is of linear form. SVMs are, therefore, learning systems that use a space of hypotheses of linear functions in a high-dimensional feature space. This learning strategy introduced by Vapnik and his co-authors (Vapnik, 2000) is a very powerful method. It has been able, in a few years since its proposal, to conquer most other learning systems in a wide variety of fields of application.

II.4.3. SVM for regression

In their origin, SVMs were developed for classification problems. However, their nature allows them to solve regression problems as well. Regression is a special case of classification where the classes of examples are not countable i.e. continue. The problem is to find, using {(x 1,y 1),…(xn,yn)}, a function $\dot{f}:\rightarrow \Re^{m}$ which approximates the *yi as closely as possible* , in other words which minimizes the difference between the $\dot{f}(x_i)$ and the y_i

Often, $\dot{f}$ is taken as a linear function: $\dot{f} = (w, x) + b$ where w is a vector and *b* is a scalar. The problem therefore amounts to finding a hyperplane characterized by w^* and b^* which minimizes the global difference between $\dot{f}$ and the y_i (equation II.7).

$$(w_*, b_*) = argmin_{w,b} \Sigma (y_i - \langle w, x_i \rangle - b)^2 \qquad \textbf{(II.7)}$$

To solve this problem, SVMs use a trick similar to that used in classification (Hamel, 2011; Farag & Mohamed, 2004; Basak et al., 2007**)** . We propose to model the regression function by a hyperplane which is located in the center of a hyper-tube of width 2 ε containing all the training examples (figure II.8).

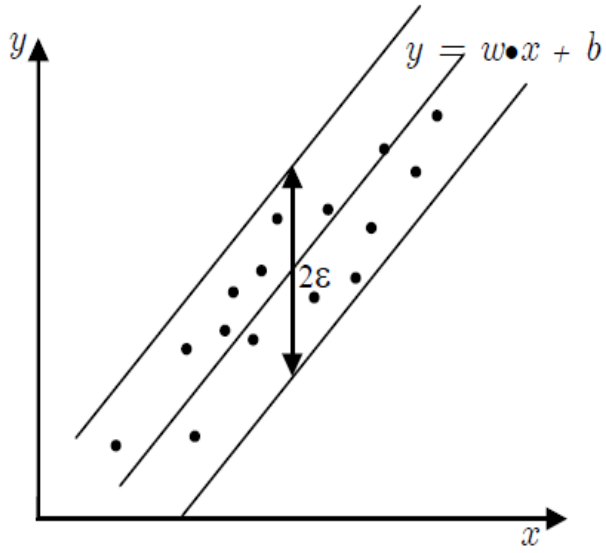


Figure II. 9. Hyper-tube modeling the regression function (Hamel, 2011).

Several hyper-tubes, of width 2 ε containing all the training examples, can exist. The optimal hyper-tube is the one which minimizes the distance between the training examples and its boundaries, in other words, which maximizes the distance of the examples from the hyperplane to the center (figure II.8).

The determination of the optimal hyper-tube is similar to the determination of the optimal hyperplane of maximum margin in the case of classification. One must therefore look for a hyper-tube of maximum margin with all the training examples inside. By an analysis similar to that of the classification problem, the solution of the regression problem is reduced to the resolution of the dual quadratic optimization problem of **equation (II.8).**

$$\begin{cases} Maximize\ f(\alpha, \alpha') = -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i - \alpha'_i)(\alpha_j - \alpha'_j)\langle x_i, x_j\rangle \\ \qquad -\epsilon\sum_{i,=1}^{n}(\alpha_i + \alpha'_i) + \sum_{i,=1}^{n} y_i(\alpha_i + \alpha'_i) \\ \sum_{i=1}^{n}(\alpha_i - \alpha'_i) = 0 \qquad\qquad (II.8) \\ 0 \le \alpha_i - \alpha'_i \le C \end{cases}$$

Where the αi and the αi, are the coefficients of the examples above and below the hyperplane respectively and C is a parameter for their penalization. The output function $\dot{f}(x_i)$ can be given by **equation (II.9).**

$$\dot{f}(x) = \sum_{i=1}^{n}(\alpha_i + \alpha'_i)\langle x_i, x\rangle + b \qquad (II.9)$$

Where b is calculated from an example where $0 < \alpha_i < C$ (support vector) by equation (II.10)

$$b = y_i - \langle w, x_i\rangle + \epsilon \qquad (II.10)$$

II.4.3.1. Using Cores

Among the strongest motivations for the development of support vector machines for regression is their simple extension to nonlinear cases, through the use of kernels.

Indeed, in a way similar to the classification case, we do a space transformation $\emptyset$ for always face a linear regression. The inverse space transformation $\emptyset^{-1}$, makes it possible to return to the original space after the resolution in the new space (figure II.10).

Figure II. 10. Use of the kernels for the resolution of the nonlinear regression

.

As in the case of classification, the transformation $\emptyset$ and its inverse are carried out thanks to a real function K (xi,xj) called Kernel. The dot product in the equations is replaced by the kernel function.

II.4.3.2. The kernels

In addition to the transition to non-linear, the use of kernels allows another very useful generalization in practice: the definition of SVMs on complex objects such as images, graphs, proteins or automata to name but a few. We therefore do not need to make any assumption about the nature of the domain of observations X and a kernel k is generally defined as a function of two variables on $\Re$:

$$(x , x ,) \rightarrow \Re$$

Table II. 3 represents some kernel functions

Most of these kernels depend on a parameter σ called "bandwidth" whose setting is often critical for the proper functioning of the method.

Table II. 3. Examples of kernels

Kind	**Last name**	**K (s,t)**
Radial	Laplacian	$\exp\ (-\frac{r}{\sigma})$
	Rational	$1-\frac{r^2}{r^2+\sigma}$
	loc gauss	$\text{Max}\ (0.1-\frac{r}{\sigma})\ \exp\ (-\frac{r^2}{\sigma})$
Not Stationary	X^2	$\exp(-\frac{r}{\sigma}), r=\sum_k \frac{S_k-t_k}{S_k+t_k}$

Projective	Polynomial	$(s^t\, t)^p$
	Refined	$(s^t\, t + \sigma)^p$
	Cosine	$\frac{s^t\, t}{\| s \| \| t \|}$
	Correlation	$\exp \frac{s^t\, t}{\|s\|\|t\|} - \sigma$

II.4.4. Model evaluation

The quality of a decision model obtained by the SVM method depends on several parameters, namely the penalty parameter on the Lagrange multipliers *C*, the kernel used and its parameters (σ in the case of the radial and projective kernels) and the examples used for training (support vectors).

The choice of the values of these parameters is currently done through several tests and evaluation to achieve satisfactory performance of the model. The optimal parameters (C^*, K^*, σ^*) for a given model are the parameters that allow it to give 100% accuracy. This would be ideal if the set of examples perfectly represented the set of all possible examples. The learned model can give very high accuracy against training examples, but behaves very badly against new examples.

This represents a well-known phenomenon in learning which is over-learning or rote learning. Over-fitting generally results in models with poor generalization ability, and therefore the measure of accuracy is not sufficient to qualify the performance of a model. The evaluation methods make it possible to draw conclusions on the behavior of a model faced with the entire space of examples by limiting the influence of the training examples and the noise that may exist there (labeling errors, errors acquisition, ...) and their order on the model (Benimam, 2019).

II.5. Dragonfly algorithm

Dragonfly Algorithm (DA), developed by Mirjalili in 2016, is a recent and interesting nature-inspired meta-heuristic optimization algorithm used to solve a wide variety of optimization problems. Dragonflies are small flying carnivorous insects that hunt and eat a wide variety of small insects like butterflies, bees, ants and mosquitoes. There are 3000 different species of dragonflies and their life cycle has two stages called nymph and adult. DA is based on the natural dynamic (migratory) and static (feeding) swarming behaviors of dragonflies. According to Reynolds, swarm behavior follows three primitive principles (Mirjalili, 2016; Tharwat et al., 2018)

- Separation, which refers to the avoidance of static collision between individuals in a population, is calculated as follows:

$$S_i = -\sum_{j=1}^{N} X - X_i \quad \text{(II.11)}$$

Where x represents the current solution, and x_i the position in the neighborhood individual i, while n is the total number of members in a neighborhood.

- The alignment, which indicates the correspondence of the speed of individuals existing in the same group and is defined as follows:

$$A_i = (\sum_{j=1}^{N} V_i / N) \quad \text{(II.12)}$$

Where V_i is the speed of neighborhood individual i.

- Cohesion, which refers to the tendency of individuals towards the neighborhood center of mass and is described by equation (II.13)

$$C_i = (\sum_{j=1}^{N} X_i / N) - X \quad \text{(II.13)}$$

The movement of dragonflies towards food in the wild can be expressed by:

$$F_i = x^+ + x \quad \text{(II.14)}$$

Where x_+ represents the position of the food, and their distance from the enemy is expressed by the equation (II.15)

$$E_i = x + x^- \quad \text{(II.15)}$$

given x_- is the position of the enemy (Mirjalili, 2016).

The location/coordinate of a dragonfly is the sum of its position x and its pitch Δx and can be expressed by equation (II.16):

$$X_{t+1} = X_t + \Delta X_{t+1} \quad \text{(II.16)}$$

Where t is the current iteration, and Δx_{t+1} is determined using equation (II.17):

$$\Delta X_{t+1} = sS_i + aA_i + cC_i + fF_i + eE_i + \omega \Delta X_t \quad \text{(II.17)}$$

where S_i is the separation for x_i, and s is its separation weight; A_i is the alignment of x_i, and a its alignment weight; C_i is the cohesion of x_i and c its cohesion weight; F_i is the food attraction of x_i, and f is the food parameter; E_i is the enemy's distraction and e its factor; and ω is the inertia weight. The factors s, a, c, f, and e are determined during optimization and have a double action, that of exploration and exploitation in the dragonfly algorithm (Dragonfly Algorithm (DA)) (Mirjalili, 2016).

The advantages of using the Dragonfly Algorithm include:

- Efficiency: The Dragonfly Algorithm efficiently explores the search space and converges towards optimal solutions by leveraging the collective behavior and movement patterns of dragonflies.
- Flexibility: The algorithm can be easily adapted and applied to various optimization problems due to its adaptive behavior and ability to search in the full range of wavelengths.
- Robustness: The Dragonfly Algorithm is less likely to get trapped in local optima due to its combination of random search and directed flight strategies, making it a robust optimization technique.
- Scalability: The algorithm can handle large-scale optimization problems with a large number of variables and constraints, making it suitable for complex biomedical engineering applications (Mirjalili, 2016).

The comparison of DA-SVM performance with existing approaches such as PSO-SVM GA-SVM algorithms, the results revealed that the DA-SVM achieved competitive results of searching the optimal SVM parameters to improve the classification accuracy (Tharwat et al., 2018).

II.6. Conclusion

The trends in Artificial Intelligence (AI) modeling approaches have been thoroughly reviewed in this chapter, with a focus on three key methods: Artificial Neural Networks (ANN), Support Vector Machines (SVM), and the innovative Dragonfly Algorithm combined with Support Vector Machines (DA-SVM). We have outlined the core aspects of ANN, delving into its definition, the crucial process of variable selection, and the architectural structure. Additionally, we provided a description of SVM, addressing its fundamental definition and its relevance in diverse applications. The incorporation of DA-SVM was introduced as a promising approach that capitalizes on the advantages of both SVM and Dragonfly Algorithm.

Chapter III: Theoretical background on the optical properties of blood

III. Chapter III: Theoretical background on the optical properties of blood

III.1. Critical reviews about optical properties

The optical properties of blood compounds have been a subject of extensive research and investigation over the past few decades. Understanding these properties is crucial for various biomedical applications, including optical diagnostics, imaging techniques, and therapeutic interventions. In recent decades, numerous studies have been achieved to explore the optical behavior of blood compounds, focusing on the measurement of experimental data, development of empirical models, and the emergence of novel approaches utilizing artificial intelligence.

In this chapter, we aim to provide an overview of the studies and works conducted between 1996 and 2022 on the optical properties of blood compounds. We will discuss the different experimental measurements employed to characterize these properties and empirical modeling techniques.

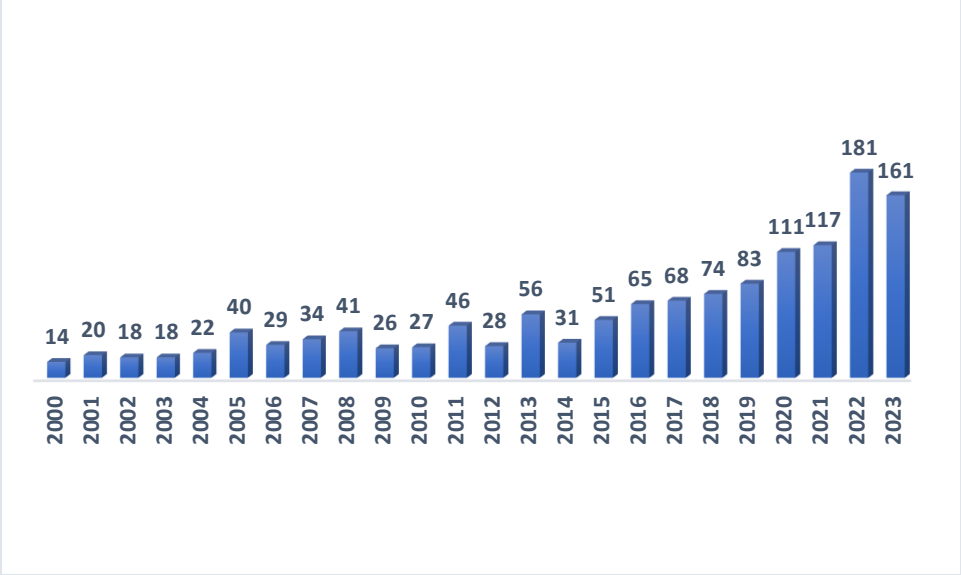


Figure III. 1. number of publications dealing with the optical properties of blood and its components (science direct 7/17/2023).

III.2. Experimental studies and empirical modeling of the optical properties of blood components

Yaroslavsky, focused on determining the optical properties of whole blood samples obtained from healthy volunteers within the near-infrared spectral range using a double integrating sphere measurement technique combined with an inverse Monte Carlo approach. They measured diffuse reflectance, total transmittance, and collimated transmittance, from which they derived the

absorption coefficient, scattering coefficient, and anisotropy factor. The spectral range investigated in this study extended from 700 nm to 1200 mm. One significant finding was that the optical properties of blood differed substantially from those of other relevant human tissues that had been previously characterized. They also explored the impact of scattering phase function approximations on the estimation of optical properties. Specifically, they considered the Henyey-Greenstein and Gegenbauer kernel phase functions and compared the calculated angular distributions of scattered light with goniophotometric measurements conducted at a wavelength of 633 nm. The results demonstrated that variations in the employed scattering phase function approximation could lead to significant discrepancies in the derived optical properties. This conclusion highlights the critical need for accurate knowledge of the scattering phase function when determining optical constants from double integrating sphere measurements (Yaroslavsky et al., 1996).

(Roggan et al., 1999), conducted a study aimed at investigating the optical properties of human blood using the integrating sphere technique and inverse Monte Carlo simulations. Their research encompassed the measurement of absorption coefficient (μ_a), reduced scattering coefficient (μ_s), and the g factor of circulating human blood. In their experiments, they noticed that at a wavelength of 633 nm, blood with a hematocrit of 10% and an oxygen saturation of 98% exhibited the following optical properties: an absorption coefficient (μ_a) of $0.210 mm^{-1}$, a reduced scattering coefficient (μ_s) of $77.3 mm^{-1}$, and a g factor of 0.9. They found, an increase in hematocrit up to 50% led to a linear increase in absorption and reduced scattering. They also observed that changes in osmolarity and wall shear rate resulted in variations in all three parameters, while variations in oxygen saturation predominantly affected the absorption coefficient. Furthermore, they conducted a comprehensive wavelength analysis (400–2500 nm) for both oxygenated and deoxygenated blood. They discovered that blood absorption closely followed the absorption behavior of hemoglobin and water. The scattering coefficient exhibited a decreasing trend for wavelengths above 500 nm, the g factor remained consistently higher than 0.9 across the entire wavelength range.

(Li, 2000), conducted a study with the aim of providing dispersive relations of the refractive index for different types of human whole blood in the visible and near-infrared ranges, as well as under various conditions. To overcome the scattering effect, they utilized an unconventional method based on total internal reflection. Their experimental setup involved a focused light source, a semicylindrical lens in contact with tissues, and a linear CCD camera (Charge-Coupled Device).
By analyzing the spatial distribution of internally reflected light, the critical angle and, consequently, the refractive index was determined. To quickly and directly ascertain the chromatic dispersion curve of materials, a monochromator was employed as the light source. The study presented a

comprehensive set of values that established the relationship between the refractive index and the wavelength and blood types of undiluted whole blood. The results of the study indicated that the refractive dispersions were nearly the same in both the visible and near-infrared ranges, regardless of the blood type. Moreover, the relationship between the refractive index and the wavelength was described using Cauchy's formula.

(Sardar, 1998) accomplished a study to investigate the optical properties of whole blood specifically at the wavelengths of 488 nm and 633 nm. The authors employed the inverse adding-doubling (IAD) method to determine the absorption and scattering coefficients, as well as the scattering anisotropy coefficient of blood. These parameters were derived from measurements of total diffuse reflectance and transmittance obtained using an integrating sphere at 488 nm. One notable aspect of the study was the comparison of the scattering anisotropy coefficient obtained through the IAD method with the results obtained independently through goniometric measurements. This comparison allowed for an assessment of the accuracy and reliability of the IAD method in determining the scattering anisotropy coefficient of blood. They also measured the refractive index of blood using the conventional minimum deviation method. The measurements were performed at a wavelength of 633 nm, and an equilateral hollow quartz prism was utilized for this purpose. It is worth noting that the blood sample used in the study was drawn from one of the authors and was collected into a vacuum tube containing the anticoagulant EDTA (ethylenediaminetetraacetic acid).

(Faber et al., 2004), carried out a study focused on the scattering properties of oxygenated and deoxygenated whole blood within the spectral range of 250 to 1000 nm. They employed a combination of Kramers-Kronig analysis and optical coherence tomography (OCT) measurements to determine the complex refractive index of oxygenated and deoxygenated hemoglobin. By incorporating Mie theory and the obtained data, the scattering properties were calculated. It is worth mentioning, the authors utilized hemoglobin solutions derived from fresh porcine blood, which were anticoagulated and maintained at 37 °C. These solutions were oxygenated to 100% or deoxygenated to 0% using a clinical oxygenator. The key finding emphasized in the study is the substantial influence of oxygen saturation on scattering effects, which necessitates careful consideration in the data analysis of optical oximetry.

In their study, (Friebel, 2005), focused on determining the complex refractive index of highly concentrated hemoglobin solutions within the wavelength range of 250 to 1100 nm. This was achieved through the utilization of transmittance and Fresnel reflectance measurements. Notably, the measured real parts of the refractive indices were found to be, on average, 0.02 units higher than the values reported in the literature. Furthermore, the wavelength dependence of the measured data in

the UV region was observed to differ from the calculated data obtained using the Kramers-Kronig relation. It is important to note that the study relied on blood samples obtained from healthy human donors. Highly concentrated hemoglobin solutions with concentrations of 28.7 and 30.6 g/dL, which closely matched the physiological hemoglobin concentration within red blood cells (31 to 36 g/dL), were prepared without the addition of any chemicals. Lower concentrations of hemoglobin were achieved by diluting the hemoglobin solution of known total hemoglobin content with physiological saline solution. All measurements were conducted at a temperature of 22 °C after adjusting the prepared hemoglobin solutions to a pH of 7.4.

(Friebel & Mienke, 2006), determined the real part of the complex refractive index of oxygenated hemoglobin solutions as a function of concentration within the wavelength range of 250 to 1100 nm through Fresnel reflectance measurements. The hemoglobin solutions used in the study were produced by physically inducing hemolysis of human erythrocytes, followed by ultracentrifugation and filtration. The study presented a model function that enables the calculation of the refractive index of hemoglobin solutions as a function of concentration within the wavelength range of 250 to 1100 nm. This model function provides insights into the dependence of refractive index on concentration and offers a useful tool for future investigations. The hemoglobin solutions, concentrates of human erythrocytes were prepared according to the methodology described in a previous study by (Friebel & Mienke, 2005). All measurements were conducted at a pH of 7.4, a temperature of 22 °C, and with a constant oxygen saturation level of 100%. The hemoglobin solutions used in the experiments were directly obtained from human erythrocytes, including salts and other organic compounds. The study revealed that the specific refractive increment calculated by the model function yielded a value 44% higher at 589 nm compared to Barer's formula. This discrepancy suggests that the presence of non-absorbing components within a red blood cell, and to purified hemoglobin, contributes to higher refractive indices.

(Jin et al., 2006), performed a comprehensive study on the refractive indices and thermo-optic coefficients of essential biomaterials, including blood plasma, hemoglobin solution, and lipid membrane. The authors utilized the method of total internal reflection at the commonly used laser wavelengths of 532 nm and 632.8 nm in the biomedical field, with the aim of investigating the effects of sample concentration and temperature on the refractive index, along with providing empirical relationships and theoretical explanations based on molecular polarization theory. Addressing the relationship between the refractive index and several variables, such as the hemoglobin concentration and the temperature. They found, as the temperature increased, the refractive index showed a linear decrease. This characteristic suggests the sensitivity of refractive

index to changes in temperature within the investigated biomaterials. The study additionally showed a linear relationship between the samples' concentrations of hemoglobin and refractive index. The degree of linear correlation between the refractive index and the concentration of hemoglobin was found to be 0.9997 at a wavelength of 632.8 nm. Similar results were found for the degree of linear correlation at 532 nm, which was 0.9996.

In the study of (Friebel et al., 2006), the absorption coefficient (μ_a), scattering coefficient (μ_s), and anisotropy factor (g) of both diluted and undiluted human blood samples with hematocrit values of 0.84% and 42.1% were determined under flow conditions. The study focused on the wavelength region between 250 and 1100 nm, including the hemoglobin absorption bands. They employed integrating sphere measurements, combined with an optimized inverse Monte Carlo simulation (IMCS) to obtain these values. The evaluation of suitable effective phase functions for both blood concentrations using the IMCS approach was made possible by the development of a new algorithm. The best results were obtained using the Reynolds-McCormick phase function with variation factors of 1.2 for hematocrit 0.84% and 1.7 for hematocrit 42.1%. The findings were then compared with the Mie theory's parameters, which define how spherical particles behave in scattering. It was allowed to take into account the non-spherical nature of erythrocytes, related absorption and scattering phenomena, and numerous scattering and interference effects through the use of IMCS in conjunction with some effective phase functions.

As a result, even in circumstances of high hematocrit and areas with extensive hemoglobin absorption, reasonable results about the optical behavior of human blood were obtained. The research also revealed the limitations of the Mie theory in precisely explaining the optical characteristics of blood.

It is worth noting that the blood preparation process employed in this study was consistent with Friebel's earlier work in 2005.

In the study conducted by (Mienke et al., 2007b), the optical properties (absorption coefficient, scattering coefficient, and anisotropy factor) of platelets (PLTs) suspended in plasma and acellular plasma were determined by utilizing diffuse reflectance and transmission measurements, as well as inverse Monte Carlo simulation, the researchers investigate the optical behavior of PLTs and red blood cells (RBCs) suspended in plasma and compare it to that of saline solution in the visible and near-infrared range. The results of the study highlight several conclusions, cell-free plasma exhibits a higher scattering coefficient and anisotropy factor than expected based on Rayleigh scattering by plasma proteins, the scattering coefficient of PLTs shows a linear increase with the concentration of PLTs, and the presence of physiological concentrations of leukocytes does not have

a measurable influence on the absorption and scattering properties of whole blood. The study showed that RBCs significantly contribute to the absorption and effective scattering properties of whole blood compared to other blood components. The substitution of saline solution for plasma leads to a significant increase in the effective scattering coefficient, and slight difference in refractive index between plasma and saline solutions can result in up to a 100% change in the effective scattering of RBCs when plasma is substituted by saline solution.

In 2009, Firebel and co-workers, studied the influence of the oxygen saturation on the optical scattering properties of human red blood cells in the spectral range 250 to 2000 nm. the optical parameters were determined using integrating sphere measurements of light transmittance and reflectance, in combination with inverse Monte Carlo simulation for two levels of oxygen saturation (SAT O2): 100% and 0%. The findings indicated that the absorption behavior is primarily influenced by hemoglobin absorption, there is no dependence on SAT O2, and the absorption approximates that of water, albeit at values 20 to 30% lower. The parameters of the scattering coefficient and anisotropy factor are significantly affected by the SAT O2-induced changes in absorption. (Friebel et al., 2009)

In the research conducted by (Meinke et al., 2007a), an extensive investigation was carried out to determine the absorption coefficient, scattering coefficient, and effective scattering phase function of human red blood cells (RBCs) in saline solution. Their study focused on analyzing the optical properties of RBCs across a range of hematocrits (Hcts) varying from 0.84% to 42.1% within the wavelength range of 250–1100 nm. The measurements were performed using integrating sphere techniques in combination with inverse Monte Carlo simulation. The research findings presented empirical model functions that allow for the calculation of Hct-dependent optical properties specifically for RBCs to understand the relationship between the optical properties of RBCs and varying Hct levels. The study also investigated the changes in optical properties when blood plasma replaced saline solution as the suspension medium. All the blood samples used in the study were oxygenated to a level exceeding 98%, at a constant temperature of 20 °C, and obtained from healthy donors (Meinke et al., 2007b; Meinke et al., 2011).

The study conducted by (Zernovaya et al., 2011), measured the refractive index of human deoxygenated and oxygenated hemoglobin across nine different wavelengths ranging from 400 to 700 nm. The hemoglobin concentrations studied were up to 140 g/ l. The results of the study revealed a linear relationship between the refractive index and concentration at all wavelengths. By analyzing the slopes of the linear relationships, the researchers determined the specific refraction increments. Based on these findings, they proposed a set of model functions to calculate the

refractive index as a function of concentration. They utilized the model functions to calculate the refractive index at the physiological concentration of 320 g/ l within erythrocytes. These results provide a basis for determining the refractive index within the visible range for arbitrary concentrations, assuming that the refractive indices exhibit a linear dependence on concentration.
In their experiments, human hemoglobin in the form of lyophilized powder was obtained from Sigma-Aldrich and dissolved in phosphate-buffered saline (PBS) to maintain a pH of 7.4, the deoxygenated and oxygenated hemoglobin solutions were prepared by adding sodium dithionite and sodium bicarbonate, respectively, to achieve concentrations ranging from 0 to 140 g /l, and the refractive index measurements were conducted at a temperature of 20 °C. The absolute measurement error was calculated and found with an average value not exceeding 0.001.

In the work of (Sydrouk et al., 2012), aimed to resolve the controversy surrounding the application of Kramers-Kronig relations to solutions of human hemoglobin. Previous attempts had yielded conflicting results and disagreements among various studies. In their research, the authors demonstrated that this controversy can be addressed by considering not only the absorption properties of hemoglobin but also the dispersion of the refractive index of the nonabsorbing solvent. To achieve this, a Kramers-Kronig analysis was employed, taking into account both contributions and considering the refractive index dispersion of the nonabsorbing solvent. The results obtained from this analysis were then compared with data from multiple studies, and showed good agreement across the visible, near-infrared, and ultraviolet regions.
In their experiments, the authors followed the hemoglobin sample preparation method described in (Zernovaya et al., 2011). All experiments were conducted at a temperature of 20°C. By employing the Kramers-Kronig relations, the researchers analyzed the refractive index of hemoglobin solutions within the wavelength range of 250 to 1000 nm. It is important to note that the total real part of a hemoglobin solution is a combination of contributions from both hemoglobin and water, which serves as the solvent in their study. Since water does not exhibit significant absorption within this wavelength range, its refractive index is not determined by the Kramers-Kronig integral evaluated at these wavelengths and should be separately considered.

(Yahya et al., 2016), achieved experimental measurements to determine the real part of the refractive index of human hemoglobin (HB) solutions at various concentrations, temperatures, and wavelengths. They utilized the Abbemat Refractometer to obtain the experimental measurements for sixteen hemoglobin samples. The researchers used dry human hemoglobin 'Lyophilized powder H7123' obtained from Sigma-Aldrich®, To create mimicking phantom hemoglobin samples for the experimental measurements. The powder was dissolved in PBS (phosphate-buffered saline) to

maintain a constant pH of 7.4 and eliminate any pH-related changes in the refractive index. Care was taken to ensure homogeneity by gradually adding the powder to the solution while stirring on a magnetic stirrer. The samples were prepared individually and used immediately following preparation to ensure measurement accuracy. Their research showed a linear relationship between temperature and refractive index, as well as between temperature and hemoglobin concentration. In addition, a nonlinear relationship between wavelength and refractive index was noted.

Based on these results, the researchers developed a mathematical model using linear regression to estimate the refractive index of HB solutions. The model included the variables of concentration, temperature, and wavelength. It was discovered that the addition of temperature and wavelength parameters to the Barer formula improved the accuracy of the refractive index estimation.

The authors reported the average accuracy of the correlation between their mathematical model and the experimental values, the comparison revealed that the average accuracy of the correlation was 1.05×10^{-3}.

(Lazareva et al., 2018), presented a comprehensive study on the calculation of blood refractive index based on experimental data for the real part of the refractive index of the main protein components, hemoglobin and albumin. The refractive index of blood proteins was measured at various wavelengths, including 480, 486, 546, 589, 644, 656, 680, 930, 1100, 1300, and 1550 nm, using a multi-wavelength Abbe refractometer at room temperature 24°C. The hemoglobin solution was prepared from whole blood obtained from a healthy human subject. Heparin was used as an anticoagulant in the collection tubes to prevent coagulation. The blood sample was centrifuged for 10 minutes at 2000 rpm to separate the blood into fractions, allowing the isolation of red blood cells from plasma and other blood corpuscles. The red blood cells were then gathered in a separate test tube, and their haemolysis was achieved through multiple freezing and defrosting cycles at a temperature of -18°C. The haemoglobin concentration in the studied samples was determined to be 260 g/l, corresponding to the average haemoglobin concentration in a red blood cell. The measurement error introduced by the instrument was reported to be within 0.0002. Their study demonstrated the feasibility of modelling the blood refractive index using the Sellmeier formula within the visible and near-infrared wavelength range.

III.3. Optical properties modelling using Artificial intelligence approaches

Despite the importance of optical properties in understanding and characterizing blood compounds, it is noteworthy that the adoption of artificial intelligence approaches in this field has been relatively limited. Specifically, the application of AI has primarily focused on estimating the

refractive index, with fewer studies considering other optical properties such as absorption coefficient, and scattering coefficient.

The work of (oyehan, 2018), focused on studying the refractive index of hemoglobin in relation to physiological and biochemical parameters such as hemoglobin concentration (Hb), wavelength, and temperature. Their objective was to estimate the real part of the hemoglobin refractive index based on experimental data obtained from (yahya at al., 2016). A machine learning model built on the Genetic Algorithm-Support Vector Machine (GA-SVM) algorithm was proposed in their work, with correlation coefficient of 0.9991, and a root mean square error (RMSE) of about 0.00034.

Additionally, during the same period of time, (Alade et al., 2018), built models using the Genetic Algorithm-Support Vector Regression (GA-SVR) algorithm using experimental data obtained from (Zernovaya et al., 2011), that included various wavelengths and different concentrations of hemoglobin (Hb). Their objective was to use these models to determine the refractive indices of oxygenated hemoglobin (HbO2) and deoxygenated hemoglobin (HB).
Their results indicated root mean square errors (RMSE) of 4.65×10^{-4} and 4.62×10^{-4} for HbO2 and HB respectively. The obtained correlation coefficients were 99.85 and 99.84 for HbO2 and HB respectively.

Shortly after, (Murat Oduncuoğlu, 2020), proposed a formula to predict the real part of the blood refractive index, the model used parameters such as wavelength, the concentration of hemoglobin, and the temperature using the genetic programming method, with a set of 492 training points obtained from (Yahiya et al., 2016) in the spectral range of 436 to 1550 nm. The temperature range considered was 20-45°C and the hemoglobin concentration range was 0-200 g/L. the predictive capacity of the program is 0.9873 for training and 0.9837 for the test.

In the study conducted by (Alizamir et al., 2022), a comprehensive investigation was carried out to establish a quantitative correlation between the refractive index (RI) of hemoglobin and its influencing parameters, namely concentration, wavelength, and temperature. Four distinct strategies were employed to develop accurate predictive models for the RI of hemoglobin.
Initially, alternating conditional expectations (ACE), a statistical approach, was utilized to generate a correlation equation for predicting the RI of hemoglobin. Subsequently, three optimized intelligent techniques, namely optimized neural network (ONN), optimized fuzzy inference system (OFIS), and optimized support vector regression (OSVR), were employed to model the RI. To enhance the performance of the intelligent models, a bat-inspired (BA) algorithm was incorporated to optimize the weights, biases, membership functions, and free parameters associated with each model. Their

results showed that, both ACE and the optimized models (ONN, OFIS, and OSVR) performed well in estimating the RI of hemoglobin. The results indicated also that concentration had the highest influence, followed by wavelength, and finally, temperature.

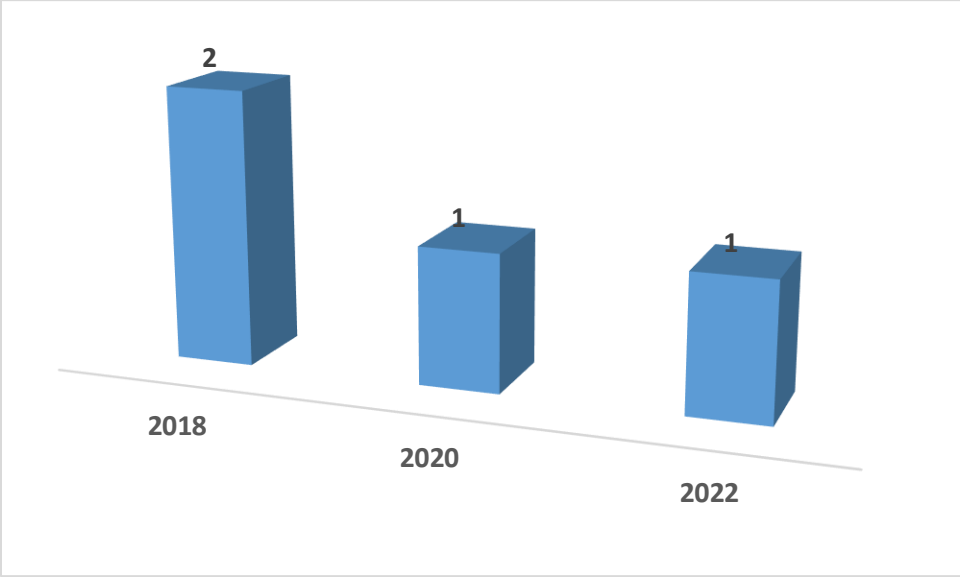


Figure III. 2. Number of publications dealing with modeling the optical properties of blood using AI methods (science direct 7/17/2023).

Chapter IV: Materials and Methods

IV. CHAPTER IV: MATERIALS AND METHODS

IV.1. Introduction

With the ever-evolving landscape of medical research, the quest for an accurate knowledge of the optical characteristics of blood is of paramount importance, as it promises to unravel crucial insights into various physiological and pathological issues.

In this chapter, we delve into the essential materials and techniques used to estimate the optical properties of blood through the ingenious application of artificial intelligence (AI) approaches. we harnessed the power of Artificial Neural Networks (ANN), Support Vector Machines (SVM), and the innovative Dragonfly algorithm coupled with SVM (DA-SVM). The results achieved by the previous method (DA-SVM) have been published in “computer methods in biomechanics and biomedical engineering” (Omari et al., 2023).

IV.2. Materials and methods

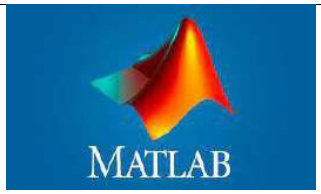
IV.2.1. Computer Tools used

Several software programs were used during this work, as shown in table IV.1:

"GetData Digitizer": is designed to extract numerical data from different types of graphs, plots, and images. It is especially helpful in collecting data from published research papers, scientific publications, reports, or any graphical data that cannot simply be copied and pasted as numerical values.

- "Microsoft Office Excel 2020", for the organization and pre-processing of the database.
- The training set of ANN models and quite a few steps, were carried out by the software called "STATISTICA Versions 10" (Stat Soft 2010), considering how frequently it appears in works of authors involving Hamadache and his co-authors (Hamadache et al., 2016), and those of Ammi and her co-authors (Ammi et al., 2015).
- All training models for Support Vector Machine (SVM) and Dragonfly algorithm coupled with SVM (DA-SVM), as well as the generation of figures and error calculations, were implemented using "MATLAB 2020b" software (Mathworks Inc 2020).

Table IV. 1. the software used in this study

Software name	LOGO	Version
Microsoft office Excel 2016		**2020**
STATISTICA		**10**
MATLAB		**2020**

IV.2.2. Data collection, pretreatment and analysis

Databases offer the main repository of structured and ordered data, providing essential information for the construction of machine learning models. Therefore, three databases (DB) were collected from the literature to predict the optical characteristics of blood such as, the refractive index (RI), the absorption coefficient (μ_a), and the scattering coefficient (μ_s) of blood, containing following inputs: for the first data includes the wavelength (λ) expressed in (nm), the hemoglobin concentration (C) in (g/dL), the temperature (T) in (K), and the level of oxygen saturation (O_2 Sat) in order to predict the refractive index **(**yahya et al., 2016; Friebel & Meinke, 2005; Friebel et al., 2006; Sydoruk et al., 2012). The remaining two databases holding the corresponding inputs: the wavelength, hematocrit percentage (HCT), and saturation of oxygen (Bosschaart et al., 2014; Roggan et al., 1999). the input variables were selected based on the interactions, correlations between the properties, and the operating conditions. The descriptive statistics of these features are shown in Tables IV (1-3). All of the optimization results of this study are presented in Tables 1-3 (Annex).

Table IV. 2. Statistical parameters of the refractive index were used in building the models

Variables of DB1	Symbol	No of Data	Unit	Minimum	Maximum	Mean	Standard deviation
Refractive Index	RI	930	(-)	1.332370	1.443060	1.362408	0.024492
wavelength	λ	930	nm	255.813	1149.660	609.305	185.146
concentration	C	930	g/dL	0.00000	28.70000	12.57591	8.57489
Temperature	T	930	K	293.1500	318.1500	300.1274	7.9926
Oxygen saturation	O_2 sat	930	%	0.0000	100.0000	96.7742	17.6780

Table IV. 3. Statistical description of the Absorption data Adjusted in this work

Variables of DB2	Unit	No of Data	Mean	Median	Minimum	Maximum	Standard deviation
Wavelength	nm	1000	999.030617	907.5	251	2220.8	524.300445
Hematocrit	%	1000	38.2888103	45	0.41352	56.7984	14.8005159
Oxygen saturation	%	1000	50.3865959	89.62765	0	100	49.318753
Absorption coefficient	mm^{-1}	1000	12.6135352	1.16	0.0105712	159.31	24.5790358

Table IV. 4. Statistical description of the Scattering Data used in building the models

Variables of DB3	Unit	No of Data	Mean	Median	Minimum	Maximum	Standard deviation
Wavelength	nm	1000	951.61995	792.5	272	1999.54	511.802475
Hematocrit	%	1000	38.7213047	45	0	69.7943	14.6956736
Oxygen saturation	%	1000	46.629	0	0	99	49.4413977
Scattering coefficient	mm^{-1}	1000	55.0827478	45.9794	0	128.16	30.2707498

The correlation coefficient between inputs, often referred to as feature correlation, plays a critical role in building models based on AI approaches. The correlation coefficient measures the strength and direction of a linear relationship between two variables. Highly correlated input features could have duplicate information and redundant features can lead to overfitting, when the model captures noise from the training data rather than learning meaningful patterns. Correlated inputs can make it challenging to interpret the model's behavior. If multiple features carry similar information, it becomes difficult to identify the features that are truly having an effect on the target variable.

