

## CERTIFICATION

This is to certify that the thesis entitled “**An Integrated Framework for COVID and Pneumonia Disease Prediction using Optimized Deep Learning Models**” submitted to the Avinashilingam Institute for Home Science and Higher Education for Woman, Coimbatore, for the award of the degree of **Doctor of Philosophy in Computer Science**, is a record of original research work done by **Ms. S. R. Kalaiselvi (19PHCSP004)**, during the period of her study in the Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, under my supervision and guidance and the thesis has not formed the basis for the award of any Degree / Diploma / Associateship / Fellowship or Seminar title to any candidate of any University.

  
Signature of the 31/3/25

**Head of the Department**

Dr. B. Kalpana, M.Sc., M.Phil., Ph.D.  
Professor and Head  
Department of Computer Science  
Avinashilingam Institute for Home Science  
and Higher Education for Women  
(Deemed University), Coimbatore - 641 043

  
Signature of the Supervisor 31/3/2025

Dr. R. VIJAYABHANU, MCA., M.Phil., Ph.D.,  
Associate Professor  
Department of Computer Science  
Avinashilingam Institute for Home Science  
and Higher Education for Women  
Coimbatore - 641 043

  
03/03/2025

**Signature of the Dean**

Dr. G. PADMAVATHI  
M.Sc., M.Phil., Ph.D., MISTE., MCSI.,  
Dean, School of Physical Science and  
Computational Sciences  
Avinashilingam Institute for Home Science  
and Higher Education for Women  
(Deemed to be University)  
Coimbatore - 641 043

## DECLARATION

I hereby declare that the thesis titled “**An Integrated Framework for COVID and Pneumonia Disease Prediction using Optimized Deep Learning Models**” is the result of investigations carried out by me in the Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, under the supervision and guidance of **Dr. R. Vijayabhanu**, Associate Professor, Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Woman, Coimbatore, and that it has not been submitted for the award of any Degree / Diploma / Associateship / Fellowship or similar title of any University or Institute.

  
2/3/2025

**Signature of the Supervisor**



**Signature of the Research Scholar**

Dr. R. VIJAYABHANU, MCA, M.Phil., Ph.D.  
Associate Professor  
Department of Computer Science  
Avinashilingam Institute for Home Science  
and Higher Education for Women  
Coimbatore - 641 043

## ACKNOWLEDGEMENT

First and foremost, I am extremely thankful to the GOD ALMIGHTY for his graces and blessings bestowed on me.

I express my respectful gratitude to Late **Ayya Dr. T.S. Avinashilingam Avl.**, Founder and the First Chancellor of Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing this temple of learning for women and my reverential gratitude to **Late Amma Dr. Rajammal P. Devadas Avl.**, M.A., M.Sc., Ph.D. (OHIO State), Hon. D.Sc., Hon D.Sc., Former Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her heavenly blessings.

I express my heartfelt thanks to **Dr. S. P. Thyagarajan**, D.Sc., M.D., Ph.D., Former Chancellor of Avinashilingam Institute for Home Science and Higher Education for Women, for providing me the moral support for the conduct of research work.

I am very grateful to **Sri. T.S.K. Meenakshisundaram**, M.A., M.Phil., Ph.D., Chancellor of Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the facilities to conduct the study.

I express my immense gratitude to **Dr. (Mrs.) Premavathy Vijayan**, M.Sc., M.Ed., Dip. Spl. Edn. (U.K.), M.Phil., Ph.D., Former Vice Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women for her constant encouragement throughout the research work.

I express my sincere gratitude to **Dr. (Mrs) V. Bharathi Harishankar** Ph.D., FRSA, Vice Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the necessary platform and support for research work.

My sincere thanks to **Dr. (Mrs) S. Kowsalya**, M.Sc., M.Phil., Ph.D., Former Registrar, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the opportunity to do this research.

I convey my thanks to **Dr. H. Indu** M.Sc., M.Ed., M.Phil., Ph.D., MBA, SLET., Registrar in-charge, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the facility to do this research

I owe my deep sense of gratitude to **Dr. (Mrs.) K. Manimozhi**, M.Sc., B.Ed., M.Phil., Ph.D., Former Controller of Examinations, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her amiable support and encouragement.

