REFERENCES

1) Abraham, S., and Li, X. A Cost-Effective Wireless Sensor Network System for Indoor Air Quality Monitoring Applications. pp. 165- 171,2014
2) Aishwarya, D. C., and Babu, C. N.,Prediction of Time Series Data Using GABPNN Based Hybrid ANN Model, IEEE 7th International Advance Computing Conference (IACC) pp. 848-853.2017.
3) Andre Esteva ,Brett Kuprel , Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. Nature, Vol 542, Issue 639, pp.115-118,2017
4) Andrew C. Kemp, Benjamin P. Horton . Climate Related Sea-Level Variations Over the Past Two Millennia. Proceedings of The National Academy of Sciences, Vol 108, Issue 27,pp. 11017-11022,2017.
5) Andrews, E., Ogren, J. A., Kinne, S., and Samset, Comparison of AoD, AAOD and Column Single Scattering Albedo from Aeronet Retrievals and In Situ Profiling Measurements. Atmospheric Chemistry and Physics, Vol 17, Issue 9,pp.6041- 6072,2017.
6) Arora, H., and Solanki, A., Prediction of Air Quality Index in Metro Cities Using Time Series Forecasting Models, Journal of Xian University of Architecture &Technology, Volume 7, Issue 06, 2020.
7) Bolstad, B. M., Irizarry, R. A., Åstrand, M., and Speed, T. P,A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Variance and Bias. Bioinformatics, Vol 19, Issue 2, pp.185-193,2003.
8) Bui, T. C., Le, V. D., and Cha, S. K, A Deep Learning Approach for Forecasting air Pollution in South Korea Using LSTM. pp.1804.07891,2018
9) Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C. A., and Coggins, J,Global Estimates of Mortality Associated With Long-Term Exposure to Outdoor Fine Particulate Matter, Proceedings of The National Academy of Sciences Vol 115,Issue 38, pp.9592-9597,2018.
10) Cai, J., Dai, X., Hong, L., Gao, Z., and Qiu, Z. An Air Quality Prediction Model Based on a Noise Reduction Self-Coding Deep Network, Mathematical Problems in Engineering, Vol 20, Issue 20, 2020.
11) Cai, S., Wang, Y., Zhao, B., Wang, S., Chang, X., and Hao, J, The Impact Of The Air Pollution Prevention and Control Action Plan on PM2. 5 Concentrations in Jingjin-Ji Region During 2012–2020, Science of The Total Environment Vol 580, pp.97-

209,2016

12) Cascio, W. E., and Long, T. C, Ambient Air Quality and Cardiovascular Health Translation Of Environmental Research for Public Health and Clinical Care, North Carolina Medical Journal, Vol 79,Issue 5 , pp.306-312,2018.
13) Chen, Y. C., Lei, T. C., Yao, S., and Wang, H. P, PM2. 5 Prediction Model Based on Combinational Hammerstein Recurrent Neural Networks. Mathematics, Vol 8, Issue 12, pp.2178,2020.
14) Cheng Cheng, Beitong Zhou, Guijun Ma, Dongrui Wu and Ye Yuan,Wasserstein Distance Based Deep Adversarial Transfer Learning for Intelligent Fault Diagnosis With Unlabeled Neurocomputing, Vol 409, pp.35- 45,2019.
15) Chin, C., and Brown, D. E., Learning in Science: A Comparison of Deep and Surface Approaches. Journal of Research in Science Teaching: The Official Journal of The National Association for Research in Science Teaching, Vol 37, Issue 2 pp.109-138,2000.
16) Choudhary, M. P., and Garg, V.Causes, Consequences and Control of Air Pollution. in All India Seminar on Methodologies for Air Pollution Control, Vol 13, Issue 13,2013.
17) Cohen AJ, Brauer M, Burnett R, Anderson HR and Forouzanfar MH., Estimates and 25-Year Trends of The Global Burden of Disease Attributable to Ambient Air Pollution: an Analysis of Data from the Global Burden of Diseases Study Lancet, pp.10082, 2017
18) Cong cong and Wen, Shufu Liu, A Novel Spatiotemporal Convolutional Long Short-Term Neural Network for Air Pollution Prediction. Science of The Total Environment, Vol 654, pp.1091- 1099, 2019.
19) Cosma, A. and Simha, R. Machine Learning Method for Real-Time Non- Invasive Prediction of Individual Thermal Preference in Transient Conditions. Build. Environment Vol. 148, pp.372-383,2018.
20) Dairi, A., Harrou, F., Khadraoui, S., and Sun, Y., Integrated Multiple Directed Attention-Based Deep Learning for Improved Air Pollution Forecasting. IEEE Transactions on Instrumentation and Measurement, Vol 70, pp.1-15,2021
21) Dhole, A., Ambekar, I., Gunjan, G., and Sonawani, S., An Ensemble Approachto Multi-Source Transfer Learning for Air Quality Prediction. In 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS),pp.

70-77,2021.

22) Disha Sharma1and Denise Mauzerall, Analysis of Air Pollution Data in India Between 2015 and 2019,Aerosol and Air Quality Research, Vol 22,Issue 2,2024.
23) Ditsuhi Iskandaryan and Francisco Ramos., Sergio Trilles, A Set of Deep Learning Algorithms for Air Quality Prediction Applications, Elsevier, Vol 17, Issue 23, pp.100562,2023.
24) Donnelly, A., Misstear, B., and Broderick, B. Real Time Air Quality Forecasting using Integrated Parametric and Non-Parametric Regression Techniques. Atmospheric Environment,Vol 103, pp.53-65,2015.
25) Du, S., Li, T., Yang, Y., and Horng, S. J, Deep Air Quality Forecasting Using Hybrid Deep Learning Framework. IEEE Transactions on Knowledge and Data Engineering, Vol 33,Issue 6,pp.2412-2424,2019
26) Duc, H., and Azzi, M. Analysis of Background Ozone in The Sydney Basin. in Proceeding,pp. 2307-2313,2009.
27) Eslami, E., Choi, Y., Lops, Y., and Sayeed,. A Real-Time Hourly OzonePrediction System Using Deep Convolutional Neural Network. Neural Computing and Applications, Vol 32, Issue 13, pp.8783-8797,2020.
28) Mishra, D, and Goyal, P. Features Extraction Model for Air Quality Forecasting. Sustainability, Vol 12,Issue 19,2015.
29) Ferreira, J., Guevara, M., Baldasano, J. M., Tchepel, O., Schaap, M., Miranda, A. I., and Borrego, C,A Comparative Analysis of Two Highly Spatially Resolved European Atmospheric Emission Inventories, Atmospheric Environment, Vol 75,pp.43-57,2013.
30) Fong, I. H., Li, T., Fong, S., Wong, R. K., and Tallon-Ballesteros, A. L.Predicting Concentration Levels of Air Pollutants by Transfer Learning and Recurrent Neural Network. Knowledge-Based Systems, Vol 192, pp.105622,2020.
31) Freeman BS, Taylor G, Gharabaghi B, Thé Forecasting Air Quality Time Series Using Deep Learning. J Air Waste Management, pp.866-886,2018.
32) Ghazali, S., and Ismail, L. H, Air Quality Prediction using Artificial Neural Network, In Proceedings of The International Conference on Civil Environmental Engineering Sustainability, Johor Bahru, Malaysia, pp. 15.2012.
33) Gokulan Ravindiran A B, Gasim Hayder A C, Karthick Kanagarathinam D, Avinash Alagumalai E, Christian Sonne,Air Quality Prediction by Machine Learning Models:A Predictive Study on The Indian Coastal City of Visakhapatnam,Elesiver, Volume 338, October 2023.

34) Graves, A.,and Schmidhuber, J. Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures. Neural Networks, Vol 18, Issue 5-6,pp.602-610,2005.

35) Han, W., Zhai, S., Guo, J., Lee, S., Chang, K., and Zhang, Analysis of No2 and O3 Air Quality Indices and Forecasting Using Machine Learning Models, pp.107-114,2018.

36) Heydari, A., Majidi Nezhad, M., and Astiaso Garcia, D., Keynia, F., and De Santoli, L, Air Pollution Forecasting Application Based on Deep Learning Model and Optimization Algorithm. Clean Technologies and Environmental Policy, Vol 1, pp.1-15.2021.

37) Hong, K. Y., Pinheiro, P. O., and Weichenthal, S. Predicting Global Variations in Outdoor Pm2. 5 Concentrations using Satellite Images and Deep Convolutional Neural Networks, pp.1906.03975,2019.

38) Hu, Q., Zhang, R.,and Zhou, Y.,Transfer Learning for Short-Term Wind Speed Prediction with Deep Neural Networks Renewable Energy, Vol 85, pp.83–95,2016.

39) Janarthanan, R., Partheeban, P., Somasundaram, K., and Elamparithi, P. N.A Deep Learning Approach for Prediction of Air Quality Index in A Metropolitan City. Sustainable Cities and Society,Vol 67, pp.102720,2021.

40) Jiang, F., He, J., and Tian, T.A Clustering-Based Ensemble Approach With Improved Pigeon-Inspired Optimization and Extreme Learning Machine for Air Quality Prediction Applied Soft Computing, Vol 85, pp.105827,2019.

41) Jiang, X. Q., Mei, X. D., and Feng, D. Air Pollution And Chronic Airway Diseases: What Should People Know and Do, Journal of Thoracic Disease, Vol 8,Issue 1,2016.

42) Jing, H., and Wang, Y. Research on Urban Air Quality Prediction Based on Ensemble Learning of Xgboost. In E3S Web of Conferences Vol. 165, 2010.

43) Joachims, T. Learning To Classify Text Using Support Vector Machines, Springer International Series in Engineering and Computer Science,Vol 668,2002.

44) Jun Ma. Identification of High Impact Factors of Air Quality on A National Scale Using Big Data and Machine Learning Techniques. Journal of Cleaner Production, 244, pp.118955,2020.

45) Kang, G. K., Gao, J. Z., Chiao, S., Lu, S., & Xie, G.. Air Quality Prediction:Big Data and Machine Learning Approaches. International Journal of Environmental Science and Development, Vol 9, Issue 1, pp.8-16,2018.

46) Kaya, K., Ş. Deep Flexible Sequential (DFS) Model for Air Pollution Forecasting.

Scientific Reports, Vol 10, Issue 1, pp.1-12,2020.

47) Kelly, F. J., & Fussell, J. C. Air Pollution and Public Health: Emerging Hazards and Improved Understanding of Risk. Environmental Geochemistry and Health, Vol 37, Issue 4 ,pp.631-649,2015.
48) Kumar, S., and Jasuja, A., Air Quality Monitoring System Based on Iot Using Raspberry Pi. In 2017 International Conference on Computing, Communication and Automation (ICCCA) pp.1341-1346,2017.
49) Kurt and A. B. Oktay, Forecasting Air Pollutant Indicator Levels with Geographic Models 3 Days in Advance Using Neural Networks, Expert System Application, Vol. 37, No. 12,2020.
50) L. Ma, Y. Gao, and C. Zhao Research on Machine Learning Prediction of Air Quality Index Based on SPSS, in Proceedings of International Conference of Computer Network, Electronics Automation, 1–5, 2020.
51) Landrigan, P. J., Fuller, R., Acosta, N. J., Adeyi, O,Arnold, R., Baldé, A. B., And Chiles, T. The Lancet Commission on Pollution and Health, The Lancet, Vol 391,Issue 10119, pp.462-512,2018.
52) Leslie, C., Eskin, E., and Noble, W. S.,The Spectrum Kernel: A String Kernel for SVM Protein Classification. in Biocomputing, pp. 564-575,2021.
53) Li, M., Wang, W. L., Wang, Z. Y, & Xue, Y, Prediction of PM2. 5Concentration Based on The Similarity in Air Quality Monitoring Network. Building and Environment, Vol 137, pp.11-17,2018.
54) Li, V. O., Lam, J. C., Chen, Y., and Gu, J. Deep Learning Model to Estimate Air Pollution Using M-BP To Fill in Missing Proxy Urban Data. in Globecom 2017-2017 IEEE Global Communications Conference, pp. 1-6,2017.
55) Li, X., Peng, L., Hu, Y., Shao, J., and Chi, T. Deep Learning Architecture for Air Quality Predictions. Environmental Science and Pollution Research, Vol 23,Issue 22,2016.
56) Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., and Chi, T, Long Short- Term Memory Neural Network for Air Pollutant Concentration Predictions: Method development and Evaluation. Environmental Pollution,Vol 231,pp. 997-1004,2017.
57) Liang Y-C, Maimury Y, Chen AH-L, Juarez JRC, Machine Learning- Based Prediction of Air Quality Applied Sciences,Vol 20,Issue 1,2020.
58) Lin, C. Y., Chang, Y. S., and Abimannan, S,Ensemble Multifeatured Deep Learning Models for Air Quality Forecasting. Atmospheric Pollution Research, Vol 12,Issue 5,

2021.

59) Lin, Y. C., Lee, S. J., Ouyang, C. S., and Wu, C. H. Air Quality Prediction by Neuro-Fuzzy Modeling Approach. Applied Soft Computing, Vol 86, pp.105898,2020.

60) Liu, F., Wang, C., Song, Y.,and Qu, J,Spatial-Temporal Analysis and Projection of Extreme Particulate Matter (PM10 And PM2. 5) Levels using Association Rules: a Case Study of The Jing-Jin-Ji Region, China. Atmospheric Environment,Vol 120,pp. 339-350,2015.

61) Lv, M., Li, Y., Chen, L., and Chen, T. Air Quality Estimation by Exploiting Terrain Features and Multi-View Transfer Semi-Supervised Regression. Information Sciences, Vol 483,pp. 82-95,2019.

62) Ma, J., Cheng, J. C., Lin, C., Tan, Y., and Zhang, J. Improving Air Quality Prediction Accuracy at Larger Temporal Resolutions using Deep Learning and Transfer Learning Techniques. Atmospheric Environment, Vol 214,2019.

63) Ma, J., Li, Z., Cheng, J. C., Ding, Y., Lin, C., and Xu, Z.Air Quality Prediction At New Stations Using Spatially Transferred Bi-Directional Long Short-Term Memory Network. Science of the Total Environment, Vol 705,2020.

64) Mao, W., Wang, W., Jiao, L., Zhao, S., and Liu, A. Modeling Air Quality Prediction using a Deep Learning Approach: Method Optimization and Evaluation. Sustainable Cities and Society, Vol 65, Issue 21,2021.

65) Masih,Machine Learning Algorithms in Air Quality Modeling,Glob. J. Environment science Management, Vol. 5, No. 4, pp. 515–534, 2019.

66) Mehdiyev, N., Lahann, J., Emrich, A., Enke, D., Fettke, P., and Loos, P. Time Series Classification Using Deep Learning for Process Planning: A Case from The Process Industry. Procedia Computer Science, Vol 114, pp.242-249,2017.

67) Mengara Mengara, A. G., Kim, Y., Yoo, Y., and Ahn, Distributed Deep Features Extraction Model for Air Quality Forecasting, Sustainability, Vol 12, Issue 9,2020.

68) Mishra, S., Mishra, D., Santra, G.H., Adaptive Boosting of Weak Regressors for Forecasting of Crop Production Considering Climatic Variability: An Empirical Assessment. J. King Saud University computing information science Issue 32, pp.949–964,2020.

69) Munk, A., & Czado, C. Nonparametric Validation of Similar Distributions and Assessment of Goodness of Fit. Journal of The Royal Statistical Society: SeriesB (Statistical Methodology),Vol 60,Issue 1, pp.223-241,1998.

70) Ni, K., Bresson, X., Chan, T., & Esedoglu, S,. Local Histogram Based Segmentation

Using the Wasserstein Distance. International Journal of Computer Vision,Vol 84,Issue 1,pp. 97-111,2009.

71) Niharika, V. M., and Rao, P. S, A Survey on Air Quality Forecasting Techniques, International Journal of Computer Science and Information Technologies, Vol 5,Issue 1,pp.103-107,2014.
72) Nikolay Laptev,Jiafan Yu and Ram Rajagopal,. Applied Timeseries Transfer Learning, Workshop Track- International Conference on Learning Representations, pp.1-4 ,2018.
73) Nikolay Laptev,Jiafan Yu and Ram Rajagopal,Deepcast: Universal Time-Series Forecaster, International Conference on Learning Representations,2018.
74) Nishant Yadav, Meytar Sorek-Hamer, Michael Von Pohle, using Deep Transfer Learning and Satellite Imagery to Estimate Urban Air Quality in Data-Poor Regions,Vol 342,Issue 24,2024.
75) Pan, S., and Yang, Q. A Survey on Transfer Learning. IEEE Transaction on Knowledge Discovery and Data Engineering, Vol 22, Issue 10,2010.
76) Pandey, G,Zhang, B, and Jian, L.Predicting Submicron Air Pollution Indicators: A Machine Learning Approach, Environmental Science: Processes and Impacts, Vol 15,Issue 5,pp. 996-1005,2013.
77) Pasero, E., and Mesin, Artificial Neural Networks to Forecast Air Pollution, In technology, pp.221-240,2010.
78) Pawul, M., and Śliwka, Application of Artificial Neural Networks for Prediction of Air Pollution Levels In Environmental Monitoring, Journal of Ecological Engineering, Vol 17,Issue 4,2016.
79) Pope , C. A, Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., And Thurston, G. D, Lung Cancer, Cardiopulmonary Mortality, And Long-Term Exposure to Fine Particulate Air Pollution, Jama,Vol 287,Issue 9,2002.
80) Qin, M., Li, Z., and Du, Z. Red Tide Time Series Forecasting By Combining ARIMA and Deep Belief Network, Knowledge-Based Systems,Vol 125,pp.39-52,2017.
81) Rachev, S. T., Stoyanov, S. V., & Fabozzi, F. J.,A Probability Metrics Approach to Financial Risk Measures. John Wiley & Sons,2011.
82) Rubner, Y., Tomasi, C., and Guibas, L. J.The Earth Mover's Distance as a Metric For Image Retrieval. International Journal of Computer Vision, Vol 40,Issue 2, pp.99-121,2000.
83) Rüschendorf, Wasserstein Distance And Approximation Theorems. Probability

Theory and Related Fields, Vol 70, Issue 1, pp.117-129,1985.

84) Ryan, W. F, The Air Quality Forecast Rote: Recent Changes and Future Challenges, Journal of The Air and Waste Management Association, Vol 66,Issue 6 pp.576-596,2016.

85) Samal, K. K. R., Babu, K. S., Panda, A. K., and Das, S. K. Data Driven Multivariate Air Quality Forecasting Using Dynamic Fine Tuning Autoencoder Layer. In 2020 IEEE 17th India Council International Conference, pp. 1- 6,2020.

86) Schurholz, D., Kubler, S., and Zaslavsky, A. Artificial Intelligence-Enabled Context-Aware Air Quality Prediction for Smart Cities. Journal of Cleaner Production,Vol 271,2020.

87) Seng, D., Zhang, Q., Zhang, X., Chen, G., and Chen, X., Spatiotemporal Prediction of Air Quality Based on LSTM Neural Network. Alexandria Engineering Journal, Vol 60,Issue 2,2021.

88) Shirali, P, Influence of Aging in The Modulation of Epigenetic Biomarkers of Carcinogenesis After Exposure to Air Pollution. Experimental Gerontology, Vol 110, pp.125-132,2018.

89) Soh, P. W., Chang, J. W,and Huang, J. W,Adaptive Deep Learning-Based Air Quality Prediction Model using the Most Relevant Spatial-Temporal Relations. Vol 18, Issue 6,2018.

90) Soh, P. W., Chen, K. H., Huang, J. W., and Chu, H. J. Spatial-Temporal Pattern Analysis and Prediction of Air Quality in Taiwan. In 2017 10th International Conference on Ubi-Media Computing and Workshops, pp. 1-6,2017.

91) Song, X., Huang, J.,and Song, D,Air Quality Prediction Based on LSTM- Kalman Model. In 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), pp. 695-699,2019.

92) Song, Y., Qin, S., Qu, J., and Liu, F, The Forecasting Research of Early Warning Systems for Atmospheric Pollutants: A Case In Yangtze River Delta Region, Atmospheric Environment, Vol 118,pp. 58-69,2015.

93) Suleiman A, Tight MR, Q. A., Applying Machine Learning Methods in Managing Urban Concentrations of Traffic-Related Particulate Matter (Pm10 and Pm2. 5), Atmospheric Pollution Research. pp. 134–44,2019.

94) Syafei, A. D., Fujiwara, A.,and Zhang, J. Prediction Model of Air Pollutant Levels Using Linear Model with Component Analysis. International Journal of Environmental Science and Development, Vol 6, Issue 7,2015.