In our study the features of DB1 are: the wavelength, hemoglobin concentration, temperature, and oxygen saturation; the output in this case is the refractive index as shown below in the figure IV.1. The correlation between wavelength and hemoglobin concentration is very close to zero (-0.0081), indicating an extremely weak linear relationship between these two variables. This suggests that changes in hemoglobin concentration are not significantly associated with changes in wavelength. Similarly, the correlation between wavelength and temperature is negative and close to zero (-0.0259), implying a very weak linear relationship. Changes in temperature are not strongly related to changes in wavelength. The positive correlation between wavelength and oxygen saturation is also very weak (0.0167). This suggests that variations in oxygen saturation do not strongly influence changes in wavelength. The correlation between wavelength and refractive index is negative and close to zero (-0.0157), implying a very weak linear relationship. Changes in refractive index do not show a strong association with changes in wavelength. The negative correlation between hemoglobin concentration and temperature is still relatively weak but slightly stronger than the other

correlations (-0.0633). This suggests that there might be a very subtle inverse relationship between hemoglobin concentration and temperature. The negative correlation between hemoglobin concentration and oxygen saturation is also weak (-0.0217). This suggests that variations in oxygen saturation are not strongly associated with changes in hemoglobin concentration. The correlation between hemoglobin concentration and refractive index is extremely close to zero (0.0016), indicating virtually no linear relationship between these two variables. Changes in hemoglobin concentration are unlikely to have any significant impact on changes in refractive index. Although there is no difference in the correlation between the features in the two other databases DB2 and DB3 (Figures IV.2 and IV.3), the correlation between the wavelength and the hematocrit, and the wavelength and oxygen saturation are both exceedingly negative and also very near to zero.

In overall, based on the provided correlation coefficient values, it appears that the relationships between the features and outcomes are generally very weak or negligible.

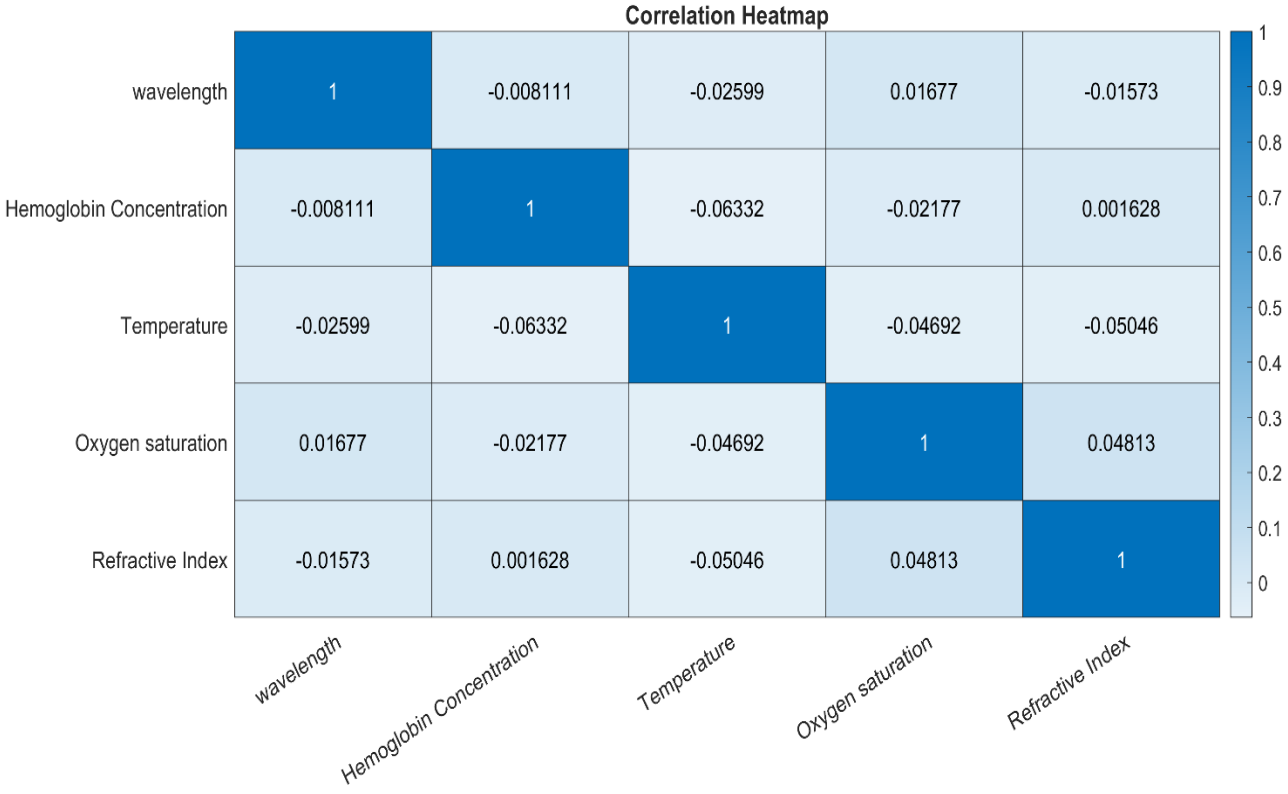


Figure IV. 1. Heatmap of correlation matrix for the RI

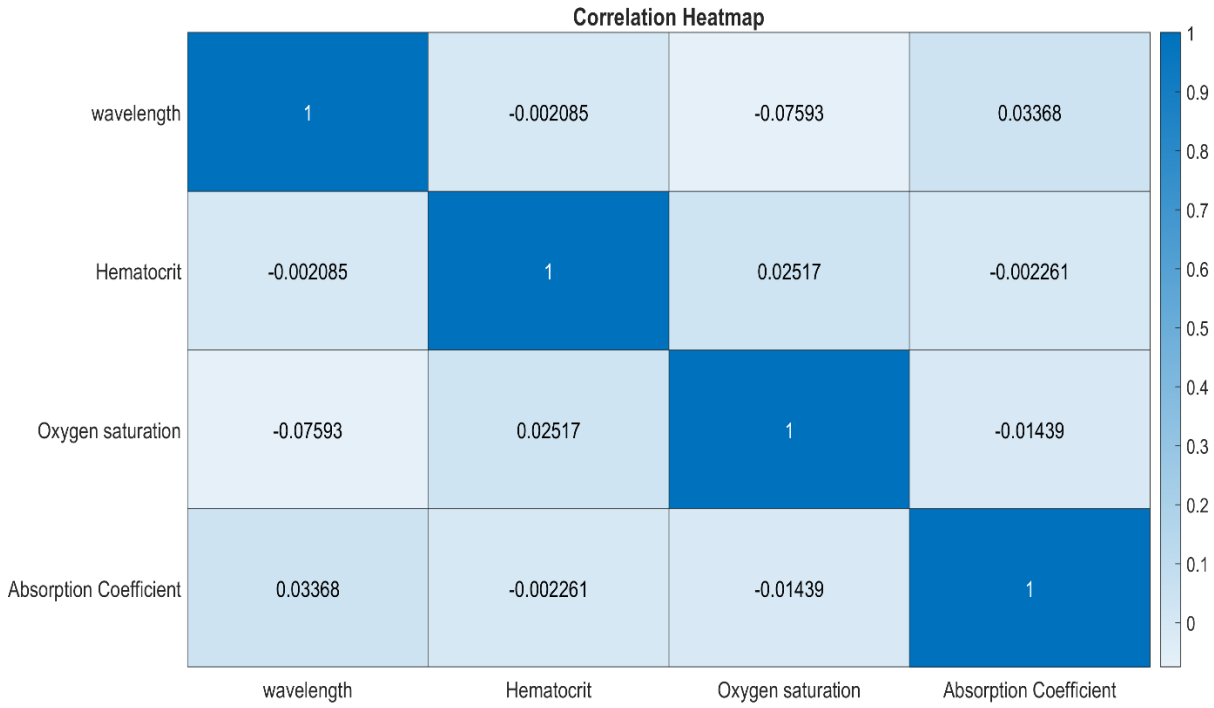


Figure IV. 2. Correlation heatmap for absorption coefficient

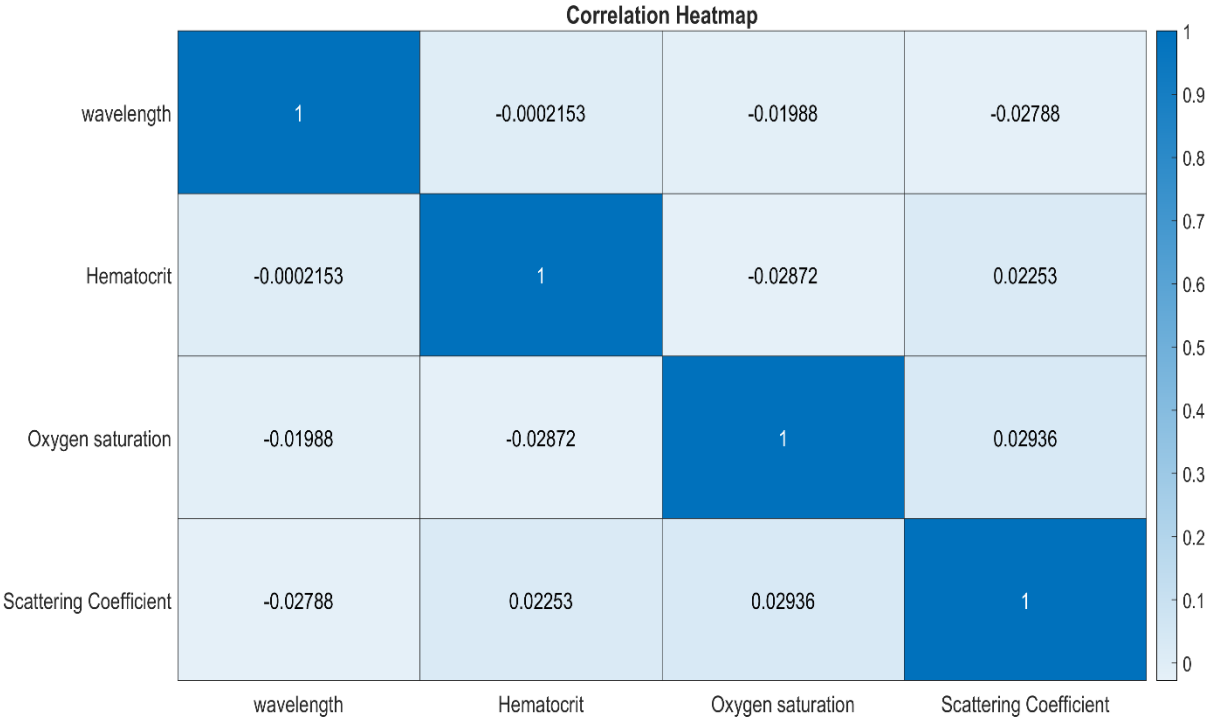


Figure IV. 3. Correlation heatmap for scattering coefficient

IV.2.3. Artificial neural networks (ANN) modeling

There are various processes involved in the construction of the models. The flowchart (Figure IV.4) shows different steps that were followed in the development of MLP-ANN models.

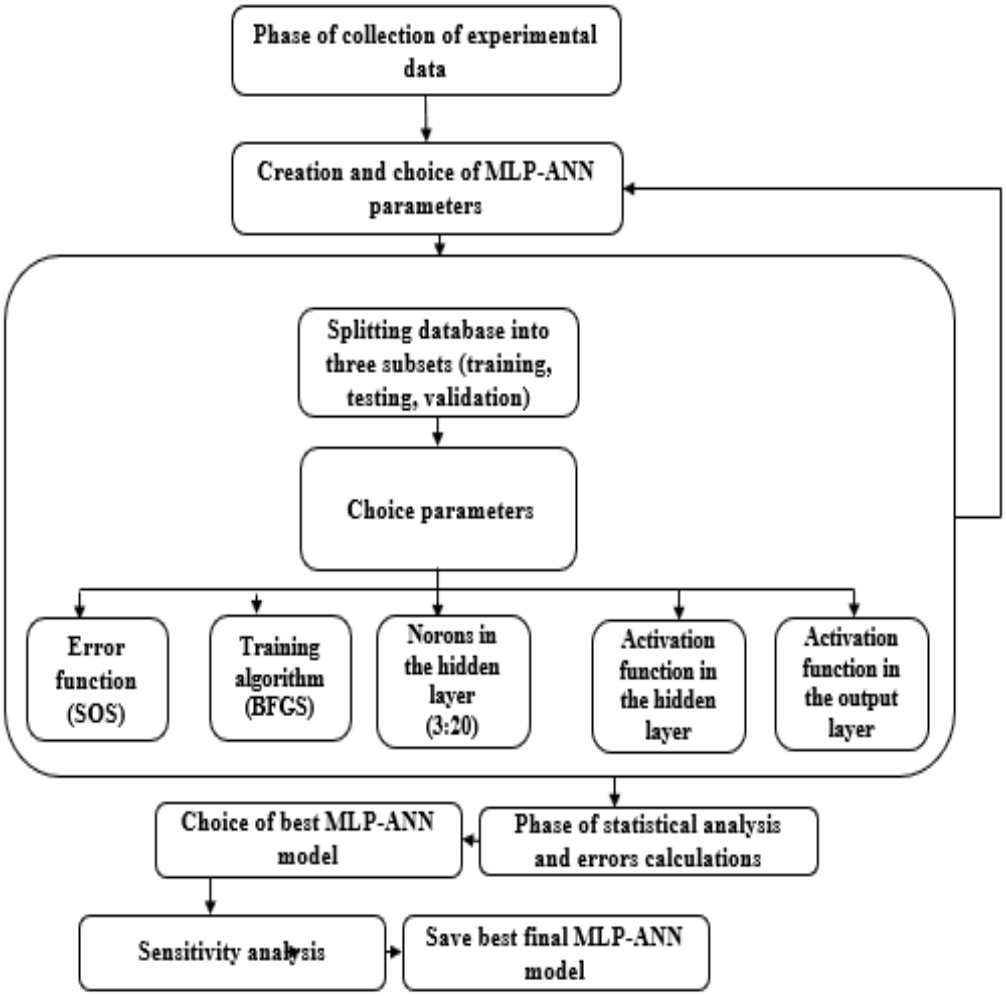


Figure IV. 4. Flowchart describing the development process of ANN used in the current work.

IV.2.3.1. Division of data base

Each database used in this work for the optimization of the MLP-ANN architecture was split into three subsets: training, testing, and validation phase, according to the following division: (60%-20%-20%), (70%-15%-15%) et (80%-10%-10%).

IV.2.3.2. Optimization of ANN architectures

The optimization of the MLP-ANN architectures was carried out by the software "STATISTICA", following these steps:

- ✓ The number of neurons in the hidden layer was varied from 2 to 30 neurons.
- ✓ Quasi Newton back propagation algorithm, specifically Broyden-Fletcher-Goldfarb-Shanno algorithm (BFGS) was used as a training algorithm (Wessel & Jurs 1994).
- ✓ Sum of Squares (*SOS*) has been chosen as an error function.
- ✓ Activation functions are performed by selecting several types of functions in the hidden and output layer. Therefore, at the end of the execution, the software provides the optimal combination of the two functions for activating the relevant layer.
- ✓ Considering the number of hidden layers, we only utilized one hidden layer for all MLP-ANN models since there is only one hidden layer in "STATISTICA'' software, version 10.
- ✓ In order to obtain the best MLP-ANN models, Multiple runs were carried out for each parameter. The execution involves a significant number of iterations between 2000 and 8000 iterations for a single execution.

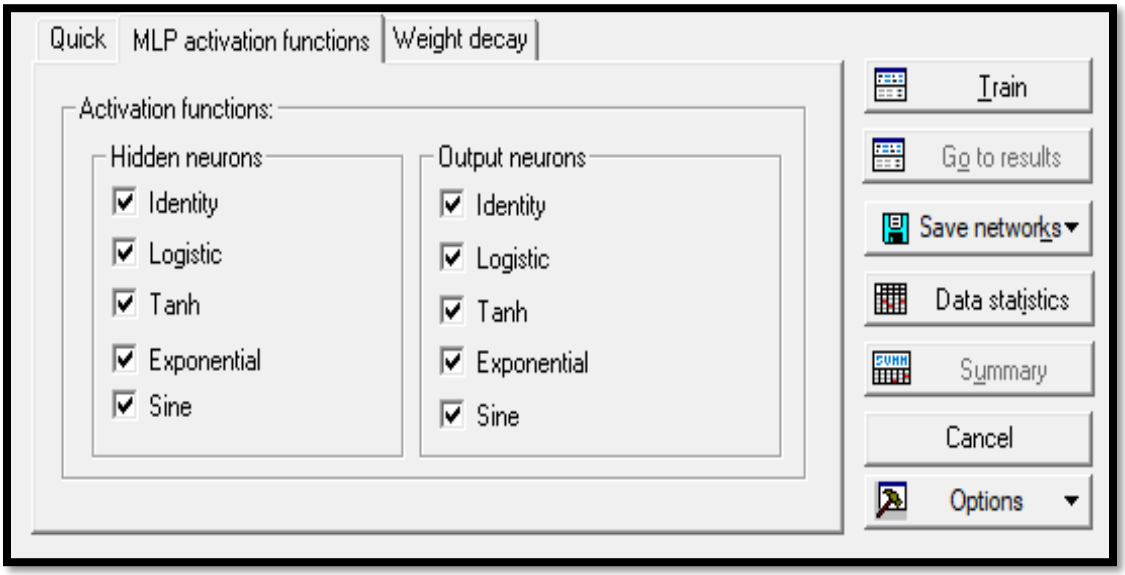


Figure IV. 5. Screenshot of STATISTICA activation functions

IV.2.3.3. Selection of the best model

The choice of best model depends on two criteria: sum of square (*SOS*) and the correlation coefficient (*R*). This choice is made so that SOS is close to 0 and R close to 1, generally values of R greater than 0.5 are considered suitable and values greater than 0.9 are considered excellent. Linear regression parameters are obtained using the MATLAB function "postreg".

IV.2.3.4. Statistical analysis and errors calculation

The performance and accuracy of the best ANN model obtained in this work, was evaluated by multiple types of errors: the root mean square error (*RMSE*), mean absolute percentage error (MAPE), and mean absolute error (MAE). the performances of the regressions were also tested using the regression coefficient (R) and the determination (R^2), in addition, the regression vectors: the values of the slope (α) and the intercept (β). This is done on all of the training, testing and validation data.

The mathematical definition of these types of errors and coefficients are given by the following equations, respectively:

$$\boldsymbol{RMSE = \sqrt{\sum_{i=1}^{N}(y^{exp} - y^{cal})^2/N}} \quad \textbf{(IV.01)}$$

$$\boldsymbol{MAPE = \left(\frac{100}{N}\right) * \sum_{i=1}^{N}\left(\left|y^{exp} - y^{cal}\right|/y^{exp}\right)} \quad \textbf{(IV.02)}$$

$$\boldsymbol{MAE = \left(\frac{1}{N}\right) * \sum_{i=1}^{N}\left(\left|y^{exp} - y^{cal}\right|/y^{exp}\right)} \quad \textbf{(IV.03)}$$

$$\boldsymbol{R = \frac{\sum_{i=1}^{N}[(y_i^{exp} - \hat{y}^{exp})(y_i^{cal} - \hat{y}^{cal}]}{\sqrt{\sum_{i=1}^{N}(y_i^{exp} - \hat{y}^{exp})^2}\sqrt{\sum_{i=1}^{N}(y_i^{cal} - \hat{y}^{cal})^2}}} \quad \textbf{(IV.04)}$$

$$\boldsymbol{R^2 = 1 - \sum_{i=1}^{N}(y_i^{exp} - y_i^{cal})^2/\sum_{i=1}^{N}(\hat{y}^{exp} - \hat{y}^{exp})^2} \quad \textbf{(IV.05)}$$

$$\boldsymbol{\hat{y}^{exp} = \sum_{i=1}^{N} y^{exp}/N} \quad \textbf{(IV.06)}$$

where N is the number of data points for each phase (training, testing, and validation), y^{exp} is the experimental value, y^{cal} is the predicted values, and $\hat{y}^{exp}$ is the mean of experimental values.

IV.2.3.5. Sensitivity analysis using weight method

Sensitivity analysis by weight method is a powerful technique used for evaluating mathematical models (Ammi et al., 2015; Roubehie Fissa et al., 2019). It is used for many purposes including model validation, evaluating the behavior of various model variables and confirming the choice of model variables by determining the contribution of each chosen variable to the model. Hence, it is important to use this analysis on the MLP-ANN model (Hamdache et al., 2016).

Sensitivity analysis was performed using the "weight" method. This method, proposed first by Garson (Garson, 1991), and resumed by Goh (Goh, 1995), gives a quantification of the relative importance (*Ri*) of the different inputs on the output of NN (Gevrey et al., 2003). It is based on the partitioning of connection weights into:

- Weight of connections input–hidden $\boldsymbol{W_{ji}^{I}}$
- Weight of connections hidden – output $\boldsymbol{W_{1j}^{H}}$

the detailed procedures of this method are described in the studies of (Khaouane, 2013; (Ammi et al., 2015; Roubehie Fissa, 2021)).

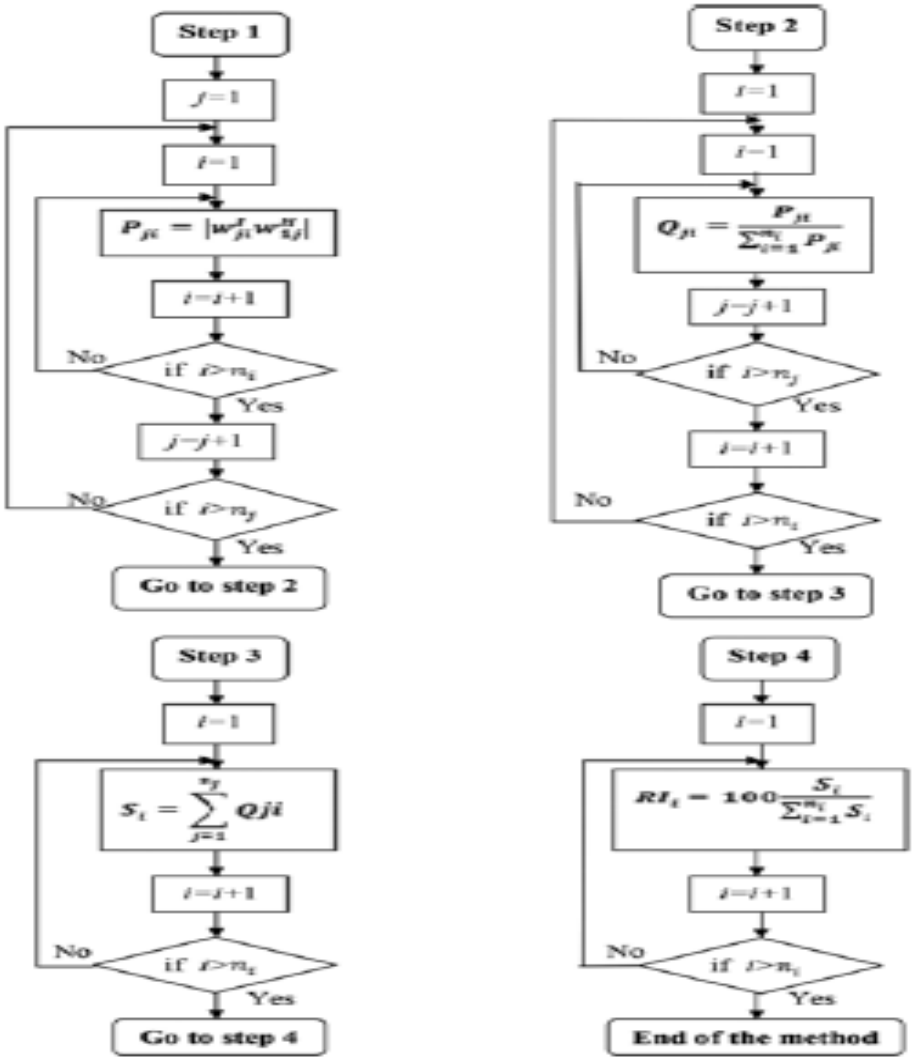


Figure IV. 6. Diagram illustrating the sensitivity analysis by weight method (Khaouane, 2013)

IV.2.4. Support Vector Machine (SVM) modeling

In this part, the SVM algorithm is used for the nonlinear modeling of each optical property of blood depending on physiological and biochemical parameters relying on the same experimental data used to develop the ANN models, the data distribution is done randomly and divided into 2 phases: 80% for training, and the remaining 20% for testing phase. The SVM models generate a non-linear relationship between the inputs and the output of every dataset. The use of a kernel function is essential to implicitly map learning patterns into the feature space. A kernel function can be any mathematical function that satisfies the Mercer's theorem, and there are a lot of kernel functions that can be found in the literature. To achieve a more reliable predictive model, the modeler systematically tests the regularization parameters during the optimization process using a Matlab program and statistical error formulas. However, this process can be challenging and may lead to overfitting the model selection criterion (usually based on cross-validation) when adjusting multiple hyperparameters. The radial basis Gaussian function (RBGF) has been selected as the kernel function for the current SVM modeling (**Scholkopf and Smola 2001; Min and Lee 2005; Keskes et al. 2020**). The mathematical expression of the kernel function under test is below:

$$k(x_i, x_j) = \exp(-\gamma . | x_i, x_j |^2)$$

where γ, is the kernel parameter

and the flowchart of the SVM method is shown in figure IV.7

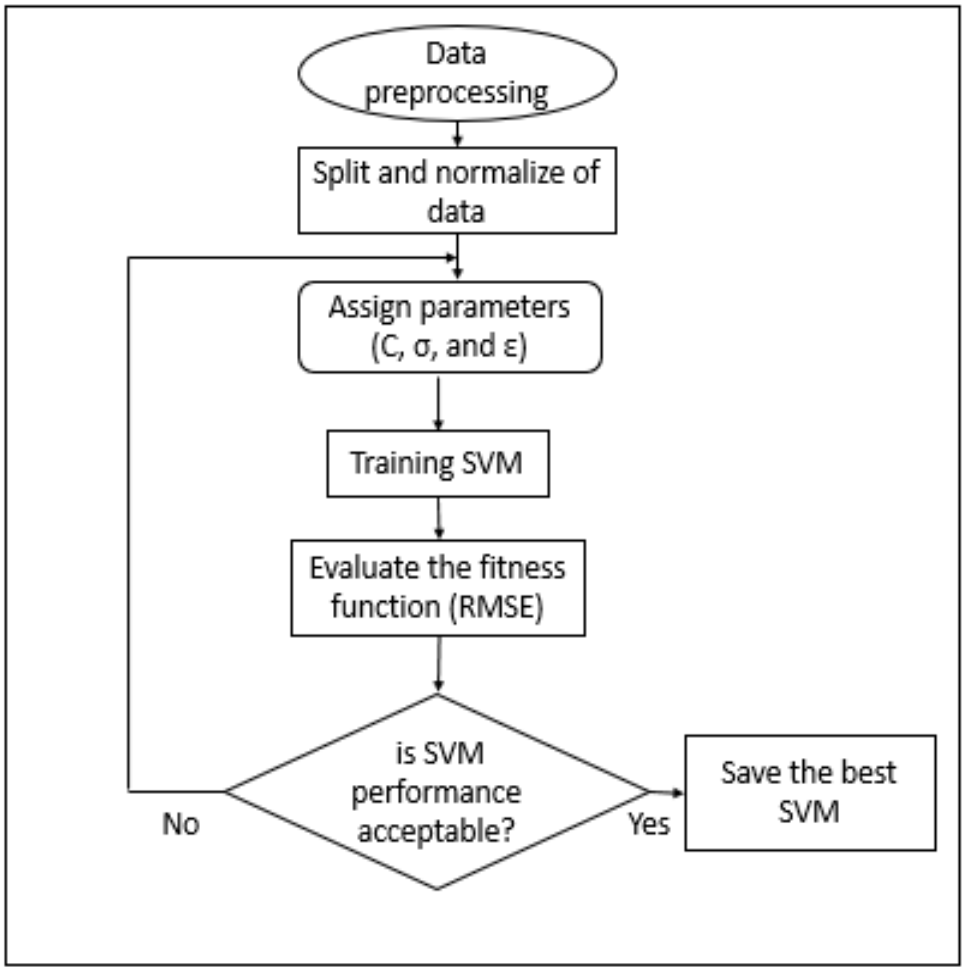


Figure IV. 7. Flowchart of SVM

There are three factors that require to be carefully considered in order achieve a reliable SVM model with strong prediction capacity. The penalty parameter C, the non-sensitivity coefficient ε (bandwidth of the Gaussian kernel RBF σ), and the kernel parameters are among these parameters, which are referred to as the hyper-parameters of the SVM model (X. Li et al. 2013; Amroune, Bouktir, and Musirin 2018). Therefore, a trial and error test is used to obtain the best values of, C, σ and ε.

The ranges of variation for the hyper-parameters employed in this modeling are shown in Table IV.5

Table IV. 5. Ranges of initial settings for SVM

	SVM hyper-parameters	C	ε	σ
	Lower bound			
	Upper bound			
SVM	Number of agents	50		
	Maximum iteration number	100		
	Training phase (%)	80		
	Testing phase (%)	20		

The statistical parameters that are used to evaluate the performance of these models are the coefficient of correlation R and the root mean square error RMSE.

IV.2.5. Modeling by Support Vector Machine coupled with Dragonfly Algorithm (DA–SVM)

The hybrid DA-SVM model is developed in this section using the same database that had been adapted to construct the ANN and SVM models, while the proposed model is based on the SVM learning algorithm coupled with the dragonfly algorithm (DA) for the optimization of its hyper-parameters. The data sharing dimensions for model training and testing, as well as the hyper-parameter ranges are the same as those described for the SVM models. The procedures followed by the proposed DA-SVM are presented in the flowchart (Figure IV.8).

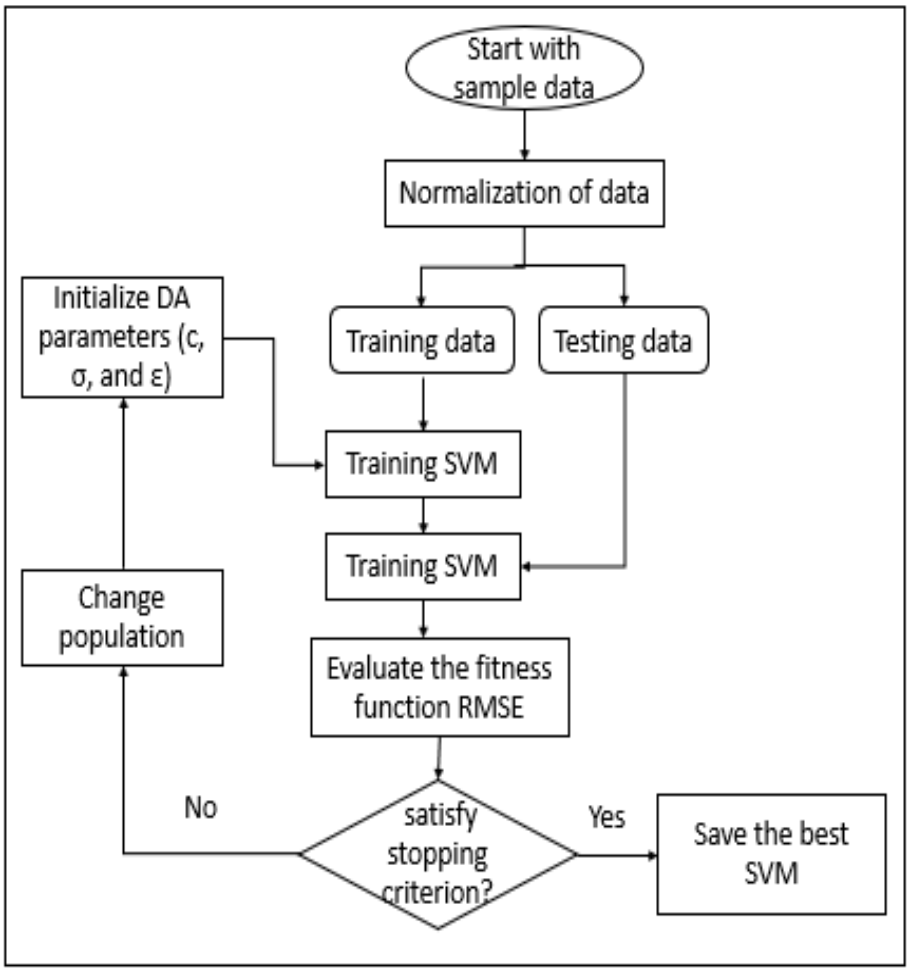


Figure IV. 8. Flowchart of DA-SVM process

The DA algorithm feeds the SVM with random hyper-parameter values within their initially defined ranges, however, the basic operation of the support vector machine learning algorithm remains the same.

At first, setting the values of the DA parameters including: number of dragonflies, maximum number of iterations, upper and lower bounds of C, σ and ε, initialize the step vectors ΔXi (i = 1, 2, ..., n), the values of s, a, c, f, e then initialize the inertia weight (w) and the values of the SVM hyper-parameters (C, σ and ε), following that train the SVM model with the training set and calculate the physical value of each dragonfly. The root mean square error (RMSE) was employed as the fitness function.

The reliability and accuracy of the models proposed in this study were evaluated using commonly used statistical performance criteria, such as the root mean square error (RMSE), mean absolute

error (MAE), mean absolute percentage (MAPE), correlation coefficient (R), and the determination coefficient (R^2).

Chapter V: Results and Discussion

V. Chapter V: Results and Discussion

V.1. Introduction

Our efforts have produced promising results in the discipline of blood optical property modeling, some of which have already been published in "Computer Methods in Biomechanics and Biomedical Engineering." This chapter marks the pivotal juncture where we dissect and engage in a comprehensive discussion of the results we have obtained. To ensure clarity and coherence, the results are thoughtfully organized into distinct sections, allowing us to delve into their implications, intricacies, and their significant contributions to the field.

In the first part, we will present the performances of ANN models. Subsequently, we will present the results of the sensitivity analysis using the "weight" method. A comparison between the experimental and ANN calculated values is then presented. Then, a discussion of the results obtained from the SVM and DA-SVM models. the last part is devoted to a comparative study of ANN models with SVM, and DA-SVM. And with the models exist in the literature.

V.2. MLP-ANN models (*Contribution I*)

In this work, each database (DB1 to predict the refractive index, DB2 for the property of absorption, and DB3 for the scattering coefficient) was randomly divided into three subsets: training, testing, and validation phase with different percentages. Various division and iteration were done in order to obtain the best ANN models, the results are available in the annex, and the best division obtained is as follow: 60% for training, 20% for testing and 20% for validation (Figure V.1). The structure of the optimized MLP-ANN models is in Table V.1.

Table V. 1. Structure of optimized MLP-ANN models

MLP-ANN models	Training Algorithm	Input layer	Hidden layer		Output layer	
		Number of neurons	Number of neurons	Activation function	Number of neurons	Activation function
MLP-ANN1	Quasi-Newton (BFGS) 645	4	20	Tanh	1	Exponential
MLP-ANN2	Quasi-Newton (BFGS) 419	3	23	Tanh	1	Sine
MLP-ANN3	Quasi-Newton (BFGS) 571	3	25	Tanh	1	Exponential

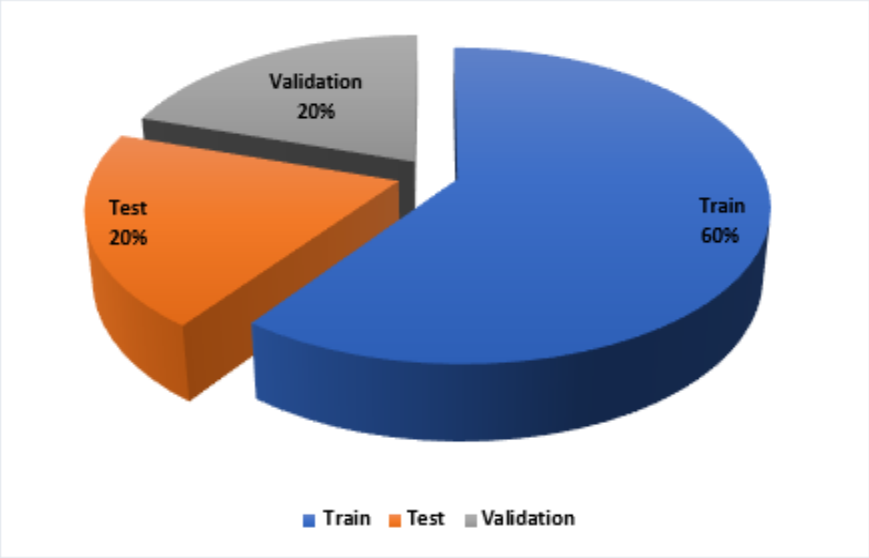
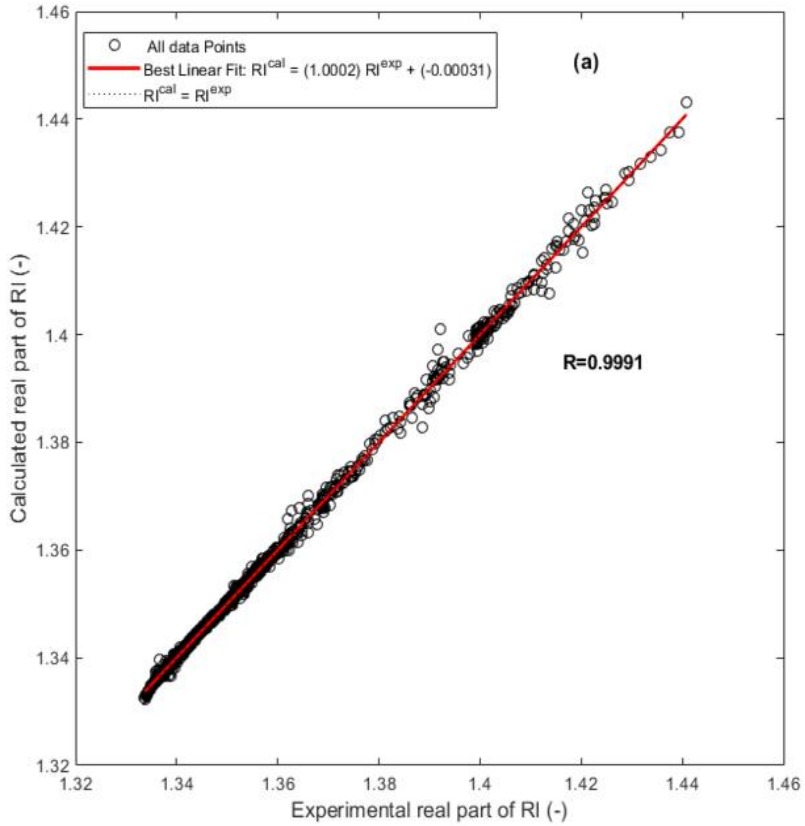
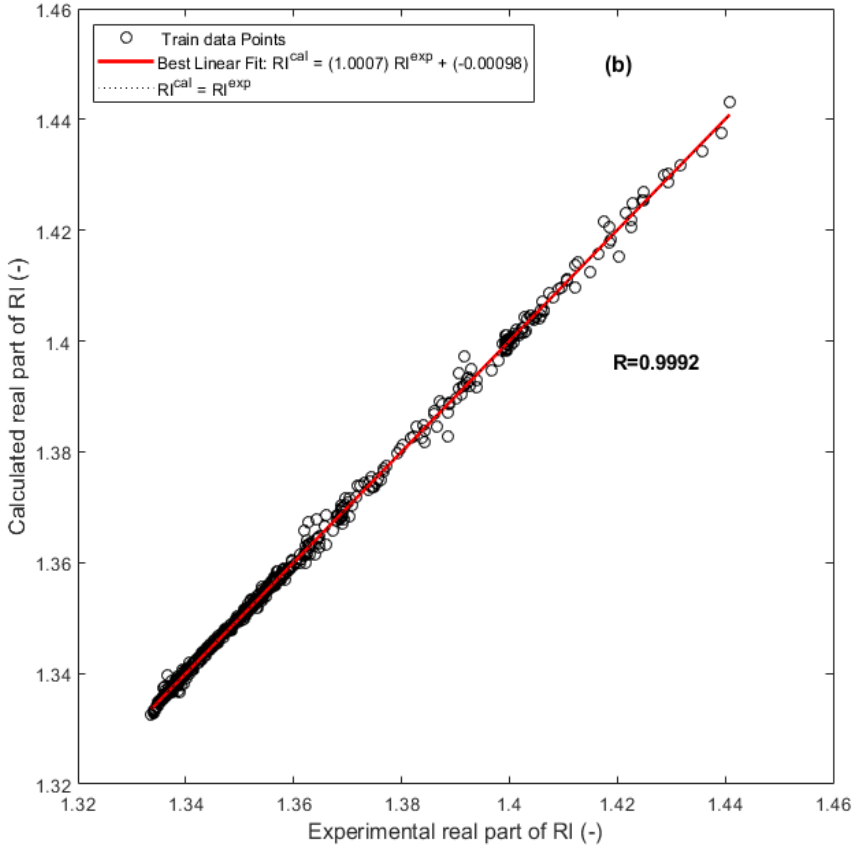


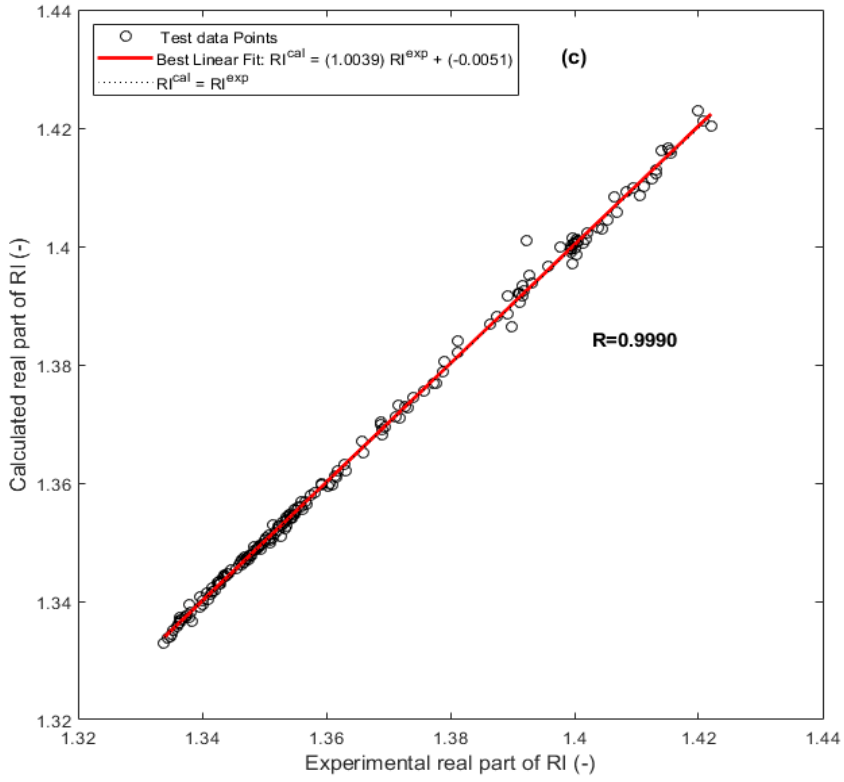
Figure V. 1. Representation of best percentages division of the data bases

As mentioned before, three feed-forward neural network models have been developed for the purpose of predicting the optical properties of human blood. The MATLAB function "postreg" is used to obtain the plot and the regression parameters: correlation coefficient (R), the slope (α), and (β) the intersection with axis. The correlation coefficient can be used to assess the prediction accuracy of the model. Generally, R value that's greater than 0.5 is regarded as good, and R value larger than 0.9 is considered as excellent (Golbraikh et Tropsha 2002).