I place my sincere thanks to **Dr. K.Sambath Rani** M.R.Sc., M.Phil., M.Ed., (MR+VI), Ph.D., NET., Controller of Examinations, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the facility to do this research

I owe deep gratitude to **Dr. G.P. Jeyanthi**, M.Sc., M.Phil., Ph.D., former Director (Research and Consultancy), Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her constant encouragement, support, and wishes.

I record my gratefulness to **Dr. (Mrs.) P.Lalitha**, M.Sc., M.Phil., Ph.D., Dean, R&D (i/c), Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for her timely help, support and encouragement in carrying out the research work.

I am also thankful to **Dr. (Mrs). K. Udaya Chandrika**, M.Sc., M.Phil., Ph.D., and **Dr. (Mrs) G. Padmavathi**, M.Sc., M.Phil., Ph.D., Former Deans of the school of Physical Sciences and Computational Sciences, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for their invaluable support and suggestions for my research work.

I express my sincere gratitude to **Dr. (Mrs). V. Radha**, M.Sc., B.Ed., PGDOR., PGDCA., M.Phil., Ph.D., Dean, school of Physical Sciences and Computational Sciences, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her continuous support, guidance, and encouragement throughout my research journey. Her valuable insights and constant motivation have been instrumental in the successful completion of my work.

My thanks are also due to **Dr. (Mrs.) Vasantha Kalyani David**, M.Sc., M.Phil. (Maths), M.Phil (Comp.Sc.), Ph.D., Professor and Former Head of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her blessings, support, timely help, encouragement and cooperation rendered towards the completion of this research. Also, I convey my sincere thanks to

**Dr. S. N. Geethalakshmi**, Ph.D., Professor and Former Head of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for her kind cooperation and support in carrying out research work

I convey my sincere thanks to **Dr. B. Kalpana**, M.Sc., MPhil., PhD., Professor and Head of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her support and mentorship of my research work.

My heartfelt thanks to **Dr. (Mrs.) R. Vijayabhanu**, MCA., M.Phil., Ph.D., Associate Professor, Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for supporting and allowing me to carry out my research on this topic, which enabled me to learn many new skills and complete my research work on time. I am grateful to her for guiding all the discussion sessions and providing opportunities to interact with international professors, which led to the publication of papers in reputed journals during my research.

I sincerely thank the Doctoral Committee expert **Dr. P. Shanmugavadivu**, M.C.A., M.Phil., Ph.D., Professor, Department of Computer Science and Applications, Gandhigram Rural Institute, Dindigal, for providing constructive ideas and valuable suggestions during the Doctoral Committee meeting that helped me to enhance the research work.

I am thankful to the Librarians of Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for utilizing various e-resources available in our university library. I also thank the technical staff members, who are the helping hands technically to carry out my research work safely and successfully.

I extend my sincere gratitude to Dr. N.G.P. Arts and Science College for their unwavering support and encouragement throughout my research work.

I am so thankful to all my beloved friends, students, research scholars, staff members, Department of Computer Science, and members of Computer Center, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for their support, cooperation, suggestions, and wishes.

I am pleased to thank all the well-wishers who helped me in data collection and rendered their timely support on several occasions throughout the study.

My heartfelt thanks to my parents, my husband, my two little children, all my family members, friends, and everyone else who has blessed me with their prayers and good wishes for the success of this endeavour.

Besides this, I would like to extend my gratitude to one, and all who have knowingly or unknowingly, directly or indirectly helped me in the successful completion of this work.

If I have failed to mention anyone who has slipped my memory by mistake, I offer my humble apologies and render my heartfelt thanks.

