CERTIFICATE

I certify that the thesis entitled "**Air Quality Prediction Using Deep Learning Techniques**" submitted for the degree of **Doctor of Philosophy (Ph.D.) in Computer Science and Engineering**, is the record of research work carried out by **Ms. P. Shree Nandhini (19PHEOF003)** during the period of her study from **July 2019 to June 2024** in the Department of Computer Science and Engineering at Avinashilingam Institute for Home Science and Higher Education for Women Coimbatore, under my guidance and supervision, and the thesis has not formed the basis for the award of any Degree/Diploma/ Associateship / Fellowship or other similar Titles of any Candidate of this Institute or any other University / Institution of Higher Learning.

10/6/2024

Signature of the
Head of the Department

10/6/2024

Signature of the Supervisor

14/6/24

Signature of the Dean

DECLARATION

I, **Ms. P. Shree Nandhini** hereby declare that the thesis entitled **"Air Quality Prediction Using Deep Learning Techniques"** submitted to the Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore in the partial fulfillment of the requirements for the award of **Doctor of Philosophy (Ph.D.) in Computer Science and Engineering** is the record of original and independent research work carried out by me during the period from **July 2019 to June 2024** under the guidance of **Dr. P. Amudha,** M.Tech., Ph.D., Professor, Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore and has not formed the basis for the award of any Degree/Diploma/ Associateship / Fellowship or any other similar titles in this Institute or any other University or other similar Institution of Higher Learning.

P. Shree Nandhini

Signature of the Research Scholar

10/6/2024

Signature of the Supervisor

LIST OF TABLES

Table No.	Title	Page No.
1.1	AQI Quantifications Based on Significant Pollutants	14
3.1	Confusion matrix	49

LIST OF FIGURES

Figure No.	Title	Page No.
1.1	Deep Learning layers	2
1.2	Architecture of ANN	3
3.1	Architecture of LSTM	36
3.2	Predictive model of Convolutional Neural Network	37
3.3	Structure of Sparse Autoencoder	39
3.4	Prediction Model of ISAE-DL	42
3.5	Architecture of Sparse Autoencoder	45
3.6	Various level of Pollutant	47
3.7	Comparison of Accuracy	50
3.8	Comparison of Precision	51
3.9	Comparison of Specificity	52
3.10	Comparison of Sensitivity	53
3.11	Comparison of AUC	54
3.12	Comparison of MCC	55
3.13	Comparison of MAER	56
3.14	Comparison of F-measure	57
4.1	Voronoi Diagram for Six Clusters	59
4.2	Architecture of VCSAE-DL	61
4.3	Comparison of Accuracy	63
4.4	Comparison of Precision	64
4.5	Comparison of Specificity	65
4.6	Comparison of Sensitivity	65
4.7	Comparison of AUC	66
4.8	Comparison of MCC	67
4.9	Comparison of F-measure	68
5.1	Structure of Model Transfer Approach	69
5.2	Proposed Model Frameworks of TSBU-LSTM	71
5.3	Architecture of T-SBU-LSTM	74
5.4	Comparison of Accuracy	79
5.5	Comparison of Precision	80
5.6	Comparison of Specificity	81
5.7	Comparison of Specificity	82
5.8	Comparison of AUC	83

Figure No.	Title	Page No.
5.9	Comparison of MCC	84
5.10	Comparison of MAER	85
5.11	Comparison of F-measure	86
6.1	Architecture of CNN –WD-DTL	88
6.2	Comparison of Accuracy	94
6.3	Comparison of Precision	95
6.4	Comparison of Specificity	96
6.5	Comparison of Sensitivity	97
6.6	Comparison of AUC	98
6.7	Comparison of MCC	99
6.8	Comparison of MAER	100
6.9	Comparison of F-measure	101
7.1	Evaluation of Accuracy	102
7.2	Evaluation of Precision	103
7.3	Evaluation of Sensitivity	104
7.4	Evaluation of Specificity	104
7.5	Evaluation of AUC	105
7.6	Evaluation of MCC	106
7.7	Evaluation of MAER	106
7.8	Evaluation of F-measure	107
7.9	Evaluation of Accuracy	109
7.10	Evaluation of Precision	110
7.11	Evaluation of Sensitivity	110
7.12	Evaluation of Specificity	111
7.13	Evaluation of AUC	112
7.14	Evaluation of MCC	114
7.15	Evaluation of MAER	114
7.16	Comparison of F-measure	115

