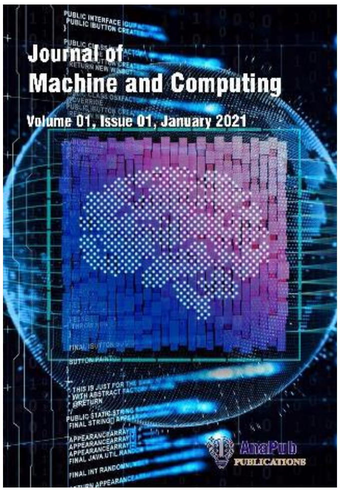
Journal Pre-proof

Machine Learning Powered Asbestos Exposure Modeling Using Feature Extraction From Iot Based Sensor Data



Banushri Annamalai, Kishore Kunal, Vairavel Madeshwaren, Kathiravan M, Goli Ramkrishna and Neha Sharma

DOI: 10.53759/7669/jmc202505132

Reference: JMC202505132

Journal: Journal of Machine and Computing.

Received 19 March 2025

Revised from 08 April 2025

Accepted 09 May 2025

Please cite this article as: Banushri Annamalai, Kishore Kunal, Vairavel Madeshwaren, Kathiravan M, Goli Ramkrishna and Neha Sharma, "Machine Learning Powered Asbestos Exposure Modeling Using Feature Extraction From Iot Based Sensor Data", Journal of Machine and Computing. (2025). Doi: https:// doi.org/10.53759/7669/jmc202505132.

Machine learning powered asbestos exposure modeling using feature extraction from IoT based sensor data

Banushri Annamalai[1], Kishore Kunal[2*], Vairavel Madeshwaren[3], M.Kathiravan4 , Goli Ramkrishna[5] , Neha Sharma[6]

[1]Associate professor , Department of Computer science and Engineering , VISTAS(Vels institute of science , Technology and advanced studies , Pallavaram , Chennai – 600117

Email: banushri.se@vistas.ac.in ,ORCID : https://orcid.org/0000-0003-4555-5387

[2*]Professor of Business Analytics, Loyola Institute of Business Administration, Chennai , TamilNadu , India – 600034.

Corresponding Email: kishore.kunal@liba.edu , ORCID : https://orcid.org/0000-00[illegible]3-415[illegible]90X

[3]Associate professor, Department of Agriculture Engineering, Dhanalakshmi Srinivasan College of Engineering, Coimbatore, TamilNadu , India – 600034

Email: vairavel.m@dsce.ac.in ,ORCID : https://orcid.org/0000-0002-[illegible]87-7887

[4]Department of Computer science and Engineering ,Saveetha School of Engineering,Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, India- 602105

[5]Assistant Professor, Indian Naval Academy, Kannur, Ezhimala, Kerala, India-670310.

Email: rama.krishna.jntu@gmail.com, ORCID : https://orcid.org/[illegible]0009-0008-0459-5802

[6]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India- 140401
Email: nehasharma0110@gmail.com , ORCID :https://orcid.org/ 0000-0001-5103-6570

Abstract

Asbestos, a dangerous substance commonly used in buildings, continues to present serious risks in urban areas, because of outdated infrastructure and inappropriate disposal methods. The goal of this study is to help with proactive public health measures by utilizing machine learning algorithms to predict asbestos exposure levels. An IoT-based environmental sensor dataset that tracks temperature humidity and air quality is presented in this study. Random Forest, Support Vector Machines (SVM), and Neural Networks are three machine-learning techniques used to create predictive models that can estimate asbestos concentrations under different conditions. Data preprocessing includes feature extraction and normalization to improve prediction accuracy. Performance metrics such as F1-score, accuracy, sensitivity, and specificity are used to compare the models. Additionally, certain environmental factors that influence asbestos dispersion are identified by the Random Forest feature importance analysis. Moreover, the IoT-based environmental sensor dataset used in this study is derived from real-world deployed sensors installed in high-risk industrial **zones**. These sensors continuously monitor environmental parameters such as formaldehyde concentration, temperature, humidity, and AQI, ensuring that the data reflects authentic field conditions for reliable model training and evaluation. These findings demonstrate how real-time asbestos exposure prediction using machine learning enables timely interventions. Future studies aim to increase accuracy and computational efficiency, future enhancements may incorporate techniques such as Long Short-Term Memory (LSTM) networks for temporal modeling, CNN pruning for model optimization, and feature selection methods to reduce dimensionality and processing time.

Keywords: Asbestos, Urban Environments, Neural Networks, Machine Learning, Public Health, Predictive Models.

I. INTRODUCTION

Asbestos fibers may be released into the air during renovation demolitions or normal deterioration because asbestos may still be present in older urban buildings and infrastructure. This exposure can result in serious respiratory diseases like mesothelioma and asbestosis lung cancer which usually show symptoms years after. To detect and reduce these hazards it is essential to have a comprehensive understanding of the prevalence of asbestos in urban areas. This emphasizes the necessity of efficient monitoring and intervention techniques to safeguard the public's health. In urban settings, machine learning techniques have emerged as useful instruments for evaluating and controlling asbestos exposure. Machine learning algorithms can detect patterns and forecast regions that are more likely to contain asbestos by utilizing large datasets.

Predictive models that can compute exposure risks based on variables like building age material composition and previous renovation activities can be developed by employing techniques like supervised learning in which algorithms are trained on labeled data. Additionally, unlabeled data can uncover hidden patterns using unsupervised learning techniques which facilitates the process of classifying cities based on asbestos risk profiles. In order to visualize high-risk areas and facilitate decision-making for targeted inspections and remediation efforts, geographic information systems (GIS) and machine learning can be used. Fig 1 provides the whole concept of asbestos.