-

**KALAISELVI S R**

## LIST OF FIGURES

FIG NO.	TITLE	PAGE NO
1.1	Types of disease	1
1.2	Recurring disease process	2
1.3	Overall Architecture Diagram of Disease Prediction	6
3.1	Overview of Proposed Methodology	44
4.1	Block diagram of Additive Log Ratio Transformed One Hot Encoding based Data Preprocessing	55
4.2	Zero Mean Feature Normalized Encoding based Preprocessing Model	58
4.3	Results of ALRTOHE Preprocessing using COVID-19 Dataset	62
4.4	Results of ALRTOHE Preprocessing using Pneumonia dataset	63
4.5	Results of ZMFNE Preprocessing using COVID-19 Dataset	64
4.6	Results of ZMFNE Preprocessing using Pneumonia Dataset	64
5.1	Structure of Nonlinear Sammon Projective Pattern Selection Tversky Similarity-Indexed Distributive Feature Embedding	68
5.2	Technique (TSIDE)	73
5.3	Process of Statistical Correlative Targeted Projection Pursuit-based Feature Selection	75
6.1	Step-by-step Classification Framework	83
6.2	Structure of Emphasis Perceptron Boosting Classification	84
6.3	Structure of Time-dependent Cox regressive Levenberg-Marquardt Convolutional Neural Learning Classifier (TCLMCNL)	87
6.4	Flow Process of Memetic Optimized U-Net Deep Learning Classifier for Disease Prediction	92
6.5	Schematic Structure of Deep U-Net Learning Model	93
6.6	Two-Point Crossover	97
6.7	Bit Flip Mutation Process	98
7.1 (a)	Ablation study visualization of classification models with pre- processing methods (ZMFNE) for COVID-19 dataset	116
7.1 (b)	Ablation study visualization of classification models with pre- processing methods (ZMFNE) for Pneumonia dataset	116

<b>FIG NO.</b>	<b>TITLE</b>	<b>PAGE NO</b>
7.1 (c)	Ablation study visualization of classification models with pre-processing (ZMFNE) methods (Time) for COVID-19 and Pneumonia datasets	117
7.1 (d)	Ablation study visualization of classification models without pre-processing methods for COVID-19 dataset	117
7.1 (e)	Ablation study visualization of classification models without pre-processing methods for Pneumonia dataset	118
7.1 (f)	Ablation study visualization of classification models without pre-processing methods (Time) for COVID-19 and Pneumonia datasets	118
8.1 (a)	Performance of Classification with preprocessing (ZMFNE) model for COVID-19 dataset	133
8.1 (b)	Performance of Classification without preprocessing model for COVID-19 dataset	134
8.1 (c)	Performance of prediction time with and without preprocessing model for COVID-19 dataset	134
8.1 (d)	Performance of Classification with Preprocessing (ZMFNE) model for Pneumonia dataset	135
8.1 (e)	Performance of Classification without Preprocessing model for Pneumonia dataset	135
8.1 (f)	Performance of Prediction time with (ZMFNE) and without Preprocessing model for Pneumonia dataset	136
9.1	Results of ROC Curve for preprocessing	144
9.2	Results of ROC Curve for feature selection	150
9.3 (a)	Results of Classification accuracy for COVID-19 dataset	153
9.3 (b)	Results of Prediction Time for COVID-19 dataset	154
9.3 (c)	Results of Precision and Recall for COVID-19 dataset	154
9.3 (d)	Results of Specificity for COVID-19 dataset	155
9.3 (e)	Results of Classification Accuracy for Pneumonia dataset	155
9.3 (f)	Results of Prediction Time for Pneumonia Dataset	156
9.3 (g)	Results of Precision and Recall for Pneumonia Dataset	156
9.3 (h)	Results of Specificity for Pneumonia dataset	157

<b>FIG NO.</b>	<b>TITLE</b>	<b>PAGE NO</b>
9.3 (i)	Results of Kappa statistics	159
9.3 (j)	Results of ROC curve	161
9.3 (k)	Results of AUC for MO-UNetDL Technique	162
9.3 (l)	Results of ANOVA test for proposed EPBC, TCLMCNL and MO-UNetDL techniques	163
9.4 (a)	Results of Classification Accuracy without Preprocessing and Feature Selection for COVID Dataset	165
9.4 (b)	Results of Classification Accuracy without Preprocessing and Feature Selection for Pneumonia Dataset	165
9.5 (a)	Results of Classification Accuracy with Preprocessing without Feature Selection for COVID dataset	167
9.5 (b)	Results of Classification Accuracy with Preprocessing and without Feature Selection for Pneumonia dataset	167
9.5 (c)	Outcomes of Classification without Preprocessing with Feature Selection for COVID-19dataset	169
9.5 (d)	Outcomes of Classification without Preprocessing with Feature Selection for Pneumonia dataset	169
9.5 (e)	Results of Classification with Preprocessing and Feature Selection for COVID-19 Dataset	171
9.5 (f)	Results of Classification with Preprocessing and Feature Selection for Pneumonia Dataset	171
9.6 (a)	Classification Results of Training/Testing Accuracy for COVID-19 dataset	177
9.6 (b)	Classification Results of Precision, Recall and Specificity for COVID-19 dataset	177
9.6 (c)	Classification Results of Prediction time for COVID-19 dataset	178
9.6 (d)	Classification Results of Training/Testing Accuracy for Pneumonia dataset	178
9.6 (e)	Classification Results of Precision, Recall and Specificity for Pneumonia dataset	179
9.6 (f)	Classification Results of Time for Pneumonia dataset	179