The linear regression curves of the refractive index calculated by ANN with the experimental refractive index are illustrated in figure V.2, where the regression vectors appeared close to the ideal ([α, β, R] = [1.0007, -0.0009, 0.9992] for training phase; [α, β, R] = [1.0039, -0.0051, 0.9990] for testing phase; [α, β, R] = [0.9942, 0.0078, 0.9991] for validation phase and [α, β, R] = [1.002, -0.0003, 0.9991] for the total phase). Giving that the learning phase, test, validation, and overall phase have correlation coefficients more than 0.9 ($0.90 \leq R \leq 1.00$), they are all considered as excellent in this case, it demonstrates the good robustness of the first neural model and the potential to accurately predict the real part of the refractive index of blood.




Train data Points
Best Linear Fit: RI^cal = (1.0007) RI^exp + (-0.00098)
RI^cal = RI^exp
(b)
R=0.9992
Calculated real part of RI (-)
Experimental real part of RI (-)
1.32
1.34
1.36
1.38
1.4
1.42
1.44
1.46


Test data Points
Best Linear Fit: RI^cal = (1.0039) RI^exp + (-0.0051)
RI^cal = RI^exp
(c)
R=0.9990
Calculated real part of RI (-)
Experimental real part of RI (-)
1.32
1.34
1.36
1.38
1.4
1.42
1.44

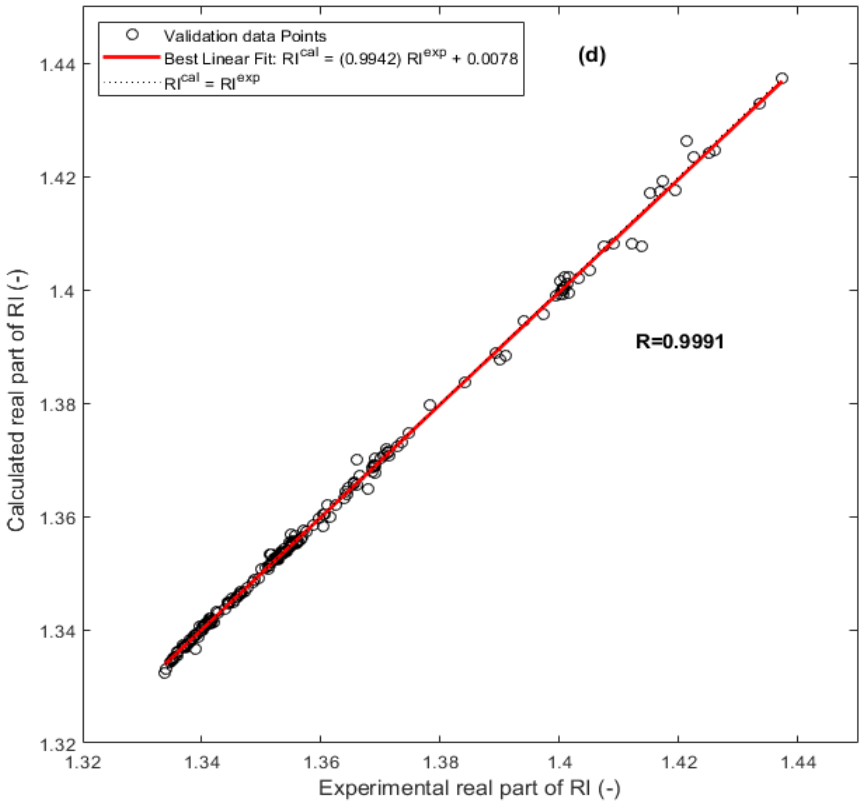


Figure V. 2. Comparison of experimental and ANN calculated values of the refractive index: (a) total phase, (b) training phase, (c) testing phase, (d) validation phase

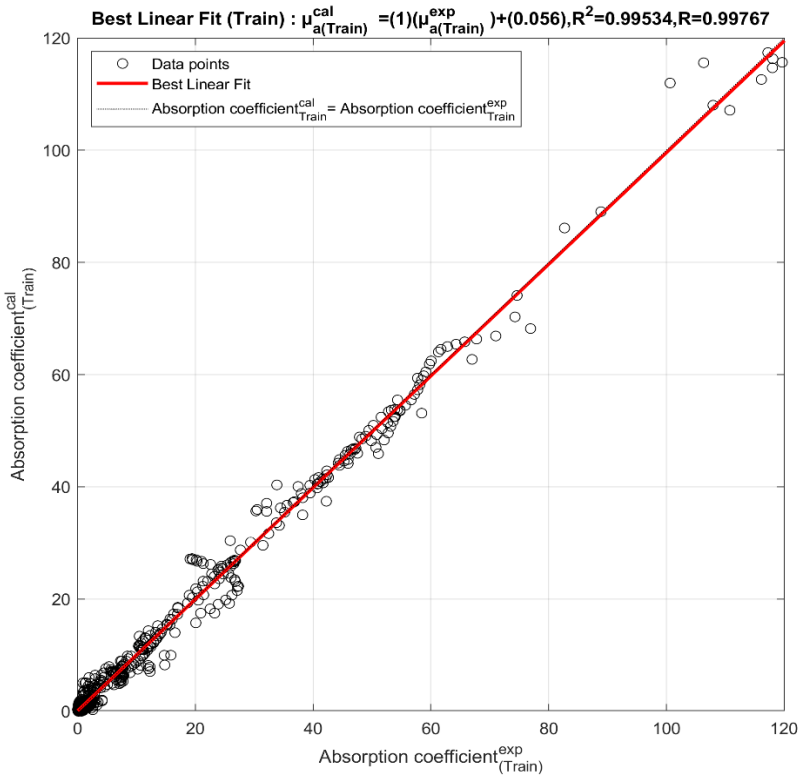
.

The figure V.3 displays the linear regression curves that compare the absorption coefficients calculated using the ANN model with the experimental absorption coefficients. The regression vectors, represented by the values [α, β, R], indicate how well the calculated absorption coefficients align with the experimental ones.

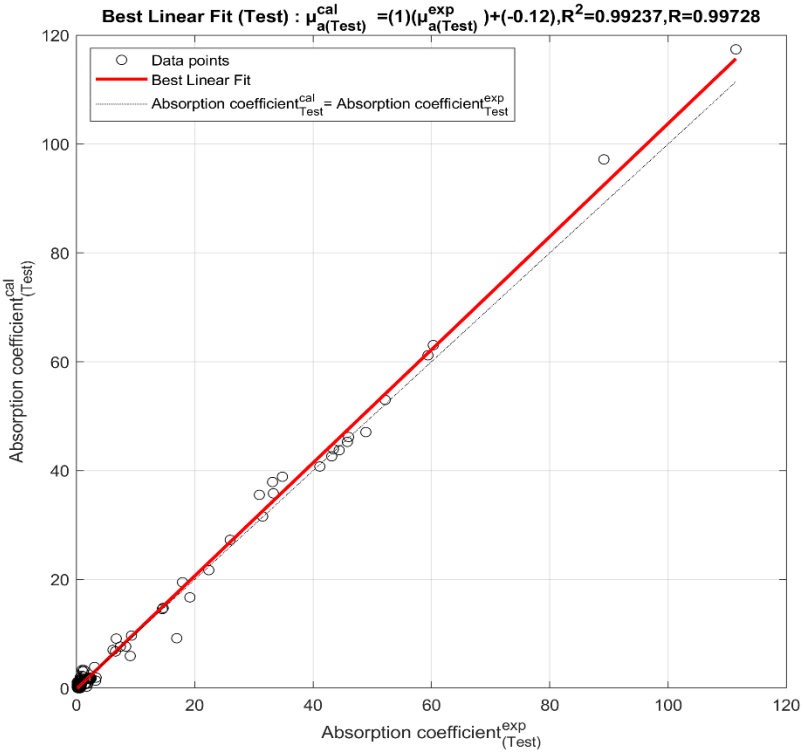
In the training phase (a), the regression vector [α, β, R] = (1, 0.056, 0.9976) is observed. The slope (α) of 1 indicates that the ANN's predicted absorption coefficients exhibit a strong linear correspondence with the experimental values. The small positive value of 0.056 for the y-intercept (β) suggests a minor upward shift in the ANN's predictions compared to the experimental data. The high correlation coefficient (R) of 0.9976 indicates that the ANN's calculated absorption coefficients are highly consistent with the experimental values during training. In the testing phase (b), the regression vector [α, β, R] = (1, -0.12, 0.9972) is observed. The slope of 1 still indicates a strong linear agreement between the ANN's predictions and the experimental data. The negative value of -

0.12 for the y-intercept suggests that the ANN's predicted values are slightly lower than the experimental data. The high correlation coefficient of 0.9972 highlights the substantial alignment between the two datasets during testing. The validation phase (c) exhibits a regression vector [α, β, R] = (1, -0.12, 0.9972), which mirrors the regression vector of the testing phase. This indicates that the relationship between the ANN's predictions and the experimental data remains consistent during validation. In the total phase (d), the regression vector [α, β, R] = (1, 0.026, 0.9974) is observed. The slope of 1 reaffirms the strong linear correspondence between the ANN's predictions and the experimental absorption coefficients. The small positive value of 0.026 for the y-intercept indicates a slight upward shift in the ANN's predictions compared to the experimental data. The high correlation coefficient of 0.9974 underscores the overall robust fit between the datasets in the total phase.

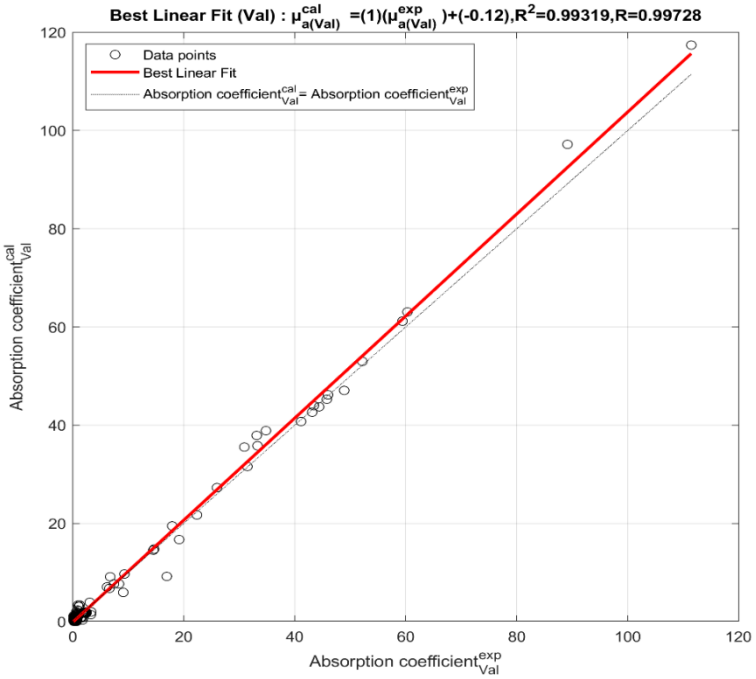
This indicates that the ANN model is performing well in estimating the absorption coefficients of blood based on the experimental data.



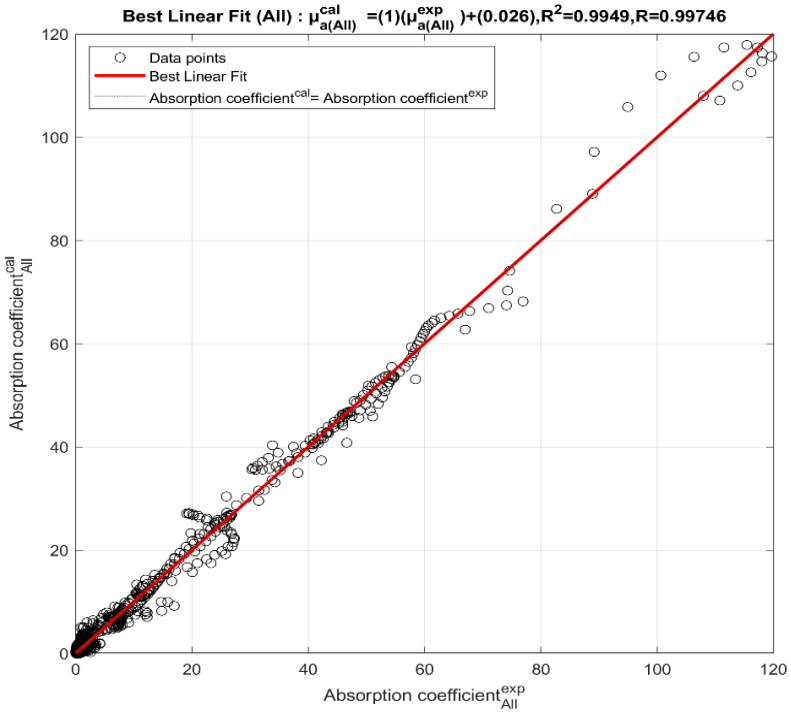
(a)


Best Linear Fit (Test) : $\mu_{a(Test)}^{cal}$ =(1)($\mu_{a(Test)}^{exp}$)+(-0.12),R^2=0.99237,R=0.99728
Data points
Best Linear Fit
Absorption coefficient$_{Test}^{cal}$ = Absorption coefficient$_{Test}^{exp}$
Absorption coefficient$_{(Test)}^{cal}$
Absorption coefficient$_{(Test)}^{exp}$
0
20
40
60
80
100
120

(b)


Best Linear Fit (Val) : $\mu_{a(Val)}^{cal}$ =(1)($\mu_{a(Val)}^{exp}$)+(-0.12),R^2=0.99319,R=0.99728
Data points
Best Linear Fit
Absorption coefficient$_{Val}^{cal}$= Absorption coefficient$_{Val}^{exp}$
Absorption coefficient$_{Val}^{cal}$
Absorption coefficient$_{Val}^{exp}$
0
20
40
60
80
100
120

(c)



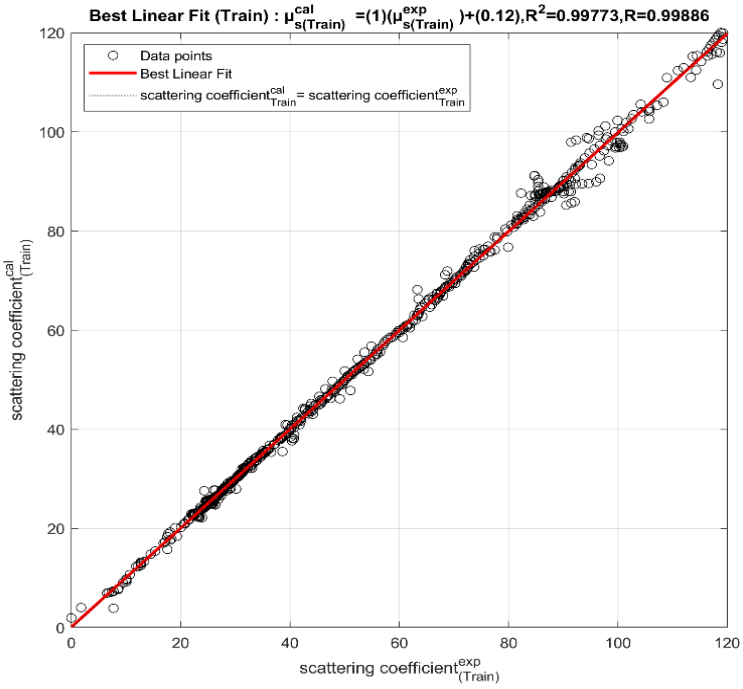
(d)

Figure V. 3. Comparison of experimental and ANN calculated values of the absorption coefficient μa: (a) training phase, (b) testing phase, (c) validation phase, (d) total phase

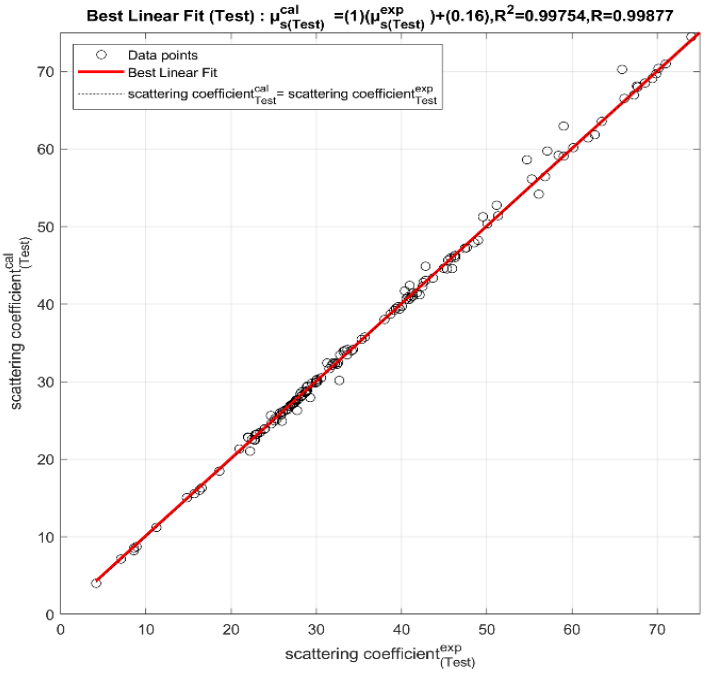
The linear regression curves of the scattering coefficient calculated by (ANN) with the experimental scattering coefficient are in the figures below (FigureV.4). The regression vector [α, β, R] = (1, 0.12, 0.9986) appears during the training phase (a). The predicted scattering coefficients generated by the ANN show a strong linear correlation with the experimental values, as indicated by the slope (α) of 1. The experimental data and the positive value of 0.12 for the y-intercept (β) indicate a minor upward change in the ANN's predictions. The two datasets had a very strong alignment during training, as indicated by the high correlation coefficient (R) of 0.9986. The results obtained during the testing phase and validation phase closely mirror those of the training phase. In both the testing (b) and validation (c) phases, we observed the regression vector [α, β, R] = (1, 0.16, 0.9987) that are nearly identical to those obtained during training. This high degree of consistency across these phases suggests that the model's performance is stable and generalizes well, demonstrating its robustness and reliability in estimating the scattering coefficients based on the provided data.

The scattering coefficient, calculated using the Artificial Neural Network (ANN) and compared to experimental values, demonstrates a remarkably strong alignment. Across all phases: training, testing, validation, and total. The correlation coefficient (R) consistently surpasses 0.998, affirming

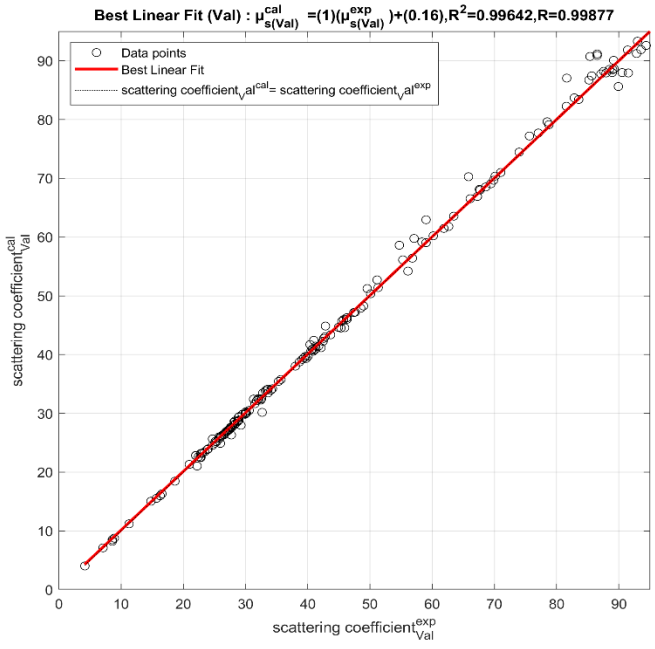
the exceptionally high degree of fit between the ANN's calculations and the experimental scattering coefficients. This signifies that the ANN model excels in accurately estimating the scattering coefficients of blood based on empirical data.



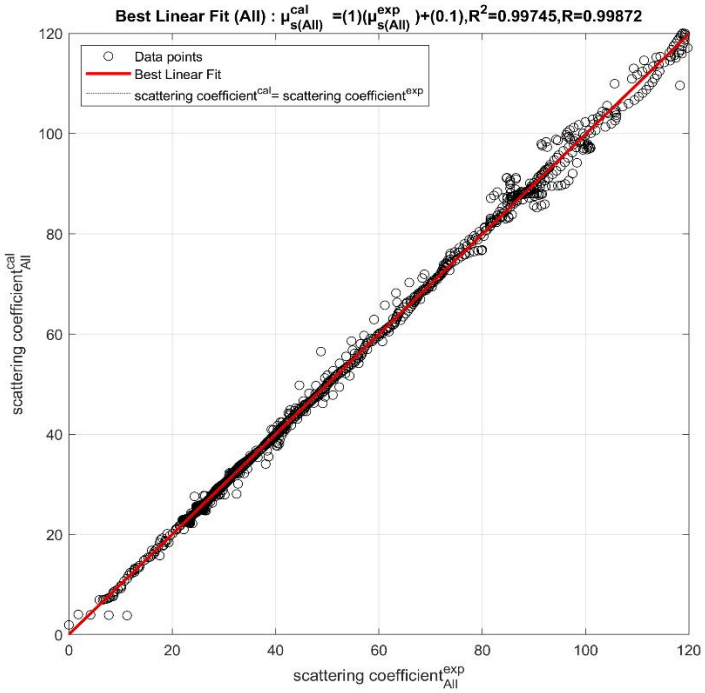
(a)


Best Linear Fit (Test) : $\mu_{s(Test)}^{cal}$ =(1)($\mu_{s(Test)}^{exp}$)+(0.16),R^2=0.99754,R=0.99877
Data points
Best Linear Fit
scattering coefficient$_{Test}^{cal}$= scattering coefficient$_{Test}^{exp}$
scattering coefficient$_{(Test)}^{cal}$
scattering coefficient$_{(Test)}^{exp}$
0
10
20
30
40
50
60
70

(b)


Best Linear Fit (Val) : $\mu_{s(Val)}^{cal}$ =(1)($\mu_{s(Val)}^{exp}$)+(0.16),R^2=0.99642,R=0.99877
Data points
Best Linear Fit
scattering coefficient$_{Val}^{cal}$= scattering coefficient$_{Val}^{exp}$
scattering coefficient$_{Val}^{cal}$
scattering coefficient$_{Val}^{exp}$
0
10
20
30
40
50
60
70
80
90

(c)



(d)

Figure V. 4. Comparison of experimental and ANN calculated values of the scattering coefficient μs: (a) training phase, (b) testing phase, (c) validation phase, (d) total phase

V.2.1. Comparison between the experimental refractive index and calculated by MLP-ANN

The three figures below provide a comprehensive evaluation of the ANN model's performance in predicting the refractive index of blood based on varying inputs.

The figure V.5 illustrates a compelling comparison between the experimental and predicted refractive index values generated by the ANN model as a function of the wavelength, a high level of agreement between the model's predictions and the actual experimental results may be observed in how closely the curves representing experimental and predicted outcomes overlap. The curves' startling resemblance demonstrates how effectively the ANN model captures the underlying patterns and trends in the data. The absence of substantial deviations or significant differences between the experimental and predicted values suggests that the model's predictions are reliable and accurate.

The experimental and calculated values in figure V.6, where the input variable is the concentration of hemoglobin, reflect a startling consistency. The curves depicting the experimental and predicted refractive index values exhibit remarkable overlap, reaffirming the model's ability to accurately estimate the refractive index of blood as a function of hemoglobin concentration.

A similar pattern can be seen in Figure V.7, which compares experimental and estimated refractive index values as a function of temperature. The distribution of the experimental and predicted values appears nearly identical, with minimal deviations between the two datasets. The strong performance of the ANN model in estimating the refractive index based on changes in temperature is emphasized by the good agreement between empirical and model-generated results.

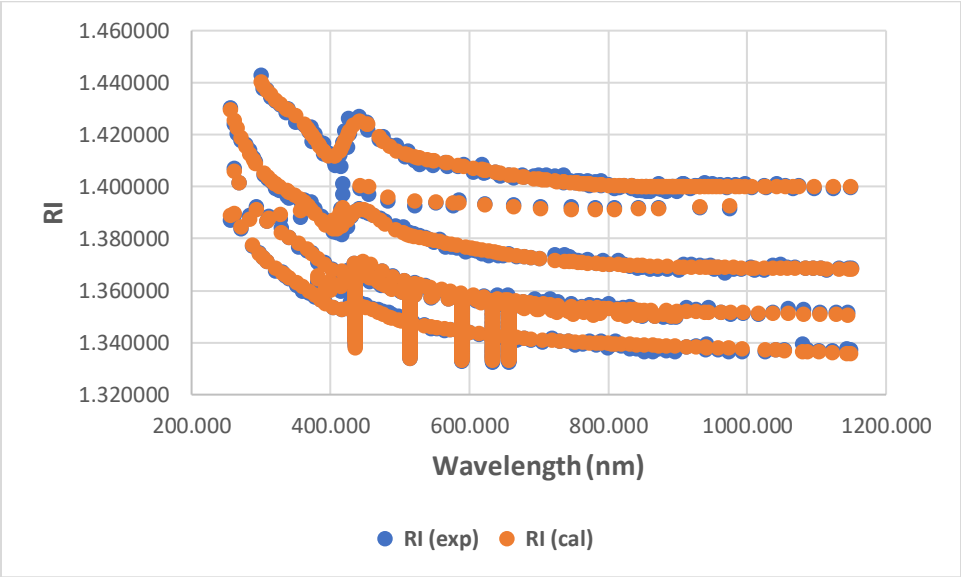


Figure V. 5. Comparison between experimental and predicted values of RI vs wavelength

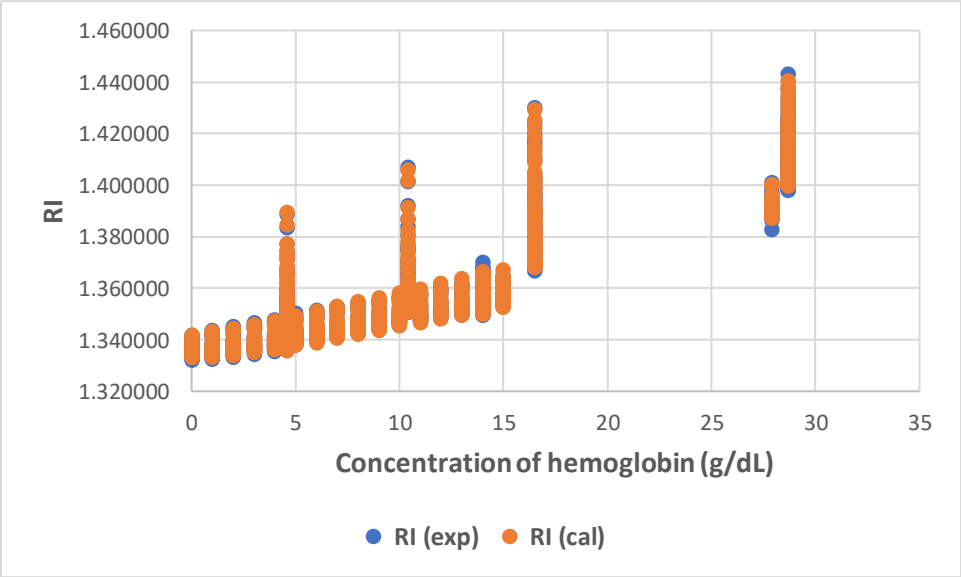


Figure V. 6. Comparison between experimental and predicted values of RI vs concentration of Hb

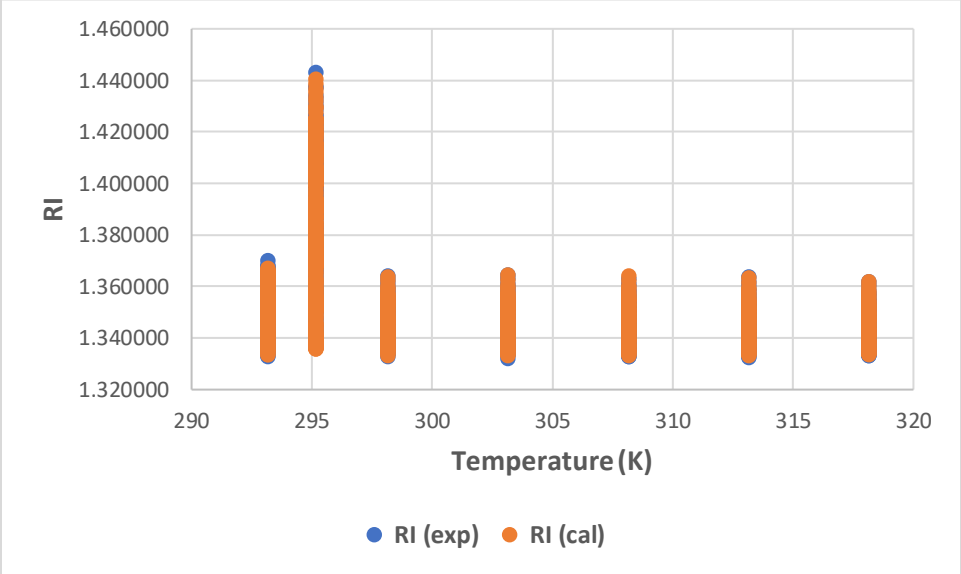


Figure V. 7. Comparison between experimental and predicted values of RI vs T

Table V. 2. Linear regression vectors from the linear

		α	β	R
MLP-ANN1	Training phase	1.0002	-0.00031	0.9992
	Testing phase	0.0007	-0.00098	0.9990
	Validation phase	1.003	-0.0051	0.9991
	Total	0.9942	0.0078	0.9991
MLP-ANN2	Training phase	1	0.056	0.9976
	Testing phase	1	-0.12	0.9972
	Validation phase	1	-0.12	0.9972
	Total	1	0.026	0.9974
MLP-ANN3	Training phase	1	0.12	0.9986

	Testing phase	1	0.16	0.9987
	Validation phase	1	0.16	0.9987
	Total	1	0.1	0.9987

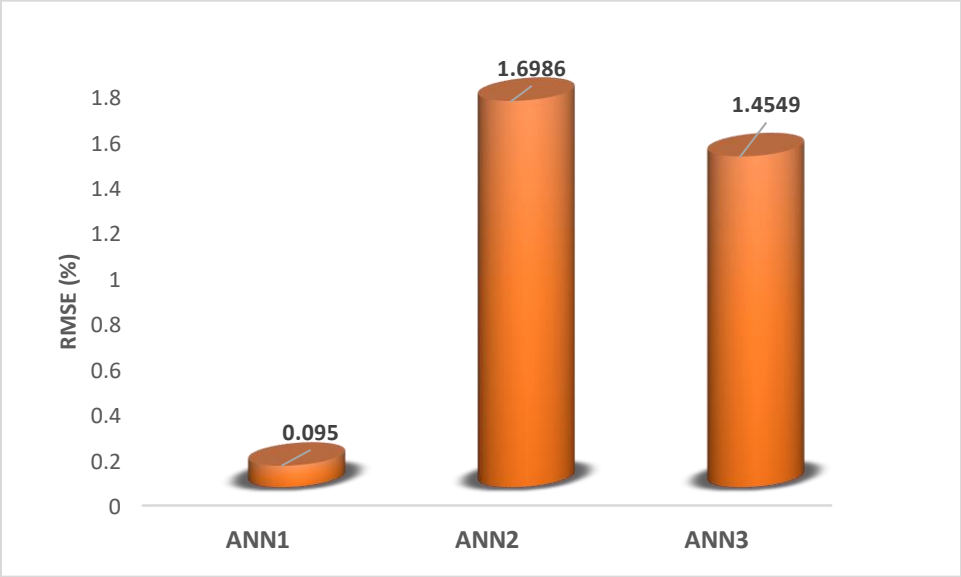
In the comparative analysis illustrated in the figure V.8, the performance of the ANN models was evaluated based on their RMSE (Root Mean Square Error) values across the different phases of this study:

ANN1, which predict the refractive index of blood, has the lowest RMSE value of 0.095 throughout the training phase, demonstrating its exceptional accuracy in training from the data. The prediction accuracy of ANN2, which focused on the absorption coefficient, was lower during training, as seen by its significantly larger RMSE of 16.986. ANN3, targeting the scattering coefficient, showed a moderate RMSE of 1.4549 in the training phase (a).

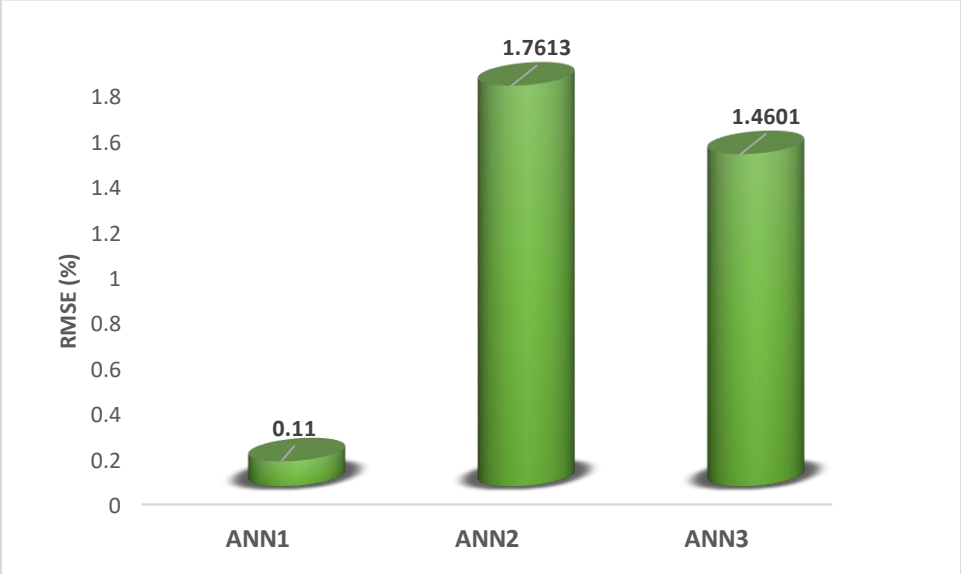
Moving to the testing phase (b), the performance of the models becomes evident. ANN1 maintained a low RMSE of 0.11, demonstrating its effectiveness in predicting the refractive index. The RMSE values for ANN2 and ANN3 were higher, at 1.7613 and 1.4601, respectively, showing that their predictions were less accurate when applied to new, testing data.

A similar pattern showed up during the validation phase (c). The RMSE of 0.11, signifying its consistency in accurate predictions. However, the RMSE values for ANN2 and ANN3 were higher, at 2.1438 and 1.7926, respectively, indicating decreased accuracy in this phase.