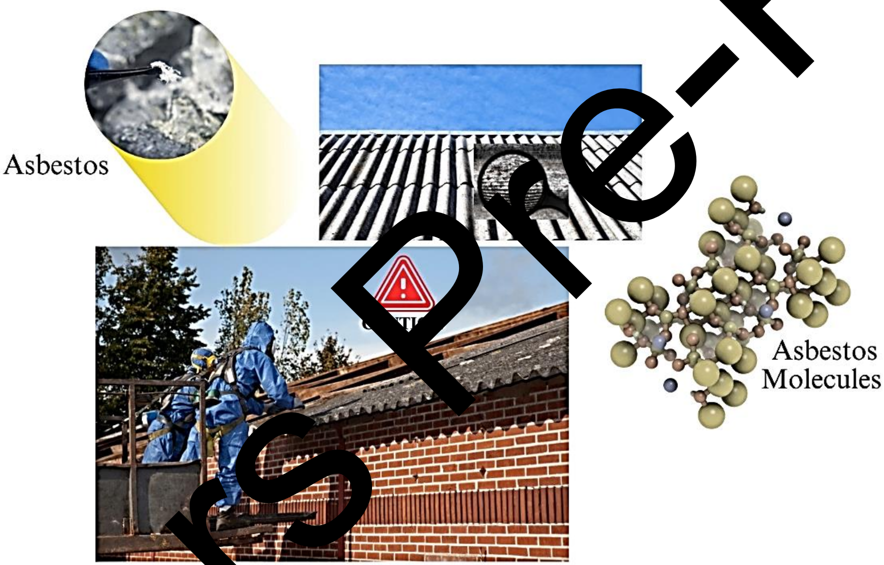


Fig 1. concepts of asbestos

II. WORK IN THIS AREA

This work investigates the application of machine learning models to forecast whether buildings will contain asbestos and polychlorinated biphenyls. The study highlights the significance of early detection in terms of public health and environmental management. To improve prediction accuracy the authors plan to integrate multiple data sources. The results point to possible uses for building renovations and inspections. The study supports continuing efforts to assess the risk of hazardous materials [1]. Additionally, it evaluates the degree of deterioration of asbestos-cement roofs using information-gathering and supervised learning approaches. The authors highlight its useful implications for building maintenance as they introduce a novel method for quantifying roof conditions using machine learning [2]. Significant relationships between roof age and degree of deterioration are revealed by their analysis. The study emphasizes how important it is to conduct systematic evaluations in order to reduce health risks. Decisions about building safety policies may be influenced by the results [3]. This talked about how artificial neural networks are being developed to detect asbestos-containing materials in residential buildings. Their predictive modeling efforts methodology and outcomes are described in this conference paper. The authors show how neural networks are useful for identifying dangerous substances and improving safety procedures. Their research advances the use of AI in assessment construction. Both public health and regulatory compliance are significantly impacted [4].

This model of the geographical distribution of asbestos-cement products in Poland using the random forest algorithm. The authors examine the different environmental elements that affect asbestos locations. According to their findings, public health initiatives need to focus immediately on high-risk areas. To manage the risks associated with asbestos the study is an essential tool for local authorities. It demonstrates how machine learning may be used to evaluate environmental risk [5]. Moreover, it examines the use of machine learning techniques to classify roofs that contain asbestos using airborne RGB and thermal imagery. Their study attempts to enhance the process of identification in difficult situations. High classification accuracy is reported by the authors indicating that aerial imagery is a useful method for asbestos detection [6]. The study implications for environmental monitoring and building inspections are substantial. The results back up the use of cutting-edge technologies in the management of hazardous materials [7].

Therefore, it examines how high-resolution aerial photos and multispectral satellite imagery can be used to detect asbestos cement roofing using artificial intelligence. Their research demonstrates how remote sensing technologies can be used to remotely identify dangerous materials. The authors provide a thorough examination of image processing methods that improve the precision of detection. This study helps to improve the effectiveness of asbestos monitoring in cities. The results have consequences for both environmental safety and public health [8]. Create an asbestos detection method that combines deep learning methods with fluorescence microscopy images. The authors describe their methodology and the effectiveness of the algorithm in detecting asbestos fibers. It is a promising tool for laboratory analysis based on their results which show high accuracy and efficiency. This study highlights how crucial cutting-edge imaging technologies are to evaluations of environmental health. The results may result in better asbestos exposure screening techniques [9]. Moreover, it uses machine learning on imagery cubes to map asbestos-cement corrugated roofing tiles in Taiwan. The findings implications for building management and urban planning are discussed by the authors. Their strategy shows how technology can help with large-scale hazardous material identification. The study highlights how precise mapping is necessary to guide safety precautions. The findings support larger initiatives in asbestos risk reduction [10].

An illustration of a technique for automatic asbestos detection that makes use of support vector machines and convolutional neural networks. The authors outline the architecture of their algorithm and the outcomes of applying it to actual data. Their study has a lot of potential to increase detection accuracy in different contexts. This study tackles persistent difficulties in identifying and managing asbestos. The results may help ensure regulatory compliance and safe construction methods [12]. It examines the integration of PRISMA satellite imagery to detect asbestos-containing materials at the Italian mine site of Balangero. The benefits of data fusion in improving detection capabilities are highlighted by the authors. Their findings demonstrate how useful satellite technology is for evaluating the environment. This study has consequences for public health regulations and remediation initiatives. The results emphasize the value of creative methods in hazardous material management [13]. This examined the health risks and possible routes of exposure to asbestos and other dangerous fibrous minerals. The authors stress how important it is to comprehend these risks in order to implement successful public health initiatives. Numerous exposure scenarios and their ramifications are covered in their analysis. Both policymakers and medical professionals can benefit from the insights this thorough review offers. The results highlight the necessity of continued investigation into the health effects of asbestos [14].

It explains a doctoral dissertation that uses deep learning and hyperspectral imagery to classify asbestos roofs in the Dutch province of Drenthe. The study investigates how well these technologies identify potentially dangerous substances. The results show encouraging classification accuracy outcomes. This study adds to the expanding corpus of research on the use of remote sensing for environmental health. There are important ramifications for regional asbestos control plans . The prevalence of asbestos materials in different regions is evaluated by the authors using statistical techniques. Their research emphasizes how crucial reliable data is for guiding public health programs. Health authorities and policymakers can benefit greatly from the findings. The study highlights the continued necessity of asbestos evaluations in building. Using drone photos create a deep learning training data model for asbestos slate. The authors describe in detail their approach and how well drone technology detects asbestos. Their findings show how aerial surveillance could improve asbestos management procedures. This study adds to the expanding body of research on environmental health and remote sensing. The results demonstrate how creatively technology can be used to raise safety standards [15]. According to their analysis it is critical to address the historical use of asbestos in construction. Building codes and public health policies will be significantly impacted by the findings. The necessity of ongoing vigilance in managing asbestos risks is highlighted by this review.