95) Tamas, W., Notton, G., Paoli, C., Nivet, M. L., and Voyant, Hybridization of Air Quality Forecasting Models Using Machine Learning and Clustering: An Original Approach to Detect Pollutant Peaks. Aerosol and Air Quality Research, Vol 16, Issue 2,2016.

96) Thompson, R. C., Allam, A. H., Lombardi, G. P., Wann, L. S., Sutherland, M. L., Sutherland, J. D.,and Thomas, G. S. Atherosclerosis Across 4000 Years of Human History: The Horus Study of Four Ancient Populations. The Lancet, Vol 381,Issue 9873,2013.

97) Tong, S,Air Pollution and Disease Burden, The Lancet Planetary Health, Vol 3,Issue 2,2019.

98) Wang, J., and Song, G.A Deep Spatial-Temporal Ensemble Model for Air Quality Prediction. Neurocomputing, Vol 314,2018.

99) Wang, J., Jiang, H., Zhou, Q., Wu, J., and Qin, S,China S Natural Gas Production And Consumption Analysis Based on The Multicycle Hubbert Model And Rolling Grey Model. Renewable And Sustainable Energy Reviews, Vol 53, pp.1149-1167,2016.

100) Wang, P., Zhang, H., Qin, Z., and Zhang, G. (2017), A Novel Hybrid Garch Model Based on ARIMA and SVM For PM2.5 Concentrations Forecasting, Atmospheric Pollution Research, Vol 8, Issue 5,pp. 850-860,2017.

101) Wen, C., Liu, S., Yao, X., Peng, L., Li, X., Hu, Y., and Chi, T,A Novel Spatiotemporal Convolutional Long Short-Term Neural Network for Air Pollution Prediction. Sci Total Environment,2019.

102) Xayasouk, T.; Lee, H.; Lee, G, Air Pollution Prediction Using Long Short-Term Memory (LSTM) and Deep Autoencoder (DAE) Models, Sustainability Vol 20, Issue 12,2020.

103) Yan, R., Liao, J., Yang, J., Sun, W., Nong, M., and Li, F.,Multi-Hour and Multi- Site Air Quality Index Forecasting in Beijing using CNN, LSTM, CNN- LSTM, and Spatiotemporal Clustering, Expert Systems With Applications, Vol 169,Issue 21,2021

104) Yang, W., Deng, M., Xu, F.,and Wang, H. (2018). Prediction of Hourly PM2. 5 using A Space-Time Support Vector Regression Model. Atmospheric Environment, Vol 181,pp.12- 19,2018.

105) Ye, R., and Dai, Q, A Novel Transfer Learning Framework for Time Series Forecasting. Knowledge-Based Systems, Vol 156, pp.74-99,2018.

106) Zafra, C, Ángel, Y, and Torres, E, ARIMA Analysis of The Effect Of Land Surface Coverage on PM10 Concentrations in A High-Altitude Megacity, Atmospheric Pollution Research, Vol 8,Issue 4, pp.660-668,2017.

107) Zhang, K., Zhang, X., Song, H., Pan, H., and Wang, B.Air Quality Prediction Model Based on Spatiotemporal Data Analysis and Meta learning, Wireless Communications and Mobile Computing, Vol 2021,Issue 1,2021.

108) Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., and Liu, J.Air Quality Predictions with a Semi-Supervised Bidirectional LSTM Neural Network. Atmospheric Pollution Research, Vol 12,Issue 1, pp.328-339,2021.

109) Zhang, S, Li, G., Tian, L., Guo, Q, and Pan, X., Short-Term Exposure to Air Pollution and Morbidity of COPD and Asthma in East Asian Area: A Systematic Review and Meta-Analysis Environmental Research, Vol 148, pp.15-23,2016

110) Zhang, Y., Wang, Y., Gao, M., Ma, Q., Zhao, J., Zhang, R,and Huang, L.A Predictive Data Feature Exploration-Based Air Quality Prediction Approach. IEEE Access,Vol 4,Issue 1,2016.

111) Zhao, P., and Zettsu, K, Decoder Transfer Learning for Predicting Personal Exposure to Air Pollution, in IEEE International Conference on Big Data Vol 19,Issue 1,pp. 5620-5629,2019.

112) Zhao, X., Zhang, R., Wu, J. L., and Chang, P. C. A Deep Recurrent Neural Network for Air Quality Classification. J. Information Hiding Multimedia. Signal Process.Vol 9,Issue 2, pp.346-354,2018.

113) Zheng, Y., Liu, F., and Hsieh, H. P, U-Air When Urban Air Quality Inference Meets Big Data in Proceedings of The 19th ACM SIGKDD International Conference On Knowledge Discovery and Data Mining,pp. 1436- 1444,2013.

114) Zheng, Y., Yi, X., Li, M., Li, R., Shan, Z., Chang, E., and Li, T, Forecasting Fine-Grained Air Quality Based on Big Data. in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

115) Zhu, D., Cai, C., Yang, T,and Zhou, X, A Machine Learning Approach for Air Quality Prediction: Model Regularization and Optimization. Big Data an Cognitive Computing, Vol 2, Issue 01,2018.

116) Zhu, S., Xia, L., Yu, S., Chen, S., and Zhang, J,The Burden and Challenges of Tuberculosis in China: Findings from the Global Burden of Disease Study Scientific Reports, Vol 7,Issue 1,2017.

Avinashilingam Institute for Home Science and Higher Education for Women
(Deemed to be University Estd. u/s 3 of UGC Act 1956, Category 'A' by
MHRDRe-accredited with A++ Grade by NAAC. CGPA 3.65/4, Category I by
UGCCoimbatore-641043,TamilNadu,India

AppendixL2

(Item No 5 of Check List)

Details of Research Publications

S.No	Article	Journal	Other Details Vol/No/Page No/Year	Published in UGC-CARE / Scopus Indexed/ Web of Science
1.	Exploitation of Advanced Deep Learning Methods and Feature Modeling for Air Quality Prediction	Revue d'Intelligence Artificielle	Vol 36, Issue No 6, pp-959- 967 December 2022	Scopus indexed
2.	Air Quality Prediction using Voronai-Based Spatial Temporal Sequence Similarity with Conjugate Gradient Enabled Sparse Auto encoder Deep Learning	International Journal on Recent and Innovation Trends in Computing and Communication	Accepted	Scopus indexed

*Proof of list of Journals from Internet to be attached along with copies of reprints.

Scholar : P.Rw

Supervisor : 8/11/2023 (Dr. P. AMUDHA)

Checked By: 10/11/23 HoD/Dean of Respective School

The Scholar Miss. Shree Nandhini, P (19PHCOF003) has published her paper in the following journals:

1. Revue d'Intelligence Artificielle - is indexed & active in Scopus from 1995 to present and
2. International Journal on Recent and Innovation Trends in Computing and Communication - is indexed and active in Scopus from 2021 to present. This may be considered.

08.11.2023.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *24 August 2023* ***Revised:*** *15 October 2023* ***Accepted:*** *27 October 2023*

Air Quality Prediction using Voronai-Based Spatial Temporal Sequence Similarity with Conjugate Gradient Enabled Sparse Autoencoder Deep Learning

P. Shree Nandhini[1], Dr. P. Amudha[2]
[1]Ph.D. Research Scholar
Department of Computer Science and Engineering, School of Engineering
Avinashilingam Institute for Home Science and Higher Education for Women
Coimbatore, India
shreenandhini2016@gmail.com
[2]Professor
Department of Computer Science and Engineering, School of Engineering
Avinashilingam Institute for Home Science and Higher Education for Women
Coimbatore, India
amudharul@gmail.com

Abstract—Air Quality Prediction (AQP) remains a difficult task because of multidimensional nonlinear spatiotemporal features. To solve this issue, an Improved Sparse Autoencoder with Deep Learning (ISAE-DL) and Enriched ISAE-DL (EISAE-DL) models were developed with the combination of concentric circle-based clustering, Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) followed by the ISAE for AQP. In EISAE-DL, concentric circle-based clustering uses Manhattan distance to efficiently split the locations into four regions using its center and cluster the spatially and temporally similar candidate locations. But it was considered a fixed structure and may struggle to find variations in several data points. Also, it accommodate clusters with regular and circular patterns, whereas irregular and non-circular cluster patterns were not handled. Similarly, the ANN inference was often offended or ignored because of complex meteorological characteristics. Hence, this paper proposes a Voronoi-based spatial-temporal sequence similarity with the Conjugate gradient-enabled SAE-DL (VCSAE-DL) model for effective AQP. First, a Voronoi clustering is performed by creating the Voronoi diagram for analogous candidate location clustering. Then, the resultant clusters of location data along with the PM2.5 and other meteorological data are given to the Improved ANN (IANN), and the target stations are given to the LSTM to capture the spatiotemporal relationship features and temporal features, respectively. Also, CNN is used to extract relationships between terrain and air quality features. These extracted features are fused in the merge layer and transferred to the ISAE for final prediction of air quality. Finally, the test outcomes demonstrate that the VCSAE-DL achieves better prediction performance compared to the existing AQP models.

Keywords-Air quality prediction, EISAE-DL, Concentric circle-based clustering, Voronoi-based clustering, Conjugate gradient.

I. INTRODUCTION

People are starting to worry more about air quality due to the health risks posed by high levels of air contamination [1]. PM2.5, PM10, and Sodium Dioxide (SO_2) monitoring stations are among the many that have been set up in various parts of the world to keep the public informed about air quality. But the needs of the public cannot be met by the data available in real-time. Predicting the air quality in a specific area using historical data, such as meteorological data and historical air quality data, is the primary goal of AQP [2].AQP has been proposed using several different approaches. At first, the prediction model is built using empirical assumptions, such as those used in air pollution dispersion models. Predictions of air quality are made using machine learning techniques like linear regression [3] and ANNs [4]. But such methods are imperfect due to a wide variety of factors, including weather, pollutants, and traffic patterns. The sensors are the source of data for the air quality dataset. There is a lot of unexpected or erratic information in these data. Anomaly information degrades the performance of a Spatio-Temporal Deep Neural Network (ST-DNN) [5] since it is not handled in the network. To learn the distribution of the data over several dimensions, the ISAE-DL and EISAE-DL models [6] were built. In these models, a mixture of continuous and discrete characteristics was used. Then, the k-Nearest Neighbor Euclidean Distance (kNN-ED) and the kNN-Dynamic Time Wrapping Distance (kNN-DTWD) were used to find spatially and temporally related data. But the delay interval is important for long-term predictions.

To deal with the time delay at the location in the prediction, a concentric circle-based distance partition approach was proposed. In this method, similar sites in space

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 24 August 2023 **Revised:** 15 October 2023 **Accepted:** 27 October 2023*

and time were clustered using the Manhattan distance. It discovered the center of the four groups of sites that are most similar to each other in terms of both space and time. Zero was used as the starting point for this center. From this central point, the rest of the distances between places with similar locations and times can be determined. The k-NN clustering, which uses Euclidean distance, and clustering based on concentric circles were two distinct approaches. In this case, a single centroid value was utilized to classify locations into distinct categories. Using standard centric and Manhattan distance measures, similar places were grouped. The PM2.5 and meteorological data, as well as the results of the concentric circles and the information about the target stations, were fed into the ANN, the LSTM, and the CNN, and the combined results were then sent to the ISAE-based Feed Forward Neural Network (FFNN) to predict the air quality.

On the other hand, it was considered a fixed structure and may struggle to find variations in many data points. Also, it accommodate clusters with regular and circular patterns, whereas irregular and non-circular cluster patterns were not handled. Similarly, the ANN inference was often offended or ignored because of complex meteorological characteristics. Therefore, this manuscript proposes the VCSAE-DL model for effective AQP. First, a Voronoi clustering is performed by creating the Voronoi diagram for clustering analogous candidate places. Then, the resultant clusters of location data along with the PM2.5, other meteorological data, and target stations are given to the IANN based on the Conjugate Gradient-enabled Artificial Backpropagation neural Network (CG-ABN), LSTM, and CNN for feature extraction. These extracted features are fused in the merge layer and transferred to the ISAE for AQP.

The residual sections are prepared as the following: Various research on AQP is the focus of Section II. The VCSAE-DL model for AQP is described in Section III. The effectiveness of its performance is shown in Section IV. Section V summarizes the article with future scope.

II. LITERATURE SURVEY

To forecast PM2.5 levels, a Spatial-Temporal Support Vector Regression (STSVR) has been presented in [7]. Spatial grouping was first used on the data to deal with the issue of spatial heterogeneity. Then, spatial autocorrelation variables were proposed to be part of input features by using the Gauss vector weight function. Finally, a sub-area of AQP depending on the regional SVR with space autocorrelation factors was built. However, the space-time clustering technique will be investigated to accurately identify space-time heterogeneity with spatial heterogeneity.

An integrated LSTM and CNN model using historical and current data on air pollution and weather conditions [5] was presented for AQP. However, due to the noise, the overall prediction performance suffers when all locations are included. A sophisticated Spatial-Temporal Ensemble (STE) model was developed [8] for AQP. Using an ensemble technique and a partitioning approach based on weather patterns, this model dynamically trains and integrates numerous independent models. After that, spatial data was generated and analyzed using Granger causalities to determine where stations were associated with one another. The long- and short-term air quality dependencies were then trained by the LSTM. However, LSTM cannot retain data for extended periods.

To predict the air quality a hybrid learning approach was developed [9]. Air quality index raw data was initially processed using Wavelet Packet Decomposition (WPD) to create an air quality index low-frequency subseries. To provide accurate predictions for the subseries, a Modified Extreme Learning Machine (MELM) was trained with pigeon-inspired weight and threshold optimization. To further establish the predictions by frequency, Multidimensional Scaling and K-means (MSK) clustering schemes were used. The ultimate result of air quality index prediction was achieved by processing the subseries together using MELM. The accuracy of future predictions will be improved with the help of DL models.

To forecast the PM2.5 absorptions at various places in Beijing [10], The Light Gradient Boosting Machine (LightGBM) approach was developed. When generating the dataset for AQP, it merged historical information with predictive information to extract temporal features that enhanced prediction accuracy. To better pinpoint features that could be crucial, multi-dimensional statistical features were considered. Also, high-dimensional time features were built by intensively mining the time association, and a sliding window was used for dimensionality reduction of training data to boost the model's prediction accuracy. But it was not suitable for large-scale datasets.

A complete prediction model using LSTM was developed to accurately monitor the AQP [11]. Particle concentration data from the existing monitoring system, data from neighboring stations, weather data, and data on gaseous pollutants in the air were integrated to create a more accurate picture of air quality. The information was translated into a supervised learning format after being normalized. Predictions of the air quality were then made using the LSTM. But it has a gradient vanishing issue that impacts the model efficiency.

A context prediction model was presented for AQP [12]. This model integrated information from the user's health profile and nearby pollution sources with an accurate air pollution prediction method utilizing LSTM-DNN, taking into account the principles of context-aware computing. To

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *24 August 2023* ***Revised:*** *15 October 2023* ***Accepted:*** *27 October 2023*

further enhance the accuracy of AQPs, this model was integrated with a My Air Quality Index (MyAQI). Inaccuracies were introduced, however, when real-world data was used due to the possibility of disruptions or faulty measurement stations.

Predicting PM2.5 for measuring the air pollution level using a semi-supervised bidirectional LSTM was proposed in [13]. This procedure needs fine PM2.5 as input and then used Empirical Mode Decomposition (EMD) to decompose the data and extract the amplitude and frequency properties. Also, the bidirectional LSTM model was used for final predictions. However, certain unusual values of Mean Absolute Percentage Error (MAPE) were recorded as a result of the uneven numerical distribution of fine particulate matter.

The Neuro-fuzzy network method was developed [14] for predicting air quality. At first, the training data were partitioned into fuzzy clusters whose membership function was specified by means and variances. Then, a four-layer fuzzy neural network was built using the fuzzy rules extracted from the fuzzy clusters. Finally, the network was trained for AQP using the steepest descent backpropagation, genetic, and particle swarm optimization algorithms. But the performance depends on the membership function of the Neuro-fuzzy network.

A transfer learning-based bidirectional LSTM was proposed for predicting air quality [15]. This method combined sophisticated DL models and transfer learning mechanisms to relocate the data from historical locations to new locations for enhancing the performance of AQP. But it did not learn spatiotemporal relationships among various data.

III. PROPOSED METHODOLOGY

This section briefly describes the proposed VCSAE-DL for predicting air quality. The place with the strongest spatio-temporal connection to the goal location is identified first, and then the location with the best air quality is predicted. To improve AQP, it is necessary to retrieve key elements by identifying temporal and spatial characteristics. Assume there is a set of locations $L = \{l_1, l_2, \dots l_n\}$ and a set of features $F = \{f_1, f_2, \dots f_m\}$ that are useful in identifying these places. The coordinates of a given location are based on its latitude and longitude data, which are given in the form of,

$$Lt_i = (lt_i, p_i, q_i) lt_i \in Lt_i (1)$$

The coordinates for the latitude and longitude at point lt_i are given by p_i and q_i respectively, in Eq. (1). Voronoi diagram is used to find the place closest to the target, while the top h relevant places are generated based on the data used to train the algorithm.

A. Voronoi Clustering

A multidimensional Voronoi diagram is a splitting of the location area $\mathbb{R}^d$ into areas $R_j$containing the possessions: all centersc_j lie exactly in one area R_j, which comprises each point $x \in \mathbb{R}^d$ that are nearer to c_j than any other center $c_k, j \neq k$:

$$R_j = \{x \in \mathbb{R}^d : \|x - c_j\| < \|x - c_k\|, \forall j \neq k\} \quad (2)$$

The centers c_j are called Voronoi points. The area related to a Voronoi point c_j is called a cell. Since a cell comprises each point, which are nearer to its centroid compared to another centroid, the cell boundaries locate accurately in the mid of 2 centroids. Fig. 1 shows the Voronoi diagram for a dataset, which comprises 6 independent clusters.

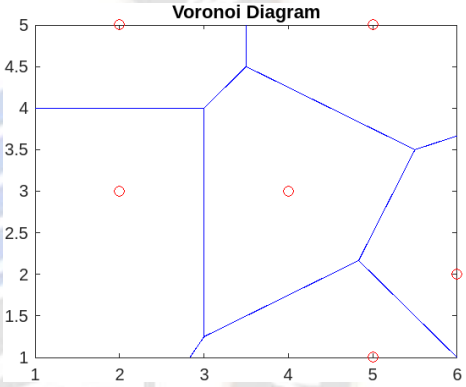


Figure 1. Voronoi Diagram for Six Clusters

Given a set of $L \subseteq \mathbb{R}^d$ to be clustered, the clustering is begun with creating the Voronoi diagram for L. Consider the 2D plane. As a given variable, the method is provided a max threshold, which represents the highest value permitted for a cell that may remainmerged into a developing cluster. Because cell volume approximates local location area density, merely that counts in deciding if the density is large sufficient for further clustering. Separate clusters can thus develop to any dimension given the local density is adequate.Once each cell of at most volume max has been considered, only comparativelyhuge cells are retained.

All points in the data Voronoi diagram form their cells. Following that, the cell volume is determined.This study calculates the appropriate cell region in the plane as the polygon region, yet in larger sizes d, it simply approximates the cell to be a hyperball with volume $\left(\pi^{\lfloor n/2 \rfloor} r^d\right) \Big/ \Gamma\left(d/2 + 1\right)$, wherein $\Gamma(z+1) = z!$.The radius $r$refers to the mean distance of cell angles from its centroid. This estimation is not more precise in low sizes, yet it

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *24 August 2023* ***Revised:*** *15 October 2023* ***Accepted:*** *27 October 2023*

enhances when increasing the size. All cells are related to a class label.

Initially, the moment at which the minimum Voronoi cell is dealt with. Because it contains no identified adjacent cells, this cell is assigned the label 1. The cell examined successivelymay be a neighbor of the initial cell, as indicated by the distribution of an angle. When the latter cell volume falls below the max threshold, they are merged and the initial cell label is considered. Or else, the 2nd cell will become unrelated to the initial, wherein it obtains its label. But it is remaining promising that the 2 locations belong to a similar cluster and will eventually be merged.

In most cases, a cell contains numerous nearby cells with multiple labels and a few that haven't been annotated. The identified adjacent is analyzed based on their dimensions (i.e., label values). The cell under examination is fused to its adjacent with the lastlabel, and its annotated adjacent considers a similar label. Cells are joined when their volume is less than the limit. There is no essential to continue by the residual adjacent cells when the maximum value has been achieved.