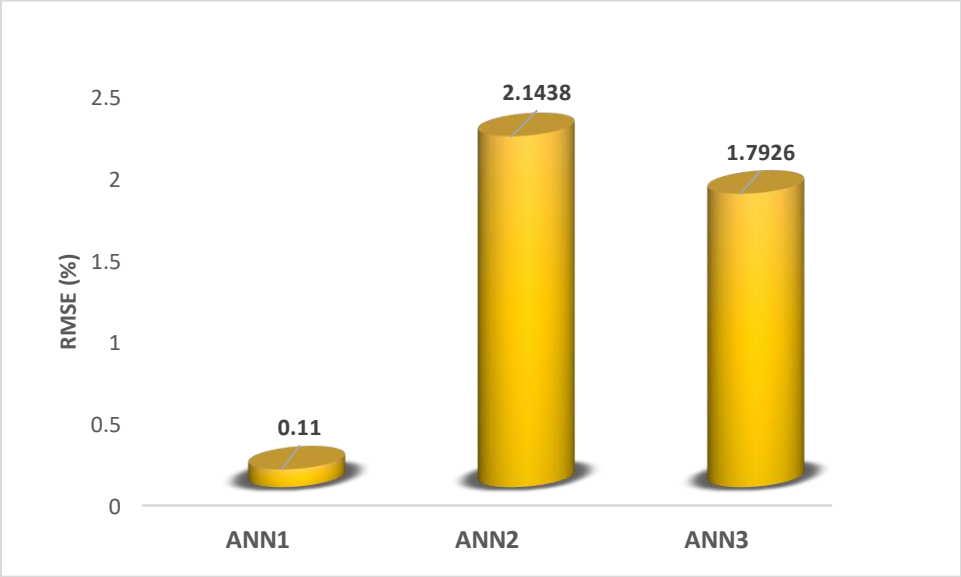
When considering the total phase (d), ANN1 once more showed the lowest RMSE of 0.1, reaffirming its strong predictive capability. While still performing moderately, ANN2 and ANN3 displayed RMSE values of 1.7544 and 1.5294, respectively. These RMSE values demonstrate that ANN1, notably in predicting the refractive index, consistently outperforms the other models across various phases of our study. However, ANN2 and ANN3 exhibit acceptable predictive capabilities.


1.6986
1.4549
0.095
RMSE (%)
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
ANN1
ANN2
ANN3

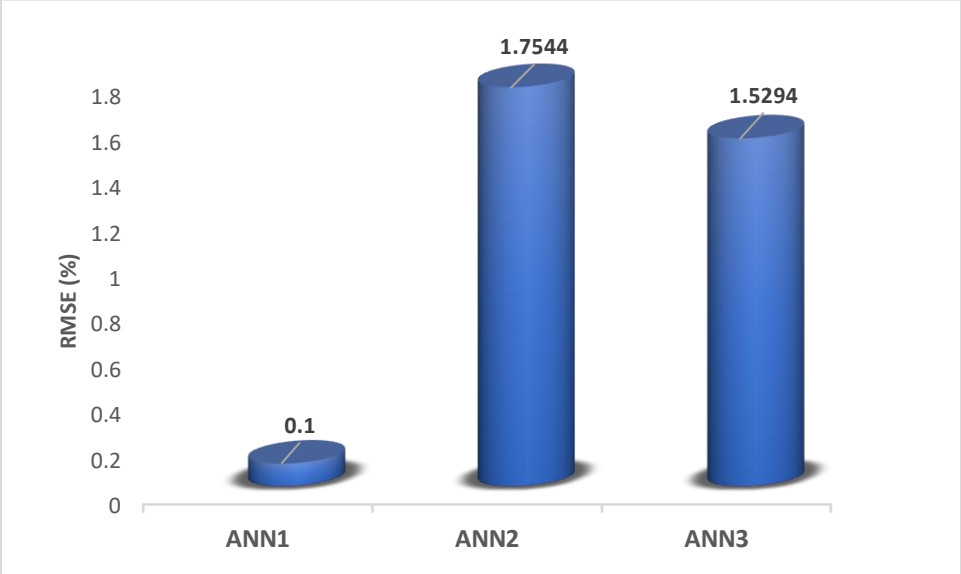
(a)


1.7613
1.4601
0.11
RMSE (%)
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
ANN1
ANN2
ANN3

(b)



(c)



(d)

Figure V. 8. Performance of MLP-ANN models for: (a) training phase, (b) testing phase, (c) validation phase, and (d) total phase

V.2.2. Mathematical expressions of the optimized neural models

$$Z_j = f^H\left[\sum_{i=1}^{H}(w_{ji}^I E_i) + B_j^H\right] = \frac{e^{\left[\sum_{i=1}^{H}(w_{ji}^I E_i)+B_j^H\right]} + e^{-\left[\sum_{i=1}^{H}(w_{ji}^I E_i)+B_j^H\right]}}{e^{\left[\sum_{i=1}^{H}(w_{ji}^I E_i)+B_j^H\right]} + e^{-\left[\sum_{i=1}^{H}(w_{ji}^I E_i)+B_j^H\right]}}$$

The output “Optical Properties”

$$S = f^{O}\left[\sum_{i=1}^{H}\left(w_{1j}^{H}Z_j\right) + B_1^{O}\right] = \frac{e^{\left[\sum_{i=1}^{H}(w_{ji}^{H}Z_j)+B_1^{O}\right]} - e^{-\left[\sum_{i=1}^{H}(w_{ji}^{H}Z_j)+B_1^{O}\right]}}{e^{\left[\sum_{i=1}^{H}(w_{ji}^{H}Z_j)+B_1^{O}\right]} + e^{-\left[\sum_{i=1}^{H}(w_{ji}^{H}Z_j)+B_1^{O}\right]}}$$

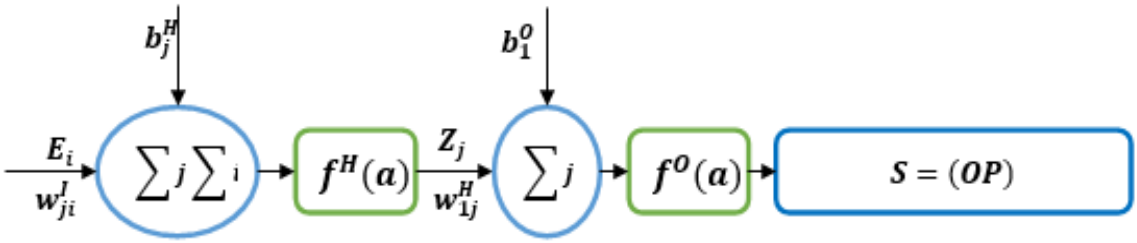


Figure V. 9. Schematic representation of optimized ANN models

V.2.3. Sensitivity analysis by "weight" method

The Relative Importance results displayed in the figure V.10 was obtained using a sensitivity analysis utilizing the weight method. The purpose of this investigation was in order to evaluate how several input parameters, such as wavelength, hemoglobin concentration, temperature, and oxygen saturation, affected the output parameter, which was the refractive index of blood. The outcomes of this sensitivity analysis provide insightful information about the relative importance of these input parameters in influencing the output.

According to the findings, the analysis reveals that the wavelength holds the highest relative importance, contributing significantly to the variation in the refractive index of blood, with a substantial weight of 36.56%. The concentration of hemoglobin following closely after, accounting for 27.39% of the of the influence on the output, indicating its significant role in determining the refractive index. Another significant factor that has an impact on the refractive index is temperature, which is found to have a relative importance of 22.20%. Lastly, oxygen saturation, while still contributing meaningfully, has a comparatively lower weight of 13.84% in determining the refractive index. Given that the impact of each input factor on the refractive index exceeds the 5% threshold, it reinforces the notion that the selected inputs, including wavelength, concentration of hemoglobin, temperature, and oxygen saturation, are indeed relevant and appropriate for our study. This thresholcriterion aligns with the idea that inputs with significant impact play a crucial role in explaining variations in the output, and their inclusion in this study is justified.

Table V. 3. Weights and bias of the best MLP-ANN model for RI (4-20-1).

Input – Hidden Layers Connections					**Hidden – Output Layers Connection**	
Wight				Bias	Wight	Bias
w_{j1}^{I}	w_{j2}^{I}	w_{j3}^{I}	w_{j4}^{I}	b_{j}^{H}	w_{1j}^{H}	b_{1}^{O}
-6.1593	-0.3349	-0.6640	-0.0294	0.7056	-3.3620	
-6.4542	-5.4726	-3.0085	1.2901	8.2998	0.3963	
2.4148	-2.1266	-4.7858	-2.2134	-1.7870	0.0082	
-0.2853	-2.4490	0.1452	-0.6949	0.8101	-2.8079	
-9.8278	-4.1306	6.6820	3.8996	3.7276	0.0539	
6.8803	1.7916	0.2254	2.4442	-4.1760	-0.4049	
1.6231	-6.6055	0.4835	-2.1633	0.1417	-6.8326	
6.2950	-0.0204	0.3619	0.0742	-0.7341	-2.7529	
-6.1396	7.8183	-3.8972	-4.1165	-2.4726	0.7753	
69.7084	0.8162	-2.0354	0.1721	-13.7531	0.1569	
3.6685	-13.0411	2.0623	4.7094	7.9092	4.0741	-0.3169
-4.6824	-3.5890	-1.4472	3.1283	3.5318	-0.5939	
0.4149	5.7637	7.3165	1.7945	1.6830	-0.4877	
2.5725	2.5342	-0.0935	-1.4761	1.2203	-2.7230	
-2.1441	24.5108	2.9271	-6.8346	-18.6351	8.0792	
-24.7775	-8.7871	-26.8217	7.2199	-0.8081	0.0220	
-10.0578	0.5830	1.1418	-0.3157	-1.0932	2.1764	
-0.2660	-1.1840	25.8752	-6.0193	4.1289	-0.4333	
-1.4721	0.5900	19.9261	-5.2320	4.8779	0.5303	
-3.3886	7.5143	-6.0585	-0.8706	-6.3010	0.6932	

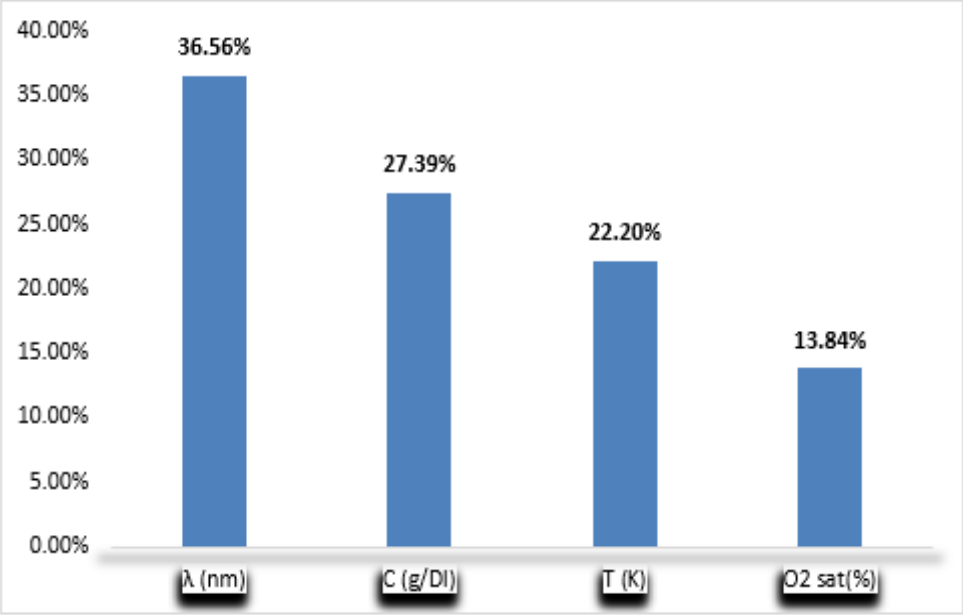


Figure V. 10. Relative importance histogram for RI

The figure V.11 provides valuable insights through Relative Importance results obtained from a sensitivity analysis utilizing the weight method. This analysis was conducted to investigate the influence of selected input parameters, in this case namely: wavelength, hematocrit, and oxygen saturation, on the output variable, specifically, the absorption coefficient of blood. The results of this analysis yield essential information regarding the relative importance of these inputs in shaping the output.

Remarkably, the sensitivity analysis reveals that the wavelength significantly dominates the relative importance, exerting a substantial influence of 60% on the absorption coefficient of blood. Hematocrit is associated with a 17% total impact on output, while oxygen saturation provides a 23% impact. These findings represent the combined effect of these variables in obtaining the absorption coefficient.

It's important to note that both hematocrit and oxygen saturation exhibit a greater impact than the 5% threshold. This finding underlines an important point: the parameters selected for this study wavelength, hematocrit, and oxygen saturation are truly appropriate and suitable. Inputs exceeding this threshold substantiate their relevance in explaining variations in the absorption coefficient, thereby justifying their inclusion.

Table V. 4. Weights and bias of the best MLP-ANN model for μa (3-23-1).

Input – Hidden Layers Connections				**Hidden – Output Layers Connection**	
Wight			Bias	Wight	Bias
w_{j1}^{I}	w_{j2}^{I}	w_{j3}^{I}	b_{j}^{H}	w_{1j}^{H}	b_{1}^{O}
-3.084	14.862	-6.276	12.177	-1.305	1.459
-229.115	46.158	-1.257	-13.624	0.517	
-17.460	2.204	0.975	-0.604	0.345	
-19.688	2.371	2.087	-4.994	-5.263	
50.092	-14.177	4.496	2.552	-3.371	
7.057	-0.165	-4.820	-0.459	-0.122	
-25.963	22.082	-13.120	0.765	0.002	
-160.640	34.473	1.149	-13.231	-0.365	
69.549	-14.718	7.331	0.238	-0.532	
8.879	0.041	0.020	-8.244	-0.191	
8.519	-0.111	0.188	-7.447	0.262	
-84.055	12.475	12.956	-8.187	0.192	
0.408	-0.002	-2.894	-0.321	-2.840	
48.623	-14.518	4.467	3.177	3.940	
15.883	-17.487	0.954	16.358	17.986	
4.301	-2.721	-10.457	3.767	-2.330	
-22.246	-4.940	27.219	0.867	0.382	
0.806	-0.007	-4.433	0.435	2.247	
-4.826	0.916	9.861	-2.047	-2.182	
-8.666	8.558	-4.544	-10.277	24.254	
-10.408	0.358	0.626	5.540	-0.032	
-5.264	-2.980	-0.563	6.210	-0.004	
29.238	-4.600	-3.195	2.292	0.379	

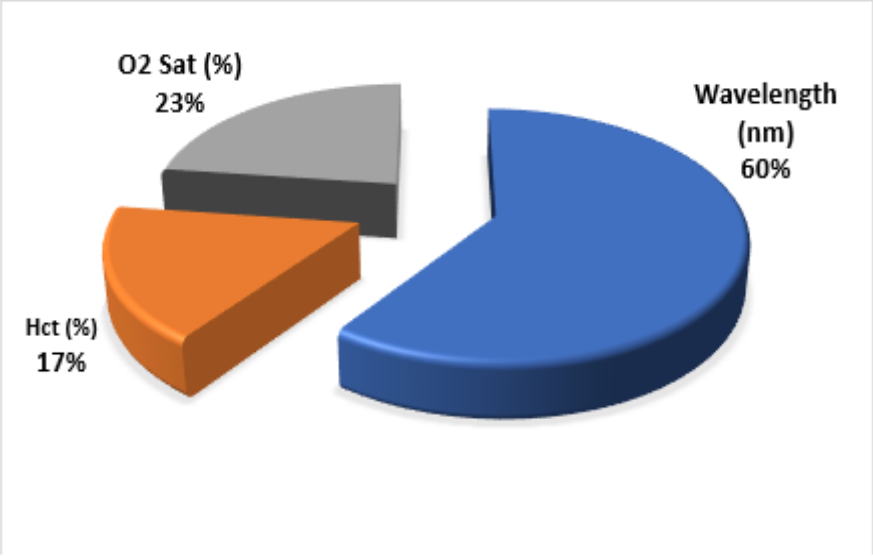


Figure V. 11. Relative importance for µa

The Relative Importance results, which were acquired by a sensitivity analysis utilizing the weight method, are presented completely in the figure V.12. This analysis was conducted to examine the impact of selected input parameters, specifically wavelength, hematocrit, and oxygen saturation, on the output variable, which in this case is the scattering coefficient of blood. The results of this analysis shed light on the relative significance of these input variables in determining the output.

The wavelength stands out as the most important input parameter, commanding a significant impact of 52.62% on the blood scattering coefficient, according to the results. The second most significant component is hematocrit, which has a considerable impact of 27.45%. While less dominant, oxygen saturation still has a significant impact on the output, accounting for 19.92% of the overall impact.

Additionally, the contribution of all three input factors is more than the 5% threshold. This result emphasizes an important point: the parameters used for the study: wavelength, hematocrit, and oxygen saturation, are indeed appropriate. The basis for their inclusion in the investigation is provided by inputs that are greater than this threshold, which confirms their importance in illuminating fluctuations in the scattering coefficient.

Table V. 5. Weights and bias of the best MLP-ANN model for μs (3-25-1).

Input – Hidden Layers Connections				**Hidden – Output Layers Connection**	
Wight			Bias	Wight	Bias
w_{j1}^{I}	w_{j2}^{I}	w_{j3}^{I}	b_{j}^{H}	w_{1j}^{H}	b_{1}^{O}
4.351	7.630	-9.930	9.969	-8.007	-1.697
-20.636	-0.977	-12.978	-14.716	3.625	
7.503	-2.186	2.646	1.289	-1.919	
1.456	0.255	-0.067	0.307	-4.021	
-2.111	12.009	-0.461	0.608	3.059	
3.668	0.442	1.803	1.024	5.291	
-37.320	4.480	-0.350	0.967	-1.634	
-34.195	9.889	2.848	-13.857	2.524	
117.671	-8.589	1.171	-6.843	-2.643	
-0.592	7.851	3.413	-0.268	-8.946	
-28.269	-0.122	-1.269	3.390	7.743	
-23.602	0.012	-1.575	2.595	-9.396	
66.275	2.269	-6.990	-7.581	-11.090	
42.901	-16.901	2.560	7.096	-3.245	
4.484	-9.980	-3.944	20.065	-25.440	
4.818	1.228	5.656	1.044	-11.290	
65.097	2.337	-6.869	-7.504	11.197	
-3.514	22.517	3.197	16.250	-6.896	
-23.010	-64.773	3.996	-4.738	-34.198	
-34.019	72.959	13.771	-6.145	3.976	
-7.903	20.553	11.830	-9.038	0.045	
1.388	-3.634	11.556	6.078	1.103	
-100.782	15.409	-0.904	0.366	-3.818	
58.585	3.348	19.848	0.075	36.450	
-1.476	-4.394	-11.311	-12.570	12.915	

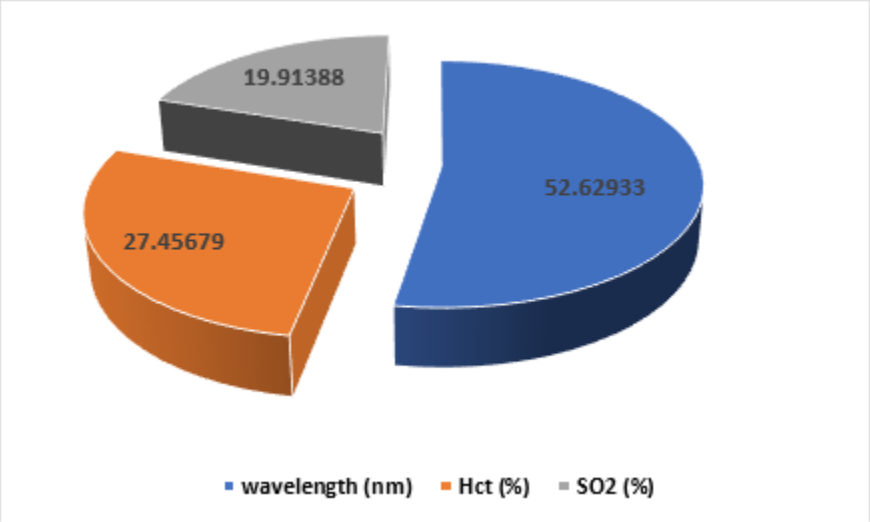
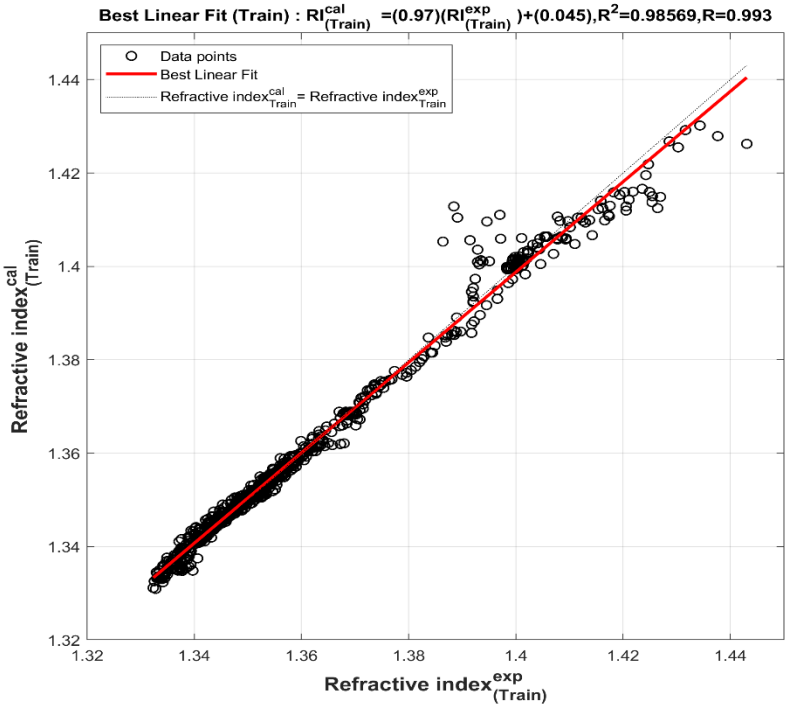


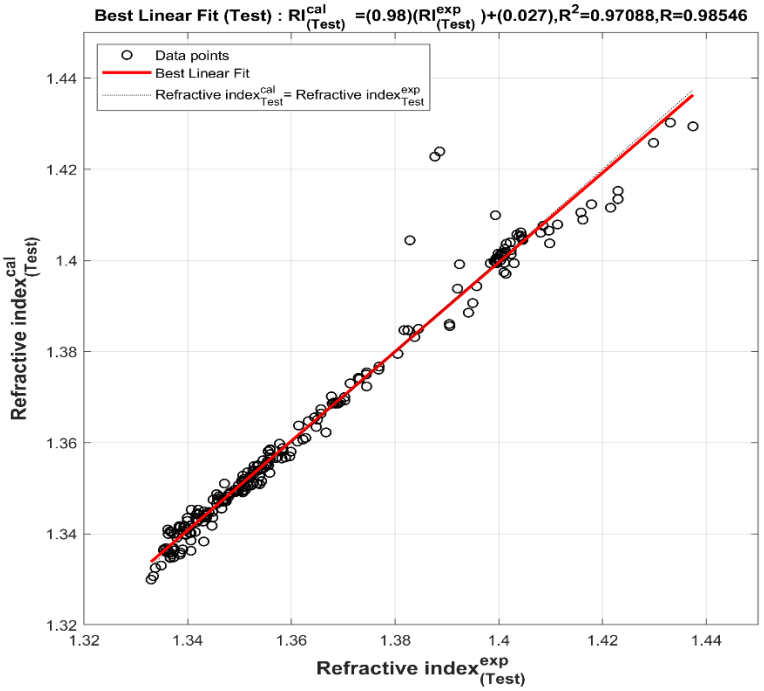
Figure V. 12. Relative importance for μs

V.3. SVM models (*Contribution II*)

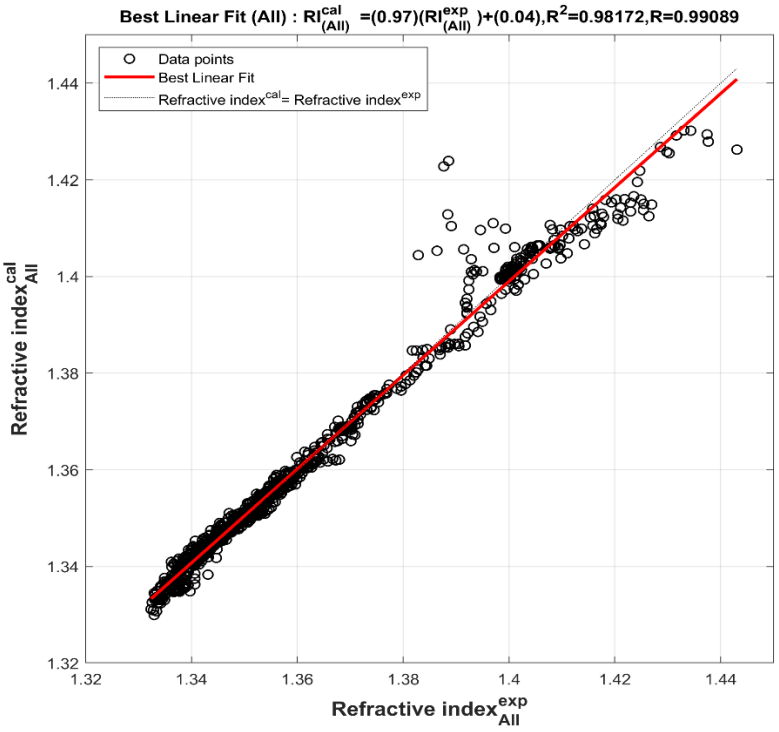
Figure V.13 presents a graphical representation of the comparison between the experimental and predicted values of the refractive index of blood, estimated using SVM. The regression vectors [α, β, R] = (0.97, 0.045, 0.9930) for the training phase; [α, β, R] = (0.98, 0.027, 0.9954) for testing phase; and [α, β, R] = (0.97, 0.04, 0.9908) for the entire dataset. The regression vectors for each phase provide valuable insights into the performance of the model. The slope (α) and the positive y-intercept (β) indicate that the model maintains its strong predictive performance. The correlation coefficients for the training phase, the testing phase, and the overall phase are all regarded as excellent ($0.90 \leq R \leq 1.00$), which demonstrates good robustness of the first established SVM model.


Best Linear Fit (Train) : RI(Train)cal =(0.97)(RI(Train)exp)+(0.045),R2=0.98569,R=0.993
Data points
Best Linear Fit
Refractive indexTraincal = Refractive indexTrainexp
Refractive index(Train)cal
Refractive index(Train)exp
1.32
1.34
1.36
1.38
1.4
1.42
1.44

(a)


Best Linear Fit (Test) : RI(Test)cal =(0.98)(RI(Test)exp)+(0.027),R2=0.97088,R=0.98546
Data points
Best Linear Fit
Refractive indexTestcal = Refractive indexTestexp
Refractive index(Test)cal
Refractive index(Test)exp
1.32
1.34
1.36
1.38
1.4
1.42
1.44

(b)



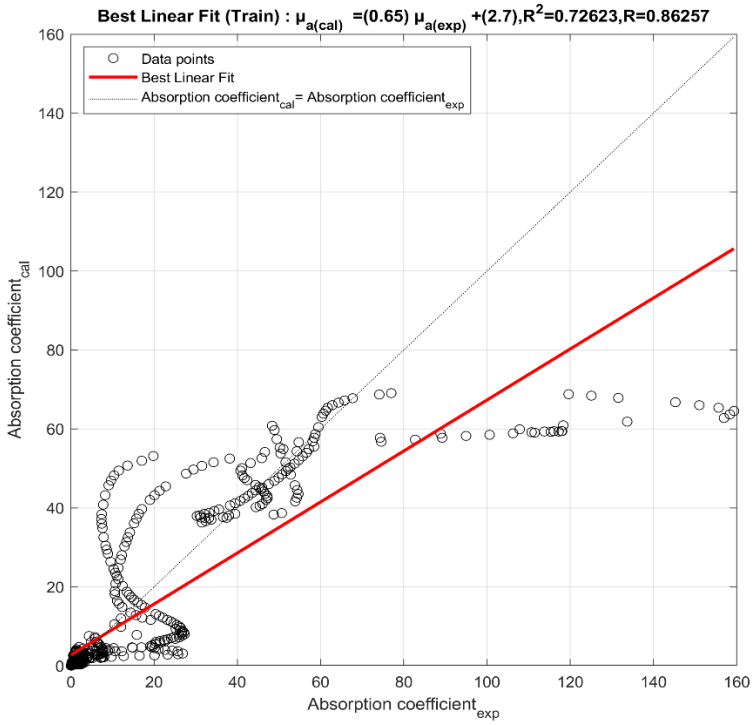
(c)

Figure V. 13. Correlation plot between the experimental and SVM predicted values of the refractive index: (a) training phase, (b) testing phase, and (c) the total phase

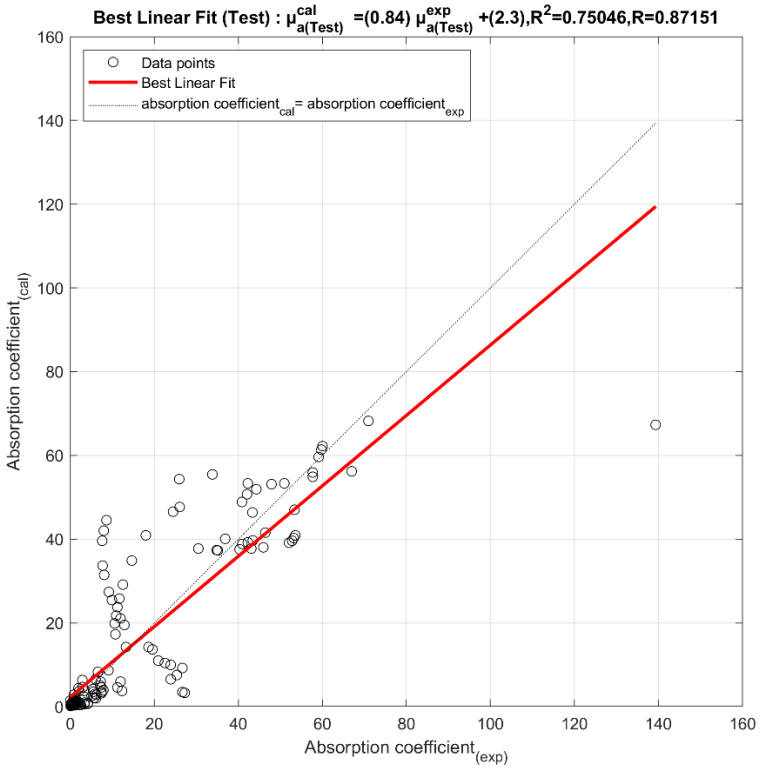
Figure V.14 visually represents the comparison between the experimental and predicted values of the absorption coefficient of blood, generated through the application of Support Vector Machines (SVM). The regression vectors obtained from each phase offer valuable insights into the model's performance. The regression vector [α, β, R] = (0.65, 2.7, 0.8625) indicates that SVM identified a modest linear relationship between the predicted and experimental absorption coefficient values during the training phase. The model appears to capture a significantly weaker correlation than a perfectly linear fit, according to the slope (α) of 0.65. The positive y-intercept (β) value of 2.7 indicates an upward shift in the model's predictions compared to the experimental data. The modest degree of fit during the training phase is indicated by the correlation coefficient (R), which is 0.8625.

The regression vector [α, β, R] = (0.84, 2.3, 0.8715) showcases an improved performance as we move into the testing phase. The slope (α) of 0.84 indicates a stronger linear alignment between the SVM's predictions and the experimental measurements. The y-intercept (β) value of 2.3 still suggests a slight positive bias in the model's predictions. The correlation coefficient (R) of 0.8715 highlights a good level of consistency between the predicted and experimental absorption coefficients during testing.

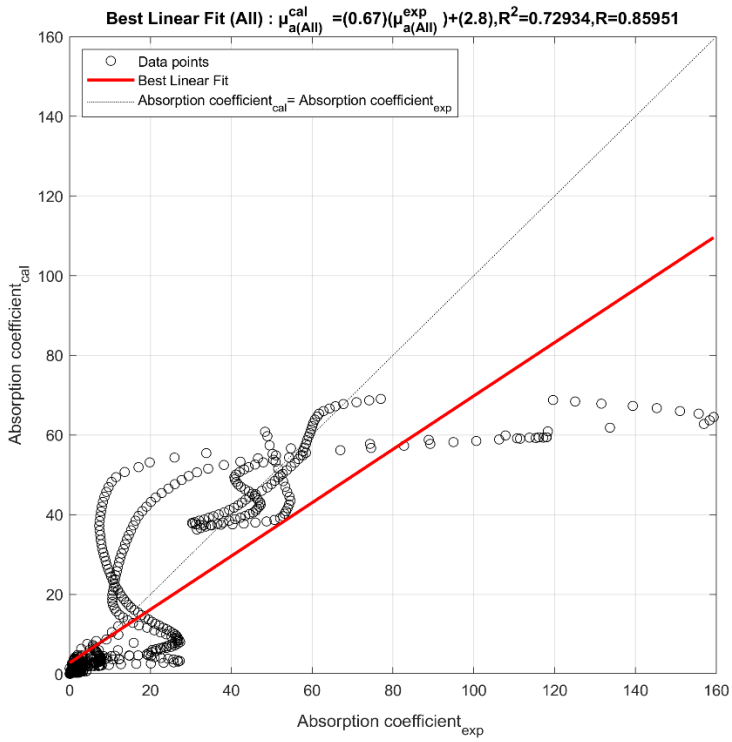
In the entire phase, represented by the regression vector [α, β, R] = (0.67, 2.8, 0.8595), the results remain consistent with those of the training and testing phases. The model maintains a moderate linear fit across the whole dataset, according to by the slope (α) value of 0.67 and the positive y-intercept (β) value of 2.8. The model's accuracy in determining the blood absorption coefficient is confirmed by the correlation coefficient (R) of 0.8595.



(a)


Best Linear Fit (Test) : $\mu_{a(Test)}^{cal}$ =(0.84) $\mu_{a(Test)}^{exp}$ +(2.3),R^2=0.75046,R=0.87151
Data points
Best Linear Fit
absorption coefficient$_{cal}$= absorption coefficient$_{exp}$
Absorption coefficient$_{(cal)}$
Absorption coefficient$_{(exp)}$
0
20
40
60
80
100
120
140
160

(b)


Best Linear Fit (All) : $\mu_{a(All)}^{cal}$ =(0.67)($\mu_{a(All)}^{exp}$)+(2.8),R^2=0.72934,R=0.85951
Data points
Best Linear Fit
Absorption coefficient$_{cal}$= Absorption coefficient$_{exp}$
Absorption coefficient$_{cal}$
Absorption coefficient$_{exp}$
0
20
40
60
80
100
120
140
160

(c)

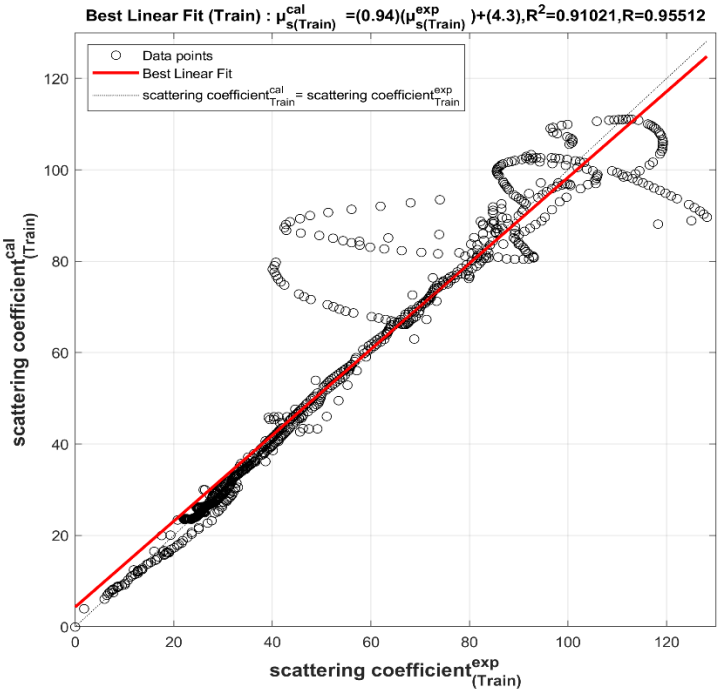
Figure V. 14. Correlation plot between the experimental and SVM predicted values of absorption coefficient: (a) training phase, (b) testing phase, and (c) the total phase

The scattering coefficient of blood, which was obtained using Support Vector Machines (SVM), is visually compared with experimental values in the figure V.15

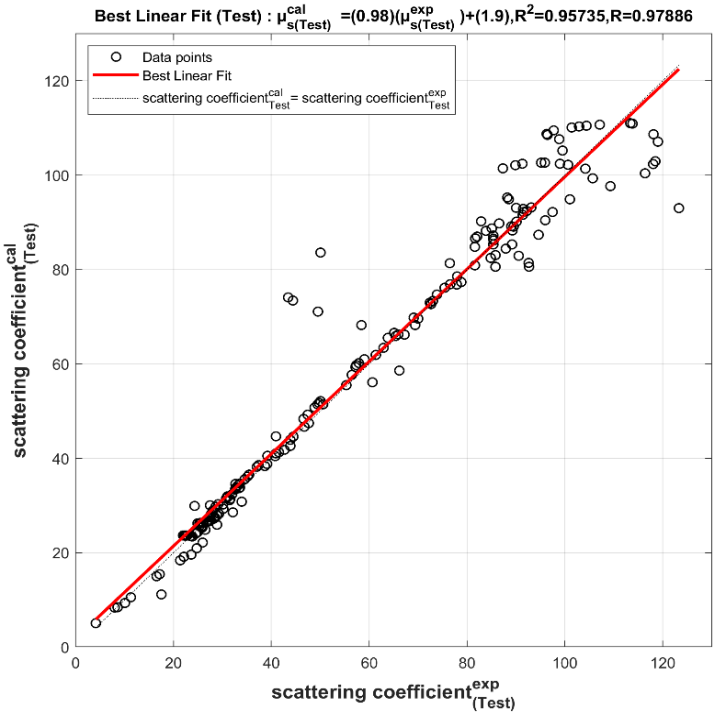
The regression vector [α, β, R] = (0.94, 4.3, 0.9551) in the training phase (a) demonstrates that the SVM proved successful in establishing a strong linear relationship between the predicted and experimental scattering coefficient values. The model captures a strong correlation that is well aligned with the experimental results, as indicated by the slope (α) of 0.94. The positive y-intercept (β) value of 4.3 suggests a minor upward shift in the model's predictions in comparison to the experimental data. The correlation coefficient (R), which is 0.9551, confirms that the training phase's fit achieved good.

In the testing phase, the regression vector [α, β, R] = (0.98, 1.9, 0.9788) highlights an enhanced performance. The SVM's predictions and the actual measurements show a strong linear alignment with a slope (α) of 0.98 putting it even closer to a perfect linear match. The y-intercept (β) value of 1.9 still suggests a slight positive bias in the model's predictions. The notably high correlation coefficient (R) of 0.9788 underscores a strong consistency between the predicted and experimental scattering coefficients during testing.

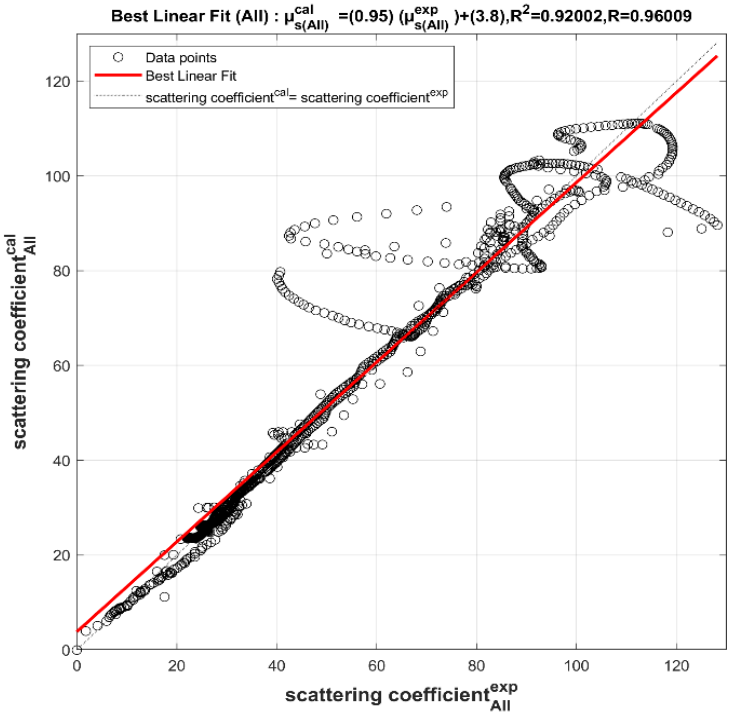
The regression vector [α, β, R] = (0.95, 3.8, 0.9600) represents the whole phase, and the results follow the same pattern as the training and testing phases. The model maintains a strong linear fit across the entire dataset, highlighted by the slope (α) value of 0.95 and the positive y-intercept (β) value of 3.8. The correlation coefficient (R) of 0.9600 confirms the model's reliability in estimating the scattering coefficient of blood.