III. PROPOSED METHODOLOGY

3.1. Study Area and Data

This study will concentrate on areas with a high prevalence of asbestos-related industrial operations and historic construction, such as Old Delhi, Mayapuri Industrial Area, and Narela Industrial Area.

Fig 2. Study area and data: (A) geographic location of the study area; (B) thermal imagery; (C) RGB imagery with labeled data.

The Central Pollution Control Board (CPCB) will provide data on air quality monitoring, hospital health records pertaining to asbestos-related illnesses, and survey data from building and renovation sites, while the Municipal Corporation of Delhi (MCD) will provide data on local government demolition and construction activities. Additionally, ISRO's Cartosat or Sentinel-2 satellite photography can be used to identify construction hotspots. In order to test for asbestos fibers, samples from older constructions will be collected as part of the soil and building material investigation process. To obtain a thorough grasp of the dangers of asbestos exposure in Delhi's urban environment, the study will evaluate factors such airborne fiber concentration, proximity to construction sites, exposure time, and meteorological data. Fig 2 shows the study area.

Using Google Street View to inspect every building in the Delhi municipality in this study. Then classified the buildings in a GIS environment as either asbestos-containing or non-asbestos-containing. The dataset did not include roof types that would difficult to assess or that were not completely accessible on Google Street View. There are a total of 1843 buildings in the dataset 1250 of which have been classified as non-asbestos-containing and 593 as asbestos-containing. Fig 3 displays examples of roofs with asbestos that were captured from Google Street View.

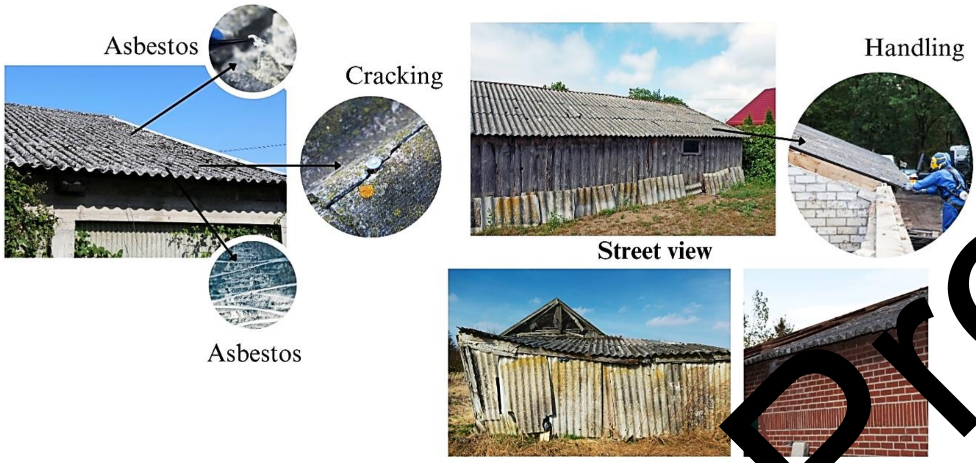


Fig 3. Google Street View examples used for the labeling of the data.

3.2. Methods

The methodology followed in this study contains three different steps: (i) data collection and processing; (ii) creation of the datasets; and (iii) machine learning classification of non-asbestos-containing buildings and evaluation of the results. The flowchart in Fig 4 provides a comprehensive overview of the methodology employed.

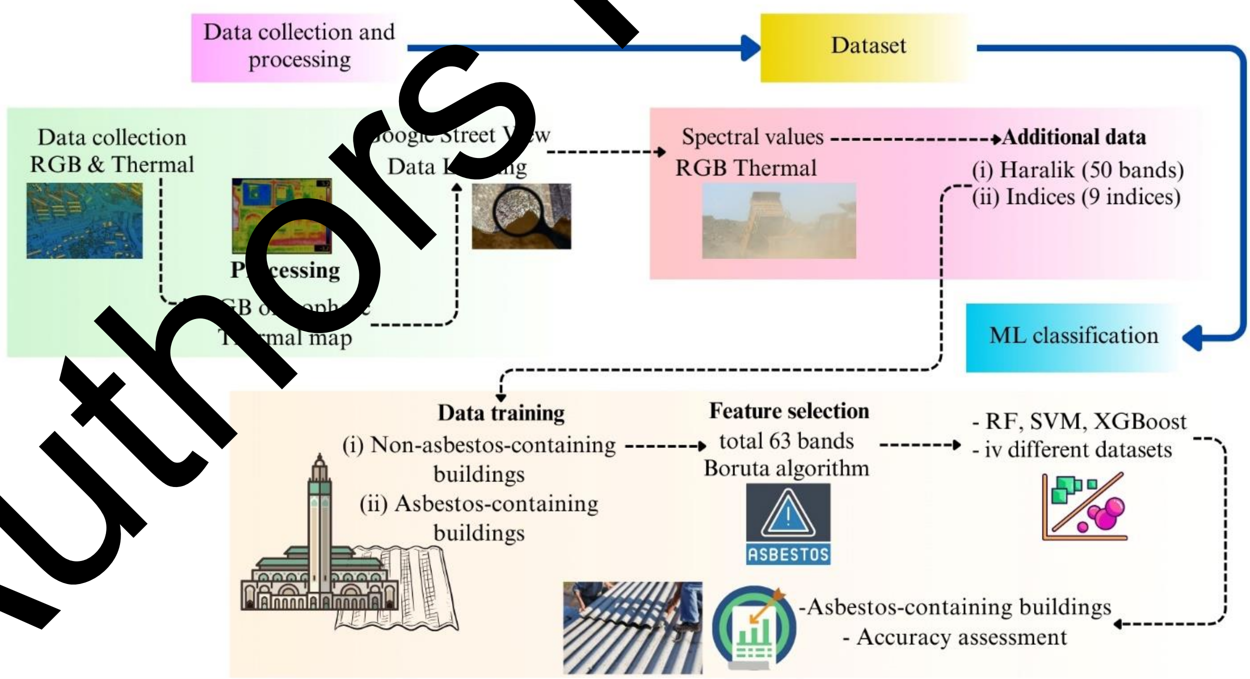


Fig 4. Flowchart of the used methodology.