After each point has been considered, every cluster combination has been executed. There is no necessity to iterate the procedure. But the clustering post-processing is needed to create it feasible. Assume an illustration in Fig. 1. It is seen that the above-explained cell merging process can identify 6 clusters for different max threshold values. But it is addressed that the farthest cluster points can't be merged, since the clusters could develop together.

Thus, in this case, the heuristics considers the cluster marginal points. Centroids with acell volume higher than the max threshold can be merged to the nearby adjacent label.

Algorithm for Voronoi Clustering:

- $\boldsymbol{function(S, max)}$
- Create the Voronoi diagram for $L = \{l_1, l_2, \dots l_n\}$;
- Estimate the Voronoi cell volumes and rank the instances. Consider the resultant rank $l_1, l_2, \dots l_n$;
- $\boldsymbol{for}(i = 1{:}\, n)$

  $\boldsymbol{If}\, R_i$ *volume related to* s_i *is maximum*

  Merge R_i with an adjacent cluster with the smallest class number;

  If one exists, or else allocate a new class number to the cell;

  $\boldsymbol{else}$

  Allocate R_i to the nearest adjacent cluster;

  $\boldsymbol{end\ if}$

  $\boldsymbol{end\ for}$
- $\boldsymbol{end\ function}$

Thus, the top h sites are chosen through a clustering method based on the Voronoi diagram, and these locations will be used to try and identify the target sequence. To predict future time series, the Temporal Relationship Extractor (TRE) takes attributes of the target historical site. The trends in the desired region are obtained by TRE LSTM, whereas LSTM models the behavior of historical time series.LSTM and IANN are used to collect data at low and high frequencies, respectively.

B. Improved Artificial Neural Network

The CG-ABN structure is split into 4 different layers as shown in Fig. 2, which comprises weight initialization, the error's forward and backward propagation, weight, and bias updation. The hidden layer comprises many neurons and all have an activation function as $f(x) = Sigmoid(x)$. The activation function obtains the total weighted input $(\omega_{11} * x_1 + \omega_{21} * x_2 + \omega_{31} * x_3 + \cdots + \omega_{n1} * x_n + 1 * b)$argument as follows:

$$f \sum_{i=1}^{n} W_{ij} X_i + b \tag{3}$$

In this IANN, forward propagation is executed initially. Then, backward propagation is prepared once theestimated result Y is compared with the observed result $\hat{Y}$ in the computation of the gradient mistake at the final layer.

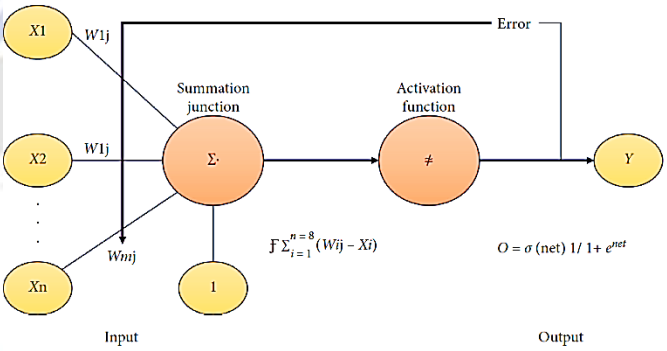


Figure 2. Architecture of CG-ABN-based IANN

The IANN is trained with air pollutants dataset utilizing the below schemes:

- Multi-Layer Perceptron (MLP)
- Bayesian normalized
- Scaled conjugate gradient

The IANN model comprises a single input, hidden, and output layer comprising $n$input, which create Loutputs. The database has (X_i, Y_i), whereinX_i denotes the input layer and Y_i denotes the estimated result. When $n$denotes the overall dimension of a whole set,

$$X_n = \{(x_i, y_i), \dots, (x_n, y_n)\} \tag{4}$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 24 August 2023 **Revised:** 15 October 2023 **Accepted:** 27 October 2023*

The variables of feed-forward are represented as θ. As learning the ANN contains the gradient of the error function $E = (X, \theta)$ related to the weight w_{ij}^k at node j in layer k for node i, biases b_i^k at i in k, therefore according totraining rate, gradient descent modifies the weights at all iterations $t$by

$$\theta_t = \theta_t - \alpha \frac{\partial E(X,\theta)}{\partial \theta_t} \quad (5)$$

In Eq. (5), θ_t is the ANN variables at t. The Mean Square Error (MSE) in backpropagation is

$$E(X,\theta) = \frac{1}{2N}\sum_{i=1}^{N}(\hat{y}_i - y_i) \quad (6)$$

In Eq. (6), $\hat{y}_i$ refers to the estimated result and y_i denotes the observed result of x_i. The derivative of $f(x)$and sigmoid function is denoted by$f'(x)$ and $\sigma'(x)$. $b(k/i)$ in k^{th} at i is added to the weight as $w_{o_i}^k$ with a result $o_o^{k-1} = 1$ at node 0 in $k-1$; so,

$$w_{o_i}^k = b_i^k \quad (7)$$

$$\overrightarrow{W} = \begin{pmatrix} w_{00} & \cdots \\ w_{0m} & \cdots \\ w_{10} & \cdots \\ w_{1m} & \cdots \\ \vdots & \vdots \\ w_{n0} & \cdots \\ w_{nm} & \cdots \\ w_{n+1,0} & \cdots \\ \vdots & \vdots \end{pmatrix} \quad (8)$$

Weight and bias ranges are set with arbitrary values; a modified weight is applied in consequent t. Eq. (8) is rewritten as:

$$\overrightarrow{W} = \begin{pmatrix} \overrightarrow{W}_{IH} \\ \vdots \\ \overrightarrow{W}_{HO} \end{pmatrix} \quad (9)$$

Eq. (9) demonstrates that $\overrightarrow{W}_{IH}$ includes the weight from the input to hidden layer, whereas $\overrightarrow{W}_{HO}$denotes the weight defining from hidden to output layer.

$$\overrightarrow{W}_{IH} = \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0m} \\ w_{00} & w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \iddots & \vdots \\ w_{n,0} & w_{n,1} & \cdots & w_{0m} \\ w_{00} & w_{11} & \cdots & \vdots \end{pmatrix} \quad (10)$$

Eq. (10) includes the weights from the input to hidden layer. It has $n \times m$components and $W_{a,b}$denotes the weight among components.

$$\overrightarrow{W}_{HO} = \begin{pmatrix} w_{n+1,0} & w_{n+1,1} & \cdots & w_{n+1,m} \\ w_{n+1,0} & w_{11} & \cdots & w_{n+2,m} \\ \vdots & \ddots & \iddots & \vdots \\ w_{n+k,0} & w_{n+k,1} & \cdots & w_{n+km} \\ w_{00} & w_{11} & \cdots & \vdots \end{pmatrix} \quad (11)$$

In this study, only one output neuron is used. Eq. (11) is rewritten as:

$$\overrightarrow{W}_{HO} = \begin{pmatrix} w_{n+1,0} & w_{n+1,1} & \cdots & w_{n+1,m} \end{pmatrix} \quad (12)$$

Corresponding to the original formulation $\mathcal{J}_{k-1}$,

$$a_i^k = b_i^k + \sum_{j=1}^{\mathcal{J}_{k-1}} w_{ji}^k o_j^{k-1} = \sum_{j=0}^{\mathcal{J}_{k-1}} w_{ji}^k o_j^{k-1} \quad (13)$$

After that, the error is determined by

$$E(X,\theta) = \frac{1}{2N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \quad (14)$$

After taking the derivative of Eq. (14),

$$\frac{\partial E(X,\theta)}{\partial w_{ji}^k} = \frac{1}{N}\sum_{d=1}^{N}\frac{\partial}{\partial w_{ji}^k}\left(\frac{1}{2}(\hat{y}_i - y_i)^2\right) = \frac{1}{N}\sum_{d=1}^{N}\frac{\partial E_d}{\partial w_{ji}^k} \quad (15)$$

The backpropagation scheme is associated with a single I/O set and each I/O in X can be generated byfusing all gradients. For derivation, $E$is

$$E = \frac{1}{2}(\hat{y} - y)^2 \quad (16)$$

Error Function Derivatives: Once the chain principle is applied,

$$\frac{\partial E_d}{\partial w_{ji}^k} = \frac{\partial E}{\partial a_j^k}\frac{\partial a_j^k}{\partial w_{ij}^k} \quad (17)$$

In Eq. (17), a_j^k acts as activation of j in k.

$$\delta_j^k = \frac{\partial E}{\partial a_j^k} \quad (18)$$

The 2nd expression from Eq. (17) for a_j^k is

$$\frac{\partial a_j^k}{\partial w_{ij}^k} = \frac{\partial}{\partial w_{ij}^k}\left(\sum_{i=0}^{\mathcal{J}_{k-1}} w_{ij}^k o_i^{k-1}\right) = o_i^{k-1} \quad (19)$$

The partial derivative of Ew.r.t w_{ij}^k is as:

$$\frac{\partial E}{\partial a_j^k} = \delta_j^k o_i^{k-1} \quad (20)$$

Output Layer: Backpropagation defines φ_1^m , wherein$m$denotes the last layer. 4-layer ANN possess $m = 3$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *24 August 2023* ***Revised:*** *15 October 2023* ***Accepted:*** *27 October 2023*

for the last layer and $m = 2$ for the 2nd to the final layer. Defining E in terms of a_1^m provides

$$E = \frac{1}{2}(\hat{y} - y)^2 = E\frac{1}{2}(g_0(a_1^m) - y)^2 \quad (21)$$

After, applying the partial derivative,

$$\delta_i^m(g_0(a_1^m) - y)g_0'(a_1^m) = (\hat{y} - y)g_0'(a_1^m) \quad (22)$$

The partial derivative of E w.r.t w_{i1}^m is

$$\frac{\partial E}{\partial w_{i1}^m} = \delta_1^m o_i^{m-1} = (\hat{y} - y)g_0'(a_1^m)o_i^{m-1} \quad (23)$$

Hidden Layer: The hidden layers' E is determined by

$$\delta_j^k = \frac{\partial E}{\partial a_j^k} = \sum_{l=1}^{J^{k+1}} \frac{\partial E}{\partial a_l^{k+1}} \frac{\partial a_l^{k+1}}{\partial a_j^k} \quad (24)$$

Error term φ_l^{k+1} gives the following equation:

$$\delta_j^k = \sum_{l=1}^{J^{k+1}} \delta_l^{k+1} \frac{\partial a_l^{k+1}}{\partial a_j^k} \quad (25a)$$

$$a_l^{k+1} = \sum_{j=1}^{J^{k+1}} w_{jl}^{k+1} g(a_j^k) \quad (25b)$$

Recalling the description of a_l^{k+1}, wherein $g(x)$ denotes the activation function.

$$\frac{\partial a_l^{k+1}}{\partial a_j^k} = w_{jl}^{k+1} g'(a_j^k) \quad (26)$$

Observe that

$$\delta_j^k = \sum_{l=1}^{J^{k+1}} \delta_l^{k+1} w_{jl}^{k+1} g'(a_j^k) = g'(a_j^k) \sum_{l=1}^{J^{k+1}} w_{jl}^{k+1} \delta_l^{k+1} \quad (27)$$

The partial derivative of E w.r.t w_{ij}^k for $1 \le k < m$ is

$$\frac{\partial E}{\partial w_{ij}^k} = \delta_l^k o_i^{k-1} = g'(a_j^k) o_i^{k-1} \sum_{l=1}^{J^{k+1}} w_{jl}^{k+1} \delta_l^{k+1} \quad (28)$$

$$\frac{\partial E}{\partial w_{ij}^k} = \delta_l^k o_i^{k-1} \quad (29)$$

For the partial derivative of E at the final layer,

$$\delta_l^m = g'(a_1^m)(\hat{y}_d - y_d) \quad (30a)$$

$$\delta_l^m = g'(a_j^k) \sum_{l=1}^{J^{k+1}} w_{jl}^{k+1} \delta_l^{k+1} \quad (31b)$$

For the hidden layers' E,

$$\frac{\partial E(X,\theta)}{\partial w_{ij}^k} = \frac{1}{N}\sum_{d=1}^{N} \frac{\partial}{\partial w_{ij}^k}\left(\frac{1}{N}(\hat{y}_d - y_d)^2\right) = \frac{1}{N}\sum_{d=1}^{N} \frac{\partial E_d}{\partial w_{ij}^k} \quad (32)$$

For merging the partial derivatives,

$$\Delta w_{ij}^k = a\frac{\partial E(X,\theta)}{\partial w_{ij}^k} \quad (33)$$

For updating the weights,

$$w_{ij}^{k+1} = w_{ij}^k + \lambda \Delta w_{ij}^k \quad (34)$$

Algorithm of IANN Model: Consider a as the training rate and weight initialization as w_{ij}^k, below procedures are utilized in the proposed IANN model.

- Perform the forward stage: For all pairs of input and outputs $(\vec{x}_d, y_d)$, save the results $(\hat{y}_d, a_j^k)$ and (o_j^k) for all nodes j in k by arranged from layer zero, input to the final layer (m).
- Perform the backward stage: For all $(\vec{x}_d, y_d)$, save the results $\left(\frac{\partial E_d}{\partial w_{ij}^k}\right)$ for w_{ij}^k linking i in $(k-1)$, to j in k by arranged from m, output to the initial layer.Compare E for the last layer using Eq. (4). Back-propagate E for δ_l^k, processing backward from $k = m - 1$, viaconstantly utilizing Eq. (5). Compare the partial derivatives of a separate error E_d regarding w_{ij}^k utilizing Eq. (3).
- Combine the individual gradients: Merge the separate gradients for all I/O pairs $\left(\frac{\partial E_d}{\partial w_{ij}^k}\right)$ to obtain the overall gradient $\left(\frac{\partial E(X,\theta)}{\partial w_{ij}^k}\right)$, for the whole collection of $X = \{(\vec{x}_1, y_1), \ldots, (\vec{x}_N, y_N)\}$ by using Eq. (6).
- Modify the weights: Based on a and $\frac{\partial E(X,\theta)}{\partial w_{ij}^k}$ and utilizing Eq. (7), the weights are updated using Eq. (29).

In this study, a is set to 0.2 and the momentum coefficient is 0.5. The network is trained for 500 epochs.

Moreover, the extracted features are fed to the FFNN-based ISAE for the final prediction of air quality.Fig.3 depicts the entire architecture of the VCSAE-DL.

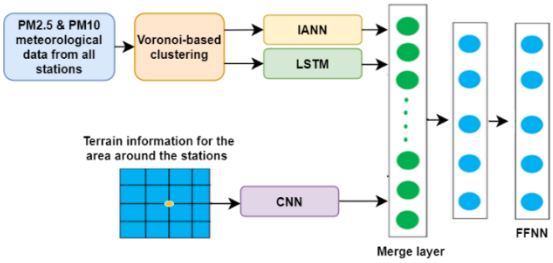


Figure 3. Overall Architecture of VCSAE-DL

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *24 August 2023* ***Revised:*** *15 October 2023* ***Accepted:*** *27 October 2023*

IV. RESULTS AND DISCUSSION

Here, the efficiency of VCSAE-DL is evaluated and compared to that of ST-SVR [7], ISAE-DL [6], and EISAE-DL [6]. Experiments are conducted using a dataset based on the prediction of PM2.5 and PM10, which are both widely reported air pollutants for predicting air quality.

A. Accuracy and Precision

It is the ratio of true forecasts amid an overall data tested.

$$Accuracy = \frac{True\ Positive\ (TP)+True\ Negative\ (TN)}{TP+TN+False\ Positive\ (FP)+False\ Negative(FN)} \quad (35)$$

Precision is determined by using the TP and FP rates.

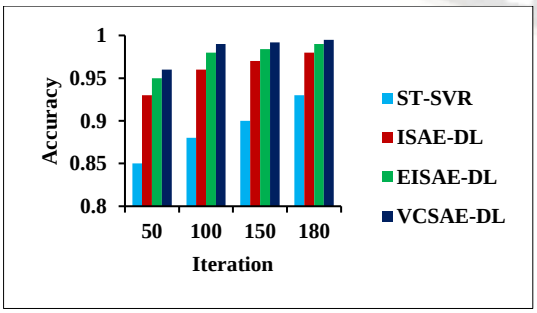
$$Precision = \frac{TP}{TP+FP} \quad (36)$$

Accuracy and precision are compared in Table 1 for existing and proposed models.

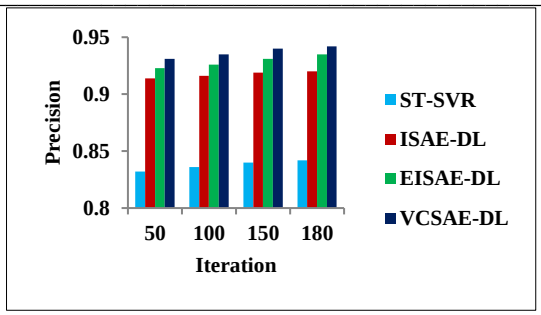
Table 1. Evaluation of VCSAE-DL in terms of Accuracy and Precision

Iteration	Accuracy				Precision			
	ST-SVR	ISAE-DL	EISAE-DL	VCSAE-DL	ST-SVR	ISAE-DL	EISAE-DL	VCSAE-DL
50	0.85	0.93	0.95	0.96	0.832	0.914	0.923	0.931
100	0.88	0.96	0.98	0.99	0.836	0.916	0.926	0.935
150	0.90	0.97	0.984	0.992	0.84	0.919	0.931	0.94
180	0.93	0.98	0.99	0.995	0.842	0.92	0.935	0.942

The accuracy (precision) of various AQP models is depicted in Fig. 4a (b) for varying iteration counts. When using 100 iterations, VCSAE-DL's accuracy (precision) in predicting air quality is higher than that of ST-SVR, ISADE-DL, and EISAE-DL by 12.5% (11.84%), 3.13% (2.07%) and 1.02% (0.97%), respectively. Thus, it demonstrates that the proposed VCSAE-DL outperforms other existing models for AQP.



(a)



(b)

Figure 4. Comparison of ST-SVR, ISAE-DL, EISAE-DL and VCSAE-DL in terms of (a) Accuracy (b) Precision

B. Sensitivity and Specificity

Sensitivity is the proportion of predicted true values. It is calculated as:

$$Sensitivity = \frac{TP}{TP+FN} \quad (37)$$

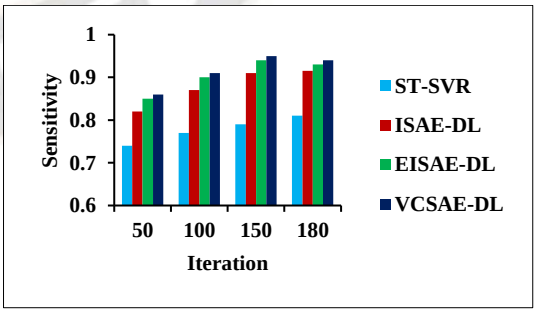
Specificity is the percentage of predicted negative values. It is calculated as:

$$Specificity = \frac{TN}{TN+FP} \quad (38)$$

The sensitivity and specificity of existing and proposed models are compared in Table 2.

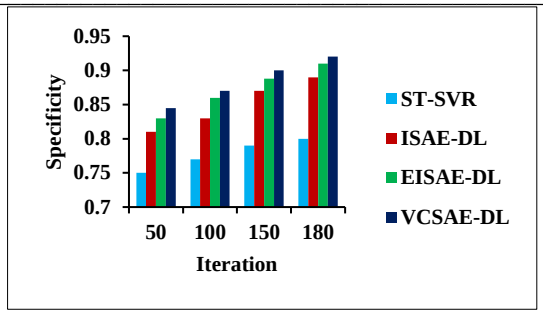
Table 2. Comparison of VCSAE-DL in terms of Sensitivity and Specificity

Iteration	Sensitivity				Specificity			
	ST-SVR	ISAE-DL	EISAE-DL	VCSAE-DL	ST-SVR	ISAE-DL	EISAE-DL	VCSAE-DL
50	0.74	0.82	0.85	0.86	0.75	0.81	0.83	0.845
100	0.77	0.87	0.90	0.91	0.77	0.83	0.86	0.87
150	0.79	0.91	0.94	0.95	0.79	0.87	0.888	0.90
180	0.81	0.915	0.93	0.94	0.80	0.89	0.91	0.92



(a)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *24 August 2023* ***Revised:*** *15 October 2023* ***Accepted:*** *27 October 2023*



(b)

Figure 5. Comparison of ST-SVR, ISAE-DL, EISAE-DL and VCSAE-DL in terms of (a) Sensitivity (b) Specificity

Fig.5a (b) compares the sensitivity (specificity) of various AQP models over iteration counts. Compared to ST-SVR, ISADE-DL, and EISAE-DL, the VCSAE-DL had a higher sensitivity (specificity) by 18.18% (12.99%), 4.6% (4.82%), and 1.11% (1.16%), respectively for predicting air quality. Thus, it demonstrates that the proposed VCSAE-DL is more accurate in predicting air quality than other existing models.