## LIST OF TABLES

<b>TABLE NO</b>	<b>TITLE</b>	<b>PAGE NO</b>
2.1	Comparative table of existing Learning Methods for Disease Prediction	<b>26</b>
2.2	Comparative table of existing Feature Selection Models for Disease Prediction	<b>28</b>
2.3	Comparative table of existing Classification Algorithms for Disease Prediction	<b>35</b>
5.1	Overall Features in COVID-19 and Pneumonia Datasets	<b>78</b>
5.2	Selected Features in COVID-19 dataset using Proposed Feature Selection Techniques	<b>79</b>
5.3	Selected Features in Pneumonia dataset using Proposed Feature Selection Techniques	<b>79</b>
6.1	Classification Accuracy – COVID-19 and Pneumonia Datasets	<b>101</b>
6.2	Results of Precision – COVID-19 and Pneumonia Datasets	<b>102</b>
6.3	Results of Recall – COVID – 19 and Pneumonia Datasets	<b>102</b>
6.4	Results of Specificity – COVID – 19 and Pneumonia Datasets	<b>103</b>
6.5	Results of Prediction Time – COVID-19 and Pneumonia Datasets	<b>104</b>
7.1	Ablation study results of proposed preprocessing models compared to the baseline models for COVID-19 and Pneumonia datasets	110
7.2	Ablation study results of proposed feature selection models with pre-processing methods for COVID-19 and Pneumonia datasets	112
7.3	Ablation study results of proposed feature selection models without pre-processing methods compared to the baseline models	112
7.4	Ablation study results of proposed classification models with feature selection (SCTPP_FS) methods for COVID-19 and Pneumonia datasets	121
7.5	Ablation study results of proposed classification models without feature selection methods for COVID-19 and Pneumonia	122

<b>TABLE</b>	<b>TITLE</b>	<b>PAGE</b>
<b>NO</b>		<b>NO</b>
	datasets	
8.1	Preprocessing Accuracy for COVID-19 and Pneumonia Dataset	127
8.2	Preprocessing Time for COVID-19 and Pneumonia datasets	128
8.3	Space Complexity for COVID-19 and Pneumonia datasets	129
8.4	Error Rate for COVID-19 and Pneumonia datasets	130
8.5	Feature selection results for COVID-19 datasets	131
9.1 (a)	Results of Preprocessing Techniques for COVID-19 dataset	141
9.1 (b)	Results of Preprocessing Techniques for Pneumonia dataset	141
9.1 (c)	Tabulation for ROC Curve for preprocessing	144
9.2 (a)	Results of Feature Selection Techniques for COVID-19 dataset	146
9.2 (b)	Results of Feature Selection Techniques for Pneumonia dataset	147
9.2 (c)	Tabulation for ROC Curve for feature selection	150
9.3 (a)	Results of Classification Techniques for COVID-19 dataset	152
9.3 (b)	Results of Classification Techniques for Pneumonia dataset	153
9.4	Kappa statistics	158
9.5	Tabulation for ROC curve	160
9.6	Classification without Preprocessing and Feature Selection	164
9.7	Classification with Preprocessing without Feature Selection Methods	166
9.8	Classification without Preprocessing with Feature Selection (SCTPP_FS) Methods	168
9.9	Classification with Preprocessing (ZMFNE) and Feature Selection (SCTPP-FS)	170
9.10	Overall comparative analysis of classification for COVID-19 dataset	172
9.11	Overall comparative analysis of classification for Pneumonia dataset	173
9.12 (a)	Results of Preprocessing Techniques for COVID-19 Dataset	174
9.12 (b)	Results of Preprocessing Techniques for Pneumonia Dataset	174
9.13 (a)	Results of Feature Selection Techniques for COVID-19 dataset	175
9.13 (b)	Results of Feature Selection Techniques for Pneumonia Dataset	175