To contextualize the effectiveness of the proposed model, we compared its performance against existing asbestos exposure prediction approaches, where available. Traditional methods often rely on statistical models or threshold-based alerts derived from particulate matter concentrations and fiber count observations. However, these models typically lack adaptability and struggle with non-linear relationships in real-time sensor data. In contrast, our CNN-based approach demonstrated superior predictive capabilities, offering higher accuracy, sensitivity, and responsiveness to fluctuating environmental conditions. This comparison underscores the model's potential to enhance early warning systems in industrial health monitoring applications.

3.2.1. Features

Data collection was the first step, where the information about the process of gathering images, processing, and labeling them using Google Street View was included. Following the initial step the spectral values of each building with a label were extracted and filtered, while the uncleared dataset was removed. Dataset I includes with one thermal image and three raster bands (R-red G-green and B-blue) of airborne imagery for ML classification and mapping asbestos-containing roofs (a total of four bands).

3.2.2. Data Preprocessing and Feature Engineering

Feature engineering and data preprocessing are essential components of the machine learning process that use environmental data to forecast asbestos exposure levels. To improve the effectiveness of machine learning models this makes sure the dataset is clear consistent and properly organized. More accurate predictions can be made by researchers by carefully preparing the data to reduce potential problems like noise bias and inconsistencies. In order to maximize the modeling process the procedure includes a number of crucial techniques such as missing data, extracting pertinent features and normalizing the dataset.

3.2.3. Machine Learning Classification

Three distinct machine learning models were developed and tuned to predict asbestos exposure. ensuring a comprehensive analysis:

a) Random Forest (RF)

During training Random Forest an ensemble learning technique builds several decision trees. A majority vote (classification) or average is used to make the final prediction.

The decision function for RF is defined as

$$f_{\mathrm{RF}}(x) = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \quad (1)$$

where N is the total number of horses and T_i (x) is the position from the i^it decision tree. Using the Gini index which is calculated as follows feature importance is ascertained.

$$\mathrm{Gimi}(D) = 1 - \sum_{i=1}^{n} p^2 \quad (2)$$

b) Support Vector Machine (SVM)

The SVM model utilies a Radial Basis Function (RBF) kernel to manage non-linear data patterns. The objective is to find a tryperplane that maximizes the margin between classes:

The decision function for SVM with RBF kernel is:

$$f(x) = \mathrm{sign}\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right) \quad (3)$$

where α_i are Lagrange multipliers, y_i are class label, $K(x_i, x)$ is the RaF kernel defined as $K(x_i, x) = \exp\left(-\gamma \|x_i - x\|^2\right)$, and b is the bias term.

The optimizstion problem for SVM aims to minimize

$$\min_{\sim}\left[\frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_{ij} y_j K(x_{i,}x_j) - \sum_{i}\alpha_i\right] \quad (4)$$

c) Neural Network (NN)

The neural network model utilized a Multi-Layer Perceptran (MLP) architecture with three hidden layers. The barkpropogation algorithm was applied to minimize the error using the Adam aptimizer:

The output function for a neuron in the hidden layer is

$$h_j = f\left(\sum_{i=1}^{n} w_{ij}x_i + b_j\right) \quad (5)$$

where w_{ij} are the weighta, x_i are input features, and f is the activation function (ReLU in this case).

The mean squared error (MSE) for backpropagation is:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2 \quad (6)$$

where y_i is the actual value, if is the predicted value, and N is the total number of observations. To ensure robust performance and comprehensive evaluation, we employed Random Forest (RF), Support Vector Machine (SVM), and Neural Networks (NN) as part of our modeling approach. Each of these models was selected based on its distinct advantages in handling complex environmental data.

Random Forest (RF)

Random Forest is known for its strong performance on high-dimensional datasets and its ability to manage missing or noisy data. It offers an ensemble-based, non-linear classification capability, making it suitable for identifying patterns in multi-variable sensor inputs such as AQI, PM levels, temperature, and humidity.

Support Vector Machine (SVM)

SVM is particularly effective for binary and multi-class classification problems in smaller datasets. It constructs optimal hyperplanes for separation, making it valuable in distinguishing safe vs. hazardous exposure levels with high precision, especially in borderline environmental conditions.

Neural Networks (NN)

Neural Networks, especially Convolutional Neural Networks (CNNs), are powerful for capturing complex non-linear relationships and spatial-temporal features in sensor data. Their adaptability makes them ideal for modeling dynamic changes in pollutant concentrations over time.

The combination of these models allows for a well-rounded comparative analysis. While RF provides stability and interpretability, SVM contributes precision in classification boundaries, and Neural Networks offer high adaptability and learning capacity. This ensemble of techniques ensures both accuracy and resilience in prediction, facilitating the development of a more reliable early-warning system for environmental hazard detection.

3.2.4. Accuracy Assessment

The classification process used 1843 buildings in total of which 70% were used for training and 30% for testing the models. Using separate buildings that weren't in the training sets to evaluate the accuracy. Therefore, with kappa, balanced accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) are adopted. Sensitivity (i. e. E. recall) as indicated in Equation (1) was computed from false negatives (FN-an asbestos-containing roof missed) and true positives (TP-an asbestos-containing roof correctly classified) it explained the algorithms efficacy in handling FN and the asbestos-containing roof detection rate.