C. Area Under Curve (AUC) and Matthew's Correlation Coefficient (MCC)

AUC is calculated as:

$$AUC = \frac{Sensitivity + Specificity}{2} \tag{39}$$

MCC is the relationship between the forecasted and observed values.

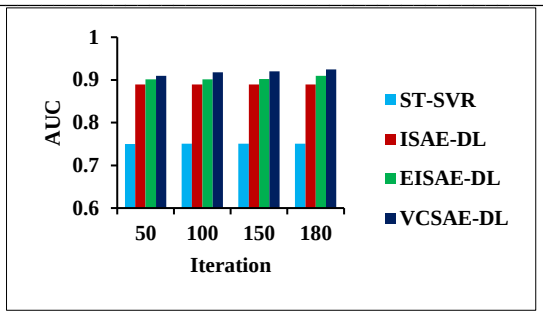
$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{40}$$

The AUC and MCC comparisons of existing and proposed models are shown in Table 3.

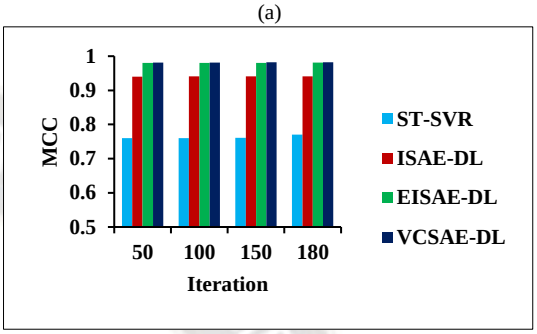
Table 3. Comparison of VCSAE-DL in terms of AUC and MCC

Iteration	AUC				MCC			
	ST-SVR	ISAE-DL	EISAE-DL	VCSAE-DL	ST-SVR	ISAE-DL	EISAE-DL	VCSAE-DL
50	0.7502	0.8892	0.9012	0.910	0.7603	0.9404	0.9802	0.981
100	0.7504	0.8894	0.9015	0.918	0.7605	0.9408	0.9804	0.9813
150	0.7508	0.8896	0.9018	0.920	0.7608	0.941	0.9806	0.9819
180	0.751	0.8897	0.91	0.925	0.7701	0.9414	0.9809	0.982

The AUC (MCC) for various AQP models at varying iteration counts is displayed in Fig. 6a (b).VCSAE-DL outperforms other existing AQP models regarding AUC (MCC) for predicting air quality by 22.33% (29.03%), 3.22% (4.3%), and 1.83% (0.09%), respectively, when the number of iterations is set to 100. Thus, it demonstrates that compared to existing AQP methodologies, the suggested VCSAE-DL has superior MCC and AUC.



(a)



(b)

Figure 6. Comparison of ST-SVR, ISAE-DL, EISAE-DL and VCSAE-DL in terms of (a) AUC (b) MCC

V. CONCLUSION

In this paper, the VCSAE-DL model is proposed, which adopts the Voronoi clustering to cluster the locations which are geographically and temporally similar. The resultant clustered locations along with the PM2.5, target stations, and other meteorological data are given to the IANN, LSTM, and CNN followed by the ISAE-based FFNN to predict air quality. Finally, the experimental results demonstrated that the proposed VCSAE-DL achieved the highest performance in contrast with the existing AQP models.

REFERENCES

[1] W. Mao, W. Wang, L. Jiao, S. Zhao and A. Liu, "Modeling air quality prediction using a deep learning approach: Method optimization and evaluation,"Sustainable Cities and Society, vol. 65, p. 102567, 2021, https://doi.org/10.1016/j.scs.2020.102567.

[2] J. Ma, J. C. Cheng, C. Lin, Y. Tan and J. Zhang, "Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques,"Atmospheric Environment, vol. 214, p. 116885, 2019, https://doi.org/10.1016/j.atmosenv.2019.116885.

[3] A. D. Syafei, A. Fujiwara and J. Zhang, "Prediction model of air pollutant levels using linear model with component analysis,"International Journal of Environmental Science and Development, vol. 6, no. 7, p. 519, 2015, https://doi.org/10.7763/IJESD.2015.V6.648.

[4] J. J. Palop, L. Mucke and E. D. Roberson, "Quantifying biomarkers of cognitive dysfunction and neuronal network hyperexcitability in mouse models of Alzheimer's disease: depletion of calcium-dependent proteins and inhibitory hippocampal remodeling," Alzheimer's disease and

Frontotemporal dementia: methods and protocols, pp. 245-262, 2011, https://doi.org/10.1007/978-1-60761-744-0_17.

[5] P. W. Soh, J. W. Chang and J. W. Huang, "Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations,"IEEE Access, vol. 6, pp. 38186-38199, 2018, https://doi.org/10.1109/ACCESS.2018.2849820.

[6] P. ShreeNandhini, P. Amudha and S. Sivakumari, "Comparative analysis of air quality prediction using artificial intelligence techniques,"ECS Transactions, vol. 107, no. 1, p. 6059, 2022, https://doi.org/10.1149/10701.6059ecst.

[7] W. Yang, M. Deng, F. Xuand H. Wang, "Prediction of hourly PM2.5 using a space-time support vector regression model,"Atmospheric Environment, vol. 181, pp. 12-19, 2018, https://doi.org/10.1016/j.atmosenv.2018.03.015.

[8] J. Wang and G. Song, "A deep spatial-temporal ensemble model for air quality prediction,"Neurocomputing, vol. 314, pp. 198-206, 2018, https://doi.org/10.1016/j.neucom.2018.06.049.

[9] F. Jiang, J. Heand T. Tian, "A clustering-based ensemble approach with improved pigeon-inspired optimization and extreme learning machine for air quality prediction,"Applied Soft Computing, vol. 85, p. 105827, 2019, https://doi.org/10.1016/j.asoc.2019.105827.

[10] Y. Zhang, Y. Wang, M. Gao, Q. Ma, J. Zhao, R. Zhang and L. Huang, "A predictive data feature exploration-based air quality prediction approach,"IEEE Access, vol. 7, pp. 30732-30743, 2019, https://doi.org/10.1109/ACCESS.2019.2897754.

[11] D. Seng, Q. Zhang, X. Zhang, G. Chenand X. Chen, "Spatiotemporal prediction of air quality based on LSTM neural network,"Alexandria Engineering Journal, vol. 60, no. 2, 2021, https://doi.org/10.1016/j.aej.2020.12.009.

[12] D. Schürholz, S. Kubler and A. Zaslavsky, "Artificial intelligence-enabled context-aware air quality prediction for smart cities,"Journal of Cleaner Production, vol. 271, p. 121941, 2020, https://doi.org/10.1016/j.jclepro.2020.121941.

[13] L. Zhang, P. Liu, L. Zhao, G. Wang, W. Zhang and J. Liu, "Air quality predictions with a semi-supervised bidirectional LSTM neural network," Atmospheric Pollution Research, vol. 12, no. 1, pp. 328-339, 2021, https://doi.org/10.1016/j.apr.2020.09.003.

[14] Y. C. Lin, S. J. Lee, C. S. Ouyangand C. H. Wu, "Air quality prediction by neuro-fuzzy modeling approach,"Applied soft computing, vol. 86, p. 105898, 2020, https://doi.org/10.1016/j.asoc.2019.105898.

[15] J. Ma, Z. Li, J. C. Cheng, Y. Ding, C. Lin and Z. Xu, "Air quality prediction at new stations using spatially transferred bi-directional long short-term memory network,"Science of The Total Environment, vol. 705, p. 135771, 2020, https://doi.org/10.1016/j.scitotenv.2019.135771.

IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Revue d'Intelligence Artificielle
Vol. 36, No. 6, December, 2022, pp. 959-967
Journal homepage: http://iieta.org/journals/ria

Exploitation of Advanced Deep Learning Methods and Feature Modeling for Air Quality Prediction

ShreeNandhini Parthiban*, Palaniswamy Amudha, Subramaniam Pillai Sivakumari

Department of Computer Science Engineering, School of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore 641108, India

Corresponding Author Email: shreenandhini2016phd@gmail.com

https://doi.org/10.18280/ria.360618

Received: 11 October 2022
Accepted: 18 December 2022

Keywords:
air quality prediction, deep learning, deep transfer learning, improved sparse auto encoder, feature modeling, stacked bidirectional and unidirectional LSTM

ABSTRACT

Air pollution is a major issue because Particulate Matter (PM) has a substantially higher effect on human health than other pollutants. Air Quality (AQ) prediction has become critical recently to take action to reduce pollution. This research introduces a unique methodology for assessing the effectiveness of PM10 and PM2.5. Enhanced spatial, temporal sequence-Improved Sparse Auto Encoder with Deep Learning (EISAE-DL) has been proposed to predict AQ affected by the prolonged dependency of air pollution congregation. However, Long Short-Term Memory (LSTM) used in EISAE-DL has suffered from the learning of a long-term dependent sequence of the training dataset. In addition, it is hard to create very reliable AQ forecasts at higher periodic frequencies, such as daily, weekly, or even monthly. This paper proposes Transfer learning (TL) in a Stacked Bidirectional and Unidirectional LSTM to solve the learning issue in LSTM for long-term dependencies. So, EISAE-DL with TL and modified LSTM model is named as EISAE-Deep Transfer Learning (EISAE-DTL). TL with a modified structure can handle large-size datasets effectively. However, training time is increased more than twice for non-transfer learning way of modeling due to TL, Wasserstein Distance-based adversarial learning is proposed in EISAE-DTL to decrease the variances among AQ data collected from any two sites. The proposed work is named EISAE- Enhanced DTL (EISAE-EDTL). The developed EISAE-DTL and EISAE-EDTL models are compared and analyzed with the performance of existing algorithms EISAE-DL, ISAE-DL, TL-BLSTM, MMSL, and ST-DNN. The experimental findings demonstrate the accuracy, precision, sensitivity, specificity, Area Under Curve (AUC), and Matthew's correlation coefficient of the proposed model performs admirably and improves present condition approaches.

1. INTRODUCTION

As AQ deterioration is becoming more of an issue, Particulate Matter (PM) has a considerable negative influence on human well-being. Since the fine PM2.5 has a smaller diameter, it can reach the alveoli and even the bronchioles more quickly, interrupting lung gas interchange. Long-term submission of PM in the air raised the risk of cardiovascular illness, respiratory problems, and lung cancer [1, 2]. AQ monitoring systems have been set up in various localities in response to rising public health awareness. On the other hand, several providers merely can expose the AQ and cannot predict it. Predicting AQ is critical for directing choices and actions aimed at minimizing PM2.5 prominence, such as evaluating whether to participate in internal or external activities. Consequently, a complicated array of elements [3, 4] such as emissions, traffic conditions, and weather data, make precise AQ prediction challenging.

In the lack of physical models, data mining enables novel approaches for analyzing AQ [5, 6] and may find hidden patterns in the acquired data. A Machine Learning (ML) based method was developed [7] that considered both spatial and regional relationships into account. The spatial classifier Artificial Neural Network (ANN) analyzes global information by utilizing mean results gathered from the surrounding regions. However, the results of this model were unaffected by local and global climatic conditions.

The three-input regression model was proposed to combine temporal and spatial predictors with regional weather information [8]. The identification features may vary to the urban scenario an extraordinary increase in people activity, and higher electricity and transport requirements. These elements play a key role in urban air pollution caused by clogged major highway networks. A linear model approach considers these factors [9].

The more factors considered for AQ prediction make decisions more precise. Also, knowledge obtained from huge datasets improves the accuracy of AQ prediction. So, DL models [10, 11] were used for AQ prediction. But, all proposed AQ prediction models suffered from data dissimilarity of collected data from various cities. Also, they cannot handle long delay-based dataset, which causes a high learning time.

So, the data dissimilarity issues were solved in ISAE-DL and EISAE-DL techniques using an auto-encoder and DL method [12]. The data collected from spatially and temporally correlated locations are grouped based on Manhattan distance. The PM-related data are fed into ANN and LSTM while topography data are fed into CNN. Then, aggregated data is fed into a sparse autoencoder for normalization. Then classifier is used for AQ prediction. The LSTM used in

EISAE-DL is suffered to learn the long-term dependencies of air pollutant concentrations. The learning issue in LSTM for long-term dependencies is solved in this paper by using Transferred Stacked Bidirectional and Unidirectional LSTM-based transfer learning in EISAE-DL which handle large size and long delay-dependent datasets. Wasserstein Distance-based DTL (WD-DTL) is employed in EISAE-DTL to minimize the disparities across the origin and destination areas through adversarial learning. In summary, the major contribution of this paper is the following:

- To develop EISAE-DTL for AQ prediction by adopting T-SBU-LSTM which extarcts temporal relation among long delay-dependent data efficiently.
- To develop EISAE-EDTL using the WD-DTL with CNN for AQ prediction, which extarcts informative data of terrain details.
- To evaluate the efficiency of the EISAE-DTL and EISAE-EDTL compared with the existing AQ prediction models.
- Thus, the utilization of TL can efficiently handle the long delay-dependent datasets and decrease the learning period of the classifier for appropriately predicting the AQ contrasted with the existing DL methods.

The remaining sections of the manuscript are structured as follows: Section 2 discusses the literature on the diagnosis of DL-based AQ prediction. Proposed EISAE-DTL and EISAE-EDTL are described in Section 3 and the results are shown in Section 4. The study is summed up and suggestions for further research are offered in Section 5.

2. RELATED WORK

An Apriori pattern mining technique [13] was employed to mine contextual spatial-temporal correlations among PMs. The temporal sequences provide high-frequency rule generation from specified intervals in each time series. Thus the techniques emerge to identify the relationship among emissions in various locations with varying timeframes. On the other hand, periodically executing the rule creation process for every data update is a time-consuming process.

As a means of improving the accuracy of the nonlinear AQ Index (AQI) series, a novel hybrid learning strategy [14] was provided with multidimensional scaling-based K-means and a Modified Extreme Learning Machine (MELM) for urban AQ prediction. Predictions for AQ were made using a neuro-fuzzy network and self-organizing clustering [15]. To fine-tune the network's settings, evolutionary techniques like steepest descent backpropagation are employed. Fuzzy rules, once learned, can be utilized to make predictions about the AQ of test data.

However, ML-based approaches are limited to processing large-size datasets. To solve this issue, DL-based approaches were proposed to handle large-size data efficiently. A Spatial-Temporal Deep Neural Network (ST-DNN) was a comprehensive predictive model [16] for AQ forecasts by incorporating data from a wide range of monitoring sites, temperature, wind features, and direction and elevation information. A temporal sliding-based Bidirectional LSTM (BLSTM) with which strong temporal connections can be preserved [17] was proposed to handle PMs and related information integrated with an appropriate time lag. AQ forecasting using a TL-BLSTM [18] was proposed which learns from PM2.5's long-term dependencies and then uses transfer learning to apply that knowledge across different time scales.

Empirical Mode Decomposition (EMD) and BLSTM were proposed [19] that consist solely of PM2.5 time-series data, which are treated as signal data. To break down the data and pull out the frequency and amplitude features, EMD is used as an unsupervised feature learning method. AQ prediction of shorter-term trends, especially for unexpected shifts, was enhanced by this method. Multi-output and Multi-index of Supervised Learning (MMSL) were developed in LSTM [20] to simultaneously learn the PM data of the current monitoring station and its nearest neighbor stations.

The underlying uncertainties associated with data necessary to run these models are another source of uncertainty in operational models. The natural sources associated, for instance, are hard to fully characterize. The disadvantages of operational models involve the usage of standard settings and the lack of information for the same geographical scale that might be used to evaluate model findings. To solve these issues, a novel AQ prediction technique is proposed in this paper.

3. PROPOSED METHODOLOGY

This model considers the Air quality data in India (2015 – 2020) database, which comprises the name of the city, date, PM2.5, PM10, NO, NO_2, NO_x, NH_3, CO, SO_2, O_3, Benzene, Toluene, Xylene, AQI, and class of different stations across 26 places in India. It is retreived from kaggle website. The suggested predictive model framework, depicted in Figure 1, is made up of four key components. T_G refers global time between all datations. T_L represent the local time history of data colelcted in every stations.Time perid is 3 years durarion from every day with every 6 hours colelctions

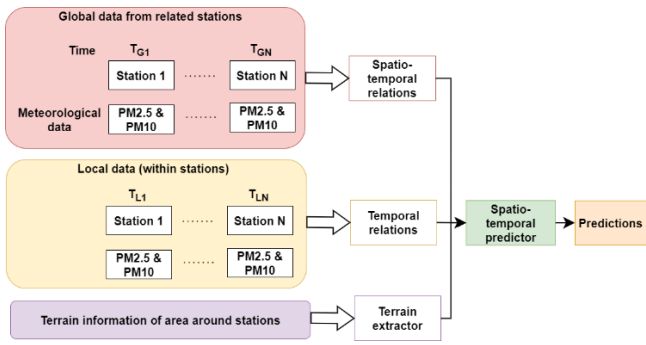


Figure 1. Overall prediction framework

To examine the sequence delays and linkages among places using past temporal patterns, as well as to analyze location feature variations for future determinants, to uncover the most significant interactions between locations. Even if they have substantial correlations with the target place, it is worthwhile to investigate nearby places. ISAE-DL and EISAE-DL connection extractors are used to find the most comparable locations to the requested location, and then create training datasets from the best k associated locations. At last, the prediction outcomes of the DL model must be trained and tested.

The suggested EISAE-DTL paradigm combines objective location temporal data, related position spatial-temporal data, and terrain data. The data stream contains statistical data from the target and related sites, such as pollution levels, weather

conditions, and objective features, as well as changes over a previous couple of hours. These data are given as input to LSTM, Adaptive Temporal Extractor (ASE), and ANN. This information was fed into the LSTM and ANN. A matrix of 121 squares was employed for topography data, equivalent to 1111 reference locations at 500 m intervals, with the principal square in the matrix reflecting the present areas. As a result, there are 120 previously unknown locations with AQI estimated using WD-based deep adversarial TL.

Concatenating the proportional altitude with the unknown points AQIs and these points are given as input to the CNN to lessen AQI influence at a greater altitude. CNN inputs can be fine-tuned later to improve accuracy. Contaminants, climatic situations, and objective attribute(s) of regions with significant resemblance (identified by ISAE-DL and EISAE-DL) were entered into the EISAE-DL and ASE without any previous training. To integrate TRE, SRE, and TE, a two-layer Feed-forwrad FNN was used. The divergence among the targeted feature value and a particular time t_{q+h} $where$ $t_{q+h} < fi_t$ was the actual prediction.

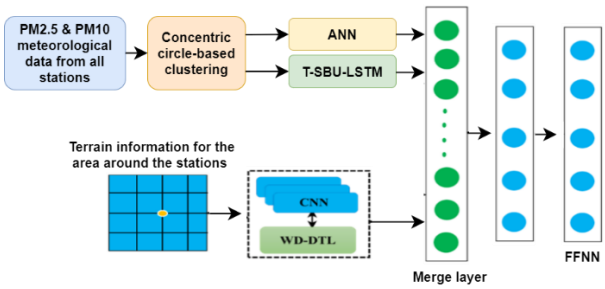


Figure 2. Architecture of EISAE-DTL framework

The model's structure is depicted in Figure 2. EISAE-DL and ASE receive data on AQ and meteorological conditions, whereas CNN receives data on terrain [21]. The models are con catenated beside each other, and the parameters are transferred onto the next layer. Because the present state evolves in terms of its impact on upcoming time intervals, the classifier is constructed hourly throughout the next 48 hours.

As a result, in specified time intervals, the inputs are integrated with the desired feature changes to learn distinct versions with identical topologies multiple times. This approach has the advantage of maintaining uniform input sizes regardless of position or time. To extract AQ feature information from linked sites identified by ISAE-DL and EISAE-DL, the spatial-temporal relationships extractor (SRE) utilizes an ANN model The temporal relationships extractor (TRE) collects AQ attributes using meteorological data obtained in the objective region during the last few hours to apply an LSTM model.