<b>TABLE NO</b>	<b>TITLE</b>	<b>PAGE NO</b>
9.14 (a)	Results of Classification Techniques for COVID-19 dataset	176
9.14 (b)	Results of Classification Techniques for Pneumonia Dataset	176
9.15	Outcome of the Training/Testing Accuracy for the Proposed Classifiers	180

## LIST OF ABBREVIATION AND SYMBOLS

ABBREVIATIONS	DESCRIPTION
ALRTOHE	- Additive Log Ratio Transformed One Hot Encoding
BAS	- Beetle Antennae Search
Bi-LSTM	- Bi-direction a Long Short-Term Memory
CKD	- Chronic Kidney Disease
CNN	- Convolutional Neural Network
DCML	- Deep Contrastive Mutual Learning
DCML	- Deep Contrastive Mutual Learning
DM	- Data Mining
DMFL_Net	- Decision-Making-Basis Federated Learning Network
DSPM	- Deep Sequential Prediction Model
DSVM	- Deep Support Vector Machine
DWS-CNN	- Depth Wise Separable Convolution Neural Network
EPBC	- Emphasis Perceptron Boosting Classification
ES	- Exponential Smoothing
IoT	- Internet of Things
LASSO	- Least Absolute Shrinkage and Selection Operator
LR	- Linear Regression
ML	- Machine Learning
MO-UNetDL	- Memetic Optimized U-Net Deep Learning
MTGP	- Multi-Task Gaussian Process
NFR	- Novel Feature Reduction
SARS-CoV-2	- Severe Acute Respiratory Syndrome Coronavirus 2
NRM	- Non-parametric Regression Model
NSPPS	- Nonlinear Sammon Projective Pattern Selection
RBV	- Routine-Blood-Values
RPN	- Region Proposal Network
SCTPP-FS	- Statistical correlative targeted projection pursuit-based feature selection
SVM	- Support Vector Machine
TCLMCNL	- Time-dependent Cox regressive Levenberg–Marquardt Convolutional Neural Learning
TSIDFE	- Tversky Similarity-Indexed Distributive Feature Embedding
ZMFNE	- Zero Mean Feature Normalized Encoding
$A_o$	- Actual results
$D_s$	- Standard deviation
$E_r$	- Error rate
$FS_{acc}$	- Feature Selection accuracy
$F_{st}$	- Feature selection time
$L_h$	- Loss
$Mem_{psd}$	- Number of patient data samples from dataset

**ABBREVIATIONS****DESCRIPTION**

$Nd_{cp}$	- Number of data correctly preprocessed
$Ndf_{cs}$	- Number of data feature correctly selected
$Ndf_i$	- Number of data feature samples involved in the simulation process for disease prediction
$Nd_{ip}$	- Number of data samples incorrectly preprocessed
$Nf_{11}, Nf_{12} \dots, Nf_{1n}$	- Normalized numerical value of the features
$P_{Acc}$	- Preprocessing accuracy
$P_o$	- Predicted classification outputs
$P_{time}$	- Preprocessing time
$S_c$	- Space complexity
$T_{df}$	- Total number of data features
$a_i$	- Input attribute
$f_i - f_j$	- Variance between the two features
$f_i(t)$	- Input features
$f_i \cap f_j$	- Mutual dependence between the two features
$f_n$	- False negative
$f_p$	- False positive
$g_o(t)$	- Covariate vector
$m_f$	- Mean of the particular feature value
$q_{input}$	- Weight
$\rho_c$	- Regression coefficient
*	- Convolutional operator
Z	- Output of pooling layer
B	- Binary representation
D	- Bias whose value is '1'
Fs	- Feature scaling
LL	- Cross-entropy loss
M	- Mean of the attribute
NBR	- Numeric to binary representation
NF	- Normalized feature matrix
O(A)	- Observed accuracy
P(A)	- Proportion of probable accuracy by chance
P(t)	- Hidden layer output
PAC	- Prediction accuracy
PTime	- Prediction time
Pre	- Precision
R	- Cramér's phi correlation function
Rec	- Recall
Td	- Total number of data
Time [sdf]	- Time for single data feature selection
Time[psd]	- Sample data
Y (t)	- Hazards function at times
argmin	- Argument of the minimum function

**ABBREVIATIONS****DESCRIPTION**

$b$	- Output in a matrix form
$n$	- Number of samples
$tn$	- True negative
$tp$	- True positive
$\delta$	- Similarity coefficient
$\delta$	- Threshold
$\sigma$	- Deviation