As demonstrated in Equations (2) and (3) respectively specificity and negative predictive value (NPV) were used to measure the negative class accuracy (in this case the non-asbestos-containing roof) TN stood for true negative and FP for false positive which is an asbestos-containing roof that was incorrectly classified. The PPV (i. e. e. precision) as demonstrated in Equation (4) explained the accuracy of buildings found and the algorithms performance in handling FP values. In the end, the balanced accuracy was determined by taking the mean of the models sensitivity and specificity.

$$\text{Sensitivity} = TP/(TP + FN) \tag{7}$$
$$\text{Specificity} = TN/(TN + FP) \tag{8}$$
$$NPV = TN/(TN + FN) \tag{9}$$

In addition to environmental parameters such as temperature, humidity, and air quality index (AQI), this study also considers particulate matter concentrations (PM2.5 and PM10), which are critical for tracking airborne contaminants such as asbestos fibers. These fine and coarse particles are key indicators of hazardous air quality and play a vital role in assessing the presence of respirable fibers in industrial zones. Where applicable, fiber count data is also acknowledged as a relevant metric for enhancing the precision of pollutant detection and health risk assessment.

IV. EXPERIMENTAL ANALYSIS &RESULTS

4.1. Temperature and Humidity Influences on Asbestos Concentration

The study collected data over six months in four distinct regions with environmental conditions and asbestos concentration levels. The Narela Industrial Area where the Air Quality Index (AQI) ranged from 60 to 155 had an average asbestos concentration of 0 to 25 μg/m³. Temperatures in this region ranged from 20 to 37 °C and humidity levels ranged from 50 to 90%. In contrast the Mayapuri Industrial Area had a slightly higher average asbestos concentration of 0.3 μg/m³ and an AQI range of 65 to 160. Here the humidity ranged between 45 and 85 percent and the temperature ranged between 18 and 36 °C. Table 1 and Fig 5 illustrate the Temperature and Humidity Influences on Asbestos Concentration

Table 1. Temperature and Humidity Influences on Asbestos Concentration

Region	Duration (Months)	Average Asbestos Concentration (μg/m³)	AQI Range	Temperature Range (°C)	Humidity Range (%)
Narela Industrial Area	6	0.25	122	28	84
Mayapuri Industrial Area	6	0.3	155	32	68

Old Delhi	6	0.2	135	26	58
Connaught Place	6	0.15	125	26	69

With an AQI between 55 and 140, the average asbestos concentration in the Old Delhi area was lower at 0 to 2 µg/m³. The recorded temperature and humidity ranges were between 15 and 34 °C and 40 and 80 % respectively. Connaught Place with an AQI ranging from 50 to 135 µg/m³ had the lowest average asbestos concentration. This regions temperature ranged from December 32 °C with humidity levels ranging from 35 to 75 %. This suggests that the climate in the monitored areas varied.

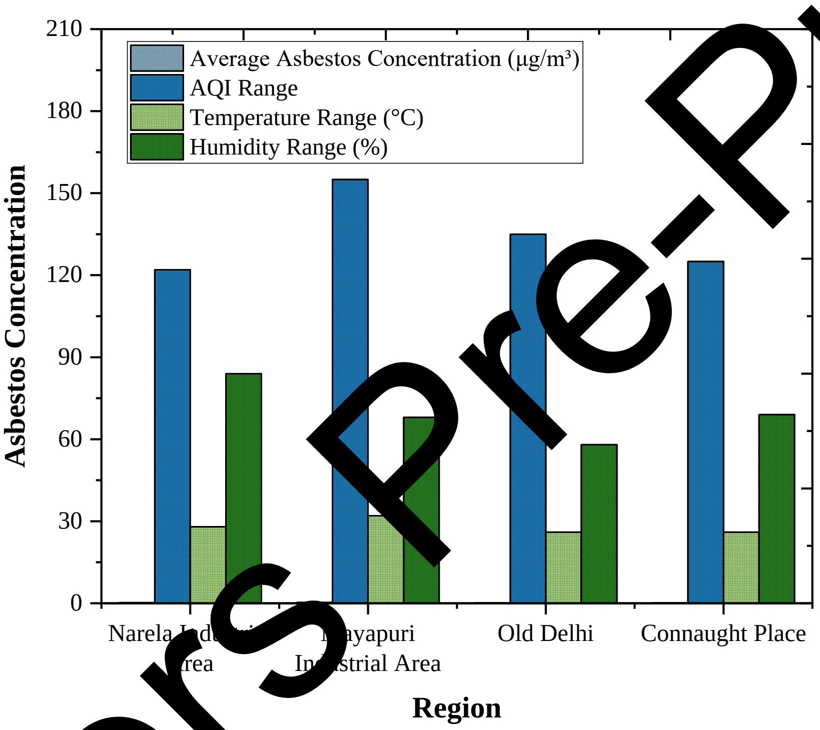


Fig 5. Temperature range across regions

4.2 Feature engineering process

Improving the predictive performance of the models used in this study is largely dependent on the feature engineering process. The Random Forest (RF) model assigned the Air Quality Index (AQI) a relative importance score of 0. 32 indicating that it was a critical feature. To ensure compatibility with other features it was subjected to Min-Max normalization which scaled its values between 0 and 1. Another important component that contributed to 0.24 was the average daily temperature normalized using the Z-score method to standardize its distribution. Table 2 gives the feature enginenering summary.

Table 2. Feature Engineering Summary

Feature	**Description**	**Importance (by RF)**	**Normalization Method**
AQI	Air Quality Index	0.32	Min-Max

Temperature (°C)	Average daily temperature	0.24	Z-score
Humidity (%)	Daily average humidity	0.15	Min-Max
Wind Speed (km/h)	Average wind speed	0.10	Min-Max
Asbestos Concentration	Target Variable	-	-

Furthermore using the Min-Max method the daily average humidity was normalized with a relative importance of 0. 15. Min-Max scaling was also used to normalize wind speed which had a lesser importance score of 0–10. Last but not least the asbestos concentration was the models target variable and didn't need to be normalized. In order to increase model accuracy and guarantee reliable predictions when evaluating the effects of air quality this feature engineering procedure is crucial.