3.1 Temporal relations

The temporal sequence input for PM2.5, PM10, and other elements are continuous and constant, and they can be classified as lower (trends) or high (fast growth) bandwidth data. EISAE-DTL is developed to acquire target location time series patterns because EISAE-DL simulates historical time series behavior; however, the ANN employs only recent data and is thus sensitive to quick changes. As a result, the EISAE-DL, and ANN obtain less and more bandwidth data from the patterns, respectively. The EISAE-DTL forecasts trends in PM2.5 and PM10 levels for the last six hours and local weather parameters (wind speed and direction, humidity, and temperature), whereas the ASE TRE enhances model sensitivity by utilizing similar attributes as the EISAE-DTL. Preceding research work has established the significance of these features in terms of AQ [22].

3.2 Spatial-temporal relations

Contaminant diffusion refers to the ability to link AQ at one area with that at other locations Because AQ in a single area is influenced by both local pollutants and pollution from neighboring locations, SRE takes into account past spatial-temporal nearby location features as inputs.

As a result, SRE is developed which leverages AQIs and meteorological data from other places to calculate AQ at the target location. Terrain factors, such as a hilltop between locations, are underestimated when partitioning spaces into sections using circles of varying diameters. Data from partitioned regions is frequently comprised of mean or mode values, which are incredibly unreliable, especially in areas with few locations. Furthermore, for an SRE position, data mining from places in the spatial-temporal proximity using ISAE-DL and EISAE-DL is required, which includes AQIs and meteorological variables (wind speed and direction) for the previous 6 hours. The spatial-temporal neighborhood period sequences such as the target location are undeniably durable and coherent. As a result, by accounting for spatial-temporal neighborhood influences, ANN-SRE is used to improve model stability [22].

3.3 Terrain extractor

Due to various barriers and elevation variances, the relationships between locations differ. As a result, most terrain information is utilized to improve position associations. Terrian is generally expressed in terms of the elevation, slope, and orientation of the area around the stations. Terrain statistics in the locations were obtained using a 121-square-section vector made up of 11*11 coordinate lines spaced 500 meters apart. To establish the correlations between terrain and PM2.5 & PM10, the approach [23] of is used for the evaluation of each location is normalized as Eq. (1):

$$El_s = \frac{ele - ele_{st}}{ele_{st}} \tag{1}$$

and transferred to equivalent altitude is expressed as Eq. (2):

$$ele_{rel} = \frac{1}{e^{El_s}} \tag{2}$$

The standardized altitude is represented as El_s, and it reduces the effect of higher altitudes, although the distribution is heavily sensitive to elevation. The TE, SRE, and TRE results are combined and given to the FFNN using an enhanced sparse autoencoder. Figure 2 depicts the suggested approach's general framework. The upgraded sparse autoencoder accomplishes the process of layer blending and outlier detection. The input data is successfully managed in a systematic manner.

Complex forms are formulated from smaller shapes with the use of an autoencoder. Autoencoders are capable and efficient detectors of important features. For the learning experience, data is combined in both continuous and categorical forms. The value obtained is multiplied by the input data in this strategy, which employs a feed-forward artificial neural

network-based perceptron. The linear function, log-sigmoid, hard limit, and hyperbolic tangent with probable saturation is used to start the process, which is then added to the total inputs and weights. The value of the consequence is calculated as Eq. (3):

$$B = f(wp + a) \tag{3}$$

In Eq. (3), $f(\cdot)$ is the activation function, w is the weight value, a is the bias, and p is the given input. Traditionally, perceptrons employ a primary function for prediction, and the commonly chosen function is provided by Eq. (4)

$$f(p) = \frac{1}{(1+e^{-x})} \tag{4}$$

The precise estimation of the weight lowers the error among the output and the assumed value is decided in the training set of data. The existence of multiple perceptrons is organized as several layers, with each layer's output data being sent to the input of the next layer. This multilayer network can solve challenging linear separable classification problems. By using the initial weights, the input data is propagated [24]. The error value can be estimated from the desired outcome, which is the variance in feed-forward propagation output. The algorithm is developed in batches, and each test is completed before the weight is updated. The mean of the elevation value is used to optimize the weights, and each input is probed by the revised weights, while the output data is categorized correspondingly.

3.4 Proposed EISAE-DTL

The proposed Enriched spatial, temporal sequence-Improved Sparse Auto Encoder with Deep TL (EISAE-DTL) algorithm is mainly used to solve the learning issue in LSTM for long-term dependencies. Deep LSTM frameworks are networks with several stacked LSTM hidden layers, with each LSTM hidden layer's output passing into the next LSTM hidden layer. To enhance prediction accuracy at higher temporal resolutions, TL is used. To transmit data from the source to the new domain, TL exploits similarities between two separate datasets, tasks, or models.

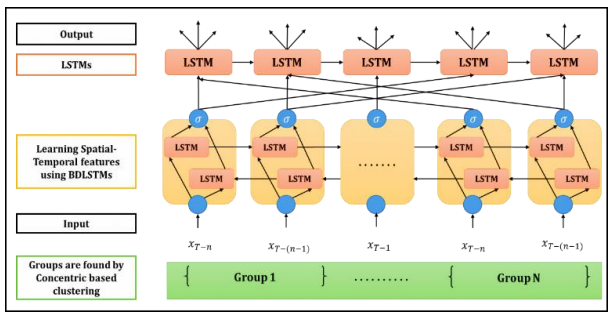


Figure 3. Architecture of T-SBU-LSTM

This research adopts the stacked-layers technique, which has been found to improve neural network efficiency. T-SBU-LSTM was used in EISAE-DL to forecast network-wide traffic speed values. Figure 3 depicts the T-SBU-LSTM design. The Transferred Stacked Bidirectional and Unidirectional LSTM (T-SBU-LSTM) is proposed as a method for learning from long-term PM2.5 dependencies in this study, and it employs TL to deliver data from lower to higher temporal resolutions. Figure 3 depicts the T-SBU-LSTM structure.

BLSTM is an improved RNN that can learn long-term dependencies from both upwards and backward sequences of time series data. SBU-LSTM (Stacked Bidirectional and Unidirectional LSTM) is an AQ prediction model developed by merging LSTM and BLSTM. The proposed paradigm can deal with both long-term and short-term dependency. This phase of research extends the AQ prediction from a single place to multiple nearby locations, with time delays ranging from short to lengthy. By learning both forward and backward dependencies, the integrated architecture improves feature learning from large-scale spatiotemporal time series data.

TL allows the model to learn and store information from samples with lower temporal resolutions, as well as increase prediction performance for data with higher temporal resolutions. The initial feature-learning layer of T-SBU-LSTM is a BLSTM layer, while the last layer is an LSTM layer. The T-SBU-LSTM can be filled with one or more LSTM/BLSTM layers in the middle to make full use of the input data and learn sophisticated and extensive features. T-SBU-LSTM predicts future temporal values for one time step using spatial time series data as input. Based on past data, the SBU-LSTM can also estimate values for numerous future time steps.

Initially, the sequences are projected for the targeted area using time sequence attributes that incorporate spatial data after selecting the places with the greatest significant spatial-temporal linkages to the specified area. Consequently, the position remains constant, but the time series might change depending on where they are. A mountain between two points, for example, could affect the sequences. As a result, the prediction model takes into account both time and spatial relationships. To extract features for prediction, the spatial and temporal connection features should be developed. Assume, the group of regions: $A = \{a1, a2, \ldots an\}$ and set of features: $FS = \{fs1, fs2, \ldots . . fsm\}$. Although each area contains geographical data like latitude and longitude, the area coordinate (AC) is described as Eq. (5):

$$AC_i = (a_i, l_i, m_i) a_i \in A \tag{5}$$

The latitude and longitude of area ai are li and mi, respectively. Considering related geographical features may enhance prediction, the distance between two locations is computed as Eq. (6):

$$D_{p,q} = dist_{area}(AC_p, AC_q) = dist_{area}\left((a_p, l_p m_p), (a_q, l_q m_q)\right), a_p, a_q \in A \; p \neq q \tag{6}$$

To discover the spatial domain's closest strongly linked areas, and the Spatial Relationships Sequence Set (SRSS) as Eq. (7):

$$SRSS = \{D_{1,2}, D_{1,3} \ldots \ldots . D_{n-1,n}\}, D_{i,j} = 0, 0 < i < n + 1 \tag{7}$$

where, n is the number of positions and the diagonal numbers $D_{i,j}$ will be 0. $SRSS\_cand(a_i, x)$ is a set of x regions with the shortest spatial proximity to a_i.To investigate the attributes of these major spatial places, as areas with identical pattern sequences may aid prediction. The Feature Sequence Interval (FSI) for a particular region is described as Eq. (8):

$$F(a_i, fs_j, t_{vt,st} = \{b(a_i, fs_j, t_{vt}), b(a_i, fs_i, t_{vt+1}), \ldots . . b(a_i, fs_i, t_{vt}), a_i \in A, fs_j \in FS, v_t < s_t\}) \tag{8}$$

where, a_i has a variable feature fs_i from start to end (v_t to s_t time ($v_t < s_t$), and b(a_i, fs_j, t_x) indicates the observed level of off fs_j at t_x. For any two sites, the proximity across feature patterns can be represented as Eq. (9):

$$DS_{p,q,t_{vt,st}} = dist_{seq}\left(F(a_p, fs_{target}, t_{v_t,s_t}), F(a_q, fs_{target}, t_{v_t,s_t})\right), \quad a_p, a_q \in A, \quad p \neq q \tag{9}$$

Apply TL to enhance the performance of the prediction model. Assume a target training set of n_t instances $Aq_t = \{(a_1^t, b_1^t), \ldots\ldots, (a_{n_t}^t, b_{n_t}^t)\}$ drawn from some probability D_i, and an origin learning set of ns samples $Aq_s = \{(a_1^t, b_1^t), \ldots\ldots, (a_{n_s}^t, b_{n_s}^t)$. The objective and source training sets maintain the identical feature space and label space, with each input feature vector $ai \in En$ and the accompanying class label $bi \in \{C_1, \ldots, C_L\}$. Further, target data Aq_t should not be inferred from the same distribution as the source data Aq_s, implying that the models derived from the source set would be unable to reliably identify the (target) test data due to the different time series. The size of Aq_t, on the other hand, is typically insufficient for training an effective classifier for the test data. The purpose of TL is to use knowledge from Aq_s to aid in the learning of the target prediction function in Aq_t.

In order to find the positions which are mostly related closely before utilizing the measures, initially, feature fs_{target} is selected to act as the objective detection sequence: in this case, PM2.5 was selected, but other objectives might be utilized if necessary. Eq. (10) can then be used to compute the collection of Temporal Relations Sequence Set (TRSS).

$$TRSS_{t_{vt,st}} = \{DS_{1,2,t_{vt,st}}, \ldots\ldots\ldots DS_{n-1,n,t_{vt,st}}\} \tag{10}$$

The set of x positions with the minimum variations from position i is then chosen as TRSS_cand (a_i,x). The Spatial-Temporal Relations (STR) cluster is introduced which is the collection of locations that are most strongly associated to ai to investigate both linkages respectively in Eq.(11):

$$STR\_cand(a_i, x) = SRSS\_cand(a_i, x) \cup TRSS\_\_cand(a_i, x), a_i \in A \tag{11}$$

Instead of using the intersection, the union of SRCS cand (a_i,x) and TRSS cand (a_i,x) is used to provide a greater value of connections for the methodology to train; the conjunction will have lesser candidates (or none at all), leading to the in the absence of significant objective quality because certain location behaviors varied from adjacent locations. The Spatial-Temporal Predictor (STP) is defined by $STR_{cand(a_i,x)}$ Eq. (12):

$$P(STR_{cand(a_i,x)})[t_{t_l,tq}] = F(a_i, fs_{target}, t_{vt',st'}), t_l < b_q < vt' \leq st' \tag{12}$$

In the above equation, *P* predicts a series of objective qualities and delivers a sequential set, *S* which provides the expected objective characteristics For the temporal range from tvt to tst,. *F* is produced from the nearest identical time series, where tl is the recalled time when comparing tl to tq.

Algorithm 1: Pseudocode for EISAE-DTL

Input: Target Terminal (T); set of geographic coordinates (LC), number of candidates (n)

Output: Set of locations L

Step 1: Start the process

Step 2: Assume a set of areas: $A = \{a1, a2, \ldots an\}$

Step 3: Assume a set of features: $FS = \{fs1, fs2, \ldots\ldots fsm\}$

Step 4: Compute Area Coordinate (AC)

$AC_i = (a_i, l_i, m_i)$

//Let latitude and longitude of area a_i are l_i and m_i

Step 5: Calculate the distance between two locations

$D_{p,q} = dist_{area}(AC_p, AC_q) = dist_{area}\left((a_p, l_p m_p), (a_q, l_q m_q)\right), a_p, a_q \in A$

//Apply TL for T-SBU-LSTM

Step 6: Initialize AQ data $\widetilde{Aq_s} = \varnothing$.

Step 7: for l = 1 to L do

7.1 Initialize a multilayer autoencoder MAE^l (W,b).

7.2 Select location-based AQ $Aq_t^{c_l}$ from Aq_t.

7.3 Train AE^l(W,b) using $Aq_t^{c_l}$

7.4 Select location-based AQ $Aq_s^{c_l}$ from Aq_s

7.5 Reconstruct data $Aq_s^{c_l} = \widetilde{MAE}^l_{Recon}(Aq_s^{c_l})$

7.6 Update the reconstructed data $\widetilde{Aq_s} \cup Aq_s^{c_l}$

Step 8: end for

Step 9: Find spatial relationships sequence

SRSS = $\{D_{1,2}, D_{1,3}, \ldots\ldots D_{n-1,n}\}$, //$D_{i,i}$= 0

// n is the integer spots, and the oblique values are $D_{i,i,}$ are 0.

Step 10: Compute FSI

$F(a_i, fs_j, t_{vt,st} = \{b(a_i, fs_j, t_{vt}), b(a_i, fs_i, t_{vt+1}), \ldots\ldots b(a_i, fs_i, t_{vt}), a_i \in A, fs_j \in FS, v_t < s_t\})$

// a_i possesses fsi that fluctuates from start to end (vt to st) time (vtst); and b(ai,fsj,tx) indicates the measured fsi level at tx.

Step 11: Find the distance between feature sequences

Step 12: $DS_{p,q,t_{vt,st}} = dist_{seq}\left(F(a_p, fs_{target}, t_{v_t,s_t}), F\begin{pmatrix} a_q, fs_{target}, \\ t_{v_t,s_t} \end{pmatrix}\right), a_p, a_q \in A$

Step 13: Compute TRSS

$TRSS_{t_{vt,st}} = \{DS_{1,2,t_{vt,st}}, \ldots\ldots\ldots DS_{n-1,n,t_{vt,st}}\}$

// The set of x positions with the minimum variance from position i is then chosen as TRSS_cand (a_i,x).

Step 14: Evaluate STR

$STR\_cand(a_i, x) = SRSS\_cand(a_i, x) \cup TRSS\_\_cand(a_i, x)$

Step 15: Each location is normalized as

$El_s = \frac{ele - ele_{st}}{ele_{st}}$

Step 16: Transferred to equivalent altitude is expressed as

$ele_{rel} = \frac{1}{e^{El_s}}$

Step 17: Calculate STP

$P(STR\_cand(a_i, x))[t_{t_l,tq}] = F(a_i, fs_{target}, t_{vt',st'})$

Step 18: ASE receives data on AQ and meteorological conditions, whereas CNN receives data on terrain

Step 19: Concatenate each other and the parameters are transferred onto the next layer

Step 20: Perform a feedforward pass using FFNN

Step 21: Determine the activations for levels L2, L3, and so on, all the way up to the output layer Lnl.

Step 22: End the process

3.5 Proposed EISAE-EDTL

EISAE-EDTL is designed to shorten the training time of TL

due to heavier data in the source domain by employing WD-DTL to optimize the differences among the actual and desired domains. Figure 4 depicts the WD-DTL architecture. The TL method is used to investigate the transferable characteristics of an AQ dataset under various meteorological situations. The atmospheric statistics are temperature, wind speed and direction, average wind speed and direction, humidity level, and elevation. Initially, a basic LSTM model is trained with appropriate data. Then, to acquire terrain features across source and target regions around the stations, the WD-DTL is developed and combined with the CNN, which minimizes the learning period. A neural network (known as domain critic) is built to calculate the actual Wasserstein proximity by improving domain critic loss. The LSTM-based feature extractor variables are then adjusted using a discriminator by lowering the predicted empirical Wasserstein distance. Using the adversarial training approach mentioned above, transferable attributes from a given area with confirmed erroneous labels can be used to diagnose a different yet similar detection objective without any tagged examples.

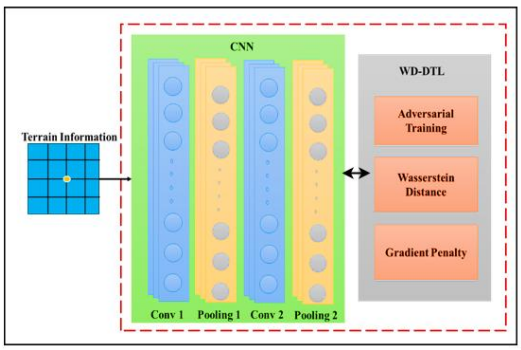


Figure 4. Architecture of CNN –WD-DTL

Suppose Aqs and Aqt are the objective datasets; the learning rate for classifier and feature learning is 2, whereas the training rate for domain conformity critic is 1. the batch size is n; the critic training step is Ct, and the balance coefficients and fe are the first CNN-based feature extractor variable, cp is the preliminary domain alignment critic parameters, and pp is the initial detective variable Gradient distance is calculated as Eq. (13):

$$ggrad \leftarrow (\|\nabla_2 g_{grad} \leftarrow (\|\nabla_f p_c(f)\|_2 - 1)^2 \quad (13)$$

The Wasserstein-1 distance can be approximated as:

$$li_{wd} = \frac{1}{N^s}\sum_{a^s \in Aq^s} p_c\left(p_f(Aq^s)\right) - \frac{1}{N^s}\sum_{b^t \in Aq^t} p_c\left(p_f(Aq^t)\right) \quad (14)$$

Algorithm 2 Pseudocode for EISAE-EDTL

Input: Target Terminal (T); Set of geographic coordinate (LC), number of candidates (n)

Output: Set of locations (L)

Step 1: Start the process

Step 2: Assume a set of areas: A={a_1,a_2,…a_n}

Step 3: Assume set of features: FS = {fs_1,fs_2,…..fs_m}

Step 4: Compute Area Coordinate (AC)

AC_i = (a_i, l_i, m_i)

//Let latitude and longitude of area a_i are l_i and m_i

Step 5: Calculate the distance between two locations

$D_{p,q}$ = $dist_{area}$ (AC_p, AC_q)

= $dist_{area}$((a_p,l_p,m_p), (a_q,l_q,m_q))

Step 6: ASE receives data on AQ and meteorological conditions, whereas CNN receives data on terrain

// Apply WD-DTL

Step 7: While θ_{fe}, θ_{cp}, and θ_{pp} have not converged do

Step 8: Assume the source dataset $Aq_t = \{(a_1^s, b_1^s)\}_{i=1}^n$

Step 9: Assume the target dataset $Aq_s = \{a_1^t\}_{i=1}^n$

Step 10: For i=0,……Ct

Step 11: $fs \leftarrow pf(Aqs), ft \leftarrow pf(Aqt),$

Step 12: $f \leftarrow \{fs, ft, fr\}$

Step 13: $g_{grad} \leftarrow (\|\nabla g_{grad} \leftarrow (\|\nabla_f p_c(f)\|_2 - 1)^2$

$\theta_{cp} \leftarrow \theta_{cp} + \beta_1 \nabla\theta_{cp} li_{wd}(a^s, b^t) - \tau g_{grad}(f)$

Step 14: end for

$\theta_{pp} \leftarrow \theta_{pp} + \beta_2 \nabla\theta_{pp} l_{ct}(a^s, b^s)$

$\theta_{fe} \leftarrow \theta_{fe} + \beta_2 \nabla\theta_{fe}[l_{ct}(a^s, b^s) + \delta li_{wd}(Aq^s, Aq^t)]$

Step 15: end while

Step 16: Find spatial relationships sequence

$SRSS = \{D1,2, D1,3, \ldots . Dn-1, n\}$, //$D_{i,i}$= 0

// n is the integer spots, and the oblique values are $D_{i,i,}$ are 0.