LIST OF ABBREVIATIONS

AdaBoost	Adaptive boosting
ANN	Artificial Neural Networks
AQI	Air Quality Index
AR	Autoregressive
AUC	Area under Curve
ARMA	Autoregressive Moving Average
BPNN	Back Propagation Neural Network
COPD	Chronic Obstructive Pulmonary Disease
CH-RNN	Combination of Hammerstein Recurrent Neural Networks
CPCB	Central Pollution Control Board
DAQI	Daily Air Quality Index
DAEDN	Denoising Autoencoder Deep Network
DL	Deep Learning
DTL	Deep Transfer Learning
EPA	Environmental Protection Agency
ESP	Electrostatic Precipitator
ELM	Extreme Learning Machine
FFNNs	Feed Forward Neural Networks
GA	Genetic Algorithm
GIS	Geographic Information System
GLCM	Grey Level Co-occurrence Matrix
GRU	Gated Recurrent Unit
INAQS	Indian National Air Quality Standards
ISAE-DL	Improved Sparse Autoencoder Deep Learning
IPIO	Improved Pigeon-Inspired Optimization
LSC	Land Surface Coverage
LSTM	Long Short-Term Memory
KL	Kullback-Leibler
kNN-DTWD	k-Nearest Neighbor Dynamic Time Warping Distance
KNN-ED	k-Nearest Neighbor Euclidean Distance

MLR	Multiple Linear Regression
MELM	Modified ELM
MICE	Multimodal Imputation by Chained Equations
ML	Machine Learning
MLEGRU	Multiple Linear Regression based GRU
MTSAE	Multiview Transfer Semi-supervised learning for Air quality Estimation
MAER	Mean Absolute Error Rate
MCC	Matthew's Correlation Coefficient
MVO	Multi-Verse Optimization Method
NAAQS	National Ambient Air Quality Standards
Pb	Lead
RF	Random Forests
RNN	Recurrent Neural Network
ROC Curves	Receiver Operating Characteristic curves
SE	Stacking Ensembles
SVR	Supporting Vector Regression
SAE	Sparse Autoencoder
STANN	Spatiotemporal Artificial Neural Network
STDL	Spatiotemporal Deep Learning
STE	Spatial-Temporal Ensemble
SVM	Support Vector Machines
ST-DNN	Spatio-Temporal Deep Neural Network
TE	Terrain Extractor
TRE	Temporal Relationship Extractor
T-SBU-LSTM	Transferred Stacked Bidirectional and Unidirectional LSTM
TSLSTME	Temporal Sliding LSTM Extended Model
TDNN	Time Delay Neural Network
UV	Ultra Violet
UCE-Normalization	Usual Case Emphasizing Normalization
VOCs	Volatile Organic Compounds
VCSAE-DL	Voronai Clustering Sparse Auto Encoder Deep Learning
WHO	World Health Organization

WD-DTL	Wasserstein Distance using Deep Transfer learning
WPD	Wavelet Packet Decomposition
	Oxides
$CFCs$	Chlorofluorocarbons
CH_4	Methane
CO	Carbon Monoxide
CO_2	Carbon dioxide
H_2CO_2	Formaldehyde
NH_3	Ammonia
$NMVOCs$	Non-Methane VOCs
NO_2	Nitrogen Dioxide
O_3	Ozone
PM	Particulate Matter
Rn	Radon
$SO2$	Sulphur Dioxide

LIST OF SYMBOLS

B_{HI}	The threshold concentration is more than or equal to the value provided.
B_{Lo}	Critical concentration is below or equal to the value given.
C	The Pollutant Attentiveness
C_{Hig}	The Attentiveness Breakpoint that is $\geq C$
C_{Low}	The Attentiveness Breakpoint that is $\leq C$
C_p	Pollutants in Concentration
d	Data Transmission Weight Matrix between the LSTM Layer's Internal Neurons
$_f(t)c$	Forget Gate
$_i(t)c$	Input Gate
(I)	Final Index
(I_i)	Sub-Index
I	The (Air Quality) index
I_{HI}	Similar to a B_{HI} AQI
I_{Lo}	AQI value equivalent to B_{Lo}
n	Several factors relating to pollution
n	Several factors relating to pollution
$_o(t)c$	Output Gate
$p_d\ (G, H)$	Probability Density Function for the Joint Distribution of Random Variables $(V,)$ with Cumulative Distributions G And H
$_s(t)c$	Cell State Node
u	Matrix that holds the Weights of the Layers that come before the Convolution Layer
v	Weight Matrix of the Convolution Layer Divided by the Weight Matrix of the Pooling Layer
w	Weight Matrix from the Pooling Layer to The LSTM Layer
w_i	Pollutant's relative importance
α	Slant of the Line
β	Intercept at X=0
ξ	The Process of using the LSTM Weight Matrix on the Complete Linkage