Best Linear Fit (Train) : $\mu_{s(Train)}^{cal}$ =(0.94)($\mu_{s(Train)}^{exp}$)+(4.3),R^2=0.91021,R=0.95512
Data points
Best Linear Fit
scattering coefficient$_{Train}^{cal}$ = scattering coefficient$_{Train}^{exp}$
scattering coefficient$_{(Train)}^{cal}$
scattering coefficient$_{(Train)}^{exp}$
0
20
40
60
80
100
120

(a)


Best Linear Fit (Test) : $\mu_{s(Test)}^{cal}$ =(0.98)($\mu_{s(Test)}^{exp}$)+(1.9),R^2=0.95735,R=0.97886
Data points
Best Linear Fit
scattering coefficient$_{Test}^{cal}$ = scattering coefficient$_{Test}^{exp}$
scattering coefficient$_{(Test)}^{cal}$
scattering coefficient$_{(Test)}^{exp}$
0
20
40
60
80
100
120

(b)



(c)

Figure V. 15. Correlation plot between the experimental and SVM predicted values of scattering coefficient: (a) training phase, (b) testing phase, and (c) the total phase

V.3.2. Comparison between the experimental optical properties and calculated by SVM

The figure V.16 below provides important context for understanding the relationship between the predicted and experimental values of blood's scattering coefficient with wavelength. The notable result shows a strong consensus: there is a significant amount of agreement between these two sets of data. This alignment suggests that our models and methodologies for predicting the scattering coefficient based on wavelength have demonstrated impressive accuracy, hence reinforcing their reliability and potential utility in the field of biomedical research.

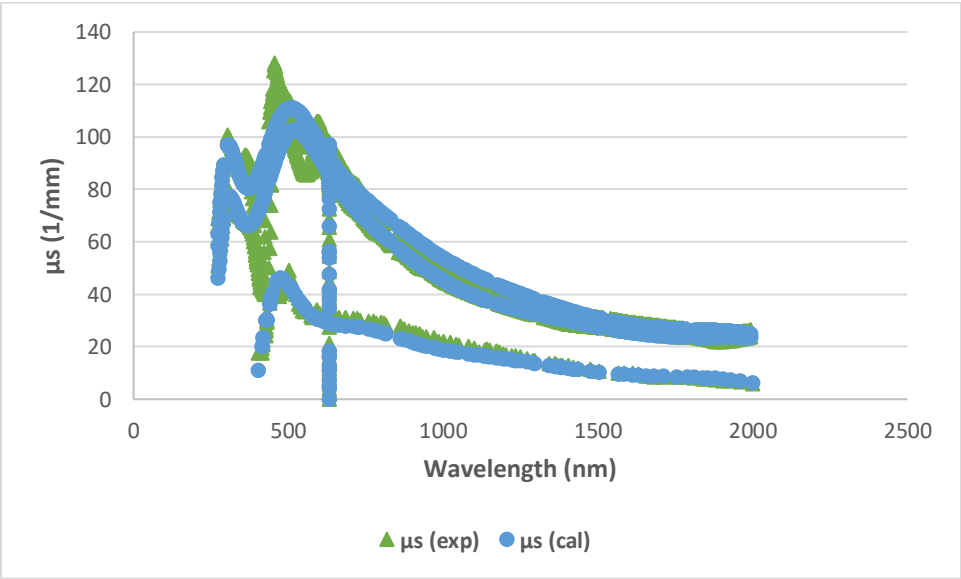


Figure V. 16. Comparison between μs (exp) and μs calculated by SVM vs wavelength

In other hand, in the figure V. 17 below, we compare the same data of experimental and calculated by SVM of scattering coefficient of blood with the hematocrit, and the results exhibited good agreement, indicating that the hematocrit is a significant input parameter.

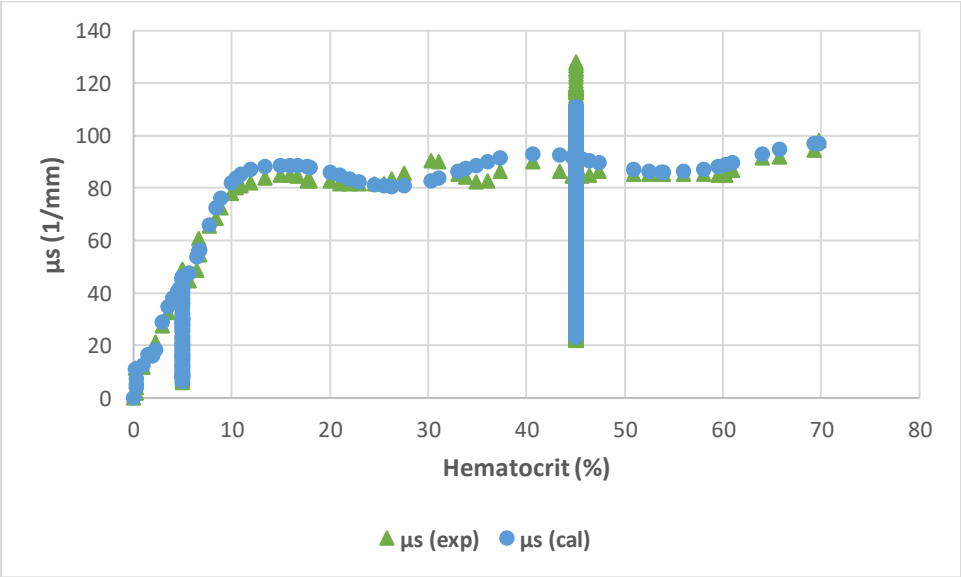


Figure V. 17. Comparison between μs (exp) and μs calculated by SVM vs hematocrit

Figure V.18 visualizes a comparison between the experimental data and the prediction results of the third SVM model as a function of wavelength. The experimental data for the optical property of absorption from the literature (shown by blue circles) and the SVM3 model's prediction results (shown by red) show a respectable degree of agreement, which suggests good predictive capabilities of the SVM3 model designed.

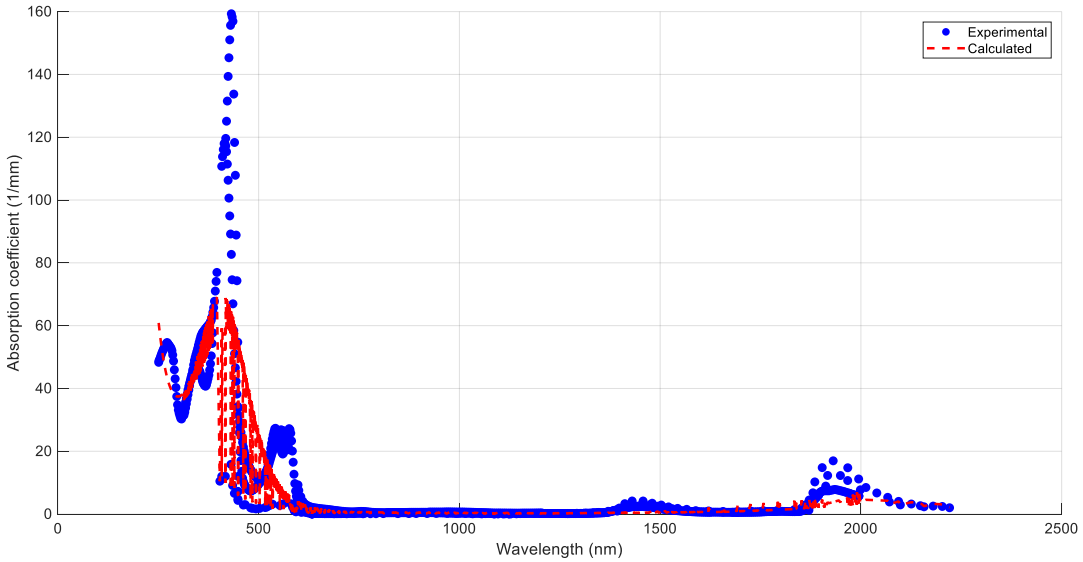


Figure V. 18. Comparison between µa experimental and µa calculated by SVM vs wavelength

The comparison between the experimental data and the prediction results of the third SVM model as a function of hematocrit is illustrated in Figure V.19. It suggests an acceptable agreement between the experimental data from the literature of blood absorption (indicated by blue circles) and the prediction results of the SVM3 model (indicated by yellow circles), which suggests good predictive capabilities of the third model. of SVM.

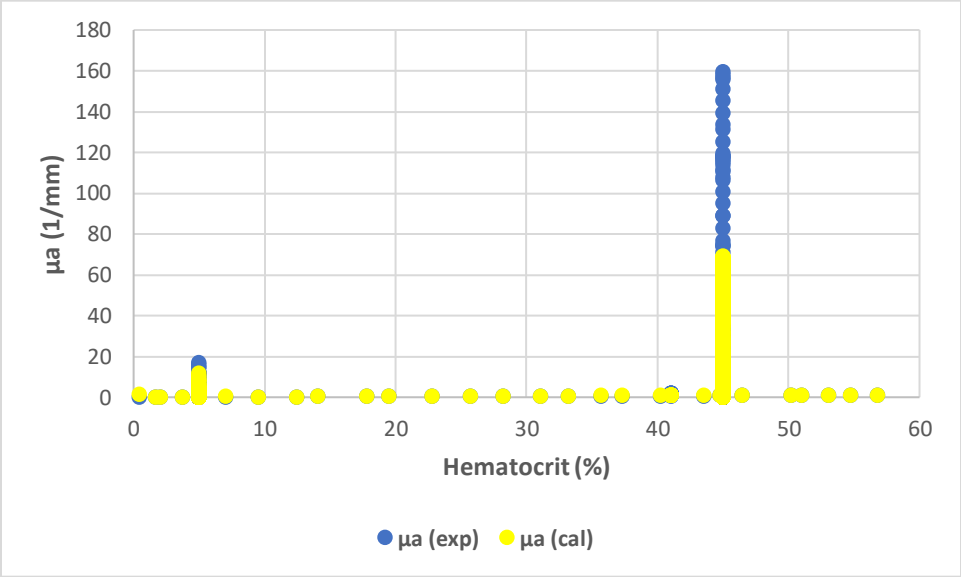


Figure V. 19. Comparison between μa (exp) and μa calculated by SVM vs hematocrit

Figure V.20 illustrates a comparison between the experimental data and the prediction results of the third SVM model as a function of oxygen saturation. It shows an acceptable agreement between the experimental data from the absorption coefficient literature (indicated by blue circles) and the prediction results of the SVM3 model (indicated by yellow circles), which suggests good predictive capabilities of the model.

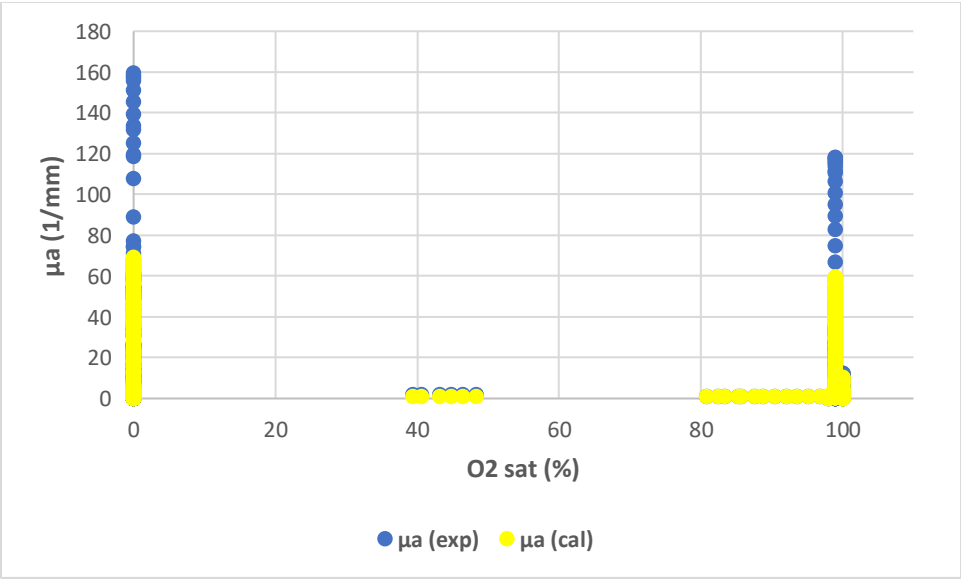


Figure V. 20. Comparison between μa (exp) and μa calculated by SVM vs O2 sat

The performance of each model was evaluated in terms of correlation coefficient (*R*) and root mean square error (*RMSE*).

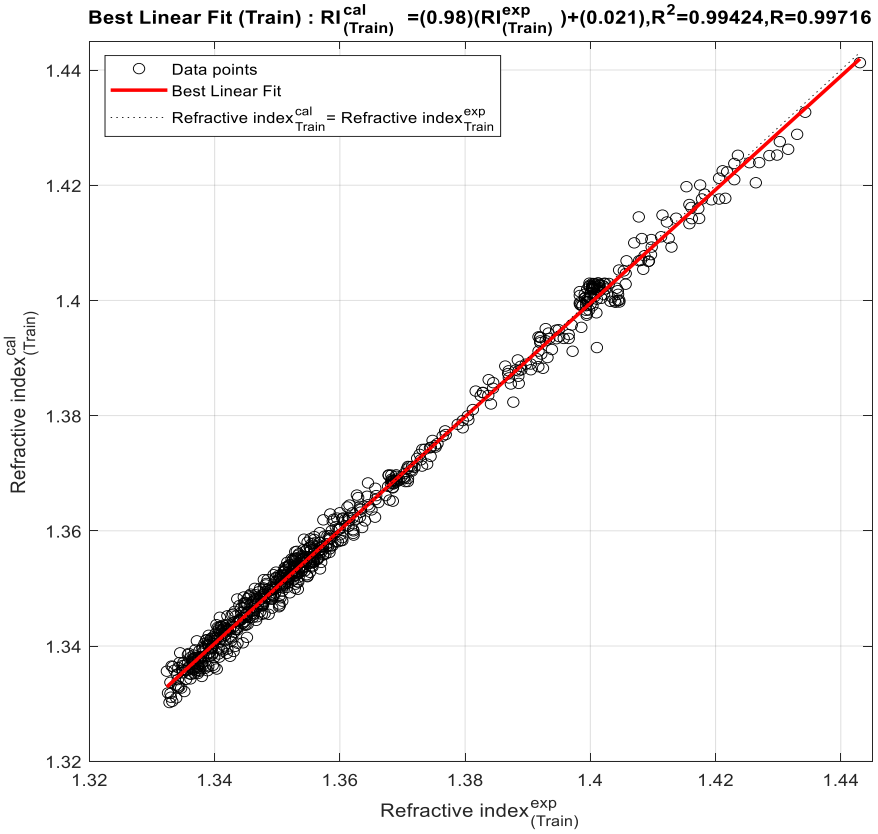
V.4. DA-SVM models (*Contribution III*)

Figure V.21 highlights the effectiveness of the Dragonfly Algorithm-Support Vector Machines (DA-SVM) in estimating the refractive index of blood and comparing these predictions against experimental data. The regression vectors disclosed notable insights into the model's performance throughout different phases.

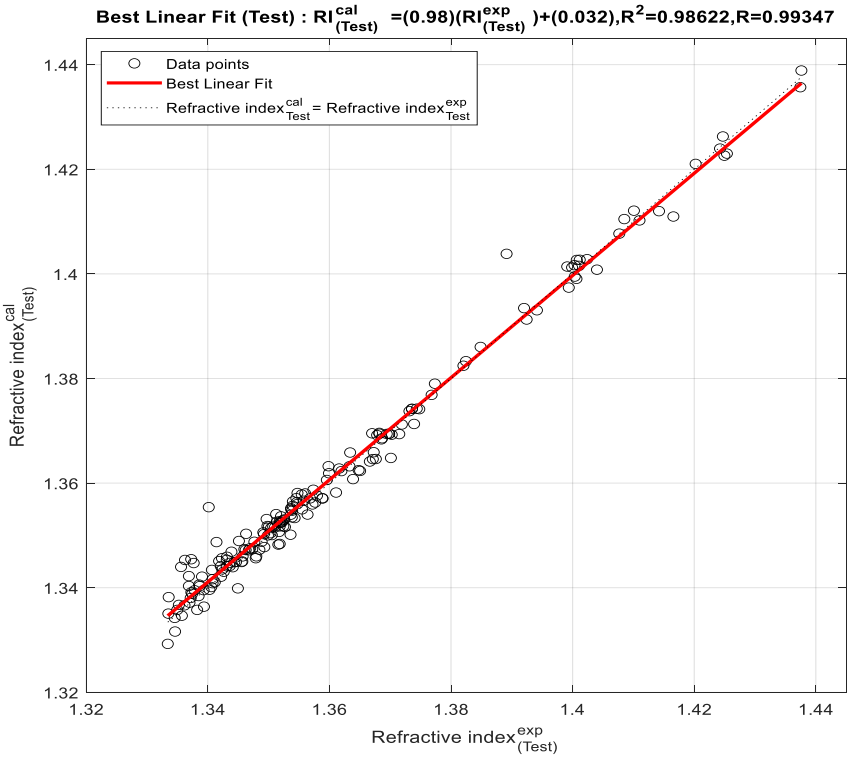
In the training phase (a), the regression vector [α, β, R] = (0.98, 0.021, 0.9971) reveals that the DA-SVM has effectively identified a strong linear relationship between the predicted and experimental refractive index values. The slope (α) of 0.98 signifies that the model captures a robust correlation closely resembling the experimental data. A slight upward change in the model's predictions with regard to the experimental data can be noticed by the small positive y-intercept (β) value of 0.021. The high correlation coefficient (R) of 0.9971 signifies an exceptional level of fit during the training phase.

Following testing (b), the regression vector [α, β, R] = (0.98, 0.032, 0.9934) reveals a sustained performance. The slope (α) of 0.98 reaffirms a strong linear alignment between the DA-SVM's predictions and the actual measurements. Even with a slightly greater positive y-intercept (β) value of 0.032, the model's predictions still show a minor positive bias. The high correlation coefficient (R) of 0.9934 shows that the experimental and predicted refractive index values matched well and consistently during testing.

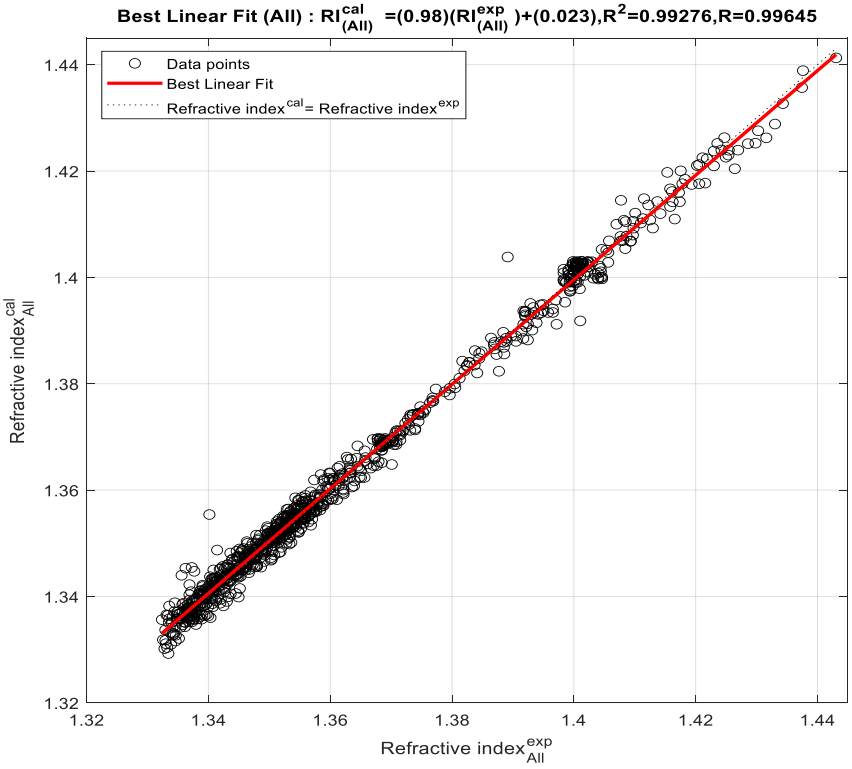
The regression vector [α, β, R] = (0.98, 0.023, 0.9964), which represents all of the phase (c), illustrates that the results seem to follow the same consistent pattern as those seen in the training and testing phases. The model's good linear fit throughout the entire dataset is reaffirmed by the slope's (α) value of 0.98 and the tiny positive y-intercept's (β) value of 0.023. The high correlation coefficient (R) of 0.9964 confirms the model's reliability in estimating the refractive index of blood.


Best Linear Fit (Train) : RI(Train)cal =(0.98)(RI(Train)exp)+(0.021),R2=0.99424,R=0.99716
Data points
Best Linear Fit
Refractive indexTrain cal = Refractive indexTrain exp
Refractive index(Train) cal
Refractive index(Train) exp
1.32
1.34
1.36
1.38
1.4
1.42
1.44

(a)


Best Linear Fit (Test) : RI(Test)cal =(0.98)(RI(Test)exp)+(0.032),R2=0.98622,R=0.99347
Data points
Best Linear Fit
Refractive indexTest cal = Refractive indexTest exp
Refractive index(Test) cal
Refractive index(Test) exp
1.32
1.34
1.36
1.38
1.4
1.42
1.44

(b)



(c)

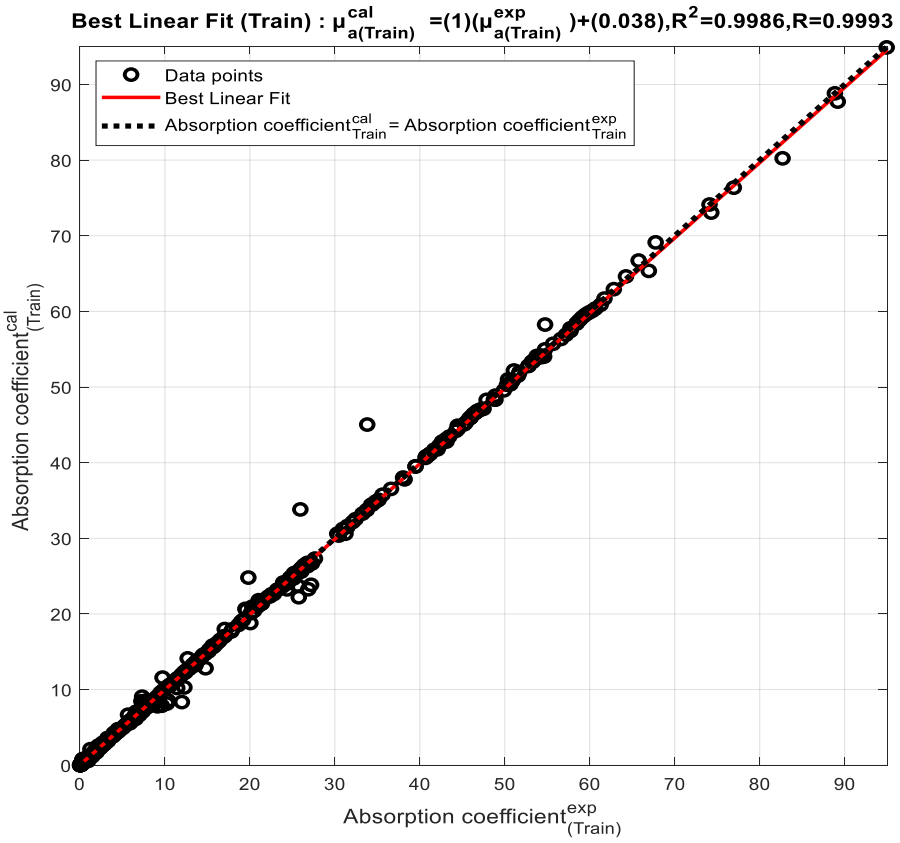
Figure V. 21. Scatter plots between experimental and DA-SVM predicted values of the refractive index: (a) training phase, (b) testing phase, and (c) the total phase

Figure V. 22 presents the results of using the Dragonfly Algorithm-Support Vector Machines (DA-SVM) to estimate the blood absorption coefficient in comparison to experimental measurements. The presented regression vectors illuminate key aspects of the model's performance across different phases.

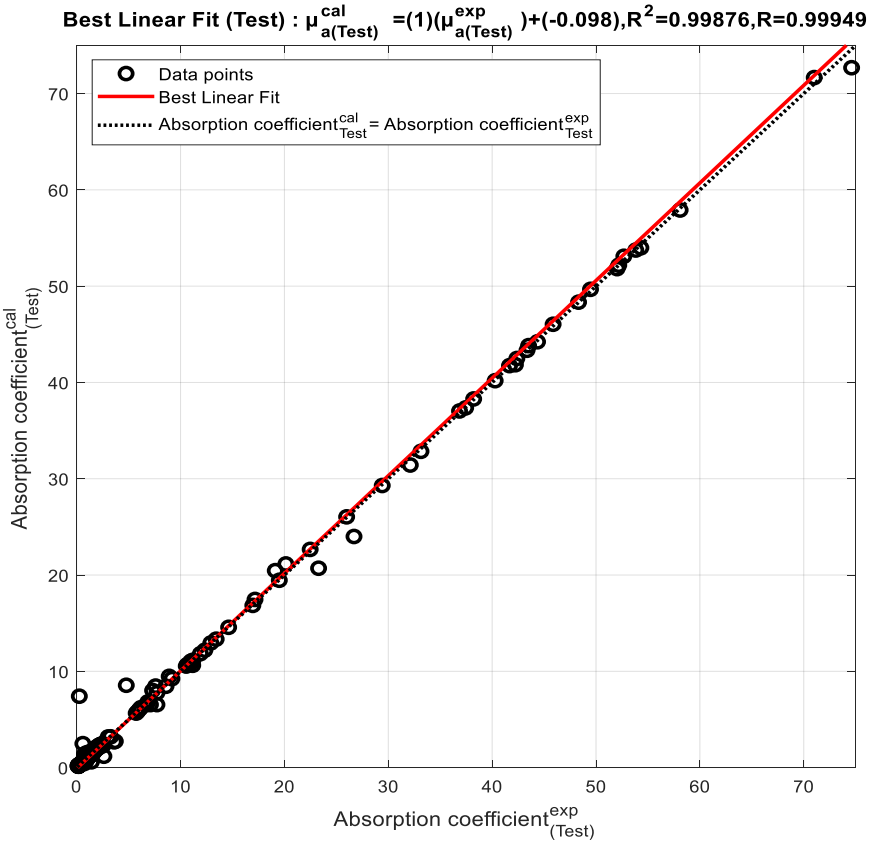
The regression vector [α, β, R] = (1, 0.038, 0.9993) in the training phase (a) indicates that DA-SVM has been effective in identifying an extremely strong linear relationship between the predicted and experimental absorption coefficient values. The model's predictions perfectly match the experimental data, as shown by the slope (α) of 1. The model's predictions are slightly shifted higher in comparison to the experimental results, as indicated by the positive y-intercept (β) value of 0.038. The exceptionally high correlation coefficient (R) of 0.9993 attests to the model's outstanding fit during the training phase.

Moving to the testing phase (b), the regression vector [α, β, R] = (1, -0.098, 0.9994) demonstrates a continuous high level performance. perfect linear correlation between the DA-SVM predicts and the actual absorption coefficient values is reaffirmed by the slope (α) of 1. The negative y-intercept (β) value of -0.098 indicates a minor downward shift in the model's predictions compared to the experimental data. The model's extraordinary accuracy in matching the anticipated and observed absorption coefficients during testing is highlighted by the extremely high correlation coefficient (R) of 0.9994.

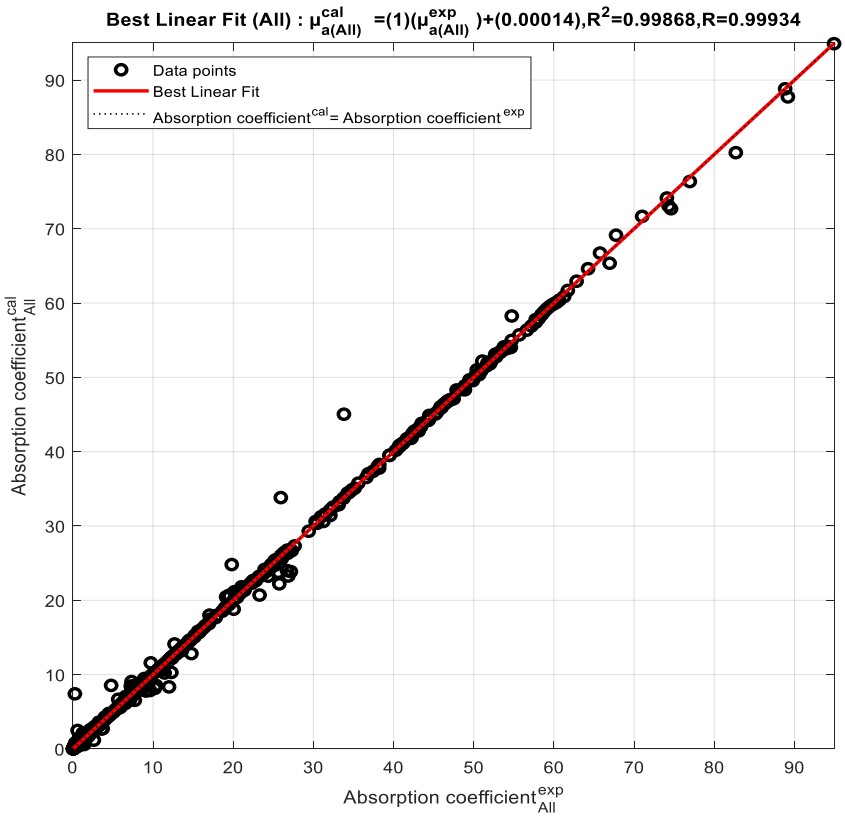
In the total phase (c), as depicted by the regression vector [α, β, R] = (1, 0.00014, 0.9993), the results remain consistently robust as observed in the training and testing phases. The slope (α) of 1 and the nearly negligible positive y-intercept (β) value of 0.00014 affirm the model's strong linear fit across the entire dataset. The very high correlation coefficient (R) of 0.9993 reaffirms the model's trustworthiness in estimating the absorption coefficient of blood.



(a)


Best Linear Fit (Test) : μ cal a(Test) =(1)(μ exp a(Test))+(-0.098),R2=0.99876,R=0.99949
Data points
Best Linear Fit
Absorption coefficient cal Test = Absorption coefficient exp Test
Absorption coefficient cal (Test)
Absorption coefficient exp (Test)
0
10
20
30
40
50
60
70

(b)



(c)

Figure V. 22. Scatter plots between experimental and DASVM predicted values of the absorption coefficient: (a) training phase, (b) testing phase, and (c) the total phase

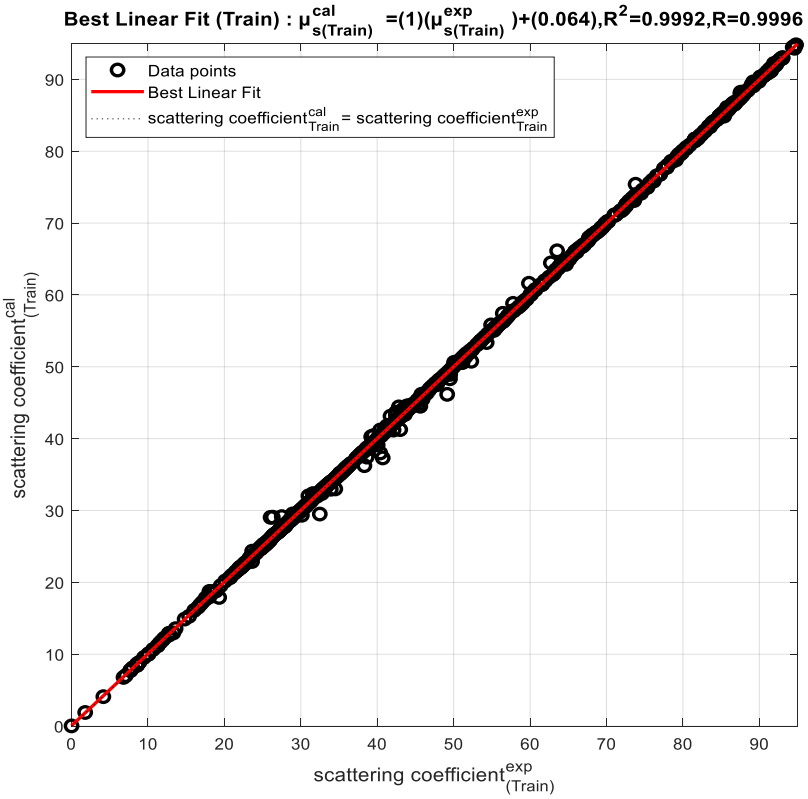
.

The figure V. 23 visually presents a comprehensive comparison between the experimental and predicted values of the scattering coefficient of blood, achieved through the implementation of the Dragonfly Algorithm-Support Vector Machines (DA-SVM). Significant aspects of the model's performance during various phases are shown by the regression vectors.

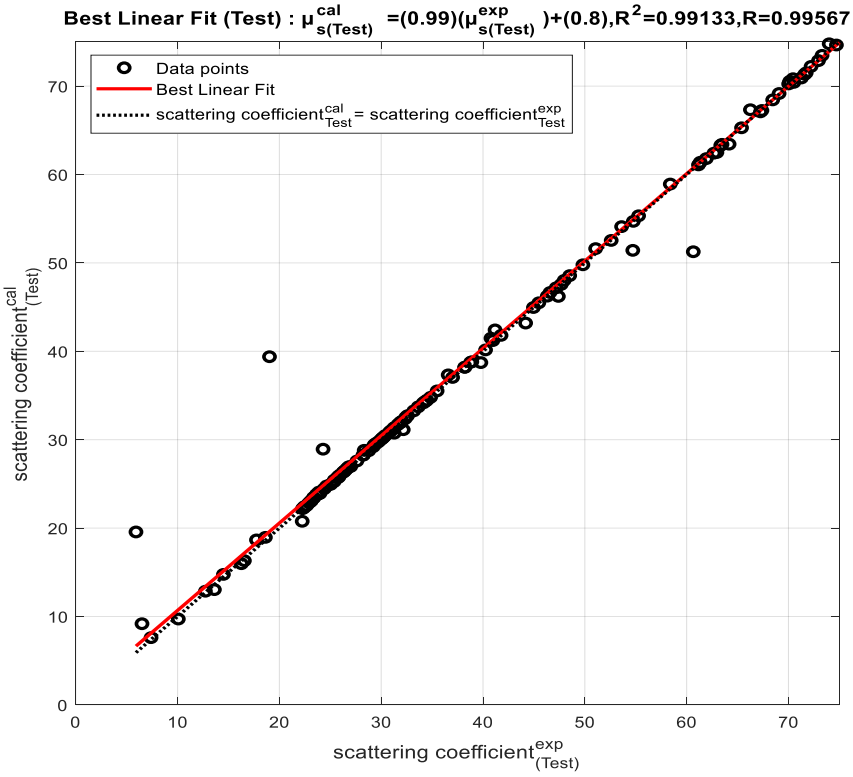
The regression vector [α, β, R] = (1, 0.064, 0.9996) from the training phase (a) emphasizes the outstanding accuracy attained by DA-SVM. The ideal linear relationship between the model's predictions and the values of the experimental scattering coefficients is denoted by a slope (α) of 1. The small positive y-intercept (β) value of 0.064 suggests a negligible upward shift in the model's predictions relative to the experimental data. An excellent level of fit during the training phase is indicated by the extraordinarily high correlation coefficient (R), which is 0.9996.

Transitioning to the testing phase (b), the regression vector [α, β, R] = (0.99, 0.8, 0.9956) reveals a consistently high-performance level. The almost perfect linear relationship between the DA-SVM estimates and the actual scattering coefficient values is confirmed by the slope (α) of 0.99. The y-intercept (β) value of 0.8, while slightly higher, indicates a minor positive bias in the model's predictions compared to the experimental data. The still notably high correlation coefficient (R) of 0.9956 underlines the model's exceptional precision in aligning the predicted and experimental scattering coefficients during testing.

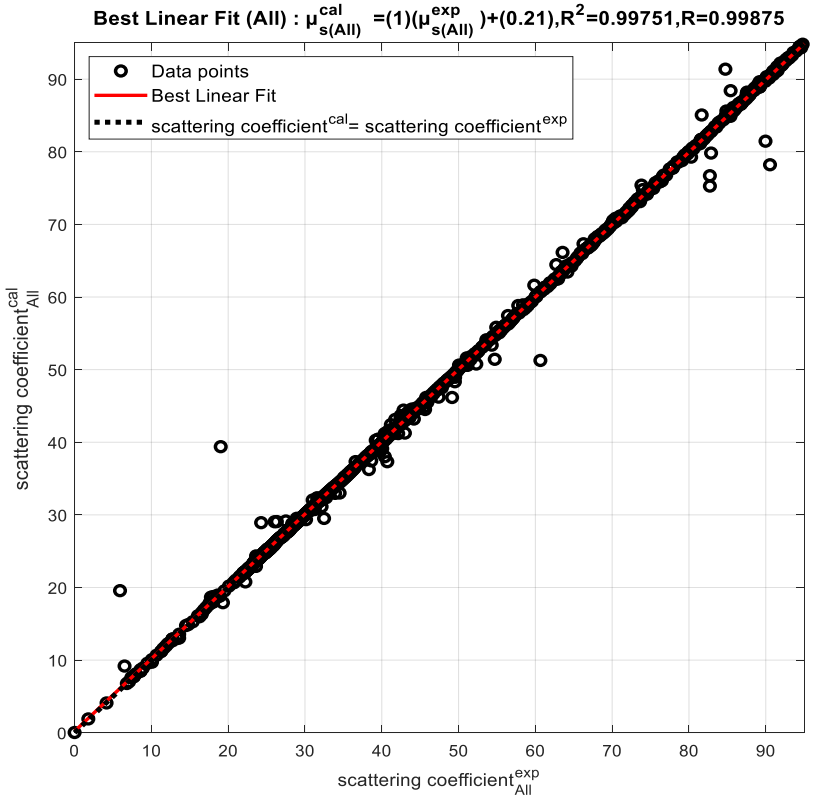
The findings preserve the similar trend observed throughout the training and testing phases in the overall phase (c), expressed by the regression vector [α, β, R] = (1, 0.21, 0.9987). The slope (α) of 1 and the positive y-intercept (β) value of 0.21 suggest that the model sustains a strong linear fit across the entire dataset, with a slight positive shift. The accuracy of the (DA-SVM) model in estimating the scattering coefficient of blood is verified by the high correlation coefficient (R), which is 0.9987.