Step 17: Compute FSI

$F(ai, fsj, tvt, st) = \{b(ai, fsj, tvt), b(ai, fsi, tvt + 1), \ldots, b(ai, fsi, tvt)\}$

//a_i has feature fs_i that changes from first to last (vt to st) time (vt<st); and $b(a_i,fs_j,t_x)$ denotes determined value of fs_i at t_x.

Step 18: Find the distance between feature sequences

$DS_{p,q,tvt,st}$= $dist_{seq}$(F(a_p,fs_{target}, $t_{vt,st}$), F(a_q,fs_{target},$t_{vt,st}$))

Step 19: Compute TRSS

$TRSS_{t_{vt,st}} = \{DS_{1,2,t_{vt,st}}, \ldots \ldots \ldots DS_{n-1,n,t_{vt,st}}\}$

// The set of x positions with the minimum variance from position i is then chosen as TRSS_cand (a_i,x).

Step 20: Evaluate STR

$STR\_cand(a_i, x) = SRSS\_cand(a_i, x) \cup TRSS\_\_cand(a_i, x)$

Step 21: Calculate STP

$P(STR\_cand(a_i, x))\ [t_{t_l,tq}] = F(a_i, fs_{target}, t_{vt',st'})$

Step 22: Execute a feedforward pass, calculating the activations for layers L2, L3, and so on all the way up to the output layer Lnl.

Step 23: End the process

4. RESULT AND DISCUSSION

The efficiency of EISAE-DTL and EISAE-EDTL is compared with ST-DNN [16], and EISAE-DL [12] on the considered dataset in terms of accuracy, precision, sensitivity, specificity, AUC, and MCC.

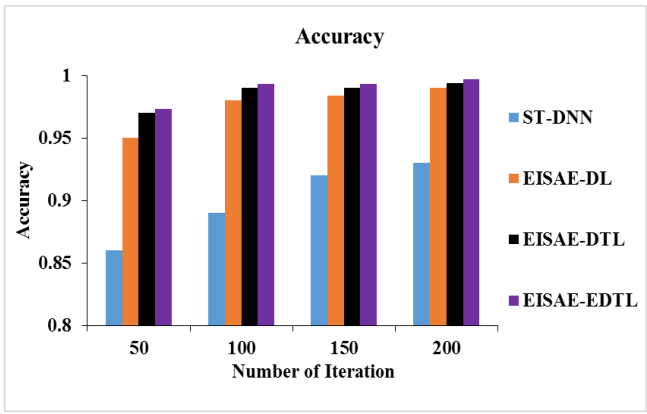
4.1 Accuracy and precision

Accuracy is the estimation of the actual value in the AQ prediction. Also, it is the identification (both valid positive and actual negative values) amongst the number of estimated classes. It is calculated as Eq. (15):

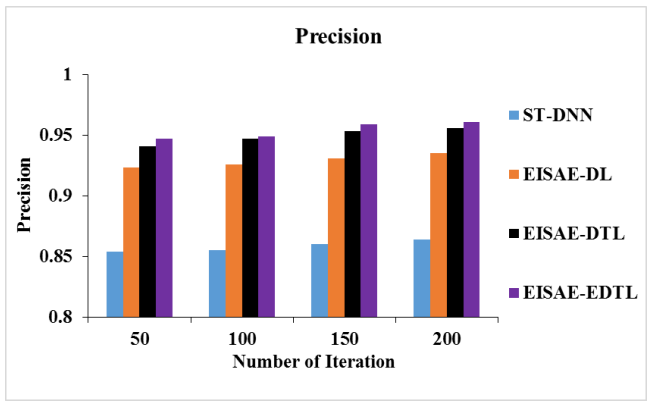
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

Precision defines the closeness of the measurement and the relevance among the values identified in the AQ prediction. It is calculated as Eq. (16):

$$Precision = \frac{TP}{TP+FP} \quad (16)$$



(a)



(b)

Figure 5. Evaluation of (a) accuracy, and (b) precision of EISAE-EDTL with existing works

Figure 5(a) & 5(b) depicts the accuracy & precision of ST-DNN, EISAE-DL, EISAE-DTL and EISAE-EDTL for various iterations. The accuracy (precision) is represented by the Y-axis, while the number of iterations is represented by the X-axis. When the range of iterations is 100, the accuracy (precision) of EISAE-EDTL for AQ prediction is 11.57% (11%), 1.33% (2.5%), and 0.303% (0.21%) higher than ST-DNN, EISAE-DL, and EISAE-DTL. This analysis demonstrates that the EISAE-EDTL exceeds other AQ detection methods in terms of accuracy and precision.

4.2 Sensitivity and specificity

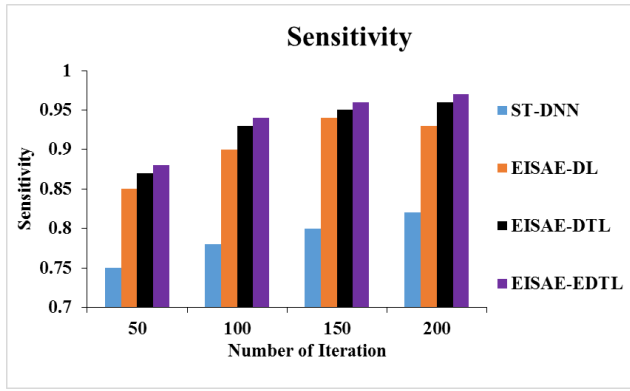
Sensitivity is the fraction of positive values that are adequately identified. It is calculated as Eq. (17):

$$Sensitivity = \frac{TP}{TP+FN} \tag{17}$$

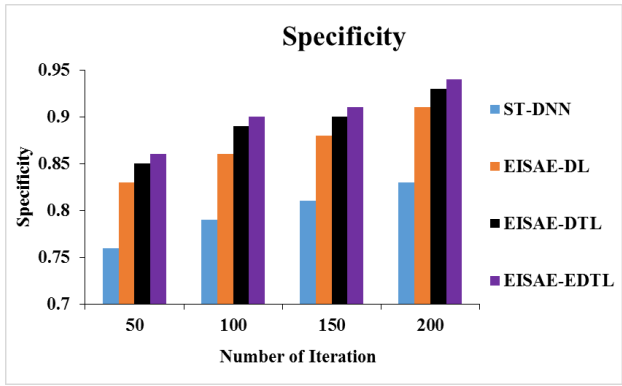
Specificity is calculated as the number of correct negative predictions divided by the total number of negatives. It measures the proportion of locations not impacted by air pollution, which is correctly predicted. It is calculated as Eq. (18):

$$Specificity = \frac{TN}{TN+FP} \tag{18}$$

Figure 6 (a) and (b) depicts the sensitivity and specificity of ST-DNN, EISAE-DL, EISAE-DTL, and EISAE-EDTL for various iterations. When the number of iteration is set to 100, the sensitivity (specificity) of EISAE-EDTL for AQ estimation is 20.51% (13.92%), and 4.44% (4.65%) and 1.08% (1.12) higher than ST-DNN, EISAE-DL, and EISAE-DTL. This evaluation demonstrates that the suggested EISAE-EDTL has higher sensitivity and specificity than traditional AQ prediction systems.



(a)



(b)

Figure 6. Evaluation of (a) sensitivity, and (b) specificity of EISAE-EDTL with existing works

4.3 AUC and MCC

AUC is the prediction indicator and is independent among the classes with distributed instances. It is computed by Eq. (19):

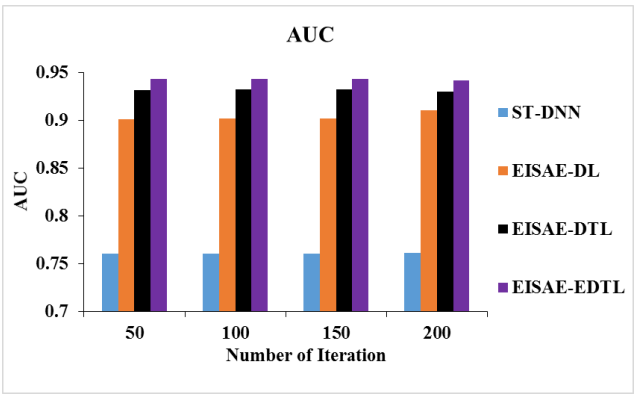
$$AUC = \frac{Sensitivity+Specificity}{2} \tag{19}$$

MCC is the correlation coefficient between the predicted and actual values. It is calculated as Eq. (20):

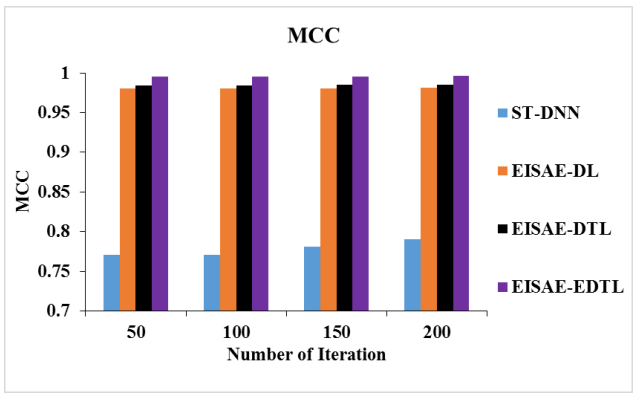
$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{20}$$

Figure 7 (a) and (b) depicts the AUC (MCC) of ST-DNN, EISAE-DL, EISAE-DTL, and EISAE-EDTL for various iterations. When the number of iterations is 100, the AUC (MCC) of EISAE-EDTL for AQ prediction is 24.12% (29.17%) 4.59% (1.52%), and 1.20% (1.138%) higher than ST-DNN, EISAE-DL, and EISAE-DTL. This investigation shows that the EISAE-EDTL has a higher AUC and MCC than classical AQ prediction systems.

Also, the proposed EISAE-EDTL is compared with the related works recently proposed TL-BLSTM [18], and MMSL[20] to show the effectiveness.



(a)



(b)

Figure 7. Evaluation of (a) AUC, and (b) MCC of EISAE-EDTL with existing works

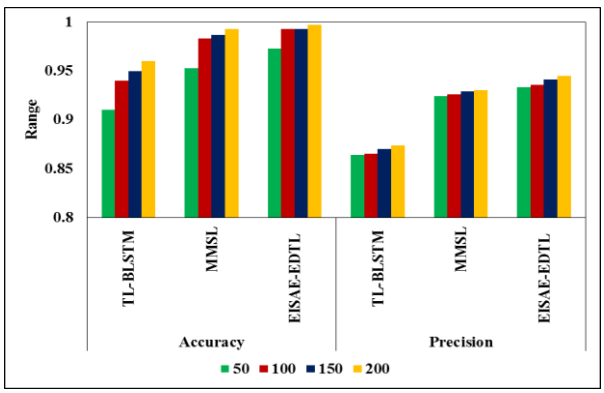


Figure 8. Evaluation of accuracy and precision of EISAE-EDTL with related work

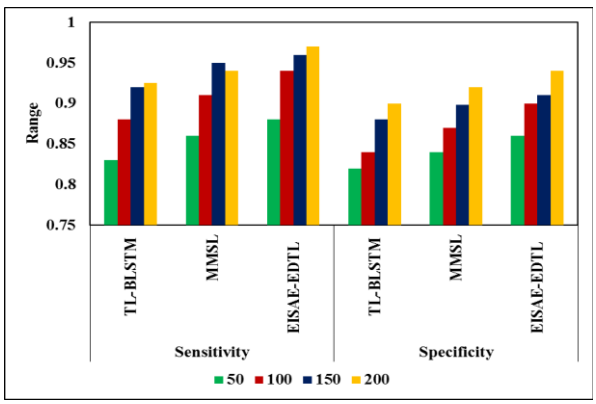


Figure 9. Evaluation of sensitivity and specificity of EISAE-EDTL with related works

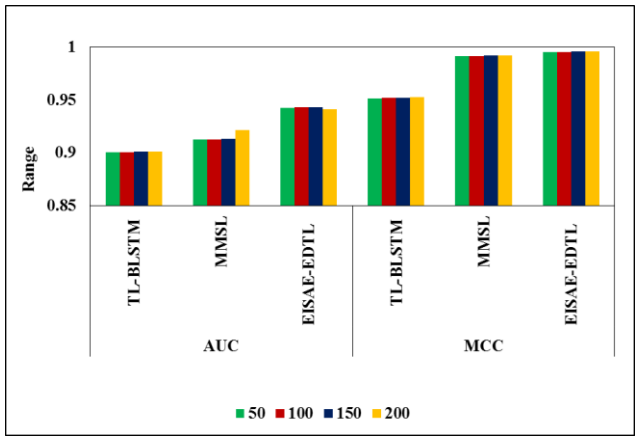


Figure 10. Evaluation of AUC and MCC of EISAE-EDTL with related works

Figure 8 depicts the accuracy (precision) of TL-BLSTM, MMSL, and EISAE-EDTL for various iterations. When the range of iterations is 100, the accuracy (precision) of EISAE-EDTL for AQ prediction is 13.45% (12.83%), and 4.12% (3.17%) higher than TL-BLSTM, and MMSL. This analysis shows that the EISAE-EDTL outperforms conventional AQ forecast systems in terms of accuracy and precision.

Figure 9 depicts the sensitivity (specificity) of TL-BLSTM, MMSL, and EISAE-EDTL for various iterations. This analysis demonstrates that the EISAE-EDTL has higher sensitivity and specificity than the classical AQ prediction models.

Figure 10 depicts the AUC (MCC) of TL-BLSTM, MMSL, and EISAE-EDTL for various iterations. When the number of iterations is 100, the AUC (MCC) of EISAE-DTL for AQ prediction is 22.24% (30.78%), and 3.41% (6.26%) higher than TL-BLSTM, and MMSL. According to this analysis, the EISAE-EDTL has a greater AUC and MCC than classical AQ prediction systems.

5. CONCLUSION

In this paper, EISAE-DTL and EISAE-EDTL are proposed to handle long-time delay-based locations for better AQ prediction. PM2.5 and PM10 datasets were used to test the proposed models. The significance of pertinent position selection was reaffirmed with the addition of all locations increasing model noise and resulting in poor predicting accuracy. The suggested method surpassed all baselines and comparator models studied. According to the suggested EISAE-EDTL, including an LSTM approach that improved first-hour forecasts with the CNN element being more successful for prolonged detections since CNN might recover the spatial latency factor from neighboring subjective qualities by employing spatial data. In terms of accuracy, precision, specificity, sensitivity, AUC, and MCC, the experimental results suggest that the suggested EISAE-EDTL surpasses current AQ prediction methods.

REFERENCES

[1] Pun, V.C., Kazemiparkouhi, F., Manjourides, J., Suh, H. H. (2017). Long-term PM2.5 exposure and respiratory, cancer, and cardiovascular mortality in older US adults.

American Journal of Epidemiology, 186(8): 961-969. https://doi.org/10.1093/aje/kwx166

[2] Zhao, Y.M. (2020). Spatial-temporal correlation-based LSTM algorithm and its application in PM2.5 prediction. Revue d'Intelligence Artificielle, 34(1):29-38. https://doi.org/10.18280/ria.340104

[3] Senthivel, S., Chidambaranathan, M. (2022). Machine learning approaches used for air quality forecast: A review. Revue d'Intelligence Artificielle, 36(1): 73-78. https://doi.org/10.18280/ria.360108

[4] Kumar, K., Pande, B.P. (2022). Air pollution prediction with machine learning: a case study of Indian cities. International Journal of Environmental Science and Technology, 1-16. https://doi.org/10.1007/s13762-022-04241-5

[5] Zhang, Y., Wang, Y., Gao, M., Ma, Q., Zhao, J., Zhang, R., Huang, L. (2019). A predictive data feature exploration-based air quality prediction approach. IEEE Access, 7: 30732-30743. https://doi.org/10.1109/ACCESS.2019.2897754

[6] Schürholz, D., Kubler, S., Zaslavsky, A. (2020). Artificial intelligence-enabled context-aware air quality prediction for smart cities. Journal of Cleaner Production, 271: 121941. https://doi.org/10.1016/j.jclepro.2020.121941

[7] Maleki, H., Sorooshian, A., Goudarzi, G., Baboli, Z., Tahmasebi Birgani, Y., Rahmati, M. (2019). Air pollution prediction by using an artificial neural network model. Clean Technologies and Environmental Policy, 21(6): 1341-1352. https://doi.org/10.1007/s10098-019-01709-w

[8] Yang, W., Deng, M., Xu, F., Wang, H. (2018). Prediction of hourly PM2. 5 using a space-time support vector regression model. Atmospheric Environment, 181: 12-19. https://doi.org/10.1016/j.atmosenv.2018.03.015

[9] Syafei, A.D., Fujiwara, A., Zhang, J. (2015). Prediction model of air pollutant levels using linear model with component analysis. International Journal of Environmental Science and Development, 6(7): 519-525. https://doi.org/10.7763/IJESD.2015.V6.648

[10] Soh, P.W., Chen, K.H., Huang, J.W., Chu, H.J. (2017). Spatial-Temporal pattern analysis and prediction of air quality in Taiwan. In 10th IEEE International Conference on Ubi-media Computing and Workshops, Pattaya, Thailand, Pattaya, Thailand, pp. 1-6. https://doi.org/10.1109/UMEDIA.2017.8074094

[11] Wang, J., Song, G. (2018). A deep spatial-temporal ensemble model for air quality prediction. Neurocomputing, 314: 198-206. https://doi.org/10.1016/j.neucom.2018.06.049

[12] ShreeNandhini, P., Amudha, P., Sivakumari, S. (2022). Comparative analysis of air quality prediction using artificial intelligence techniques. ECS Transactions, 107(1): 6059. https://doi.org/10.1149/10701.6059ecst

[13] Qin, S., Liu, F., Wang, C., Song, Y., Qu, J. (2015). Spatial-temporal analysis and projection of extreme particulate matter (PM10 and PM2.5) levels using association rules: a case study of the Jing-Jin-Ji region, China. Atmospheric Environment, 120: 339-350. https://doi.org/10.1016/j.atmosenv.2015.09.006

[14] Jiang, F., He, J., Tian, T. (2019). A clustering-based ensemble approach with improved pigeon-inspired optimization and extreme learning machine for air quality prediction. Applied Soft Computing, 85: 1-30. https://doi.org/10.1016/j.asoc.2019.105827

[15] Lin, Y.C., Lee, S.J., Ouyang, C.S., Wu, C.H. (2020). Air quality prediction by neuro-fuzzy modeling approach. Applied Soft Computing, 86: 1-13. https://doi.org/10.1016/j.asoc.2019.105898

[16] Soh, P.W., Chang, J.W., Huang, J.W. (2018). Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations. IEEE Access, 6: 38186-38199. https://doi.org/10.1109/ACCESS.2018.2849820

[17] Mao, W., Wang, W., Jiao, L., Zhao, S., Liu, A. (2021). Modeling air quality prediction using a deep learning approach: method optimization and evaluation. Sustainable Cities and Society, 65: 1-25. https://doi.org/10.1016/j.scs.2020.102567

[18] Ma, J., Cheng, J.C., Lin, C., Tan, Y., Zhang, J. (2019). Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques. Atmospheric Environment, 214: 1-10. https://doi.org/10.1016/j.atmosenv.2019.116885

[19] Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., Liu, J. (2020). Air quality predictions with a semi-supervised bidirectional LSTM neural network. Atmospheric Pollution Research, 12(1): 328-339. https://doi.org/10.1016/j.apr.2020.09.003

[20] Seng, D., Zhang, Q., Zhang, X., Chen, G., Chen, X. (2021). Spatiotemporal prediction of air quality based on LSTM neural network. Alexandria Engineering Journal, 60(2): 2021-2032. https://doi.org/10.1016/j.aej.2020.12.009

[21] Cheng, C., Zhou, B., Ma, G., Wu, D., Yuan, Y. (2020). Wasserstein distance based deep adversarial transfer learning for intelligent fault diagnosis with unlabeled or insufficient labeled data. Neurocomputing, 409: 35-45. https://doi.org/10.1016/j.neucom.2020.05.040

[22] Ferrero, L., Bolzacchini, E., Petraccone, S., Perrone, M. G., Sangiorgi, G., Porto, C.L., Ferrini, B. (2007). Vertical profiles of particulate matter over Milan during winter 2005/2006. Fresenius Environmental Bulletin, 16(6): 697-700.