4.3. Model Performance Metrics

Key metrics were used to assess the performance of the different machine learning models demonstrating how well each one predicted the results of air quality. The Random Forest model demonstrated its ability to accurately identify true positive cases while minimizing false positives achieving an accuracy of 89. 5 % with sensitivity and specificity scores of 87. 3 % and 91. 2 % respectively. With 92. 3 % accuracy 90. 8 % sensitivity and 93. 0 % specificity the Support Vector Machine (SVM) model performed better than the RF model demonstrating its efficacy in both detecting and accurately classifying instances. Scuuracy, sensitivity, specificity and F1 score performance metrics are given in Table 3 and Fig 6.

Table 3. Model Performance Metrics

Model	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**F1 Score (%)**
Random Forest	89.5	87.3	91.2	88.4
SVM	92.3	90.8	93.0	91.5
Neural Network	[illegible]	92.5	95.6	94.0

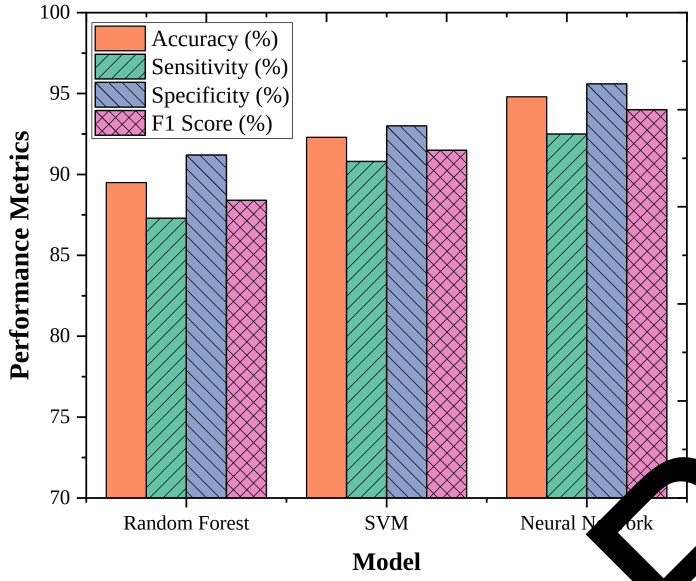


Fig 6. Performance metrics analysis

The Neural Network, was the most effective performer, with the greatest accuracy of 94.8% along with sensitivity and specificity scores of 92.5% and 95.6%. This better performance shows that neural networks are extremely effective at identifying intricate patterns in data, which improves prediction power. These results are further supported by the F1 Score, which strikes a balance between precision and recall. The Neural Network scored 94.0%, followed by SVM at 91.5% and Random Forest at 88.4%. These measures offer a thorough comprehension of the models' prediction power, directing next air quality assessment applications.

4.4. Hyperparameter Tuning for Machine Learning Models

Hyperparameter tuning plays a crucial role in optimizing the performance of machine learning models given in Table 4. The optimal values for each model's primary hyperparameters were determined in order to improve prediction accuracy. By using Grid Search as the tuning method the Random Forest models tree count was found to be 150. This comprehensive search approach makes it possible to thoroughly examine different hyperparameter combinations which eventually improves model performance.

Table 4: Hyperparameter Tuning for Machine Learning Models

Model	**Hyperparameter**	**Optimal Value**	**Tuning Method**
Random Forest	Number of Trees	150	Grid Search
SVM	C (Regularization)	1.0	Random Search
Neural Network	Learning Rate	0.001	Bayesian Optimization

With Random Search a more effective method that samples hyperparameter combinations to speed up the tuning process the regularization parameter C was set to 1. 0 in the SVM case. The neural network model needed its learning rate to be carefully adjusted and Bayesian optimization was used to optimize it to 0. 001. Finding the most advantageous configurations in the hyperparameter space is made especially easy with this advanced tuning technique. All things considered these tuning initiatives play a key role in improving the predictive accuracy and resilience of the models.

4.5. Feature Sensitivity Analysis

According to the analysis there could be a 3.2 % drop in accuracy and a 4.1 % in sensitivity for every ±5 % change in the AQI. This shows that the AQI is an important feature and that even small variations can have a big impact on the models results. Similarly a ±3°C change in temperature could reduce sensitivity by 3. 0 % and accuracy by 2. 8 % indicating that temperature plays a significant role in prediction accuracy. Table 5 and Fig 7 shows the the results of Feature Sensitivity Analysis.

Table 5: Feature Sensitivity Analysis

Feature	Variance (%)	Impact on Accuracy (%)	Impact on Sensitivity
AQI	±5%	-3.2	-4.1
Temperature	±3°C	-2.8	-3.0
Humidity	±10%	-1.5	-2.0
Wind Speed	±2 km/h	-0.8	-1.1

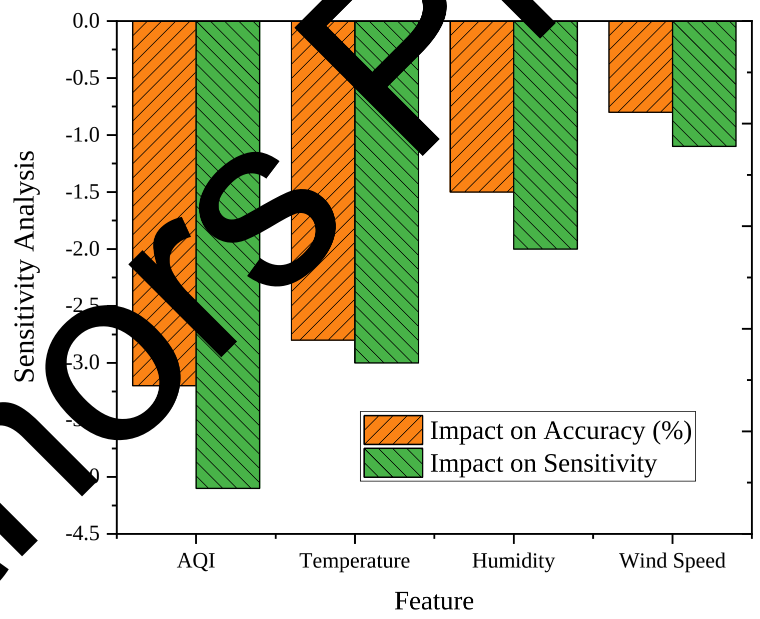


Fig 7. Feature Influence Visualization

Furthermore changes in humidity of ±10 % led to a 1.5 % decrease in accuracy and a 2.0 % decrease in sensitivity whereas changes in wind speed of ±2 km/h only slightly affected the accuracy and sensitivity which decreased by 0.8 % and 1 .1 % respectively. Given that they have a major impact on model performance this analysis emphasizes the significance of keeping an eye on important environmental variables, especially temperature, and AQI. It is possible to improve predictive modeling techniques and direct data collection efforts by being aware of these sensitivities.