(a)


Best Linear Fit (Test) : $\mu_{s(Test)}^{cal}$ =(0.99)($\mu_{s(Test)}^{exp}$)+(0.8),R^2=0.99133,R=0.99567
Data points
Best Linear Fit
scattering coefficient$_{Test}^{cal}$ = scattering coefficient$_{Test}^{exp}$
scattering coefficient$_{(Test)}^{cal}$
scattering coefficient$_{(Test)}^{exp}$
0
10
20
30
40
50
60
70

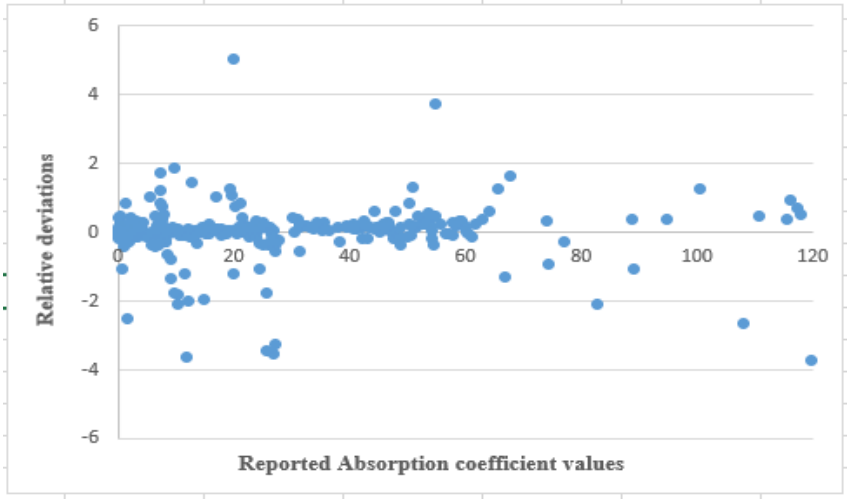
(b)


Best Linear Fit (All) : $\mu_{s(All)}^{cal}$ =(1)($\mu_{s(All)}^{exp}$)+(0.21),R^2=0.99751,R=0.99875
Data points
Best Linear Fit
scattering coefficient^{cal}= scattering coefficient^{exp}
scattering coefficient$_{All}^{cal}$
scattering coefficient$_{All}^{exp}$
0
10
20
30
40
50
60
70
80
90

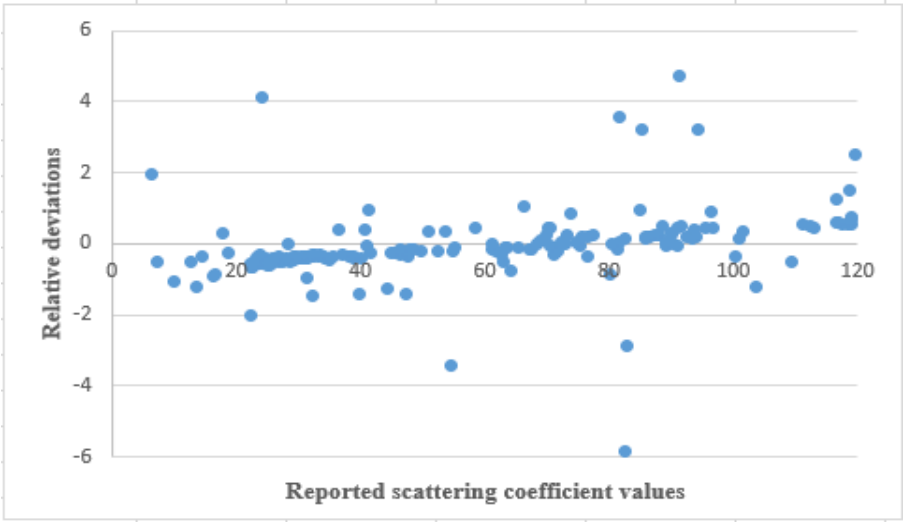
(c)

Figure V. 23. Scatter plots between experimental and DASVM predicted values of the scattering coefficient: (a) training phase, (b) testing phase, and (c) the total phase

To affirm the stochastic nature of any inherent errors, we conducted residual analyses, as depicted in Figures V 24 and 25, during both the training (a) and testing (b) phases for absorption and scattering coefficients. These residual plots reveal a consistent distribution of residuals, evenly scattered around the zero line, with deviations typically confined within the range of ±6. This pattern underscores the robustness and credibility of the DA-SVM models we developed

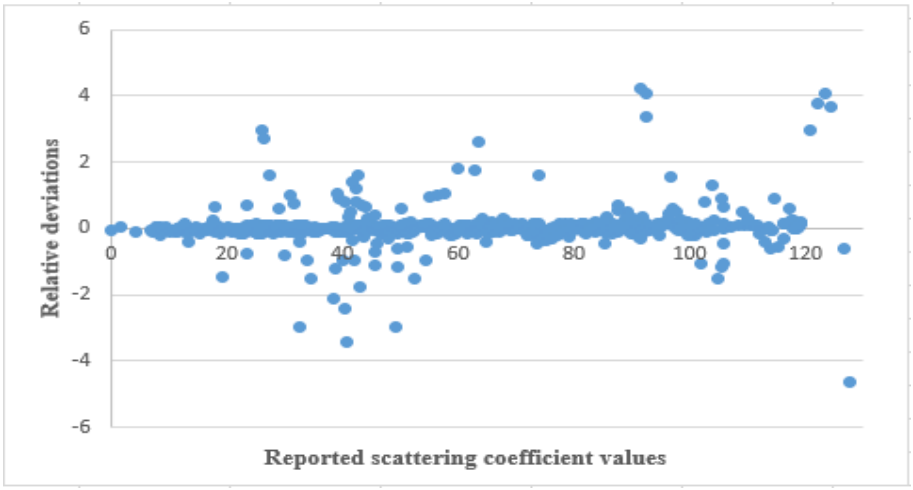


(a)

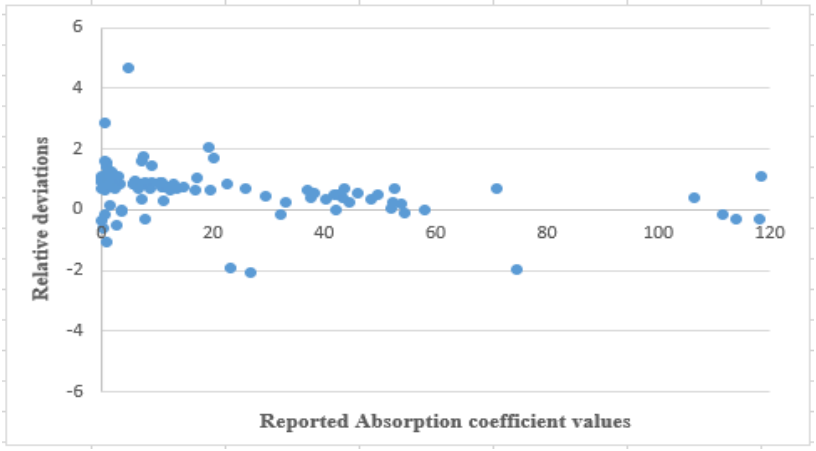


(b)

Figure V. 24. Residual plots for: (a) training dataset, and (b) testing dataset for μa



(a)



(b)

Figure V. 25. Residual plots for: (a) training dataset, and (b) testing dataset for µs

Establishing the applicability domain (AD) of a prediction model is a crucial step in verifying its reliability. To interpret the AD, we employed William's plot, as illustrated in figures V. 26 and 27. These plots reveal a noteworthy observation: the majority of samples within the dataset fall comfortably within the model's applicability domain. Specifically, the warning leverage limit is nearly 0.013, and out of the total 1000 experimental data points for both µa and µs, no more than 36 data points are considered outliers. This underscores the robustness and trustworthiness of our model.

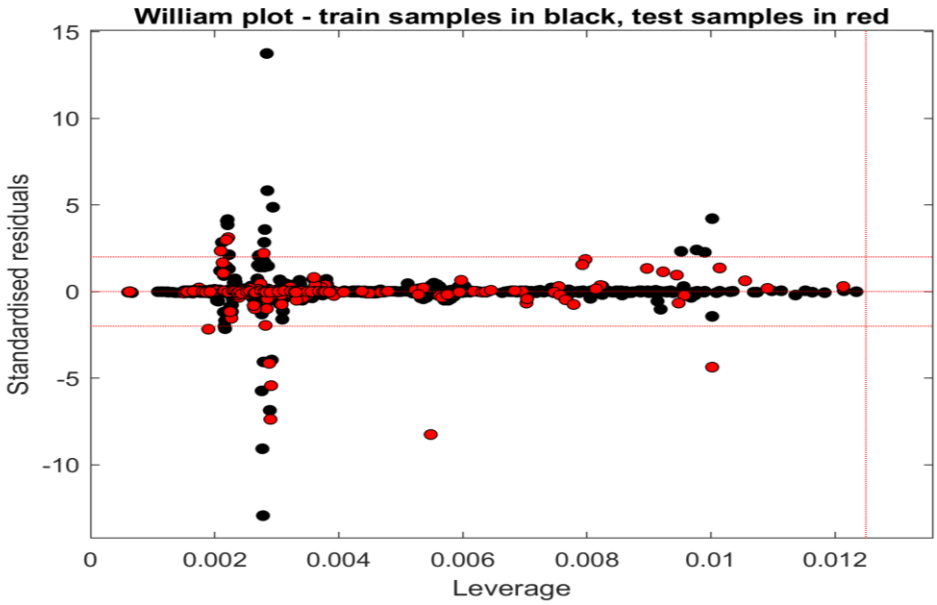


Figure V. 26. William's diagram for detection outliers for DA-SVM model (for µa).

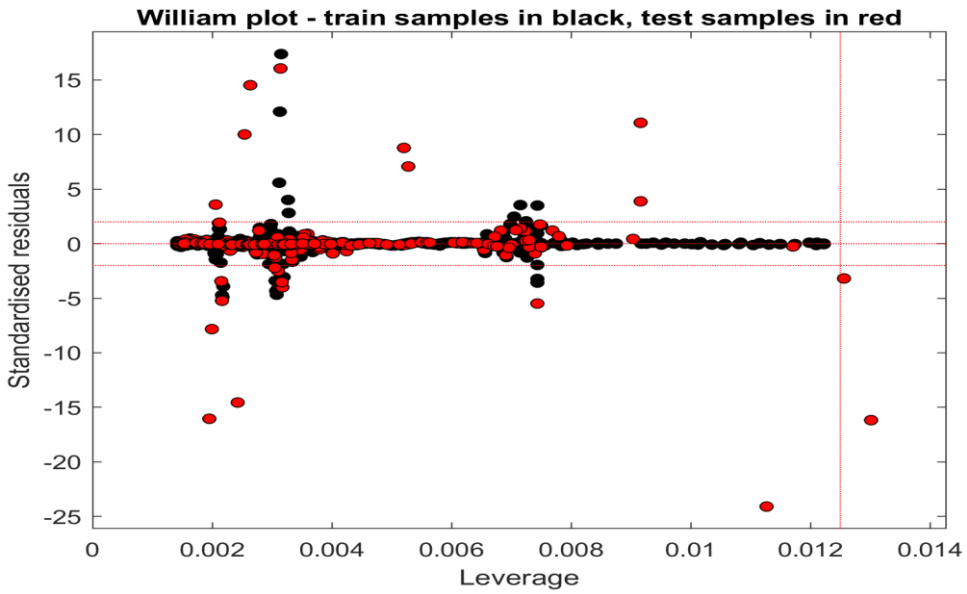


Figure V. 27. William's diagram for detection outliers for DA-SVM model (for µs)

V.5. Comparison between MLP-ANN, SVM and DA-SVM models

In figure V. 28, a comprehensive comparison of three predictive models (ANN1, SVM1, and DA-SVM1) used to predict the refractive index of blood is presented, the evaluation is based on several key performance metrics, including correlation coefficient (R), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) in the testing phase.

In terms of correlation coefficient (R) = (0.9990, 0.9854, and 0.9934) for ANN1, SVM1, and DA-SVM1, respectively. The linear strength of the relationship between the predicted and experimental values is evaluated by the correlation coefficient. The model is more accurate when R is higher. In this context, ANN1 demonstrates an exceptional correlation coefficient, indicating a strong alignment between predicted and actual refractive index values. In this aspect, the DA-SVM1 also exhibits outstanding performance with a strong correlation. Although SVM1 displays a commendable correlation, it falls slightly behind the other two models.

In terms of Root Mean Square Error (RMSE), the results were: 0.0011, 0.0043, and 0.0021, for ANN1, SVM1, and DA-SVM1, respectively. The average magnitude of prediction errors is

determined by DA-SVM1, lower RMSE values indicate more precise predictions. Given that ANN1 has the lowest RMSE, it is the most accurate at predicting the refractive index, with a low RMSE and a great prediction accuracy, DA-SVM1 follows closely. Although it is still reasonable, SVM1 exhibits a slightly higher RMSE, which denotes somewhat larger prediction errors.

In terms of Mean Absolute Error (MAE), the values were: 0.0505, 0.0021, and 0.0018 for ANN1, SVM1, and DA-SVM1, respectively. MAE measures the average absolute difference between calculated and experimental values. More accurate predictions are reflected by lower MAE values. Here, SVM1 and DA-SVM1 both perform noticeably better than ANN1 in terms of MAE, demonstrating their accuracy in reducing absolute prediction errors. SVM1, in particular, excels in this aspect, closely followed by DA-SVM1.

In terms of Mean Absolute Percentage Error (MAPE), the values were: 0.0504, 0.1541, and 0.1338 for ANN1, SVM1, and DASVM1, respectively. More accurate predictions are indicated by a lower MAPE. Once more, ANN1 outperforms the competition with the lowest MAPE, demonstrating its strong performance in reducing prediction errors as a percentage of actual values. DA-SVM1 follows closely behind, While SVM1 displays a higher MAPE, which indicates relatively greater percentage errors in its predictions.

When these models are analyzed and compared, it becomes clear that ANN1 performs best in terms of correlation coefficient, RMSE, and MAPE, demonstrating its strong predictive capabilities. However as seen by their low MAE values, SVM1 and DA-SVM1 stand out for their remarkable precision.

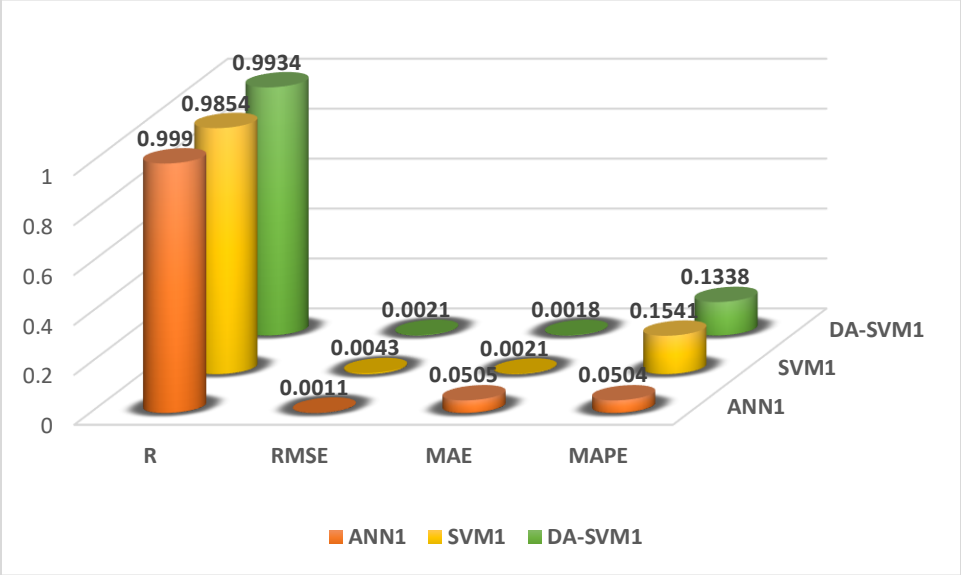


Figure V. 28. comparison between the three modes developed for the refractive index

Three prediction models (ANN2, SVM2, and DA-SVM2) used to estimate the optical property of blood, named absorption coefficient, during the testing phase, are discussed in detail in figure V. 29. The evaluation is based on the same performance metrics as shown previously (Correlation coefficient (R), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE)).

the correlation coefficients (R) were found to be 0.9972, 0.8715, and 0.9994 for ANN2, SVM2, and DA-SVM2, respectively. With a remarkable correlation coefficient in this case, the DA-SVM2 shines out and shows that the predicted and actual values of the absorption coefficient are closely aligned. ANN2 also demonstrates a robust correlation, while SVM2 lags behind significantly in this aspect.

The RMSE results for ANN2, SVM2, and DA-SVM2 were 1.7613, 9.7414, and 0.9792, respectively. With the lowest RMSE, DA-SVM2 is in the lead, emphasizing its higher accuracy in predicting the absorption coefficient. ANN2 follows with a reasonably low RMSE, indicating good predictive accuracy. SVM2, on the other hand, exhibits a substantially higher RMSE, indicating that the errors in its predictions are significantly greater.

The MAE values were 1.0073, 4.6213, and 0.3509, for ANN2, SVM2, and DA-SVM2, respectively. DA-SVM2 demonstrates exceptional accuracy in reducing absolute prediction errors. ANN2 follows with a reasonably low MAE, while SVM2 shows a significantly higher MAE, suggesting that its predictions have larger absolute errors.

For ANN2, SVM2, and DA-SVM2, the MAPE values were (71.324, 125.433, and 34.6588), respectively. DA-SVM2 outperforms the other models with the lowest MAPE, signifying its strong performance in reducing prediction errors as a percentage of actual values. ANN2 follows with a considerably higher MAPE, while SVM2 exhibits the highest MAPE, indicating relatively larger percentage errors in its predictions.

The analysis and comparison of these models for the testing phase show that DA-SVM2 excels in terms of correlation coefficient, RMSE, MAE, and MAPE, demonstrating its strong predictive capabilities and precision in estimating the absorption coefficient of blood. Although it has slightly bigger errors, ANN2 also makes predictions that are fairly accurate. In contrast, SVM2 lags significantly behind in terms of prediction accuracy and precision.

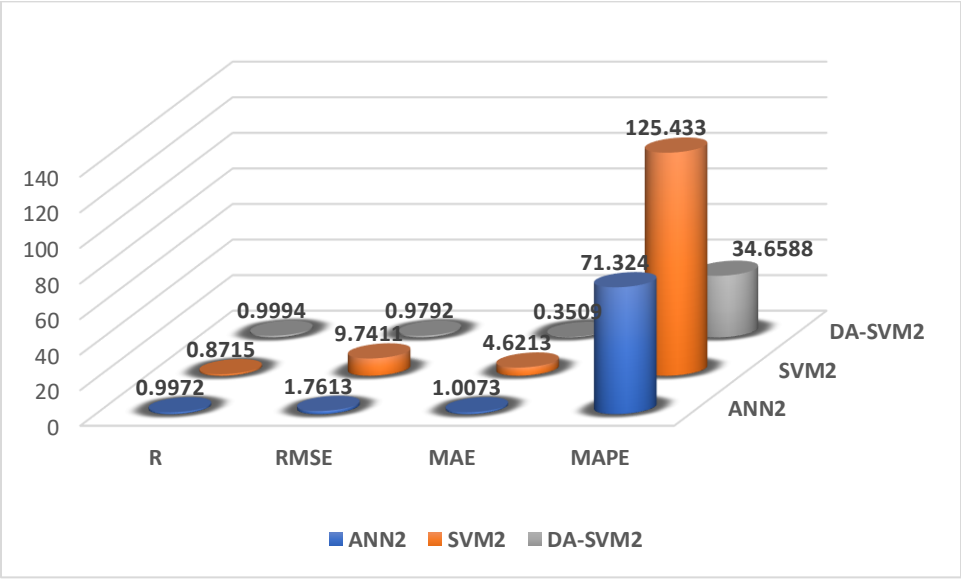


Figure V. 29. comparison between the three modes developed for the absorption coefficient

Figure V. 30 provides a comprehensive comparison of three predictive models: ANN3, SVM3, and DA-SVM3 used for estimating the scattering coefficient of blood during the testing phase. The evaluation is always based on the correlation coefficient (R), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

For ANN3, SVM3, and DA-SVM3, respectively, the correlation coefficients were 0.9987, 0.9788, and 0.9956. The noteworthy correlation coefficient provided by ANN3 indicates an excellent agreement between the values of the predicted and measured scattering coefficients. DA-SVM3 also performs impressively in this regard, with a strong correlation. SVM3 lags slightly behind but still maintains a reasonable correlation.

The RMSE results for ANN3, SVM3, and DA-SVM3 were 1.4601, 6.3632, and 2.9193, respectively. With the lowest RMSE, ANN3 stands out for having excellent predicting accuracy for the scattering coefficient. DA-SVM3 follows closely with a low RMSE, signifying strong predictive accuracy.

The MAE values for the ANN3, SVM3, and DA-SVM3 models were 0.8299, 5.8679, and 0.9692, respectively. Both ANN3 and DA-SVM3 perform much better than SVM3, demonstrating their accuracy in reducing absolute prediction errors. The SVM3 model has the highest MAE, which indicates that its predictions have considerably bigger absolute errors.

The MAPE values were 1.4408, 3.3594, and 3.1221 for ANN3, SVM3, and DA-SVM3, respectively. The lowest MAPE is achieved by ANN3, indicating that it performs well in minimizing prediction errors as a percentage of actual values. DA-SVM3 follows closely, while SVM3 displays a higher MAPE, indicating a larger percentage of inaccuracy in its predictions.

the analysis and comparison of these models for the testing phase reveal that ANN3 excels in terms of correlation coefficient, RMSE, MAE, and MAPE, highlighting its great predictive powers and accuracy in predicting the scattering coefficient of blood. In comparison to SVM3, DA-SVM3 also offers precise predictions with reduced MAE and MAPE values. While still reasonable, SVM3 is less accurate and precise in its predictions.

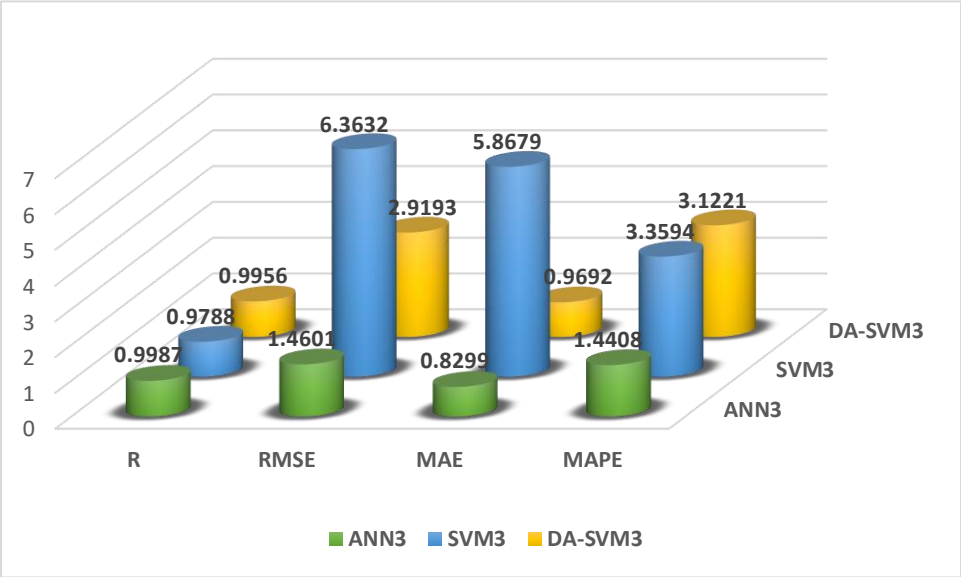


Figure V. 30. comparison between the three modes developed for the scattering coefficient

V.6. Comparison between our models and the models available in the literature

To assess the significance of the results obtained, we attempted to compare them with similar considers, and particularly, with the models that have the same inputs as we used.

Table V. 6 shows the comparison between the MLP-ANN model and three other models finding, the approach was compared in terms of correlation coefficient (R), the root mean squared error (RMSE), and the mean absolute error (MAE), by comparing these results, there was no significant difference with Oyehan's results, and we are able to say that our model performed better than the other models.

Table V. 6. Comparison with the existing models in the literature

Statistical parameters	MLP-ANN model			GA-SVM model		Yahia model	Genetic programming model	
Number of data	Train	Test	Val	Train	Test	All	Train	Test
	558	186	186	385	96	481	445	47
Correlation coefficient (R)	0.9992	0.9990	0.9991	0.9994	0.9991	0.9878	0.9746	0.9712
RMSE	9.50010^{-4}	0.0011	0.0011	0.00028	0.00034	0.0046	1.8410^{-3}	
MAE	0.0463	0.0504	0.0458	0.00023	0.00027	0.0039	2.0810^{-3}	

V.7. Conclusion

In this chapter, we present the results derived from the development of various models employing artificial intelligence techniques, specifically Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Dragonfly Algorithm-aided Support Vector Machines (DA-SVM), for the prediction of blood's optical properties. Our study encompasses the construction of three ANN models, which demonstrated remarkable performance by yielding minimal errors and exhibiting high correlation coefficients. Sensitivity analysis results corroborated the selection of appropriate input variables.

Furthermore, the DA-SVM models showcased their robustness and precision in predicting all three optical properties. This was evidenced not only by the strong correlation observed between experimental and calculated values but also by the consistently low error rates. When compared to previous models, SVM, while delivering results of acceptable quality, presents slightly lower accuracy. Nevertheless, the overall outcomes affirm the reliability and precision of our developed models in predicting these critical optical properties of blood. Importantly, our models have been found to compete effectively with existing models in the literature, particularly excelling in predicting the refractive index, underlining their competence and advancement in this field.

GERAL CONCLUSION AND PERSPECTIVES

General conclusion and perspectives

This thesis explored the crucial significance of the optical properties of blood in medical diagnostics and therapeutic applications within the field of laser medicine. However, due to the inherent complexity of blood tissue, experimental measurement of these properties is a formidable challenge.

Understanding the optical properties of blood is important for both routine medical diagnostics as well as various diagnostic and therapeutic applications in laser medicine. The evaluation of the optical characteristics of human blood offers significant implications for biology, medical research, and healthcare. These optical measurements can provide invaluable insights into the composition and function of blood, aiding in the detection, monitoring, and treatment of various medical conditions.

Artificial intelligence techniques, such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and the Dragonfly Algorithm combined with Support Vector Machine (DA-SVM), have become indispensable in biomedical research due to their capacity for complex data analysis, predictive capabilities, time efficiency, automation, scalability, and constant learning. Their widespread adoption in the field is driven by their ability to reveal hidden patterns and insights, increase diagnostics, and improve medical research.

In our study, we harnessed the power of ANN, SVM, and DA-SVM to predict the optical properties of blood which are the refractive index, the absorption coefficient, and the scattering coefficient. Contributing to the advancement of medical science and the quest for accurate, efficient, and data-driven biomedical research methodologies.

For this purpose, three feed-forward neural models are constructed, each intended to predict one of three optical properties. Each model follows a similar architecture, featuring an input layer with specific variables, a hidden layer with an optimized number of neurons determined during the learning process, and an output layer. The first model, designed for refractive index prediction, incorporates four input neurons representing wavelength, hemoglobin concentration, oxygen saturation, and temperature. Subsequent models share a three-input neuron structure, representing wavelength, hematocrit, and oxygen saturation, with one output neuron dedicated to either absorption coefficient or scattering coefficient prediction. The three ANN models showed a good agreement between the experimental and predicted values (the root mean square errors for the test

databases were: 0.11 for the ANN1 model, which is describing the refractive index of blood; 1.7613 for the model describing the optical propriety of absorption (ANN2); and 1.4601 for ANN3 describing the scattering coefficient).

On the other hand, the sensitivity analysis conducted in our study revealed the most influential variables for each optical property. For the refractive index, the key factors are the wavelength, hemoglobin concentration, temperature, and oxygen saturation. Regarding the absorption coefficient, the dominant variable is the wavelength, followed by oxygen saturation and hematocrit. For the scattering coefficient, the primary influencer is the wavelength, with hematocrit and oxygen saturation also playing significant roles.

In our second contribution, we introduced three SVM models aimed at predicting the refractive index, absorption coefficient, and scattering coefficient. Notably, these models demonstrated commendable predictive capabilities, exhibiting strong correlations between the experimental and predicted values.

In our third contribution, we delved into the innovative realm of the Dragonfly Algorithm Support Vector Machine (DA-SVM) and constructed models to estimate the same optical properties. The results were particularly impressive, characterized by high correlation coefficients and minimal error values.

Upon comparing these models, we found that ANN excelled in predicting the refractive index, closely followed by DA-SVM. For the absorption coefficient, the DA-SVM model emerged as the most accurate predictor, with ANN in a respectable second place. When we focus on the scattering coefficient, ANN showcased its exceptional predictive powers and precision, again followed by DA-SVM. In contrast, SVM exhibited relatively lower accuracy and precision in its predictions. The results from this study highlight the significance of selecting the appropriate model with particular optical behaviors and offer useful insights for future investigations and applications in the field of blood's optical properties.

We are able to make the following conclusions considering the results obtained:

- Artificial Intelligence approaches are statistical tools, which make it possible to adjust very general nonlinear functions to sets of points like any statistical method;

- The three AI approaches used in this study are based only on experimental and/ or semi-experimental data (deduced from analytical, empirical models, etc.), therefore the quality and quantity of data decide the accuracy and the robustness of the model;
- Neural networks, Support vector machines, Dragonfly Algorithm can be very robust and efficient even in the presence of little experimental data, provided that the latter are statistically well distributed in the input domain of the models,

Numerous avenues of research emerged from this thesis work which will be the subject of future work.

- ✓ ANN, SVM, and DA-SVM approaches have already demonstrated their efficacy in medical diagnostics. Future research can focus on refining these models for more accurate and early disease detection, personalized treatment plans, and improved patient care.
- ✓ As technology advances, there is room for improving the existing models further. We can fine-tune the algorithms, incorporate additional variables, or develop hybrid models for even more accurate predictions.
- ✓ The accuracy of these models is closely tied to the quality and quantity of data. As a future perspective, we may involve efforts to gather larger and more diverse datasets to further enhance model performance.
- ✓ As a natural progression, in the future research, we can extend beyond blood to explore the optical properties of various biological tissues. Investigating the behavior of light in tissues such as skin, muscle, or organs can offer insights for applications in dermatology, surgery, and non-invasive diagnostics.
- ✓ To further enhance predictive accuracy and adapt to different challenges, future research may delve into the application of advanced AI techniques beyond ANN, SVM, and DA-SVM. This includes deep learning models like Convolutional Neural Networks (CNNs) for image analysis and recurrent neural networks (RNNs) for time-series data.
- ✓ This study can also inspire educational initiatives, promoting the integration of AI and machine learning into medical and scientific curricula, ensuring that the next generation of researchers and healthcare professionals are well-versed in these technologies.

Bibliographic References

References

Alade, I. O., Bagudu, A., Oyehan, T. A., Rahman, M. A. A., Saleh, T. A., & Olatunji, S. O. (2018). Estimating the refractive index of oxygenated and deoxygenated hemoglobin using genetic algorithm – support vector regression model. *Computer Methods and Programs in Biomedicine*, *163*, 135–142. https://doi.org/10.1016/j.cmpb.2018.05.029

Aldhyani, T. H. H., Al-Yaari, M., Alkahtani, H., & Maashi, M. (2020). Water Quality Prediction Using Artificial Intelligence Algorithms. *Applied Bionics and Biomechanics*, *2020*. https://doi.org/10.1155/2020/6659314

Amari, S., & Wu, S. (1999). Improving support vector machine classifiers by modifying kernel functions. *Neural Networks*, *12*(6), 783–789. https://doi.org/10.1016/S0893-6080(99)00032-5

Ammi, Y., Khaouane, L., & Hanini, S. (2015). Prediction of the rejection of organic compounds (neutral and ionic) by nanofiltration and reverse osmosis membranes using neural networks. *Korean Journal of Chemical Engineering*, *32*(11), 2300–2310. https://doi.org/10.1007/s11814-015-0086-y

Andersson-engels, S. (2016). *Yavari (2016) - Diffuse reflectance spectroscopy Master Thesis.pdf*. *May*.

Baillot, D. (2018). Why are neuron axons long and spindly? *Medical Xpress*, *July*, 11–13. https://ucsdnews.ucsd.edu/pressrelease/why_are_neuron_axons_long_and_spindly%0Ahttps://medicalxpress.com/news/2018-07-neuron-axons-spindly-theyre-optimizing.html%0Ahttps://medicalxpress.com/news/2018-07-neuron-axons-spindly-theyre-optimizing.html%0A

Benimam, H., Moussa, C. S., Hentabli, M., Hanini, S., & Laidi, M. (2020). Dragonfly-Support Vector Machine for Regression Modeling of the Activity Coefficient at Infinite Dilution of Solutes in Imidazolium Ionic Liquids Using σ-Profile Descriptors. *Journal of Chemical and Engineering Data*, *65*(6), 3161–3172. https://doi.org/10.1021/acs.jced.0c00168

Boser, B. E., Vapnik, V. N., & Guyon, I. M. (1992). Training Algorithm Margin for Optimal Classifiers. *Perception*, 144–152.

Bosschaart, N., Edelman, G. J., Aalders, M. C. G., Van Leeuwen, T. G., & Faber, D. J. (2014a). A literature review and novel theoretical approach on the optical properties of whole blood. *Lasers in Medical Science*, *29*(2), 453–479. https://doi.org/10.1007/s10103-013-1446-7

Bosschaart, N., Edelman, G. J., Aalders, M. C. G., Van Leeuwen, T. G., & Faber, D. J. (2014b). A literature review and novel theoretical approach on the optical properties of whole blood. *Lasers in Medical Science*, *29*(2), 453–479. https://doi.org/10.1007/s10103-013-1446-7

Brar, S. K., & Verma, M. (2011). Measurement of nanoparticles by light-scattering techniques. *TrAC - Trends in Analytical Chemistry*, *30*(1), 4–17. https://doi.org/10.1016/j.trac.2010.08.008

Byvatov, E., Fechner, U., Sadowski, J., & Schneider, G. (2003). Comparison of Support Vector Machine and Artificial Neural Network Systems for Drug/Nondrug Classification. *Journal of Chemical Information and Computer Sciences*, *43*(6), 1882–1889. https://doi.org/10.1021/ci0341161

Calin, M. A., & Parasca, S. V. (2010). In vivo study of age-related changes in the optical properties of the skin. *Lasers in Medical Science*, *25*(2), 269–274. https://doi.org/10.1007/s10103-009-0725-9

Chatterjee, S., Biswas, S., Majee, A., Sen, S., Oliva, D., & Sarkar, R. (2022). Breast cancer detection from thermal images using a Grunwald-Letnikov-aided Dragonfly algorithm-based deep feature selection method. *Computers in Biology and Medicine*, *141*(November 2021), 105027. https://doi.org/10.1016/j.compbiomed.2021.105027

Diagnostic strategies for blood borne infections in Sweden. (n.d.).

Drucker, H., Surges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support vector regression machines. *Advances in Neural Information Processing Systems*, *1*, 155–161.

Elblbesy, M. A. (2021). The refractive index of human blood measured at the visible spectral region by single-fiber reflectance spectroscopy. *AIMS Biophysics*, *8*(1), 57–65. https://doi.org/10.3934/biophy.2021004

Faber, D. J., Aalders, M. C. G., Mik, E. G., Hooper, B. A., Van Gemert, M. J. C., & Van Leeuwen, T. G. (2004). Oxygen saturation-dependent absorption and scattering of blood. *Physical Review Letters*, *93*(2), 2–5. https://doi.org/10.1103/PhysRevLett.93.028102

Friebel, M., Do, K., Hahn, A., Mu, G., Berlin, D.-, Medizin-, L.-, & Universita, F. (1999). O Ptical P Roperties of C Irculating H Uman B Lood in the W Avelength R Ange 400 – 2500 Nm. *Journal of Biomedical Optics*, *4*(1), 36–46.