[23] Ma, J., Li, Z., Cheng, J.C., Ding, Y., Lin, C., Xu, Z. (2020). Air quality prediction at new stations using spatially transferred bi-directional long short-term memory network. Science of the Total Environment, 705: 1-10. https://doi.org/10.1016/j.scitotenv.2019.135771

[24] Kurt, A., Oktay, A.B. (2010). Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. Expert Systems with Applications, 37(12): 7986-7992. https://doi.org/10.1016/j.eswa.2010.05.093

Comparative Analysis of Air Quality Prediction Using Artificial Intelligence Techniques

ECS Transactions, 107 (1) 6059-6066 (2022)
10.1149/10701.6059ecst ©The Electrochemical Society

Comparative Analysis of air quality prediction using artificial intelligence Techniques

P.ShreeNandhini[1]Dr.P.Amudha[2]Dr.S. S.Sivakumari[3]

PhD. Research Scholar [1] Professor [2] Professor and Head of department[3]
Department of Computer Science and Engineering, School of Engineering,
Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore
shreenandhini2016@gmail.com[1] amudharul@gmail.com[2] prof.sivakumari@gmail.com[3]

Abstract

Air population is the primary concern in most urban areas because of its notable impact on the economy and health across the universe. The emergence of industry and automobiles made air pollution worldwide, which causes a highly critical issue and a more significant impact on humans' health than the contaminants. It causes health-related Problems like lung-related diseases, namely respiratory problems and cardiovascular disease, and increases cancer. Accurate monitoring of air quality is of great importance to daily human life. The consuming time delay long-term delay is essential for long-term predictions. In this article proposed Enriched spatial-temporal sequence (EISAE-DL) improves the prediction accuracy by considering the long time delay based on locations. And Compared for Experimental algorithm Improved Sparse Auto encoder with Deep Learning (ISAE- DL) The experimental results show the effectiveness of the proposed EISAE-DL in terms of accuracy, precision, sensitivity, specificity, Area Under Curve (AUC), and Matthew's correlation coefficient (MCC).

Introduction

Air population is the primary concern in most urban areas because of its notable impact on the economy and health across the universe. The emergence of industry and automobiles had made the contamination in air measured with the assistance of sensors. Air quality statistics have made the precise and vital real-time monitoring device, whereas the acquired statistics played effective, ineffective monitoring. The non-scalability, inadequate data access, and massive cost of conventional air quality monitoring approach impose the researchers to establish the system by incorporating the advanced technologies wireless sensor network (WSN), cost-efficient ambient sensors, machine learning (ML), and internet of things (IoT) (1) Air quality is reported based on theair quality index (AQI) and its estimation by the parameters, namely PM10, PM2.5, and other components (2).

Degeneration of air quality has grabbed attention recently, with the particulate matter (PM), and it has a significant impact on humans' health. The minor diameter of delicate particulate matter (PM 2.5 & PM 10) permits it to pierce deep into the alveoli portion that obstructs the exchange gas within the lungs. Exposure to these particulate matters will increase the possibility of lung cancer, cardiovascular disease, and respiratory issues (3). Most of the available approaches display the air quality, and it doesn't forecast the air quality. The air quality forecasting system helps limit the actions that cause air pollution due to indoor or outdoor activities.

Literature Survey

The extra information is added and combined through the system, and that features learned from an improved sparse auto encoder can be leveraged in networks to establish a complete system model. Temporal Prediction is not to alter over time because it takes more time or failure to complete the Process, such as degradation or loss. To evaluate the change between distributions, detailed state data is a fuse between the distribution. The Metadata Combination of discrete and Continuous features Concepts deal with the learning, not dealing with The network. The improved sparse autoencoder can prove an effective and efficient feature detector. The feed-forward neural network-based perceptron method uses The improved sparse encoder method to interface with the Process's development. The proposed system attains effective forecasting with better accuracy when compared with the existing ST-DNN approach.

The spatial, temporal deep neural network uses numerous data from various monitoring stations that consist of PM2.5.PM10, humidity temperature, and wing, etc.,(4). The Process ST-DNN trains the air pollutant data. ST-DNN is established based on the temporal and geographical association integration approach between the monitoring locations. In this approach, anomaly data is not deal with during the Prediction. Approach to monitoring air quality.

A stand-alone and real-time air quality monitoring system with Rasberry Pi's assistance begins with based ARM-based minicomputer (5). Through this Method, the Measured air pollutants are across Delhi, and the results were displayed. In this, the Analyzation results have not carried the Process. k-Nearest Neighbor-based classification (6) approach was used to classify the air location with high pollution. This approach uses the machine learning-based classification scheme that is ineffective in the classification when the incoming data is high.

The observed air quality and Prediction increases with machine learning or mobile microscopy This approach is highly economical and portable, whereas the system gives microscope images that are digital holographic. During the processing of images, Complication increases the images used to measure the particulate matter. The greedy layer-wise Method is used and trained by the deep spatiotemporal learning (STDL) approach to monitoring air quality. The significant features of air quality are predicated. From the stacked auto encoder assistance system

Proposed Methodology

Initially, the spot is the most effective Spatio-temporal relationship to the objective location. It then forecasts the target location based on the spatial data's time-series features. The forecasting process retrieves the significant feature for Process the temporal and spatial features are determined. Suppose set of locations Lt={lt1,lt2,….,ltn} and set of relevant features Ft={f1,f2,…,fm}. Every location has geographical data, namely longitude and latitude, the location coordinate determined as,

$LtC\text{i} = (lt\text{i}, p\text{i}, q\text{i}) lt\text{i} \in Lt$

Pi signifies the latitude, and qi signifies longitude at its location, whereas the prediction process enriches the relevant location features. The Distance among two location coordinate is Equation as

$$Ds_{y,r} = dist_{loca}(LtC_y, LtC_r)$$
$$= dist_{loca}((lt_y, p_y, q_y), (lt_r, p_r, q_r))$$

to identify the correlated locations in the spatial region and sequence of relationship (SRS) is equated as,

$$SRS = \{Ds1{,}2, Ds1{,}3, \ldots \ldots, Dsn{-}1{,}n\},$$
$$Ds\text{i,i} = 0,\ 0 < \text{i} < n + 1$$

The location coordinates. The matrix location's count elements denote as n; the diagonal elements represent as

$D_{si,\ i}$are assigned as zero. The feature sequence interval (FSI) for a definite location is Equa

$$(lt\text{i}, ft\text{i}, t_{str.fini}) = \{\text{e}(lt\text{i}, ft\text{i}, t_{str}), (lt\text{i}, ft\text{i}, t_{str+1}), \ldots \ldots \ldots, \text{e}(lt\text{i}, ft\text{i}, t_{fini})\},\ lt\text{i} \in Lt, ft\text{i} \in F, str < fini$$

The lti has some feature fti that diverse from the start str to the finish fini, and the time constraint is str<fini and e(lti,fti,tk) denote the measured value of fti to tk. The Distance among the feature sequences for any desired two locations is the Equation as

$$Dt_{y,r,t_{str.fini}} = dist_{seq}\ (S(l_y, f_{tar}, t_{str.fini}), S(l_r, f_{tar}, t_{str.fini})),\quad l_y, l_r \in Lt \quad y \neq r$$

to acquire the most related locations in the temporal region. The target sequence prediction f_{tar} and the necessary parameters passed into this. The most related locations are related strongly, and the spatial, temporal relation is Equation as,

$$STRS_{candi(lt\text{i,k})} = STRS_{candi(lt\text{i,k})} \cup TRS_{candi(lt\text{i,k})},\quad lt\text{i} \in Lt$$

The two sequences minimize the scaling errors, and shifting among the line among the Distance is estimated by the DTW algorithm. The sequence distortion occurs by the Euclidian space, which the DTW rectifies and maps the data points to the appropriate intervals. The Distance is chronological, and The robust temporal time series are identified by the kNN-DTWD Method. In contrast, the top is elected as a candidate to spot the sequence of the target location. These approaches cannot estimate the similarity among the time series with the missing information. The most specific interval and issue deal with the short

filter method that uses the lt min threshold, and these values are regenerated as meaningful unit DTW. Finally, the candidate item for the predictive model is used and established by the Algorithm kNN-DTW

The temporal relationship extractor (TRE) utilizes the historical target location features that forecast the imminent time series. The input data of PM2.5, PM2.10, and other concentrations are coherent and continuous, categorized as low frequency and rapid changes that are high frequency of data. Since the LSTM formulates historical time series

behavior, TRE LSTM acquires the trends in the target location. The current data used by the Artificial Neural Network (ANN) and is susceptible to quick alterations. Thus, the ANN acquires the high frequency of data and low frequency of information reached by the LSTM, and the ANN receives the high frequency of data. The relevancy of the feature is diverse concerning the quality of air is learned via previous studies

The air quality in specific locations is influenced by pollutant dispersal, and the emission is positively affected by the surrounding areas. Therefore, a new approach predicts the target location's air quality based on meteorological information and AQIs. The sensitivity is improved byintroducing the ANN SRE approach and considering the neighbourhood data. The improved, sparse auto encoder effectively handles the noise information using the feed-forward neural network (FFNN).

The relationship among the location is diverse due to the numerous altitude variation andbarriers. The information related to the terrain includes enriching the site's Correlation with the terrain extractor (TE). The Correlation among particulate matter and terrain determinewhereby the elevation of every point normalize as,

$$El_s = \frac{ele - ele_{st}}{ele_{st}}$$

and the transformation to the relevant elevation is given as,

$$ele_{rel} = \frac{1}{e^{El_s}}$$

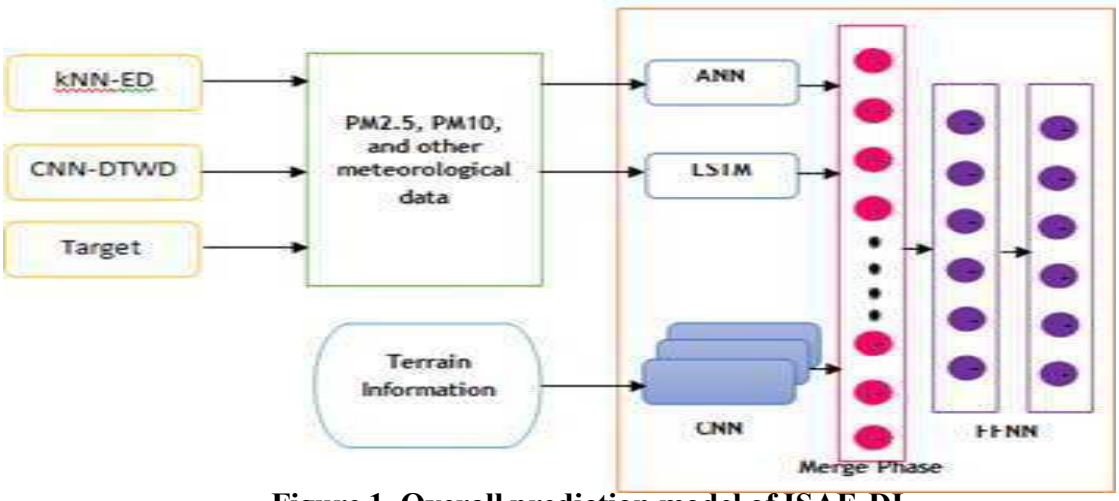


Figure 1. Overall prediction model of ISAE-DL

The current data is used by Artificial Neural Network (ANN), and it is susceptible to sudden changes. LSTM and ANN, respectively, obtain the low frequency and high frequency of data. The relevancy of the feature is diverse, for the quality of air is learned. An improved sparse autoencoder effectively handles the noise information using the Feed Forward Neural Network (FFNN). The terrain information is included to enrich the Terrain Extractor (TE) correlation. The TE, SRE and TRE results are combined and transferred to the improved sparse autoencoder-based FFNN for air quality prediction. The overall framework of the EISAE-DL is shown in Figure 4.

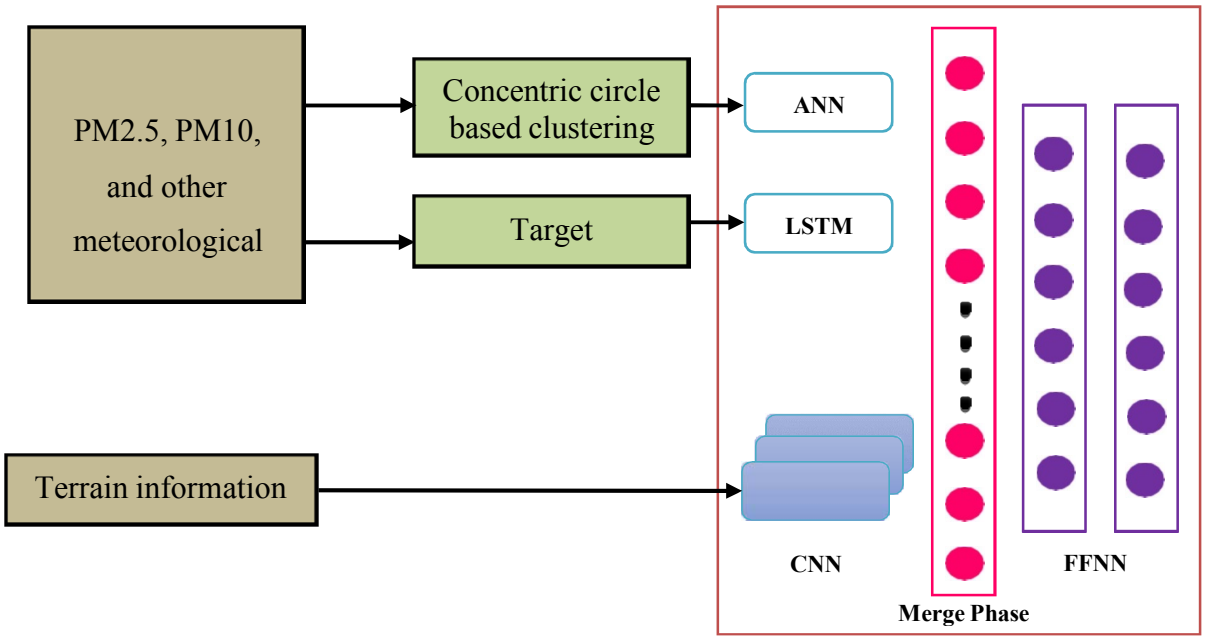


Figure.2 Overall framework of EISAE-DL

Result and Discussion

Performance Metrics Evaluation

The prediction accuracy and the classification performance of the proposed ISAE-DL and the existing ST-DNN are evaluated and shown.

Accuracy

Accuracy postulates the nearness of a specific value from the categorized instances. The accuracy depicts systematic errors and statistical bias. It estimates the actual value and identifies (TP and TN values) amongst the estimated classes' count. The occurrence of the least accuracy causes variation among the resultant and actual resultant values. It is the ratio of accurate disease detection over the total amount of instances evaluated [22]. It computes as,

$$Acy = \frac{TruePositive+TrueNegative}{TruePositive+TrueNegative+FalsePositive+FalseNegative}$$

Precision

The positive analytical value or Precision denotes the measurement's closeness and relevance among the values identified. The random errors that determine the statistical variables are the state as Precision. The importance of Precision and accuracy are similar terms. They are typical. They represent the value of Precision. Binary or decimal digits are used in the system. It is measured based on True Positive (TP) and False Positive (FP) rates. The value of Precision directly relies on the per cent of positive values in the total population. In the classification process, the precision value for a particular problem is the count of the valid positive values (i.e., the count of the item correctly labelled as positive classes). The high precision algorithm signifies that the resultant value achieves more needed information than irrelevant information [22]. It calculates as

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Table 1, Compares the outcomes of the proposed approach ISAE-DL and the existing approach ST- DNN. In this comparison, the algorithms' accuracy and precision values are shown.

Iteration	Accuracy		Precision	
	ST-DNN	ISAE-DL	ST-DNN	ISAE-DL
50	0.9	0.93	0.8867	0.914
100	0.91	0.96	0.8871	0.916
150	0.95	0.97	0.8873	0.919
200	0.97	0.99	0.8875	0.921

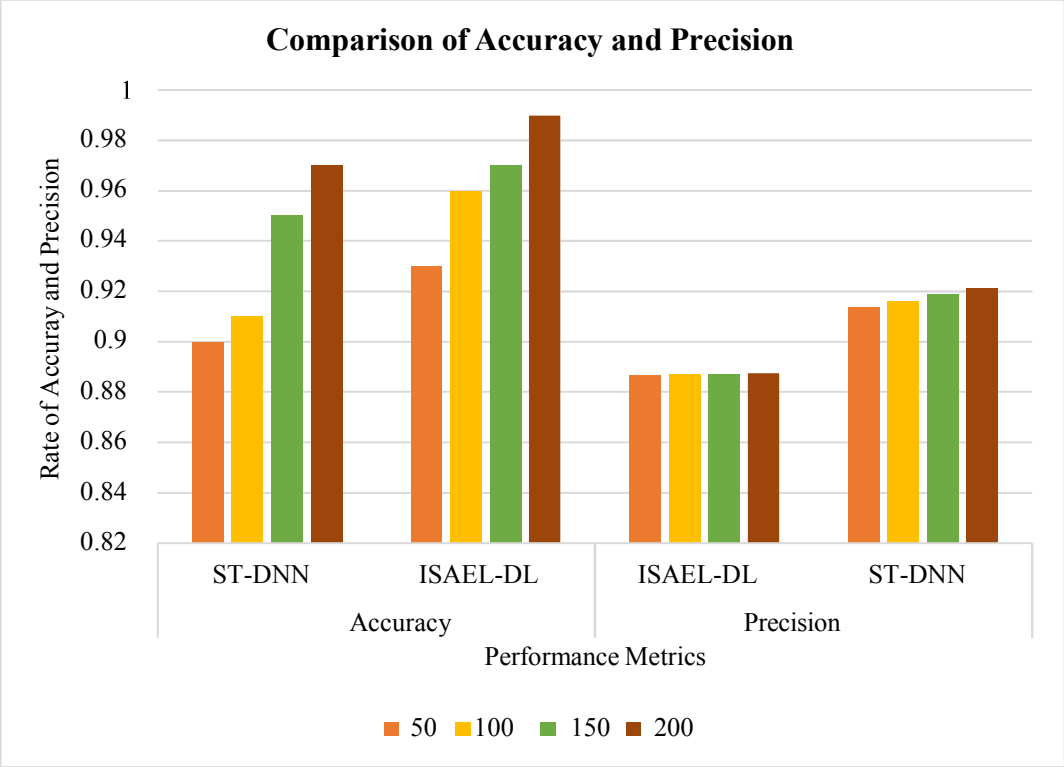


Figure 3. Comparison of Accuracy and Precision

Figure 3 identifies that the prediction accuracy of the proposed approach ISAE-DL is high compared to the existing Method. The prediction accuracy is improved by 0.36%, and the removal of noise information effectively attains it. The proposed approach's precision rate is more effective than the existing Method.

Table 2 shows the comparison between ST-SVR, ISAE-DL, and EISAE-DL in terms of accuracy and Precision.

Table.2 Evaluation of EISAE-DL in terms of Accuracy and Precision

Iteration	Accuracy			Precision		
	ST-SVR	ISAE-DL	EISAE-DL	ST-SVR	ISAE-DL	EISAE-DL
50	0.85	0.93	0.95	0.832	0.914	0.923
100	0.88	0.96	0.98	0.836	0.916	0.926
150	0.90	0.97	0.984	0.84	0.919	0.931
180	0.93	0.98	0.99	0.842	0.92	0.935

Figure 4. Comparison of ST-SVR, ISAE-DL, and EISAE-DL in terms of (a)Accuracy (b) Precision

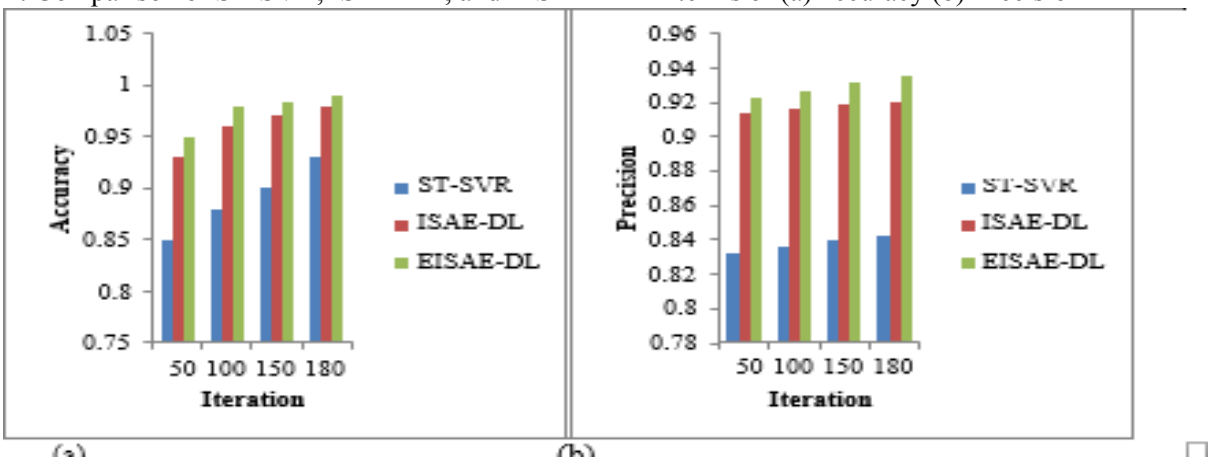


Figure 4a (b) shows the accuracy (Precision) of ST-SVR, ISAE-DL, and EISAE-DL for the different number of iteration. The X-axis denotes the iteration number, and the Y-axis denotes the accuracy (Precision). When the number of iteration is 100, the accuracy (Precision) of EISAE-DL is 11.36% (10.77%) and 2.08% (1.09%) than ST-SVR and ISADE-DL for air quality prediction. This analysis proves that the proposed EISAE-DL has high accuracy and Precision than state-of-the-art methods for air quality prediction.