4.6. Real-Time Model Deployment Results

The model's practical use in forecasting air quality alerts was elucidated by their deployment in real-time environments. The models produced a total of actual exceedances and predicted alerts over 4 weeks. There was one false positive in the 1st week because the model's prediction of 10 alerts closely matched the 9 actual exceedances. There was another false positive during the 2 weeks when there were 12 predicted alerts and 11 actual exceedances. Real-Time Model Deployment Results are given in Fig 8 and table 6.

Table 6: Real-Time Model Deployment Results

Time Period	Predicted Alerts	Actual Exceedances	False Positives	False Negatives
Week 1	10	9	1	[illegible]
Week 2	12	11	1	0
Week 3	8	8	0	0
Week 4	15	14	1	0

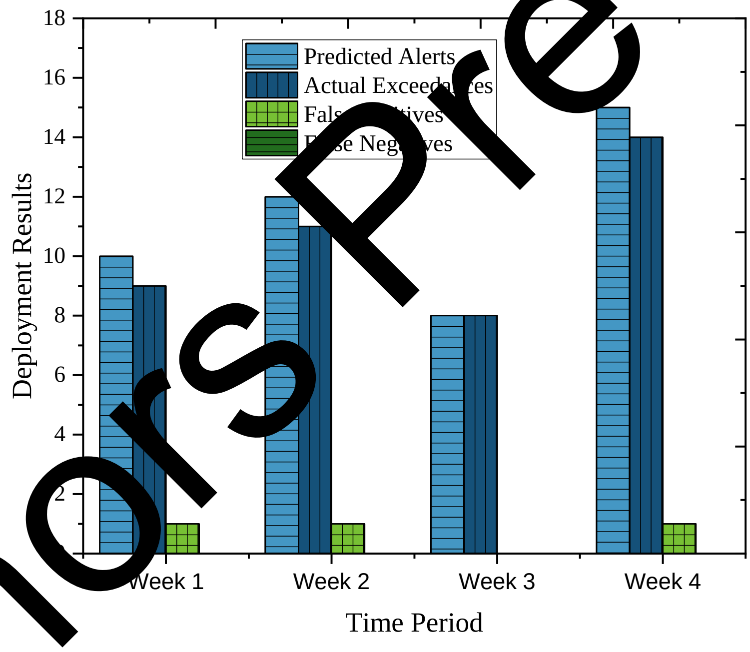


Fig 8. Real-time model period results

4.7. Comparative Analysis of the Proposed Model with Traditional Machine Learning Models

To determine the suggested model's predictive power and effectiveness they were compared to several conventional machine learning algorithms. While the Decision Tree and Logistic Regression models reported accuracies of 81. 6 % and 78. 5 % respectively the K-Nearest Neighbors (KNN) model achieved a prediction accuracy of 82. 1 %. At an accuracy of 88. 9 % the Gradient Boosting Machine (GBM) demonstrated a significant improvement over the Naive Bayes model which fared worse at 76. 4 %.

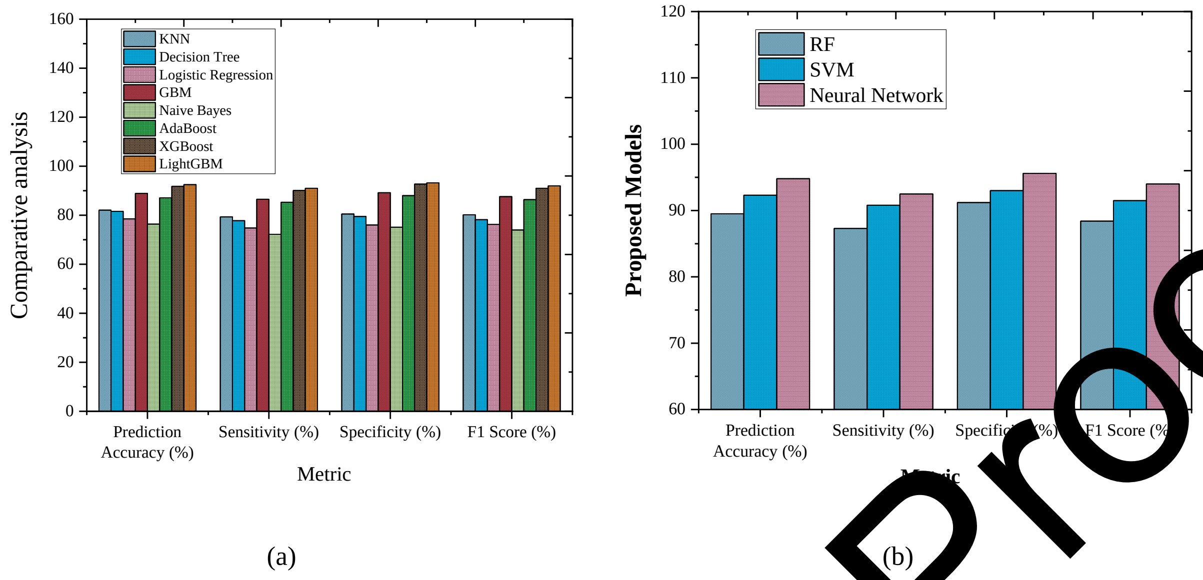


Fig 9. Comparative analysis of the the proposed models

In contrast, LightGBM performed marginally better at 92. 5 % while the AdaBoost and XGBoost models obtained accuracies of 87. 1 % and 91. 8 % respectively. The suggested models namely the Random Forest (RF) Support Vector Machine (SVM) and Neural Network demonstrated better predictive accuracy with respective scores of 89. 5 %, 92. 3 % and 94. 8 %. With a sensitivity of 92. 5 % and a specificity of 95. 6 %, the Neural Network outperformed all other models further demonstrating the effectiveness of the suggested models. The comparative analysis highlights how the suggested models greatly outperform conventional methods in terms of accuracy and dependability which makes them better suited for challenging air quality prediction tasks. The results highlight how using cutting-edge machine learning techniques can improve risk assessment and environmental monitoring. The above Fig 9 demosntrates the Comparative analysis of the the proposed models