Friebel, M., Helfmann, J., Netz, U., & Meinke, M. (2009). Influence of oxygen saturation on the optical scattering properties of human red blood cells in the spectral range 250 to 2000 nm. *Journal of Biomedical Optics*, *14*(3), 034001. https://doi.org/10.1117/1.3127200

Friebel, M., & Meinke, M. (2005). Determination of the complex refractive index of highly concentrated hemoglobin solutions using transmittance and reflectance measurements. *Journal of Biomedical Optics*, *10*(6), 064019. https://doi.org/10.1117/1.2138027

Friebel, M., & Meinke, M. (2006). Model function to calculate the refractive index of native hemoglobin in the wavelength range of 250-1100 nm dependent on concentration. *Applied Optics*, *45*(12), 2838–2842. https://doi.org/10.1364/AO.45.002838

Friebel, M., Roggan, A., Müller, G., & Meinke, M. (2006). Determination of optical properties of human blood in the spectral range 250 to 1100 nm using Monte Carlo simulations with hematocrit-dependent effective scattering phase functions. *Journal of Biomedical Optics*, *11*(3), 034021. https://doi.org/10.1117/1.2203659

Gienger, J., Smuda, K., Müller, R., Bär, M., & Neukammer, J. (2019). Refractive index of human red blood cells between 290 nm and 1100 nm determined by optical extinction measurements. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-38767-5

Gupta, D., Bleakley, B., & Gupta, R. K. (2007). Dragon's blood: Botany, chemistry and therapeutic uses. *Journal of Ethnopharmacology*, *115*(3), 361–380. https://doi.org/10.1016/j.jep.2007.10.018

Gupta, P. K. (1999). Tissue optics. *Current Science*, *76*(10), 1341–1347. https://doi.org/10.1117/3.684093

Hamadache, M., Hanini, S., Benkortbi, O., Amrane, A., Khaouane, L., & Moussa, C. S. (2016). Artificial neural network-based equation to predict the toxicity of herbicides on rats. *Chemometrics and Intelligent Laboratory Systems*, *154*, 7–15. https://doi.org/10.1016/j.chemolab.2016.03.007

Hammer, M., Schweitzer, D., Michel, B., Thamm, E., & Kolb, A. (1998). Single scattering by red blood cells. *Applied Optics*, *37*(31), 7410. https://doi.org/10.1364/ao.37.007410

Hariharan, M., Sindhu, R., Vijean, V., Yazid, H., Nadarajaw, T., Yaacob, S., & Polat, K. (2018). Improved binary dragonfly optimization algorithm and wavelet packet based non-linear features for infant cry classification. *Computer Methods and Programs in Biomedicine*, *155*, 39–51. https://doi.org/10.1016/j.cmpb.2017.11.021

He, Z., Wen, X., Liu, H., & Du, J. (2014). A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *Journal of Hydrology*, *509*(February 2014), 379–386. https://doi.org/10.1016/j.jhydrol.2013.11.054

Howard, F. H. (1920). the Determination of the Refractivity of Hemoglobin in Solution. *Journal of Biological Chemistry*, *41*(4), 537–547. https://doi.org/10.1016/s0021-9258(18)87181-2

https://chemistrytalk.org/spectrophotometer-chemistry-instrument/

https://images.app.goo.gl/rNQMX1XcS1PiMtAm9

https://images.app.goo.gl/JbSbawN8TBXzUjAs7

https://microbenotes.com/blood-components-formation-functions/

http://www.clker.com/clipart-monocyte.html

https://www.cedars-sinai.org/newsroom/research-reveals-blood-platelets-play-important-role-in-kawasaki-disease/

https://www.microscopemaster.com/granulocytes.html

https://www.researchgate.net/publication/296478720_Albuminbased_drug_delivery_Harnessing_nature_to_cure_disease

https://www.shutterstock.com/search/lymphocyte

Ibrir, A., Kerchich, Y., Hadidi, N., Merabet, H., & Hentabli, M. (2021). Prediction of the concentrations of PM1, PM2.5, PM4, and PM10 by using the hybrid dragonfly-SVM algorithm. *Air Quality, Atmosphere and Health*, *14*(3), 313–323. https://doi.org/10.1007/s11869-020-00936-1

Jacques, S. L. (2013). Erratum: Optical properties of biological tissues: A review (Physics in Medicine and Biology (2013) 58). *Physics in Medicine and Biology*, *58*(14), 5007–5008. https://doi.org/10.1088/0031-9155/58/14/5007

Jedrzejewska-Szczerska, M. (2013). Measurement of complex refractive index of human blood by low-coherence interferometry. *European Physical Journal: Special Topics*, *222*(9), 2367–2372. https://doi.org/10.1140/epjst/e2013-02018-7

Jin, Y. L., Chen, J. Y., Xu, L., & Wang, P. N. (2006). Refractive index measurement for biomaterial samples by total internal reflection. *Physics in Medicine and Biology*, *51*(20). https://doi.org/10.1088/0031-9155/51/20/N02

Khaouane, L., Ammi, Y., & Hanini, S. (2017). Modeling the Retention of Organic Compounds by Nanofiltration and Reverse Osmosis Membranes Using Bootstrap Aggregated Neural Networks. *Arabian Journal for Science and Engineering*, *42*(4), 1443–1453. https://doi.org/10.1007/s13369-016-2320-2

Laidi, M., & Hanini, S. (2013). Optimal solar COP prediction of a solar-assisted adsorption refrigeration system working with activated carbon/methanol as working pairs using direct and inverse artificial neural network. *International Journal of Refrigeration*, *36*(1), 247–257. https://doi.org/10.1016/j.ijrefrig.2012.09.016

Lazareva, E., & Tuchin, V. (2018). Blood refractive index modelling in the visible and near infrared spectral regions. *Journal of Biomedical Photonics & Engineering*, *4*(1), 010503. https://doi.org/10.18287/jbpe18.04.010503

Li, H., Lin, L., & Xie, S. (2000). Refractive index of human whole blood with different types in the visible and near-infrared ranges. *Laser-Tissue Interaction XI: Photochemical, Photothermal, and Photomechanical*, *3914*, 517. https://doi.org/10.1117/12.388073

Li, L. L., Zhao, X., Tseng, M. L., & Tan, R. R. (2020). Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. *Journal of Cleaner Production*, *242*, 118447. https://doi.org/10.1016/j.jclepro.2019.118447

Magazine, M.H. (2021). Word press http://webphysique.fr/indice-de-refraction

Matiatou, M., Giannios, P., Koutsoumpos, S., Michalopoulos, N. V., Toutouzas, K. G., Zografos, G. C., & Moutzouris, K. (2022). Complex refractive index of freshly excised human breast tissue as a marker of disease. *Lasers in Medical Science*. https://doi.org/10.1007/s10103-022-03524-0

Meinke, M. C., Friebel, M., & Helfmann, J. (2011). Optical Properties of Flowing Blood Cells. *Advanced Optical Flow Cytometry: Methods and Disease Diagnoses*, *Mcv*, 95–132. https://doi.org/10.1002/9783527634286.ch5

Meinke, M., Müller, G., Helfmann, J., & Friebel, M. (2007a). Empirical model functions to calculate hematocrit-dependent optical properties of human blood. *Applied Optics*, *46*(10), 1742–1753. https://doi.org/10.1364/AO.46.001742

Meinke, M., Müller, G., Helfmann, J., & Friebel, M. (2007b). Optical properties of platelets and blood plasma and their influence on the optical behavior of whole blood in the visible to near infrared wavelength range. *Journal of Biomedical Optics*, *12*(1), 014024. https://doi.org/10.1117/1.2435177

Meraihi, Y., Ramdane-Cherif, A., Acheli, D., & Mahseur, M. (2020). Dragonfly algorithm: a comprehensive review and applications. *Neural Computing and Applications*, *32*(21), 16625–16646. https://doi.org/10.1007/s00521-020-04866-y

Mirjalili, S. (2016). Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Computing and Applications*, *27*(4), 1053–1073. https://doi.org/10.1007/s00521-015-1920-1

Mordon, S., Rochon, P., Dhelin, G., & Lesage, J. C. (2005). Dynamics of temperature dependent modifications of blood in the near-infrared. *Lasers in Surgery and Medicine*, *37*(4), 301–307. https://doi.org/10.1002/lsm.20238

Nilsson, A. M. K., Lucassen, G. W., Verkruysse, W., Andersson-Engels, S., & Van Gemert, M. J. C. (1997). Changes in optical properties of human whole blood in vitro due to slow heating. *Photochemistry and Photobiology*, *65*(2), 366–373. https://doi.org/10.1111/j.1751-1097.1997.tb08572.x

ODUNCUOĞLU, M. (2021). Refractive Index Formula of Blood as a Function of Temperature and Concentration. *Anais Da Academia Brasileira de Ciências*, *93*(suppl 4), 1–8. https://doi.org/10.1590/0001-3765202120201634

Oyehan, T. A., Alade, I. O., Bagudu, A., Sulaiman, K. O., Olatunji, S. O., & Saleh, T. A. (2018). Predicting of the refractive index of haemoglobin using the Hybrid GA-SVR approach. *Computers in Biology and Medicine*, *98*, 85–92. https://doi.org/10.1016/j.compbiomed.2018.04.024

Paper, I. (n.d.). optical properties ofblood in the near infrared spectral range A.N.Yaroslavsky, I.V.Yaroslavsky, T. Goldbach, H.-J. Schwarzmaier Department ofLaser Medicine, Heinrich-Heine-University, Universitätstr. 1, D40225 Dusseldorf Germany. *Library*, *2678*.

Patterson, M. S., Wilson, B. C., & Wyman, D. R. (1991). The propagation of optical radiation in tissue I. Models of radiation transport and their application. *Lasers in Medical Science*, *6*(2), 155–168. https://doi.org/10.1007/BF02032543

Pekel, E., & Kara, S. S. (2017). a Comprehensive Review for Artifical Neural Network. *Sigma Journal of Engineering and Natural Sciences*, *35*(1), 157–179.

Petersson, K. (2018). *Clinical applications of acoustophoresis in blood based diagnostics*.

Prince, S., & Malarvizhi, S. (2010). Estimation of optical properties of normal and diseased tissue based on diffuse reflectance spectral model. *WCE 2010 - World Congress on Engineering 2010*, *1*(July 2014), 578–582.

Properties, A. O., & City, W. (n.d.). *Aerosol Optical Properties and Contribution to Di ff erentiate Haze and Haze-Free Weather in Wuhan City*.

Raith, S., Vogel, E. P., Anees, N., Keul, C., Güth, J. F., Edelhoff, D., & Fischer, H. (2017). Artificial Neural Networks as a powerful numerical tool to classify specific features of a tooth based on 3D scan data. *Computers in Biology and Medicine*, *80*(November 2016), 65–76. https://doi.org/10.1016/j.compbiomed.2016.11.013

Ramesh, A. N., Kambhampati, C., Monson, J. R. T., & Drew, P. J. (2004). Artificial intelligence in medicine. *Annals of the Royal College of Surgeons of England*, *86*(5), 334–338. https://doi.org/10.1308/147870804290

Rawle, A. F. (2016). Refractive index measurements. In *Encyclopedia of Spectroscopy and Spectrometry* (3rd ed.). Elsevier Ltd. https://doi.org/10.1016/B978-0-12-409547-2.12090-6

Rezrazi, A., Hanini, S., & Laidi, M. (2016). An optimisation methodology of artificial neural network models for predicting solar radiation: a case study. *Theoretical and Applied Climatology*, *123*(3–4), 769–783. https://doi.org/10.1007/s00704-015-1398-x

Roubehie Fissa, M., Lahiouel, Y., Khaouane, L., & Hanini, S. (2019). QSPR estimation models of normal boiling point and relative liquid density of pure hydrocarbons using MLR and MLP-ANN methods. *Journal of Molecular Graphics and Modelling*, *87*, 109–120. https://doi.org/10.1016/j.jmgm.2018.11.013

Sardar, D. K., & Levy, L. B. (1998). Optical properties of whole blood. *Lasers in Medical Science*, *13*(2), 106–111. https://doi.org/10.1007/s101030050062

Sardar, Dhiraj K., Swanland, G. Y., Yow, R. M., Thomas, R. J., & Tsin, A. T. C. (2007). Optical properties of ocular tissues in the near infrared region. *Lasers in Medical Science*, *22*(1), 46–52. https://doi.org/10.1007/s10103-006-0421-y

Si-Moussa, C., Hanini, S., Derriche, R., Bouhedda, M., & Bouzidi, A. (2008). Prediction of high-pressure vapor liquid equilibrium of six binary systems, carbon dioxide with six esters, using an artificial neural network model. *Brazilian Journal of Chemical Engineering*, *25*(1), 183–199. https://doi.org/10.1590/S0104-66322008000100019

Singh, S. (2002). Refractive Index Measurement and its Applications. *Physica Scripta*, *65*(2), 167–180. https://doi.org/10.1238/physica.regular.065a00167

Smola, A. J., Schölkopf, B., & Müller, K. R. (1998). The connection between regularization operators and support vector kernels. *Neural Networks*, *11*(4), 637–649. https://doi.org/10.1016/S0893-6080(98)00032-X

Soh, D. C. K., Ng, E. Y. K., Jahmunah, V., Oh, S. L., Tan, R. S., & Acharya, U. R. (2020). Automated diagnostic tool for hypertension using convolutional neural network. *Computers in Biology and Medicine*, *126*(August). https://doi.org/10.1016/j.compbiomed.2020.103999

Steimann, F. (2001). On the use and usefulness of fuzzy sets in medical AI. *Artificial Intelligence in Medicine*, *21*(1–3), 131–137. https://doi.org/10.1016/S0933-3657(00)00077-4

Steinke, J. M., & Shepherd, A. P. (1992). Effects of temperature on optical absorbance spectra of oxy-, carboxy-, and deoxyhemoglobin. *Clinical Chemistry*, *38*(7), 1360–1364. https://doi.org/10.1093/clinchem/38.7.1360

Stoddard, J. L., & Adair, G. S. (1923). the Refractometric Determination of Hemoglobin. *Journal of Biological Chemistry*, *57*(2), 437–454. https://doi.org/10.1016/s0021-9258(18)85506-5

Sydoruk, O., Zhernovaya, O., Tuchin, V., & Douplik, A. (2012). Refractive index of solutions of human hemoglobin from the near-infrared to the ultraviolet range: Kramers-Kronig analysis. *Journal of Biomedical Optics*, *17*(11), 115002. https://doi.org/10.1117/1.jbo.17.11.115002

Talbi, M. L. (2011). *Analyse et Traitement du Signal Électrocardiographique (ECG)*.

Tuchin, V. V. (n.d.). *Introduction to Biophotonics Emerging Technology Platforms for Stem Cells Emerging Tools for Single-Cell Analysis Fluorescence Methods and Applications Flow Cytometry Fundamentals of Light Microscopy and Electronic Imaging*.

Tzanakou, E., Phan, F., & Sideman, S. (2000). Speaker identification using neural networks and wavelets. *IEEE Engineering in Medicine and Biology Magazine*, *19*(1), 92-101.

Van Dao, D., Ly, H. B., Trinh, S. H., Le, T. T., & Pham, B. T. (2019). Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete. *Materials*, *12*(6).

https://doi.org/10.3390/ma12060983

Version, D., & Engineering, B. (2021). *Acoustofluidic preparation of whole blood components*.

Wang, J., Deng, Z., Wang, X., Ye, Q., Zhou, W., Mei, J., Zhang, C., & Tian, J. (2015). Measurement of the refractive index of hemoglobin solutions for a continuous spectral region. *Biomedical Optics Express*, *6*(7), 2536. https://doi.org/10.1364/boe.6.002536

Wang, X., Liu, Y., & Xin, H. (2021). Bond strength prediction of concrete-encased steel structures using hybrid machine learning method. *Structures*, *32*(April), 2279–2292. https://doi.org/10.1016/j.istruc.2021.04.018

Wang, Z., & Srinivasan, R. S. (2017). A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and Sustainable Energy Reviews*, *75*(September 2015), 796–808. https://doi.org/10.1016/j.rser.2016.10.079

Wray, S., Cope, M., Delpy, D. T., Wyatt, J. S., & Reynolds, E. O. R. (1988). Characterization of the near infrared absorption spectra of cytochrome aa3 and haemoglobin for the non-invasive monitoring of cerebral oxygenation. *BBA - Bioenergetics*, *933*(1), 184–192. https://doi.org/10.1016/0005-2728(88)90069-2

Xie, T., Yao, J., & Zhou, Z. (2019). DA-based parameter optimization of combined kernel support vector machine for cancer diagnosis. *Processes*, *7*(5). https://doi.org/10.3390/pr7050263

Yahya, M., & Saghir, M. Z. (2016). Empirical modelling to predict the refractive index of human blood. *Physics in Medicine and Biology*, *61*(4), 1405–1415. https://doi.org/10.1088/0031-9155/61/4/1405

Yahya, Mohammed, & Saghir, M. Z. (2015). Prediction and Experimental Measurement of Refractive Index in Ternary Hydrocarbon Mixtures. *Journal of Chemical and Engineering Data*, *60*(8), 2329–2342. https://doi.org/10.1021/acs.jced.5b00251

Yang, Z., Shang, J., Liu, C., Zhang, J., & Liang, Y. (2022). Classification of oral salivary gland tumors based on texture features in optical coherence tomography images. *Lasers in Medical Science*, *37*(2), 1139–1146. https://doi.org/10.1007/s10103-021-03365-3

Yetilmezsoy, K., Ozkaya, B., & Cakmakci, M. (2011). *ARTIFICIAL INTELLIGENCE-BASED PREDICTION MODELS*. 193–218.

Yıldırım, Ö., Pławiak, P., Tan, R. S., & Acharya, U. R. (2018). Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Computers in Biology and Medicine*, *102*, 411–420. https://doi.org/10.1016/j.compbiomed.2018.09.009

Zhernovaya, O. S., Tuchin, V. V., & Meglinski, I. V. (2008). Monitoring of blood proteins glycation by refractive index and spectral measurements. *Laser Physics Letters*, *5*(6), 460–464. https://doi.org/10.1002/lapl.200810007

Zhernovaya, O., Sydoruk, O., Tuchin, V., & Douplik, A. (2011). The refractive index of human hemoglobin in the visible range. *Physics in Medicine and Biology*, *56*(13), 4013–4021. https://doi.org/10.1088/0031-9155/56/13/017

Zhernovaya, Olga S., Bashkatov, A. N., Genina, E. A., Tuchin, V. V., Meglinski, I. V., Churmakov,

D. Y., & Ritchie, L. J. (2007). *<title>Investigation of glucose-hemoglobin interaction by optical coherence tomography</title>*. *6535*, 65351C-65351C – 7. https://doi.org/10.1117/12.740985

Zijlstra, W. G., Buursma, A., & Meeuwsen-van der Roest, W. P. (1991). Absorption spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin, and methemoglobin. *Clinical Chemistry*, *37*(9), 1633–1638. https://doi.org/10.1093/clinchem/37.9.1633

Zybin, A., Koch, J., Wizemann, H. D., Franzke, J., & Niemax, K. (2005). Diode laser atomic absorption spectrometry. *Spectrochimica Acta - Part B Atomic Spectroscopy*, *60*(1), 1–11. https://doi.org/10.1016/j.sab.2004.10.001

Annex

Annex

Table 1. ANN1 architecture optimisation results

Div	Iteration	Best ANN-MLP	Train Algorithm (BFGS)	Error function SOS			Correlation coefficient R			Activation function	
				Train	Test	Val	Train	Test	Val	Hidden activation	Output activation
60-20-20	2000	4-19-1	566	$5.091.10^{-7}$	$7.486.10^{-7}$	$6.688.10^{-7}$	0.9991	0.9988	0.9989	Tanh	Exponential
	4000	1-18-1	712	$5.176.10^{-7}$	$7.727.10^{-7}$	$6.450.10^{-7}$	0.9991	0.9987	0.9989	Tanh	Exponential
	6000	4-20-1	645	$4.512.10^{-7}$	$6.059.10^{-7}$	$5.568.10^{-7}$	0.9992	0.9990	0.9991	Tanh	Exponential
	8000	4-18-1	673	$5.588.10^{-7}$	$7.823.10^{-7}$	$6.101.10^{-7}$	0.9990	0.9987	0.9989	Tanh	Exponential
70-15-15	2000	4-14-1	368	$1.144.10^{-6}$	$8.537.10^{-7}$	$1.653.10^{-6}$	0.9980	0.9987	0.9970	Tanh	Exponential
	4000	4-15-1	823	$1.019.10^{-6}$	$1.067.10^{-6}$	$1.336.10^{-6}$	0.9982	0.9984	0.9975	Tanh	Sin
	6000	4-19-1	328	$9.002.10^{-7}$	$7.298.10^{-7}$	$1.350.10^{-6}$	0.9984	0.9989	0.9975	Tanh	Exponential
	8000	4-8-1	323	$1.057.10^{-6}$	$7.813.10^{-7}$	$1.260.10^{-6}$	0.9982	0.9988	0.9976	Tanh	Exponential
80-10-10	2000	4-17-1	499	$7.081.10^{-7}$	$6.493.10^{-7}$	$1.126.10^{-6}$	0.9988	0.9990	0.9980	Tanh	Exponential
	4000	4-20-1	481	$6.418.10^{-7}$	$7.100.10^{-7}$	$1.143.10^{-6}$	0.9989	0.9989	0.9981	Tanh	Exponential
	6000	4-14-1	2096	$5.904.10^{-7}$	$4.542.10^{-7}$	$7.547.10^{-7}$	0.9990	0.9993	0.9987	Tanh	Exponential
	8000	4-19-1	670	$6.536.10^{-7}$	$6.683.10^{-7}$	$9.266.10^{-7}$	0.9989	0.9990	0.9985	Tanh	Exponential

Table 2. ANN2 architecture optimisation results

Division	Iteration	Best ANN-MLP	Training Algorithm (BFGS)	Error Function SOS			Correlation Coefficient R			Activation function	
				Train	Test	Validation	Train	Test	Validation	Hidden activation	Output activation
60-20-20	2000	3-20-1	525	2.7513	2.7376	3.2318	0.9959	0.9934	0.9943	Tanh	Identity
	4000	3-9-1	564	4.0586	4.1680	3.3736	0.9940	0.9900	0.9941	Logistic	Exponential
	6000	3-14-1	551	2.8870	3.6350	3.5420	0.9957	0.9909	0.9937	Tanh	Identity
70-15-15	2000	3-16-1	340	3.2270	3.5820	3.0260	0.9950	0.9920	0.9939	Tanh	Identity
	4000	3-15-1	467	3.0720	4.1581	3.3301	0.9955	0.9915	0.9937	Tanh	Exponential
	6000	3-6-1	1011	4.7255	3.4623	3.3526	0.9927	0.9924	0.9933	Tanh	Exponential
80-10-10	2000	3-11-1	521	4.2787	3.0645	3.8867	0.9931	0.9955	0.9951	Tanh	Exponential
	4000	3-18-1	452	4.6987	3.1848	4.7977	0.9934	0.9953	0.9929	Tanh	Exponential
	6000	3-14-1	304	4.5221	4.1650	4.1453	0.9930	0.9935	0.9938	Tanh	Identity

Division	Iteration	Best ANN-MLP	Training Algorithm (BFGS)	Error Function SOS			Correlation Coefficient R			Activation function	
				Train	Test	Validation	Train	Test	Validation	Hidden activation	Output activation
60-20-20	2000	3-24-1	570	1.6018	2.0184	2.5673	0.9982	0.9976	0.9971	Tanh	Identity
	4000	3-25-1	571	1.0584	1.0709	1.6067	0.9988	0.9987	0.9982	Tanh	Exponential
	6000	3-30-1	612	1.3705	1.3184	2.3489	0.9985	0.9984	0.9974	Logistic	Exponential
70-15-15	2000	3-28-1	517	1.3380	2.0120	2.0755	0.9985	0.9976	0.9976	Tanh	Exponential
	4000	3-29-1	1160	1.2035	1.5718	2.7670	0.9987	0.9981	0.9969	Logistic	Exponential
80-10-10	2000	3-27-1	914	1.3373	2.5488	1.4824	0.9985	0.9970	0.9984	Logistic	Exponential
	4000	3-30-1	898	1.5241	2.6377	1.7153	0.9983	0.9969	0.9981	Logistic	Exponential

Table 3. ANN3 architecture optimisation results

Table 4**.** The activation functions appeared in this study

	Activation functions		
	Hyperbolic Tangent	Exponential	Sinus
Notation in STATISTICA	***Tanh***	***Exponential***	***Sine***
Notation in Matlab	***tanh***	***exp***	***sin***
Mathematical Formula	$a = \frac{e^{+n}-e^{-n}}{e^{+n}+e^{-n}}$ (1)	$a = e^{-n}$ (2)	$a = sin(n)$ (3)

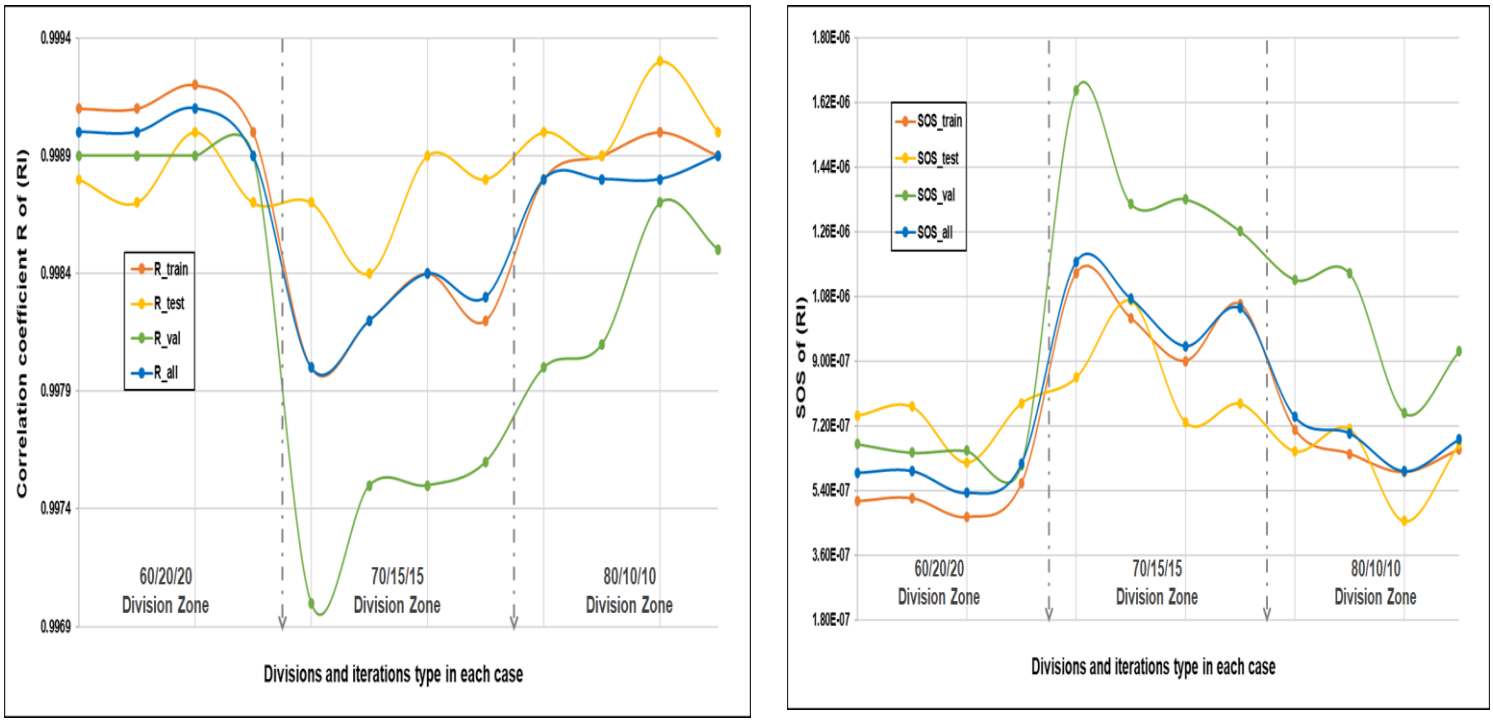


Figure 1. (a) Correlation coefficients, (b) SOS values for the refractive index

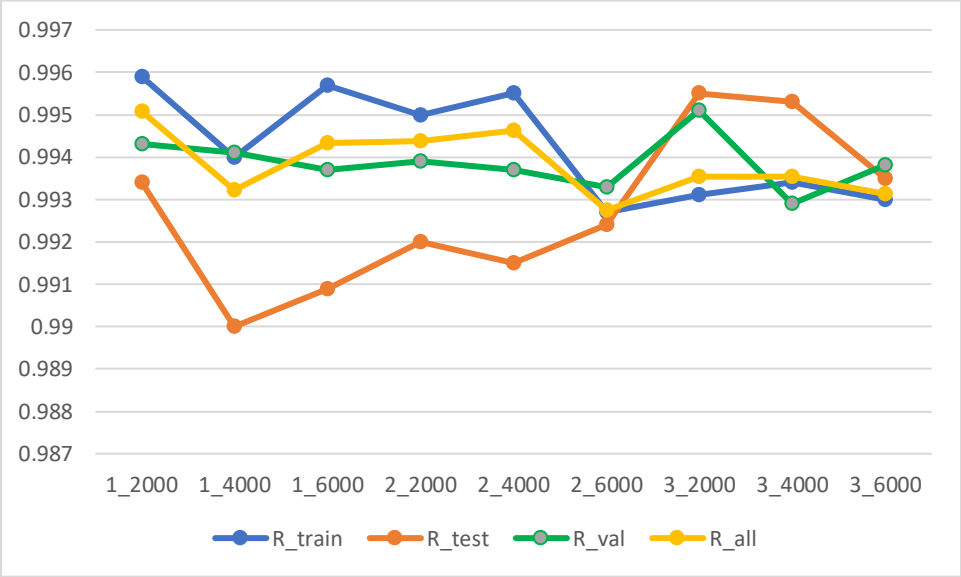


Figure 2. (a) Correlation coefficients for absorption coefficient

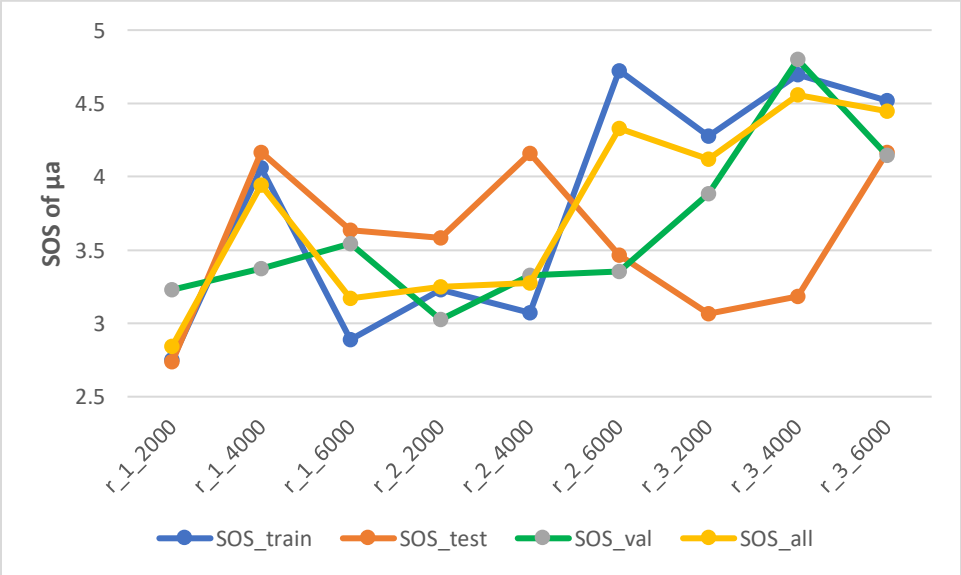


Figure 2. (b) SOS of absorption coefficient

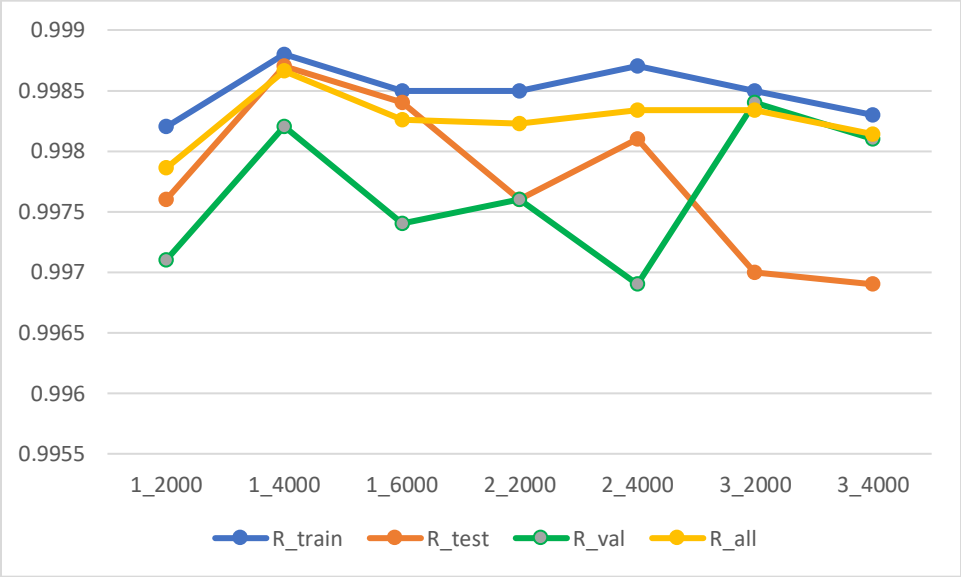


Figure 3. (a) Correlation coefficients for scattering coefficient

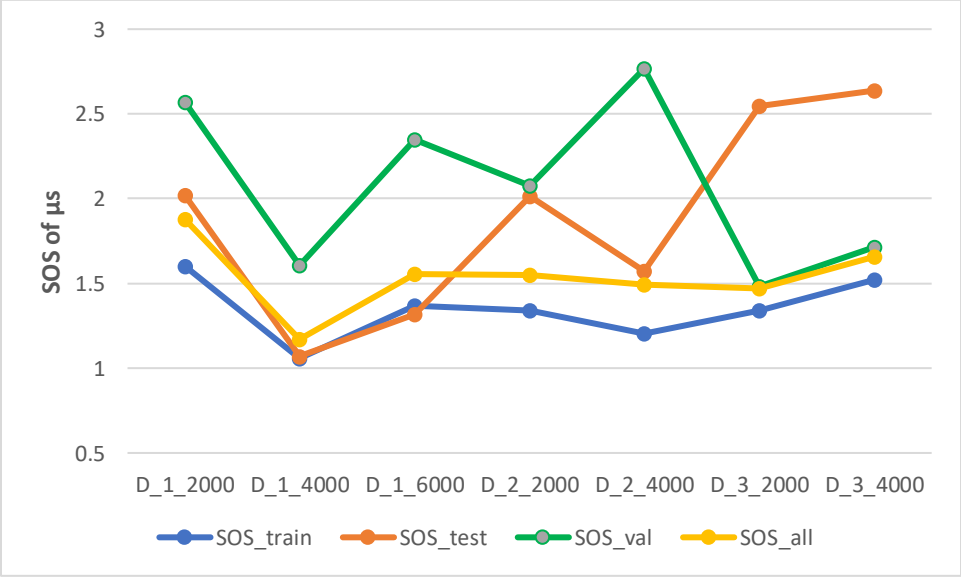
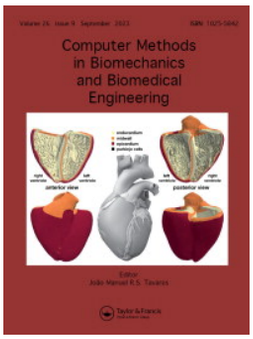


Figure 3. (b) SOS of scattering coefficient



Taylor & Francis
Taylor & Francis Group

Computer Methods in Biomechanics and Biomedical Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gcmb20

Dragonfly algorithm–support vector machine approach for prediction the optical properties of blood

Faiza Omari, Latifa Khaouane, Maamar Laidi, Abdellah Ibrir, Mohamed Roubehie Fissa, Mohamed Hentabli & Salah Hanini

To cite this article: Faiza Omari, Latifa Khaouane, Maamar Laidi, Abdellah Ibrir, Mohamed Roubehie Fissa, Mohamed Hentabli & Salah Hanini (2023): Dragonfly algorithm–support vector machine approach for prediction the optical properties of blood, Computer Methods in Biomechanics and Biomedical Engineering, DOI: 10.1080/10255842.2023.2228957

To link to this article: https://doi.org/10.1080/10255842.2023.2228957

Published online: 28 Jun 2023.

Submit your article to this journal

Article views: 24

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=gcmb20

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING
https://doi.org/10.1080/10255842.2023.2228957

Taylor & Francis
Taylor & Francis Group

Check for updates

Dragonfly algorithm–support vector machine approach for prediction the optical properties of blood

Faiza Omari[a], Latifa Khaouane[a], Maamar Laidi[a], Abdellah Ibrir[a,b], Mohamed Roubehie Fissa[a], Mohamed Hentabli[a,c] and Salah Hanini[a]

[a]Laboratory of Biomaterials and Transport Phenomena (LBMTP), Yahia Fares University, Medea, Algeria; [b]Materials and Environment Laboratory (LME), Faculty of Technology, Yahia Fares University, Medea, Algeria; [c]Quality Control Laboratory, SAIDAL Complex of Medea, Medea, Algeria

ABSTRACT

Knowledge of the optical properties of blood plays important role in medical diagnostics and therapeutic applications in laser medicine. In this paper, we present a very rapid and accurate artificial intelligent approach using Dragonfly Algorithm/Support Vector Machine models to estimate the optical properties of blood, specifically the absorption coefficient, and the scattering coefficient using key parameters such as wavelength (nm), hematocrit percentage (%), and saturation of oxygen (%), in building very highly accurate Dragonfly Algorithm-Support Vector Regression models (DA-SVR). 1000 training and testing sets were selected in the wavelength range of 250-1200 nm and the hematocrit of 0-100%. The performance of the proposed method is characterized by high accuracy indicated in the correlation coefficients (*R*) of 0.9994 and 0.9957 for absorption and scattering coefficients, respectively. In addition, the root mean squared error values (*RMSE*) of 0.972 and 2.9193, as well as low mean absolute error values (*MAE*) of 0.2173 and 0.2423, this result showed a strong match with the experimental data. The models can be used to accurately predict the absorption and scattering coefficients of blood, and provide a reliable reference for future studies on the optical properties of human blood.

ARTICLE HISTORY
Received 21 March 2023
Accepted 19 June 2023

KEYWORDS
Absorption coefficient; scattering coefficient; human blood; dragonfly algorithm; support vector machine

1. Introduction

The optical properties of blood, are important parameters in biomedical optics, and have been extensively studied in the literature. Hemoglobin (HB), the main component of blood, allows the transport of oxygen from the lungs to the human tissues and plays a crucial role in the optical properties of blood. The absorption and scattering properties of hemoglobin depend on numerous physiological and biochemical parameters, including its concentration, hematocrit HCT (the volume fraction of red blood cells in blood), oxygen saturation (the percentage of hemoglobin that is bound to oxygen), and other parameters. (Faber et al. 2004; Friebel and Meinke 2005; Friebel et al. 2009).

Various experimental techniques have been used to measure the optical properties of blood, such as spectrophotometry, spectroradiometry, and reflectance/transmission measurements. These techniques typically involve measuring the intensity of light before and after passing through or interacting with blood, and then using mathematical models to extract the absorption and scattering properties of blood.

Friebel and Meinke (2005) determined the absorption coefficient of blood by transmittance measurements that have given by Beer-Lambert's law: $T = \exp(-\mu_a d)$ where d is the thickness of the sample. The integrating sphere measurements and inverse Monte Carlo simulations have been made by Friebel et al. (2006) for highly diluted blood. Roggan et al. (1999) and Yaroslavsky et al. (1996) measured μ_a and μ_s of human blood, their results showing that blood absorption followed the absorption behavior of Hb and water. Some studies performed the scattering and absorption cross sections using Mie theory, and Rayleigh–Gans approximation (Hammer et al. 1998).

The effect of optical properties of platelets and plasma on the optical behavior of whole blood was investigated by Meinke et al. (2007a, b), their results showed that the optical properties of red blood cells (RBCs) are greater compared to the other blood components; it is important to take this into account

CONTACT Faiza Omari omari.faiza@gmail.com
© 2023 Informa UK Limited, trading as Taylor & Francis Group

when making statements concerning the optical parameters of whole blood.

Until the present, it is regrettable that only a few theoretical and experimental studies on the optical properties of blood have reliable results as evidenced by the discrepancies in the findings reported in the literature, these might be caused by the various experimental techniques and measurement devices used in these works. (Zhernovaya et al. 2011; Yahya and Saghir 2016). In addition, there are several limitations and challenges in measuring the optical properties of blood. One major challenge is the potential issues related to aggregation of red blood cells at high concentrations, which can alter the scattering properties of blood and affect the accuracy of measurement results. Another challenge is the absorption of light by other chromophores in blood, such as water, lipids, and other proteins, which can interfere with measurement of optical properties of blood is also affected by the wavelength range of interest. As well, physiological and biochemical factors, such as pH, temperature that can cause changes in optical behavior of blood, (as higher temperature leads to damage to the shape of the red blood cells RBCs from disc-shaped to spherical, denaturation, and aggregation of blood), and the presence of other chemicals, can also affect the optical properties of blood and introduce variability in the measurement results. (Nilsson et al. 1997; Mordon et al. 2005; Steinke and Shepherd 1992). Higher hemoglobin concentration, hematocrit, and oxygen saturation generally result in higher absorption of light by blood, particularly in the visible range. In the near infrared range, the absorption is lower, but scattering is still significant, which can also impact the accuracy of measurement results (Nilsson et al. 1997; Mordon et al. 2005; Fu et al. 2011).

The detailed understanding of the interaction between light and blood is pivotal in various medical applications, including routine medical diagnostics, hematology, cardiac surgery, and therapeutic applications in laser medicine (Sardar and Levy 1998; Sardar et al. 2007; Jacques 2013). Additionally, optical techniques such as optical tomography, optical biopsy, diaphanoscopic, laser-induced thermotherapy, and others, heavily rely on the optical properties of blood to assess light distribution in blood-perfused tissues (Roggan et al. 1999; Jin et al. 2006). The absorption (μ_a) and scattering (μ_s) properties of hemoglobin are particularly valuable in these techniques as they provide insights into biophysical cell parameters and shed light on blood cell diseases, including cancer (Prince and Malarvizhi 2010). However, despite their significance, there is limited scientific research available in the literature on the absorption and scattering properties of blood, in contrast to the widely studied and estimated the refractive index using various machine learning approaches. (Yahya and Saghir 2016; Oyehan et al. 2018; Alade et al. 2018; Oduncuoğlu 2021; Rakhshani and Mansouri-Birjandi 2017; Alizamir et al. 2022).

Additionally, some of these techniques do not consider the most important factors that influence the optical properties, key examples include the oxygen saturation (SO2), and hematocrit (HCT). To fill this gap, it is necessary to develop new and accurately strategy that can be reliably used to estimate the optical properties of blood.

According to what addressed before, machine learning (ML) approaches have demonstrated their predictive power in solving challenging issues in diverse fields including chemical and biomedical engineering, petroleum industry, water purification, and environment protection (Al-Jamimi et al.2019; Alade et al.2019; Wang et al. 2012; Al-Jamimi et al. 2018; Saleh 2021).

The aim of this work is to develop Dragonfly Algorithm/Support Vector Regression (DA-SVR) models to predict the optical properties of human blood, specifically, the absorption coefficient (μ_a) and the scattering coefficient (μ_s), using the previous physiological and biochemical parameters for the first time, which makes the models can be easily used in the full range of wavelengths with flexibility, and have the potential to be used for the real world applications in various fields, ranging from non-invasive blood monitoring, medical diagnostics (anemia, diabetes, and cancer) (Zhernovaya et al.2007; Prince and Malarvizhi 2010) which are associated with changes in blood optical properties, drug development, biomedical research, to clinical decision support, and can contribute to advancements in healthcare and medical technologies.