CONCLUSION

The proposed ISA-DL outperforms the existing ST-DNN. Typically, the proposed ISA-DLapplies to other pollutant forecasting approaches, and it devises proficient solutions to minimize pollution. Future It can improve and handle the enormous data accumulation and generation and pollution propagation. The experimental results prove that the proposed EISAE-DL has better accuracy, Precision, specificity, sensitivity, AUC, and ECC than other methods for air quality prediction.

References:

(1) Abraham, S., & Li, X. Elsevier , 34, (2014) 165 – 171
DOI: https://10.1016/j.procs.2014.07.090

(2) Li, M., Wang, W. L., Wang, Z. Y., &Xue, Y Elsevier ,137,(2018) 11-17
DOI: https://doi.org/10.1016/j.buildenv.2018.03.058

(3). Xing, Y. F., Xu, Y. H., Shi, M. H., &Lian, Y. X.. Journal of thoracic disease, *8*(1),2016 ,E69-74. doi: https://10.3978/j.issn.2072-1439.2016.01.19

(4). Kumar, S., &Jasuja, A. IEEE,International Conference on Computing, Communication andAutomation (ICCCA) ,pp. 1341- 1346 ,(2017)
DOI:https://10.1109/CCAA.2017.8230005

(5) Rosero-Montalvo, P. D., Caraguay-Procel, J. A., Jaramillo, E. D., Michilena-Calderón, J. M., Umaquinga- Criollo, A. C., Mediavilla-Valverde, M., ... &Peluffo, D.H.. International Conference on Information Systems and Computer Science (INCISCOS) pp. 75-80, (2018). DOI:https://10.1109/INCISCOS.2018.00019

(6).Wu, Y. C., Shiledar, A., Li, Y. C., Wong, J., Feng, S., Chen, X., ..&Ozcan, A. Light Science Application (2017) Sep 8;6(9):e17046.DOI: https://doi.org/10.1038/lsa.2017.46.

Air Quality Prediction using Artificial Intelligence

P. ShreeNandhini, P. Amudha, S. Sivakumari

Abstract*: Due to the critical impacts of air pollution, prediction and monitoring of air quality in urban areas are essential tasks. However, because of the dynamic nature and high Spatio-temporal variability, the prediction of the air pollutant concentrations is a complex Spatio-temporal problem. The data is collected in specific area such as climate condition and vehicular pollutant occurring in the peak hours. the predication process is used to compare the algorithm artificial neural network and support vector machine process. This paper presents a survey on Air quality prediction using artificial intelligence

Keywords: Spatio-temporal, Ambient air pollution, Support vector machine,

I. INTRODUCTION

Due to human activities air pollution is more in the surrounding environment and cause many health issues the weather forecasting changed day by day process due to air pollutant s and damage the atmosphere process. The predication of pollutant brings together a monitor system of air pollutant and It makes use of global meteorological and crop data such as type of agriculture, country, fertilization, etc. Before data modeling, during data preprocessing, several measures were taken into consideration to improve the overall performance of process including the removal of extreme values in the dataset, exclusion of boreal ecosystem ranking of essential features in the dataset; and supply of controlled number input variables. For a fair assessment, the linear and nonlinear regression models are the results.

The result shows that the accuracy performance of the RF-based model was 20-24% higher than that of both regression models. Whereas another RF-based model It is used to predict the AQI by using urban public data based on road information, air quality and meteorological datasets of all the regions of Shenyang, China. Over proficient assessment of the model, it was verified that RF with its highest precision (R2 = 0.81) and lowest error value (RAE = 36.9%) had outperformed the state of the art classifiers such as Naïve Bayes, Logistic Regression, single decision tree and ANN.A simplified regression technique based on Quito Ecuador.

II. LITERATURE SURVEY

Y. Jiang, K .et.al [1] Proposed the crowd sensing-based method [1] is a predicate air pollutant such as vehicular area in the pollutant The period of sensing air pollutant is hard .and the predication of air quality take more time of process.

Grange et al. (2018) Proposed RF as a classification technique to build a model by using meteorological conditions, atmospheric pollutant data, and temporal factors to analyze the trend of the particular matter of air pollution as well to make a long term prediction. The methodology of the study involves the creation and training of many out of the bag (OOB) samples to grow Decision Tree, which later was combined to make a prediction. The study makes use of daily input variables named: wind speed and direction, threshold and time, and synoptic-scale, were obtained from 31 monitoring stations of Switzerland for 20 years. The average accuracy calculated in terms of correlation coefficient (R2) at all 31 stations is to be 0.62, with wind speed and boundary layer being the best input predictors and synoptic-scale the worst. The prediction performance obtained under the model at different locations varied from R2 = 0.53 to R2 = 0.71. The lower values were important recorded at stations located near rural mountainous areas.

Diez et al. [3] proposed the method of statistical approach for predicate the air pollutant in the international airport the fast and accurate method is analyzed for observed data analytical methods. for reducing air quality issue in aircraft operation in airport and hard to find the actual operation.

K. D. Kuhn[4] Proposed the machine learning method for predicate international airport. The Air quality index is a substitute for particular pollutant. The international civil aviation configuration and actual performance for meteorological data The developing and predicate the unsupervised learning method is the best model and accuracy.

Li et al. [5] Proposed the weather forecasting datasets. The first is the influence of meteorological conditions such as atmospheric currents, and the second is the change in the average astigmatism coefficient caused by the absorption and scattering of light due to fine particles and significant air pollutants.

Xiaoyu.et.al[6][7]Proposed the Comprehensive Pollution Index Method determines the air quality level on based the calculated pollution index value and finally calculated the average level o air pollutant and the damage level of a single contaminant in

Revised Manuscript Received on February 06, 2020.
***** Correspondence Author
P.ShreeNandhini, Research scholar, Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, School of Engineering, Coimbatore, India.
Dr.P.Amudha,Associate Professor Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, School of Engineering, Coimbatore, India.
Dr.S.Sivakumari, Professor and Head, Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, School of Engineering, Coimbatore, India

the calculation, which is simple and easy to conduct. The air pollution is increasing due to weather, and wind that should be cut off value and finally evaluate the result. Meanwhile, the calculation result depends on the ratio of the observed highest pollutant concentration value to the corresponding standard cost, which would result in a higher Comprehensive Pollution Index if the perceived amount of a particular pollutant is relatively higher.

The whistleblowers proposed the method air pollution is a direct victim in the environment protection. For long term process air pollutant is affected by whistleblowers due to issue in air pollution exposure and finally report the process of whistleblowers.

Website of Urban Air [11], is an icon on the map stands for a checking and monitoring all area in the station. The number consider with image air quality index and the number lesser than and better than air quality The color of an image is denoted by air quality, e.g., "green" means a "good" and "yellow" means “moderate" by Chinese AQI standards. The time line is used to analyze air quality prediction particular time interval.

K. V. Hamilton Ron .et.al due to risk in the environment problem cause many health diseases.in the urban air quality process. Such as respiratory diseases and lung cancer and asthma and many issue in the building tools [12].

R.C.Abernethy .et.al. Proposed the method of environment pollution prediction. Day by day pollution level are changed because pollutant is caused by vehicular areas in the meteorological data [13–14]. The environment data is many datasets are used for statistical tool. The multivariate analysis methods are used many methods for high amount of data.

L. V. Perez-Arribas .et.al Proposed the Concept of clustering data process in the weather forecasting data such as wind direction in the hours and monthly data. The environment data is dependent upon another process data. Different parameter is used to multidimensional analysis in the same time end with many Pollutant. [15].

III. COMPARATIVE ANALYSIS OF METHODOLOGY

Ong et al. [16] proposed the method of time series air quality prediction tasks for automatic encoder pre-training method. The VENUS system accuracy goes beyond the training process. The sites of pollutant avoid the previous machine learning process. There is no input variable detecting and cause impact of air quality prediction

Cho K, Merrienboer [17] proposed the method of recurrent neural network (RNN) deals with spatial –temporal data are used to analyses the air quality process. The recurrent neural network model of GRU in the encoder decoder process to overcome the problem of overfitting data in the GRU and used less parameter it provides more predicated result and accuracy is good [17].

Kim PS [18] Tang Y, HuangY[19] Wang Y[20] proposed the method of prediction air quality process. Hours data predication is calculated in the first stage process the prediction data in the meteorological condition Due to forecasting sequence learning method are known in the deep learning technique. The real time point value. Denoted in the point process.

Several health problem is caused by air pollutant such as asthma dying people in the world.it performs European has performs thirty less than Percentage pollution. Due to respiratory infection cause 13 %percentage death because of increasing vehicular pollutant.

The spatial-temporal feature proposed the method of air quality prediction is the variable of internal tasks. It focused on method of time-series data are used to design for by air pollutant interactions and transmissions as well.

Due to climate process and forecast the process is a framework for improve accuracy prediction.

Due to forecast air quality data increases is used to improve to predict data in air quality pollution Air planning is a method is used to develop the forecasting. [8][9][10].

The real-time emission data, is an input, for decision-makers to analyze the performance of air planning decisions.

The tree model is a decision tree is used to represent branch node and represent decision tree.fig1.the main node describe the process and end node are leaves.

In S. Deleawe et al. proposed the method decision tree model for classification process and select the predict the values in the air, and air quality class are used to the attribute process. The confidence factor is 0.25 are used in decision tree process. Implementation of process is Weka algorithm.

Besides a fair performance during regular days, it was interesting to see that improved results during the weather conditions. The concept of lazy learning technique captures the difference PM10 emissions and AQI Hourly dataset of SO2, NOx, CO, PM10, and Ammonia (NH3) for one year period from an area called Lombard in Italy

(Carnevale et al., 2016). Other specifications of the study include the application of the Dijkstra algorithm for large scale data processing. The ration between data 80% and 20% is used for training and validation purposes, respectively. The validation phase results and compared with deterministic models instead of comparing it with the targeted values. The results obtainable through this approach in the form of R2were nearly the same as achieved by the Transport Chemical Aerosol Model (TCAM), i.e., R2=0.99. By the way, TCAM is a costly computational method commonly used in decision making.

Bougoudis et al. (2016) built a hybrid system by using ANN, RF, fuzzy logic along with unsupervised clustering for the prediction of various pollutants. The experiment conducted makes use of 12 years' hourly dataset of air pollutants the performance of R2 model is evaluated using Unsupervised clustering Mamdani proposed the method of fuzzy logic system was applied to enquire about factors affecting the quality of air. The model performed well to estimate the concentration of CO and NO with R2 =0.95 when FIS ensemble with RF and ANN respectively while using RF alone it could predict NO and O3 with accuracy equal to 0.91

MartínezEspaña et al. (2018) conducted work in Murcia, Spain. The study aimed at the prediction accuracy of ground-level O3 by using five different classification models, namely Bagging, RF, Decision Tree, k-Nearest Neighbor, and Random Committee. The work applied two years' air pollution dataset of NO, NOx, PM10, O3, SO2, C6H6, C7H8, XIL, and environmental parameters such as pressure, solar radiations, temperature, relative humidity, wind speed, and direction. Two phases in the experimental design. At first, the prediction accuracy of all the models was tested out and compared against each other, whereas to study the number of models required for O3 modeling in the Murcia region, during the second phase, the work adopted a hierarchical clustering approach. Results obtained in terms of the correlation coefficient are as follow: the performance of RF has been superior, having an R2 value equal to 0.85 as compared to Random Committee (0.83), Bagging (0.82), and Decision Tree (0.82). Among the five models, CNN performed the worst (0.78), whereas NOx, threshold, climate condition is the top predictors. In the end, using a clustering approach suggested that the study region only requires two models for accurate modeling of O3.

Sayegh et al., (2014) aimed at capturing the variability of PM10 by employing several statistical and machine learning models such as Boosted Regression Tree (BRT), Generalized Additive Model, Linear, and Quantile Regression models (QRMs). The (observed/ predicted values) is the performance of the coefficient of determination. Considering the role of different quantiles instead of the central tendency of PM10, the QRM had performed better as compared to other data mining tools to predict the hourly concentrations PM10.

Nieto et al. (2015) conducted research using different air pollutant dataset. is collected. The experimental setup employs MLP and Multivariate Adaptive Regression Splines as modeling tools. The proficient assessment of modeling results inferred that the estimation of NO2 (R2 = 0.85) has been pretty good, followed by a slightly lower accuracy obtained for predicate parameter. Meanwhile, to predict the concentration of a greenhouse gas N2O, another study adopts RF algorithm

Kleine Deters et al. (2017) Proposed method predictions used to understand the effects of meteorological factors on the precise prediction of PM2.5. During collection of metrological data training and tested datasets it performance of the technique cross-validation and coefficient of determination R2. The meteorological parameter is a best choice of predicating air quality.

Eldakhly et al. (2018) to make 1 hour ahead forecast of PM10 concentrations. The study uses fuzzy logic and chance weight value to handle fuzziness of the data; the target minimizes an outliner point in the training process. The different results obtained are one of the rare examples, as the study establishes that the proposed approach can outperform ensemble learning algorithms.

Xu et al. (2017). It proposed to estimating ozone profile shapes. In this work, the author developed a KNN based algorithm, namely Physics Inverse Learning Machine. The working principle of the model involves five significant steps: 1) application of k-means clustering to group different ozone profiles based on its concentration values; 2) generation of simulated UV spectra from each cluster of respective ozone profiles; 3) input predictive feature selection by using PCA to enhance classification effectiveness; 4) spectrum of uv rays are classification models (5) the shape of ozone process are fit for ozone column. The predicted and observed values in the tested. Results obtained indicate that, altogether, a total of 11 clusters and estimation error lower than 10%.different technique based on less reliable dense mobile sensors data collected from sparse monitoring stations a fine granularity (Hu et al., 2017). For this Sydney, Australia based study, seven regression models are method to predict the concentration of CO. Before modeling and validation, and the study involves three main steps. During first regression models is used ten years of data, including previous years off data is collected. Comparison and validation in the Proficient model during the second and third steps, respectively. Out of seven modeling techniques, Support Vector Machine for regression (SVR), RF, and Decision Tree regression achieved the best results, with SVR having the highest spatial resolution and precise demarcation of pollution boundaries as compared to other models.

IV. RESULT OF LITERATURE REVIEW

In this Section Process of Literature Review shown in this Table1

Table 1. Process of Literature Review

Author Name	Machine learning method
Y. Jiang, K .et.al [1]	Crowd based sensing method
Grange et al[2]	Decision Tree
Diez et al. [3]	Statistical approach
Ong et al. [16]	Venus system accuracy
Cho K, Merrienboer [17]	Recurrent neural network
Eldakhly et al. (2018)	Fuzzy logic
Sayegh et al., (2014)	Boosted Regression Tree (BRT), Generalized Additive Model
Xu et al.	KNN based algorithm

V. CONCLUSION

Here the paper survey the air quality prediction using artificial intelligence method, and compares present and existing research work on air quality evaluation based on machine learning algorithm and techniques.

REFERENCES

1. Y. Jiang, K. Li, L. Tian, R. Piedrahita, X. Yun, O. Mansata, Q. Lv, R. P. Dick, M. Hannigan, and L. Shang. Maqs: A personalized mobile sensing system for indoor air quality. In Proc. of UbiComp 2011.

2. Grace .Et.Al Machine Learning Algorithm Using Air Quality Modeling 2018
3. D.M.Diez,F.Dominici,D.Zarubiak,andJ.I.Levy,"Statically approaches for identifying air pollutant mixtures associated with aircraft departures at Los Angeles international airport," Environmental Science& Technology, vol.46, no.15, pp.8229– 8235,2012
4. K. D. Kuhn, "A methodology for identifying similar days in air trafficflowmanagementinitiativeplanning,"TransportationResearchPartC:EmergingTechnologies,vol.69,pp.1–15,2016.
5. Li, H.; Tan, X.; Guo, J.; Zhu, K.; Huang, C. Study on an Implementation Scheme of Synergistic Emission Reduction of CO2 and Air Pollutants in China's Steel Industry. Sustainability 2019, 11, 352.
6. Xiaoyu, L.; Jun, Y.; Pengfei, S. Structure and Application of a New Comprehensive Environmental Pollution Index. Energy Procedia 2011, 5, 1049–1054.
7. Ga¸siorek, M.; Kowalska, J.; Mazurek, R.; Paja¸k, M. Comprehensive assessment of heavy metal pollution into soil of historical urban parkonan example of the Planty ParkinKrakow(Poland). Chemosphere2017,179, 148–158
8. \Ministry of Environmental Protection of the People's Republic of China. The Ministry of Environmental Protection Will Open the "010-12369" Environmental Whistleblowing Hotline on the World EnvironmentDay(5June).Availableonline:http://www.mee.gov.cn/gkml/sthjbgw/qt/200910/t20091023_179602.htm (accessed on 17 November 2018).
9. \Ministry of Environmental Protection of the People's Republic of China. The "12369 Environment Whistleblowing Platform" on WeChat Has Been Working Well Since Its Opening in 2015. Availableonline:http://www.mee.gov.cn/gkml/sthjbgw/qt/201603/t20160322_334063.htm (accessed on 17 November 2018).
10. Environmental Emergency and Accident Investigation Center of the Ministry of Environmental Protection. 12369: Promoting Environmental Information Disclosure and Public Participation. World Environ. 2017, 4, 76–77.
11. Urban Air Website: http://urbanair.msra.cn/
12. K. V. Hamilton Ron, T. Johan, and J. Watt, The Eects of Air PollutiononCulturalHeritage,Springer-Verlag,Boston,MA, USA, 2009.
13. R.C.Abernethy,R.W.Allen,I.G.McKendry,andM.Brauer,"A land use regression model for ultrafine particles in vancouver, Canada," Environmental Science and Technology, vol. 47, no. 10, pp. 5217–5225, 2013.
14. M.Rivera,X.Basagaña,I.Aguileraetal.,"Spatial distribution of ultrafine particles in urban settings: a land useregressionmodel,"AtmosphericEnvironment,vol.54,pp.657–666,2012.
15. L. V. P´erez-Arribas, M. E. Le´on-Gonz´alez, and N. RosalesConrado, "Learning principal component analysis by using data from air quality networks," Journal of Chemical Education, vol. 94, no. 3, pp. 458–464, 2017.
16. Ong B T, Sugiura K, and Zettsu K, "Dynamic pretraining of Deep Recurrent Neural Networks for predicting environmental monitoring data," IEEE International Conference on Big Data. IEEE, 2015, 760765.
17. Cho K, Merrienboer B V, Gulcehre C, Bahdanau D, Bougares F, and Schwenk H, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," Computer Science, 2014.
18. Kim P S, Lee D G, and Lee S W, "Discriminative Context Learning with Gated Recurrent Unit for Group Activity Recognition," Pattern Recognition, 2017, 76.
19. Tang Y, Huang Y, Wu Z, Meng H, Xu M, and Cai L, "Question detection from acoustic features using recurrent neural network with gated recurrent unit," IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2016:6125-6129
20. Wang Y, Liu M, Bao Z, and Zhang S, "Short-Term Load Forecasting with Multi-Source Data Using Gated Recurrent Unit Neural Networks", Energies, 2018.

AUTHORS PROFILE

P.ShreeNandhini,Research scholar, Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, School of Engineering, Coimbatore, India.

Dr.P.Amudha,Associate Professor , Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, School of Engineering, Coimbatore, India.

Dr.S.Sivakumari, Professor and Head, Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, School of Engineering, Coimbatore, India