2. Materials and methods

2.1. Description of the support Vector machine regression

The SVM is one of the powerful machine learning techniques, based on statistical learning theory (SLT) which is proposed by Vapnik and Corte (Vladimir Vapnik 1995; Boser et al. 1992) it is a reliable algorithm for the classification and regression of linear and non-linear problems (Drucker et al. 1997) in the principle of structural risk minimization (SRM)

instead of empirical risk minimization based Neural networks (Byvatov et al. 2003), this improves the ability of SVR in prediction even when the dataset and their descriptors are few (Benimam et al. 2020; Ibrir et al. 2021; Hariharan et al. 2018; Hung and Chiu 2017).

Given a training dataset (x_i, y_i) with $i = 1, 2 \ldots n$, x, and y represent the model inputs and output, and n is the number of data points.

The objective of support vector regression is to find a regression function to produce the target $\hat{y}$ it can be estimated by

$$\hat{y} = \omega^{T}.\emptyset(x) + b \tag{1}$$

where $\emptyset$ represents the nonlinear function, ω is a weight vector that can be adjusted, and b is the bias which is the scaler. To find the minimum error or deviation ε, the smallest weight vector must be found by minimizing the $\|\omega\|^2$ and measuring the same permissible deviation on non-negative slack variable ε and ε^* representing upper and lower learning errors, this leads to the following expression:

$$\text{Minimise}: \frac{1}{2} \quad \|\omega\|^2 + C \sum_{i=1}^{N} (\varepsilon_i + \varepsilon_i) \tag{2}$$

$$\text{subject to} \begin{cases} y_i - \hat{y}_i = y_i - (\omega^T.\emptyset(x) + b) \leq \varepsilon + \varepsilon_i \\ \hat{y}_i - y_i = (\omega^T.\emptyset(x) + b) - y_i \leq \varepsilon + \varepsilon_i \end{cases}$$

The constant $C > 0$ is called penalty parameter or regularized parameter that measures the trade-off between the complexity of the model and empirical error (He et al. 2014), according to the method of Lagrange multiples, (L) the above optimization problem in Equation (2) can be solved to transform it into dual variables (ω; C; ε; ε^*):

Min

$$\sum_{i=1}^{N} y_i \left(\alpha_i - \alpha_i'\right) - \varepsilon \sum_{i=1}^{N} (\alpha_i - \alpha_i') - \frac{1}{2} \sum_{i=1}^{N} \sum_{i=1}^{N} (\alpha_i - \alpha_i') \left(\alpha_j - \alpha_j'\right) \emptyset(x)^{T}.\emptyset(x) \tag{3}$$

$$\text{Subject to} \sum_{i=1}^{N} (\alpha_i - \alpha_i') = 0 \quad (i = 1, 2 \ldots, N)$$

$$0 \leq (\alpha_i - \alpha_i') \leq C$$

where α and α' are Lagrange multipliers. Which are solved by a standard quadratic programming algorithm, then the regression function becomes

$$f(x) = \sum_{i=1}^{N} (\alpha_I - \alpha_I') k(x_i, x_j) + b \tag{4}$$

where (α_i - α_i') $\neq 0$, α_i and α_i' are support vectors, N is the number of support vectors, b is the bias, and the kernel function k, which facilitates solving nonlinear problems linearly, it is a measure of similarities between input vectors in feature space (Smola et al. 1998)

$$k(x_i, x_j) = \emptyset(x_i)^{T} \emptyset(x_j) \tag{5}$$

There are three different kernel functions that have a significant effect on the SVM's effectiveness (Amari and Wu 1999) which are;

- Linear, $k(x_i, x_j) = (x_i, x_j)$ (6)
- Polynomial, $k(x_i, x_j) = [\gamma (x_i, x_j) + 1]^{d}$, $\gamma > 0$ (7)
- Gaussian or radial basis function (RBF)

$$k(x_i, x_j) = \exp(-\gamma . |x_i, x_j|^2) \tag{8}$$

where d, is the degree of the polynomial, and γ is the kernel parameter

2.2. Dragonfly algorithm

Recently, Meta-heuristic algorithms have been used to optimize the SVM's parameters, such as genetic algorithm (GA) (Xie et al. 2019), particle swarm optimization (PSO) (Li et al. 2020), and the dragonfly algorithm (DA) (Moussa et al. 2022), which is proposed by Mirjalili (2016) based on the natural swarming behaviors of dragonflies, both static and dynamic (Chatterjee et al. 2022).

The Dragonfly Algorithm is based on the collective behavior and movement patterns of dragonflies, which exhibit both exploration and exploitation behaviors. The algorithm begins with an initial population of dragonflies, each representing a potential solution to the optimization problem. These dragonflies move in the search space, exploring different solutions and exploiting promising areas to find the optimal solution. The population of dragonfly demonstrates five behaviors: separation, alignment, cohesion, foraging and evading enemies or predators, the following mathematical models with Eqs (9) -(13) describe these behaviors.

- The separation (S), is to avoid collision in the neighborhood:

$$S_i = -\sum_{j=1}^{N} X - Xj \tag{9}$$

  where X represents the position of the current individual, Xj is the position in the neighborhood individual j, and N is the total number of close individuals.
- Alignment principal (A), represents the identical speed of the members existing in the same group, and it is expressed as follows:

$$A_i = \left(\sum_{j=1}^{N} Vj/N\right) \quad (10)$$

Where Vj is the velocity of the j^{th} neighboring individual.

- The cohesion (C): is relied on the Attraction of individuals towards the center of the mass of the neighborhood, and it is described as follows:

$$C_i = \left(\sum_{j=1}^{N} Xj/N\right) - X \quad (11)$$

- The foraging behavior (F), the following equation determines the attraction to a food source, with X^+ representing the location of the food.

$$Fi = X^+ - X \quad (12)$$

- The behavior of eluding enemies (E) is described as follows:

$$Ei = X + X^- \quad (13)$$

X^- Represents the position of the enemy.

The localization of an artificial dragonfly is the sum of its positions X and it step ΔX, expressed as:

$$X_{t+1} = X_t + \Delta X_{t+1} \quad (14)$$

and it is shown in the following equation:

$$\Delta X_{t+1} = sS_i + aA_i + cC_i + fF_i + eE_i + \omega \Delta X_t \quad (15)$$

Where s is the separation weight and S_i is the separation for X_t noting a is the alignment weight, A_i is the alignment of i^{th} individual, c indicates the cohesion weight, C_i is the cohesion of the i^{th} individual, F_i is the food attraction of i^{th} individual and f is the food parameter, e is the enemy factor and F_i is the enemy's location of individual dragonfly, ω is the inertia weight, and t is the iteration counter (Mirjalili 2016).

The advantages of using the Dragonfly Algorithm include:

- Efficiency: The Dragonfly Algorithm efficiently explores the search space and converges towards optimal solutions by leveraging the collective behavior and movement patterns of dragonflies.
- Flexibility: The algorithm can be easily adapted and applied to various optimization problems due to its adaptive behavior and ability to search in the full range of wavelengths.
- Robustness: The Dragonfly Algorithm is less likely to get trapped in local optima due to its combination of random search and directed flight strategies, making it a robust optimization technique.
- Scalability: The algorithm can handle large-scale optimization problems with a large number of variables and constraints, making it suitable for complex biomedical engineering applications (Mirjalili 2016).

The comparison of DA-SVM performance with existing approaches such as PSO-SVM GA-SVM algorithms, the results revealed that the DA-SVM achieved competitive results of searching the optimal SVM parameters to improve the classification accuracy (Tharwat et al. 2017).

2.3. Description of the datasets collection and computational methodology

In this work, experimental data of the absorption and scattering coefficients of human blood were compiled averages from literature, in case tabulated data, and using the Get Data Graph Digitizer program (Bosschaart et al. 2014; Roggan et al. 1999) to build the models. These data were considered sufficient in building a reliable model because of repeatability of the results. Their blood samples were obtained from adults and healthy donors, fresh erythrocytes were centrifugated three times and washed in phosphate buffer solution (300mosmol/L, pH7.4). The hematocrit, which serves as an essential marker for the concentration of red blood cells, was precisely measured by employing dilution techniques with phosphate buffer to yield a diverse range of values up to 70%. They used Light microscopy to make sure that the samples were not hemolyzed or deformed, and the blood was mixed gently with a laminar flow, and the oxygen saturation was controlled by a blood oximeter and varied by a continuous flow of a mixture of O_2, N_2, and CO_2 through a modified membrane oxygenator, it is noteworthy that the experimental set up was carried out at constant temperature 20 C°. It is known that the optical properties of scattering and absorption media cannot be measured directly, they were determined by a double integrating sphere technique, evaluating the diffuse backscattering R_d, the total transmission T_t.

It should be mentioned that all computations for this paper were done on MATLAB 2019b, each of the two datasets was split into two parts; 80% of the randomly chosen data points were used to train the DA-SVR model, while the remaining 20% were used to validate the performance of the models developed. 800 out of the 1000 data points were used to train the machine and the remaining 200 data points were

used to test the performance of the models. We used the radial basis function (RBF), particularly the Gaussian kernel type, renowned for its higher performance (Wang et al. 2003). The procedures followed by the proposed DA-SVM are presented in Figure 1. Before starting iterations, initialization DA parameters, the position vector and step vector is required, and setting the maximum number of iteration (100 iterations), population size (30), upper and lower boundaries of particles (500 and 300, respectively). The data normalized as follows, $f' = \frac{f - \min}{\max - \min}$, where f is the original value, *min* and *max* are the lower and upper bounds of the feature value, respectively, and f' is the scaled value. The DA is determined when the maximum number of iterations are attained. this unique optimization strategy was able to effectively optimize the parameters of our model and achieve superior performance compared to other traditional optimization techniques. After the full search, the optimal parameters of DA-SVM models are shown in Table 1.

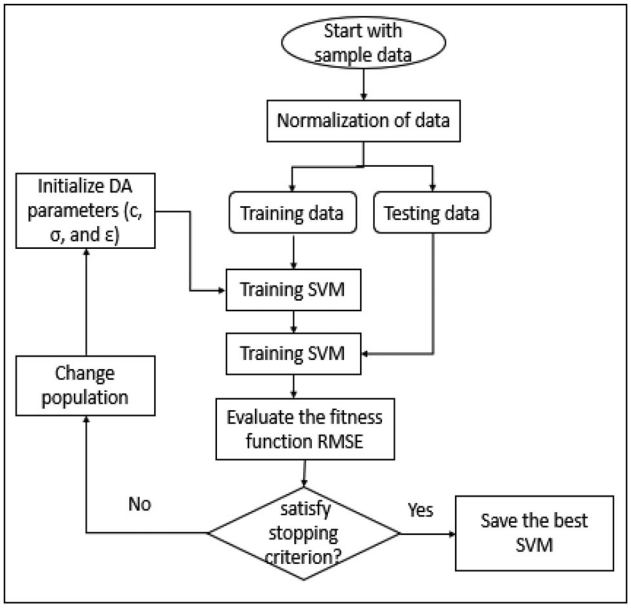


Figure 1. Flowchart of the DA-SVM process.

Table 1. Optimal DA-SVM parameters obtained.

Optimized SVR parameters	Absorption coefficient	Scattering coefficient
C	435	435
Kernel parameter	0.1	0.1046
Epsilon	0.6	0.6
Kernel function	Gaussian	Gaussian

3. Results

Proper use and interpretation of statistical parameters are crucial in research as they provide quantitative evidence to support research findings, facilitate comparisons, and help draw valid conclusions. The correlation coefficients (r) of the relationship between the features (independent variables) and the outputs (dependent variable) is an important statistical measure that can aid in feature selection, model performance assessment, and feature engineering, ultimately leading to better model interpretability, stability, and predictive accuracy (Table 2). As indicated, there is a negative relationship between the wavelength and the targets (μ_a and μ_s), and an extremely positive correlation coefficient (r) between the hematocrit and the target in both cases (μ_a and μ_s). The (r) values prove that there is a significant linear relationship between the input features variables and the targets (μ_a and μ_s) as long as *p*value are less than 0.05 and 0.01 which is a measure of the strength of evidence against the null hypothesis in a statistical hypothesis test. This indicates that these data could be used to build a reliable model for estimating the absorption and scattering coefficients of human blood. The mean or average, it is statistical parameter that provides a measure of central tendency, indicating the typical value in a dataset, is calculated by adding up all the values in a dataset and dividing by the number of values. The means and medians values of wavelength and hematocrit are very close (for both, absorption and scattering behaviors), it is an indication that the data are approximately symmetrically distributed, and the median value of absorption coefficient is closer to the lower quartile ($1.16\,mm^{-1}$), it suggested that the data may be skewed towards lower values, while the oxygen saturation value is closer to the upper quartile (89.62%), therefore, it may be skewed towards higher values.

The standard deviation measures the amount of variation in a dataset, there is larger standard

Table 2. Correlation coefficients and the respective *p*value of the relationship between the outputs and their descriptors.

	Absorption coefficient (μ_a)			Scattering coefficient (μ_s)		
Features	Wavelength vs μ_a	Hematocrit vs μ_a	Saturation of oxygen vs μ_a	Wavelength vs μ_s	Hematocrit vs μ_s	Saturation of oxygen vs μ_s
Correlation coefficient(r)	−0.4959	0.1872	0.0972	−0.7303	0.3623	0.26005
p-value	3.27×10^{-63}	2.44×10^{-09}	0.002	2.09×10^{-167}	2.15×10^{-32}	6.40×10^{-17}

Table 3. Statistical description of the absorption data adjusted in this work.

Variable	Unit	No of Data	Mean	Median	Minimum	Maximum	Standard deviation
Wavelength	nm	1000	999.030617	907.5	251	2220.8	524.300445
Hematocrit	%	1000	38.2888103	45	0.41352	56.7984	14.8005159
Oxygen saturation	%	1000	50.3865959	89.62765	0	100	49.318753
Absorption coefficient	mm^{-1}	1000	12.6135352	1.16	0.0105712	159.31	24.5790358

Table 4. Statistical description of the scattering data used in building the model.

Variable	Unit	No of Data	Mean	Median	Minimum	Maximum	Standard deviation
Wavelength	nm	1000	951.61995	792.5	272	1999.54	511.802475
Hematocrit	%	1000	38.7213047	45	0	69.7943	14.6956736
Oxygen saturation	%	1000	46.629	0	0	99	49.4413977
Scattering coefficient	mm^{-1}	1000	55.0827478	45.9794	0	128.16	30.2707498

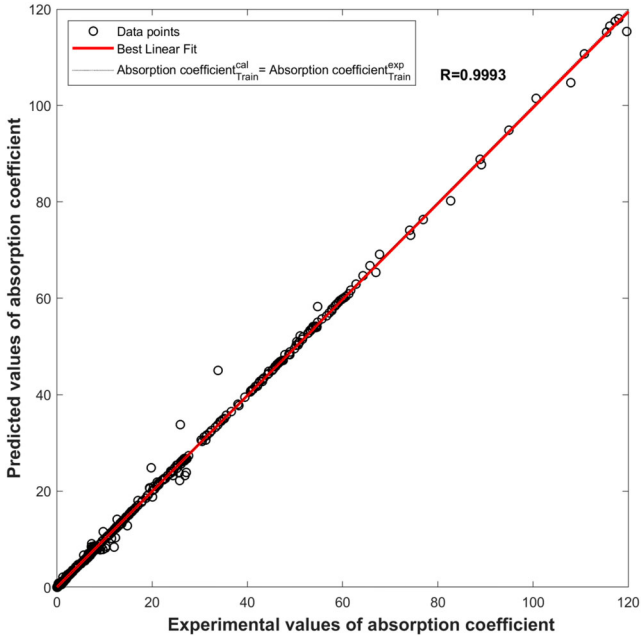


Figure 2. Scatter plots between experimental and DA-SVM predicted values for the training phase.

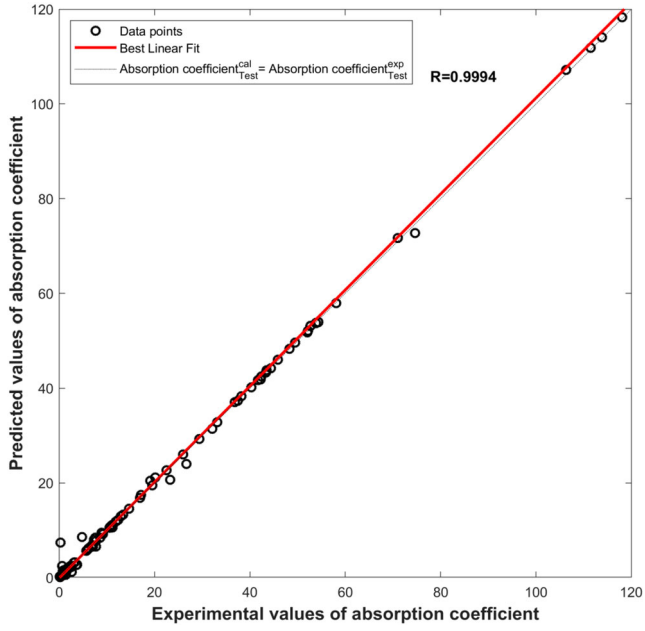


Figure 3. Scatter plots between experimental and DA-SVM predicted values for the testing phase.

deviation in the wavelength and hematocrit values, that showed greater variability in the data, while the oxygen saturation has a small standard deviation.

The descriptive statistics of these features are shown in Tables 3 and 4.

4. Discussion

To provide a comprehensive evaluation of the optimized DA-SVM models, we utilized the MATLAB function 'postreg' to plot the correlation between experimental and predicted values of absorption coefficient for both training and testing datasets. The resulting correlation coefficients were 0.9993 and 0.9994, respectively, as shown in Figures 2 and 3. These plots demonstrate the excellent performance of the models during both the training and testing phases. While the cross plots of measured and predicted values of scattering coefficient for the training and testing phases are shown in Figures 4 and 5. The plots showed correlation coefficients (R) of 0.9996 and 0.9957 for the training and testing datasets, respectively. Notably, the models displayed comparable prediction performance for the training phase, with correlation coefficients (R) values of 0.9993 and 0.9996. The correlation coefficients being close to unity indicate a strong agreement between the predicted and experimental results. This is further demonstrated by the good alignment of the predicted data set with all the experimental data on the first bisector regression, confirming the high level of accuracy achieved by the models.

The applicability domain (AD) of a prediction is also an important step in proving the reliability of the model, therefore the applicability domain is interpreted using William's plot (Figures 6 and 7). It is obvious that most samples in the dataset are within

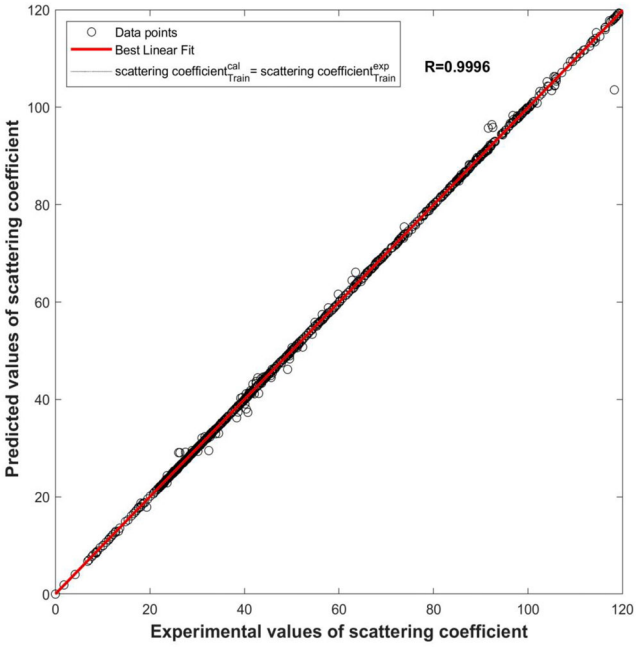


Figure 4. Correlation plot between the experimental and DA-SVM predicted values of training dataset.

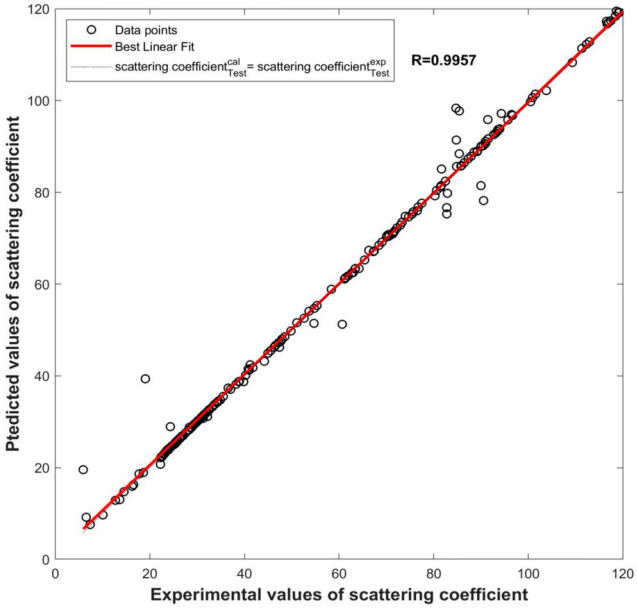


Figure 5. Correlation plot between the experimental and DA-SVM predicted values of testing dataset.

the applicability domain of the model with warning leverage limit is almost 0.013, and no more than 36 data points are outliers out of the 1000 overall experimental data for both of μ_a and μ_s.

In addition, assessing the model's predictive accuracy using statistical parameters such as the correlation coefficient (R), the root mean squared error ($RMSE$), and the mean absolute error (MAE). The mathematical formulas for these parameters are given below;

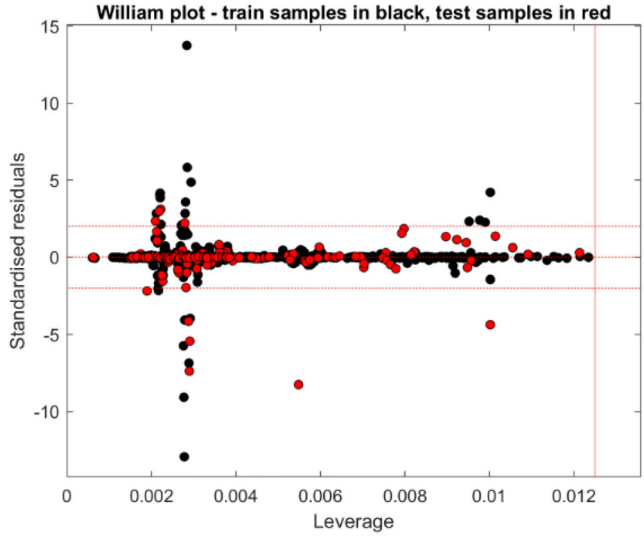


Figure 6. William's diagram for detection outliers for DA-SVM model (for μ_a).

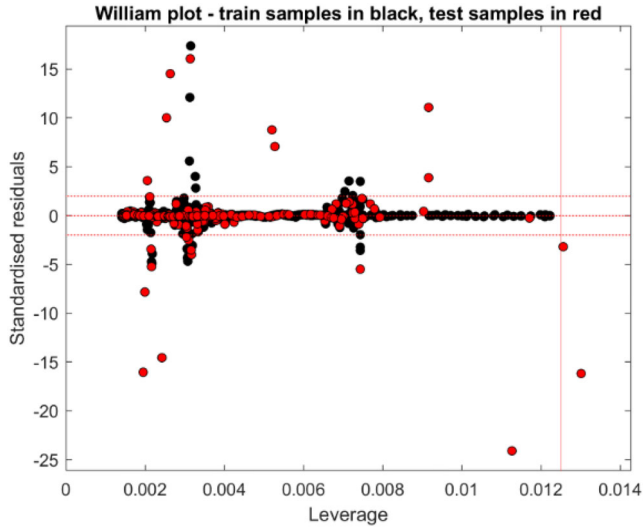


Figure 7. Williams plots to evaluate the applicability domain of the DA-SVM model (for μ_s).

$$R = \frac{\sum_i^N \left[\left(y_i^{exp} - \hat{y}^{exp}\right)\left(y_i^{cal} - \hat{y}^{cal}\right)\right]}{\sqrt{\sum_{i=1}^N \left(y_i^{exp} - \hat{y}^{exp}\right)^2}\sqrt{\sum_{i=1}^N \left(y_i^{cal} - \hat{y}^{cal}\right)^2}} \quad (16)$$

$$RMSE = \sqrt{\sum_{i=1}^N \left(y_i^{\exp} - y_i^{cal}\right)^2 / N} \quad (17)$$

$$MAE = (100/N) * \sum_{i=1}^N \left(\left|\left(\left(y_i^{\exp} - y_i^{cal}\right)/y_i^{\exp}\right)\right|\right) \quad (18)$$

Where N is the number of the dataset, y^{exp}, y^{cal} are the experimental and predicted values, respectively, and $\hat{y}^{\exp}$ is the mean of the experimental values, with

Table 5. Statistical results for the absorption and scattering coefficients of blood.

	Absorption coefficient		Scattering coefficient	
Statistics	Training	Testing	Training	Testing
Correlation Coefficient (*R*)	0.9993	0.9994	0.9996	0.9956
RMSE	0.8678	0.9792	0.8484	2.9193
MAE	0.2173	0.3509	0.2423	0.9692

$\hat{y}^{\exp} = \sum_{i=1}^{N} y_i^{\exp}/N$, the statistical results of the training and testing phases obtained during the development of the models (Table 5) revealed high correlation coefficients (close to 1). Additionally, the models demonstrated low root mean square error (*RMSE*) values of 0.8678 and 0.8484, as well as low mean absolute error (*MAE*) values of 0.2173 and 0.2423, for the training phases, respectively. These results showcase the robustness of the computational models and confirm the validity of the optimized DA-SVM models.

5. Conclusion

In this paper, we have successfully developed a new strategy to estimate the absorption (μ_a) and scattering properties of blood (μ_s) using computational intelligence. The proposed DA-SVR models used both physiological and biochemical factors such as the wavelength, the hematocrit percentage, and the oxygen saturation as descriptors in building the models The results of the Dragonfly Algorithm/Support Vector Regression models showed good correlative (*R*) with an accuracy of 0.9994 and 0.9956 for μ_a and μ_s, respectively, and low *RMSE* values were achieved. This study is pioneering the use of artificial intelligence to estimate the absorption and scattering properties of human blood, because of the excellent performance and high accuracy, we expect that this work will be helpful and opens up the horizons to the researchers to use computational intelligence approach to other optical properties of blood and even other human tissues.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors received no funds or grants for this work.

References

Alade IO, Bagudu A, Oyehan TA, Rahman MAA, Saleh TA, Olatunji SO. 2018. Estimating the refractive index of oxygenated and deoxygenated hemoglobin using genetic algorithm – support vector regression model. Comput Methods Programs Biomed. 163:135–142. doi:10.1016/j.cmpb.2018.05.029.

Alade IO, Abd Rahman MA, Saleh TA. 2019. Modeling and prediction of the specific heat capacity of Al2 O3/water nanofluids using hybrid genetic algorithm/support vector regression model. Nano-Structures Nano-Objects. 17:103–111. doi:10.1016/j.nanoso.2018.12.001.

Al-Jamimi HA, Al-Azani S, Saleh TA. 2018. Supervised machine learning techniques in the desulfurization of oil products for environmental protection: a review. Process Safety Environmental Protection. 120:57–71. doi:10.1016/j.psep.2018.08.021.

Al-Jamimi HA, Bagudu A, Saleh TA. 2019. An intelligent approach for the modeling and experimental optimization of molecular hydrodesulfurization over AlMoCoBi catalyst. J Molecular Liquids. 278:376–384. doi:10.1016/j.molliq.2018.12.144.

Alizamir A, Gholami A, Bahrami N, Ostadhassan M. 2022. Refractive Index of Hemoglobin Analysis: a Comparison of Alternating Conditional Expectations and Computational Intelligence Models. ACS Omega. 7(38):33769–33782. doi:10.1021/acsomega.2c00746.

Amari S, Wu S. 1999. Improving support vector machine classifiers by modifying kernel functions. Neural Netw. 12(6):783–789. doi:10.1016/s0893-6080(99)00032-5.

Benimam H, Moussa CS, Hentabli M, Hanini S, Laidi M. 2020. Dragonfly-Support Vector Machine for Regression Modeling of the Activity Coefficient at Infinite Dilution of Solutes in Imidazolium Ionic Liquids Using σ-Profile Descriptors. J Chem Eng Data. 65(6):3161–3172. doi:10.1021/acs.jced.0c00168.

Boser BE, Vapnik VN, Guyon IM. 1992. Training Algorithm Margin for Optimal Classifiers. Perception 144–152.

Bosschaart N, Edelman GJ, Aalders MCG, Van Leeuwen TG, Faber DJ. 2014. A literature review and novel theoretical approach on the optical properties of whole blood. Lasers Med Sci. 29(2):453–479. doi:10.1007/s10103-013-1446-7.

Byvatov E, Fechner U, Sadowski J, Schneider G. 2003. Comparison of Support Vector Machine and Artificial Neural Network Systems for Drug/Nondrug Classification. J Chem Inf Comput Sci. 43(6):1882–1889. doi:10.1021/ci0341161.

Chatterjee S, Biswas S, Majee A, Sen S, Oliva D, Sarkar R. 2022. Breast cancer detection from thermal images using a Grunwald-Letnikov-aided Dragonfly algorithm-based deep feature selection method. Comput Biol Med. 141:105027. doi:10.1016/j.compbiomed.2021.105027.

Drucker H, Surges CJC, Kaufman L, Smola A, Vapnik V. 1997. Support vector regression machines. Adv. Neural Inf. Process. Syst. 1:155–161.

Faber DJ, Aalders MCG, Mik EG, Hooper BA, Van Gemert MJC, Van Leeuwen TG. 2004. Oxygen saturation-dependent absorption and scattering of blood. Phys Rev Lett. 93(2):028102. doi:10.1103/PhysRevLett.93.028102.

Friebel M, Meinke M. 2005. Determination of the complex refractive index of highly concentrated hemoglobin solutions using transmittance and reflectance measurements. J Biomed Opt. 10(6):064019. doi:10.1117/1.2138027.

Friebel M, Roggan A, Müller G, Meinke M. 2006. Determination of optical properties of human blood in the spectral range 250 to 1100 nm using Monte Carlo simulations with hematocrit-dependent effective scattering phase functions. J Biomed Opt. 11(3):34021. doi:10.1117/1.2203659.

Friebel M, Helfmann J, Netz U, Meinke M. 2009. Influence of oxygen saturation on the optical scattering properties of human red blood cells in the spectral range 250 to 2000 nm. J Biomed Opt. 14(3):034001. doi:10.1117/1.3127200.

Fu F, Lin J, Xu X. 2011. Effect of pH on optical properties during blood coagulation progress by optical coherence tomography. 4th International Conference on Biomed Eng Inf (BMEI) 1: p. 268–271.

Hammer M, Schweitzer D, Michel B, Thamm E, Kolb A. 1998. Single scattering by red blood cells. Appl Opt. 37(31):7410–7418. doi:10.1364/ao.37.007410.

Hariharan M, Sindhu R, Vijean V, Yazid H, Nadarajaw T, Yaacob S, Polat K. 2018. Improved binary dragonfly optimization algorithm and wavelet packet based non-linear features for infant cry classification. Comput Methods Programs Biomed. 155:39–51. doi:10.1016/j.cmpb.2017.11.021.

He Z, Wen X, Liu H, Du J. 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. J Hydrol. 509: 379–386. doi:10.1016/j.jhydrol.2013.11.054.

Hung FH, Chiu HW. 2017. Cancer subtype prediction from a pathway-level perspective by using a support vector machine based on integrated gene expression and protein network. Comput Methods Programs Biomed. 141:27–34. doi:10.1016/j.cmpb.2017.01.006.

Ibrir A, Kerchich Y, Hadidi N, Merabet H, Hentabli M. 2021. Prediction of the concentrations of PM1, PM2.5, PM4, and PM10 by using the hybrid dragonfly-SVM algorithm. Air Qual Atmos Health. 14(3):313–323. doi:10.1007/s11869-020-00936-1.

Jacques SL. 2013. Erratum: optical properties of biological tissues: a review. Phys Med Biol. 58(14):5007–5008. doi:10.1088/0031-9155/58/14/5007.

Jin YL, Chen JY, Xu L, Wang PN. 2006. Refractive index measurement for biomaterial samples by total internal reflection. Phys Med Biol. 51(20):N371–N379. doi:10.1088/0031-9155/51/20/N02.

Li LL, Zhao X, Tseng MLT, Tan RR. 2020. Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. J Clean Prod. 242:118447. doi:10.1016/j.jclepro.2019.118447.

Meinke M, Müller G, Helfmann J, Friebel M. 2007a. Optical properties of platelets and blood plasma and their influence on the optical behavior of whole blood in the visible to near infrared wavelength range. J Biomed Opt. 12(1):014024. doi:10.1117/1.2435177.

Meinke M, Müller G, Helfmann J, Friebel M. 2007b. Empirical model functions to calculate hematocrit-dependent optical properties of human blood. Appl Opt. 46(10):1742–1753. doi:10.1364/ao.46.001742.

Mirjalili S. 2016. Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. Neural Comput & Applic. 27(4):1053–1073. doi:10.1007/s00521-015-1920-1.

Mordon S, Rochon P, Dhelin G, Lesage JC. 2005. Dynamics of temperature dependent modifications of blood in the near-infrared. Lasers Surg Med. 37(4):301–307. doi:10.1002/lsm.20238.

Moussa H, Dahmoune F, Hentabli M, Remini H, Mouni L. 2022. Optimization of ultrasound-assisted extraction of phenolic-saponin content from Carthamus caeruleus L. rhizome and predictive model based on support vector regression optimized by dragonfly algorithm. Chemom Intell Lab Syst. 222:104493. doi:10.1016/j.chemolab.2022.104493.

Nilsson AMK, Lucassen GW, Verkruysse W, Andersson-Engels S, Van Gemert MJ. 1997. C. Changes in optical properties of human whole blood in vitro due to slow heating. Photochem Photobiol. 65(2):366–373. doi:10.1111/j.1751-1097.1997.tb08572.x.

Oduncuoğlu M. 2021. Refractive Index Formula of Blood as a Function of Temperature and Concentration. An Acad Bras Cienc. 93:1–8.

Oyehan TA, Alade IO, Bagudu A, O Sulaiman K, Olatunji SO, Saleh TA. 2018. Predicting of the refractive index of haemoglobin using the Hybrid GA-SVR approach. Comput Biol Med. 98:85–92. doi:10.1016/j.compbiomed.2018.04.024.

Prince S, Malarvizhi S. 2010. Estimation of optical properties of normal and diseased tissue based on diffuse reflectance spectral model. WCE 2010 - World Congr Eng 2010 1: 578–582.

Rakhshani MR, Mansouri-Birjandi MA. 2017. High sensitivity plasmonic refractive index sensing and its application for human blood group identification. Sensors Actuators, B Chem. 249:168–176. doi:10.1016/j.snb.2017.04.064.

Roggan A, Friebel M, Do Rschel K, Hahn A, Mu Ller G. 1999. Optical properties for circulating human blood in thewavelength range 400– 2500 nm. J Biomed Opt. 4(1): 36–46. doi:10.1117/1.429919.

Saleh TA. 2021. Protocols for synthesis of nanomaterials, polymers, and green materials as adsorbents for water treatment technologies. Environmental Technology Innovation. 24:101821. doi:10.1016/j.eti.2021.101821.

Sardar DK, Levy LB. 1998. Optical properties of whole blood. Lasers Med Sci. 13(2):106–111. doi:10.1007/s101030050062.

Sardar DK, Swanland GY, Yow RM, Thomas RJ, Tsin ATC. 2007. Optical properties of ocular tissues in the near infrared region. Lasers Med Sci. 22(1):46–52. doi:10.1007/s10103-006-0421-y.

Smola AJ, Schölkopf B, Müller KR. 1998. The connection between regularization operators and support vector kernels. Neural Netw. 11(4):637–649. doi:10.1016/s0893-6080(98)00032-x.

Steinke JM, Shepherd AP. 1992. Effects of temperature on optical absorbance spectra of oxy-, carboxy-, and deoxy-hemoglobin. Clin Chem. 38(7):1360–1364. doi:10.1093/clinchem/38.7.1360.

Tharwat A, Gabel T, Hassanien AE. 2018. Parameter optimization of support vector machine using dragonfly algorithm. In: Hassanien A, Shaalan K, Gaber T, Tolba M, editors. Proceedings of the International Conference on Advanced Intelligent Systems and Informatics. p. 309–319. Cham: Springer International Publishing. doi:10.1007/978-3-319-64861-3_29

Vladimir NVapnik. 1995. The nature of staristical learning theory, Discov Data Min Knowl 1–47.

Wang J, Shi Z, Fan J, Ge Y, Yin J, Hu G. 2012. Self-assembly of graphene into three-dimensional structures promoted by natural phenolic acids. J Mater Chem. 22(42): 22459–22466. doi:10.1039/c2jm35024f.

Wang W, Xu Z, Lu W, Zhang X. 2003. Determination of the spread parameter in the Gaussian kernel for classification and regression. Neurocomputing. 55(3-4):643–663. doi:10.1016/S0925-2312(02)00632-X.

Xie T, Yao J, Zhou Z. 2019. DA-based parameter optimization of combined kernel support vector machine for cancer diagnosis. Processes. 7(5):263. doi:10.3390/pr7050263.

Yahya M, Saghir MZ. 2016. Empirical modelling to predict the refractive index of human blood. Phys Med Biol. 61(4):1405–1415. doi:10.1088/0031-9155/61/4/1405.

Yaroslavsky AN, Yaroslavsky IV, Goldbach, T, Schwarzmaier H J. 1996. Optical properties of blood in the near-infrared spectral range. Proceedings Volume 2678, Optical Diagnostics of Living Cells and Biofluids. San Jose, CA: Photonics West '96. doi:10.1117/12.239516.

Zhernovaya O, Bashkatov AN, Genina EA, Tuchin V, Meglinski IV, Churmakov DY, Ritchie LJ. 2007. Investigation of glucose-hemoglobin interaction by optical coherence tomograph. In. Saratov Fall Meeting: optical Technologies in Biophysics and Medicine. VIII. 6535:423–429.

Zhernovaya O, Sydoruk O, Tuchin V, Douplik A. 2011. The refractive index of human hemoglobin in the visible range. Phys Med Biol. 56(13):4013–4021. doi:10.1088/0031-9155/56/13/017.

Université Saad Dahleb Blida 1
Faculté des Sciences-
Département de Chimie

-Attestation de participation-

1er Séminaire National de Chimie Appliquée

USD BLIDA 1, 9-10 Mai 2023

Le comité d'organisation du 1er Séminaire National de Chimie Appliquée (SNCA2023), atteste que :
M. / Mme/ Melle : Faiza OMARI

A présenté une communication intitulée : Modélisation de Coefficient d'absorption du Sang par Réseau de Neurones Artificiel.

Sous forme : Orale

Co-auteurs : Samira Tared, Latifa Khaouane et Maamar Laidi.

Présidente du comité d'organisation :
Dr. Zakia ZEFFOUNI

CERTIFICATE Of PARTICIPATION

Faiza Omari

has participated in 4th International Conference on Applied Engineering and Natural Sciences on 10-13 November in 2022 at Konya/Turkey.

PAPER TITLE *Artificial Intelligence to Predict the Real Part of the Refractive Index of Hemoglobin*

PRESENTATION TYPE *Oral*

ICAENS 2022

ICAENS 2022 CONFERENCE CHAIRMAN

Asst. Prof. Dr. Umut ÖZKAYA