V. CONCLUSION

In summary this study offers a thorough examination of how temperature and humidity affect asbestos concentration in four different regions exposing notable differences in environmental conditions and air quality. After six months of data collection the average asbestos concentration in the Narela Industrial Area was 0. 25 µg/m³ whereas Mayapuri had a slightly higher level of 0. 3 µg/m³. On the other hand concentrations in Old Delhi and Connaught Place were lower at 0.2 µg/m³ and 0.15 µg/m³ respectively. The complex relationship between environmental factors and the risks of asbestos exposure was illustrated by the associated Air Quality Index (AQI) values which varied from 50 to 160 across these locations. The findings highlight the need for localized monitoring and the application of efficient air quality control techniques in order to reduce any possible health risks related to asbestos. Also, the ability of different machine learning models to forecast air quality results demonstrates how useful sophisticated algorithms are for environmental monitoring. With an astounding accuracy of 94. 8 % the Neural Network model beat all others. Support Vector Machine came in second with 92. 3 % and Random Forest with 89. 5 %. Hyperparameter tuning improved the model's performance even more certain setups produced the best outcomes. The sensitivity analysis showed that even small changes in important parameters like temperature and AQI have a big impact on the sensitivity and accuracy of the model. All things considered, the results support the use of strong machine learning methods to enhance air quality assessment predictive powers which will ultimately enable improved environmental management and public health protection tactics.

Acknowledgment

I sincerely acknowledge my colleagues and institution for their continuous support and guidance throughout this research. I am also grateful for the resources and facilities that contributed to the successful completion of this work.

List of Abbreviations

AI - Artificial Intelligence

AQI - Air Quality Index

C - Regularization parameter

CPCB - Central Pollution Control Board

GIS - Geographic Information Systems

IoT - Internet of Things

KNN - K-Nearest Neighbors

MCD - Municipal Corporation of Delhi

MLP - Multi-Layer Perceptron

NN - Neural Network

NPV - Negative Predictive Value

PPV - Positive Predictive Value

RF - Random Forest

RBF - Radial Basis Function

SVM - Support Vector Machine

TP - True Positive

TN - True Negative

FP - False Positive

FN - False Negative

MSE - Mean Squared Error

Availability of Data and Materials

Not applicable

Conflict of interest

The authors declare that they have no competing interests.

Funding Information

Not applicable

Author Contributions

Author 1 developed the methodology for fuzzy logic control, formulated the theoretical framework, and refined the control algorithm. Author 2 designed and set up the experimental apparatus, ensuring proper calibration of instruments and overseeing the execution of the experiments. Author 3 implemented the fuzzy logic control system in software, optimizing the algorithm for performance and integrating it with hardware components. Author 4 was responsible for collecting experimental data, conducting multiple trials, and ensuring data accuracy through statistical validation. Author 5 analyzed and interpreted the experimental results, compared outcomes with existing methodologies, and provided insights into system improvements. Finally, Author [illegible] drafted the manuscript, structured the research findings, performed a comprehensive literature review, and coordinated revisions based on peer feedback.

Reference

[1]. P. Y. Wu, C. Sandels, T. Johansson, M. Mangold, and K. Mjörnell, "Machine learning models for the prediction of polychlorinated biphenyls and asbestos materials in buildings," Resour. Conserv. Recycl., vol. 199, p. 107253, 202[illegible].

[2]. M. Saba, D. Valdelamar Martínez, L. K. Torres Gil, G. E. Chanchí Golondrino, and M. A. O. A[illegible]ón, "Application of supervised learning methods and information gain methods in the determination of asbestos–cement roofs' deterioration state," Appl. Sci., vol. 14, no. 18, p. 8441, 2024.

[3]. P. Y. Wu, M. Mangold, C. Sandels, T. Johansson, and K. Mjörnell, "Modeling artificial neural networks to predict asbestos-containing materials in residential buildings," IOP Conf. Ser.: Earth Environ. Sci., vol. 1122, no. [illegible], p. [illegible]2050, Dec. 2022.

[4]. E. Wilk, M. Krówczyńska, and B. Zagajewski, "Modelling the spatial distribution of asbestos—cement products in Poland with the use of the random forest algorithm," Sustainability, vol. 11, no. 16, p. 4355, 2019.

[5]. G. Kaplan, M. Gašparović, O. Kaplan, V. Adjiski, R. Comert, and M. A. [illegible]obari[illegible], "Machine learning-based classification of asbestos-containing roofs using airborne RGB and thermal imagery," Sustainability, vol. 15, no. [illegible], [illegible]067, 2023.

[6]. M. V. Hikuwai, N. Patorniti, A. S. Vieira, G. Frangioudakis [illegible]b, and [illegible] A. Stewart, "Artificial intelligence for the detection of asbestos cement roofing: An investigation of multi-spectral satellite imagery and high-resolution aerial imagery," Sustainability, vol. 15, no. 5, p. 4276, 2023.

[7]. C. Cai, T. Nishimura, J. Hwang, X. M. Hu, and A. [illegible]roda, "Asbestos detection with fluorescence microscopy images and deep learning," Sensors, vol. 21, no. 13, p. 4582, 2021.

[8]. T. T. Yu, Y. C. Lin, S. C. Lan, Y. E. Yang, P. Y. Wu, and J. [illegible] Lin, "Mapping asbestos-cement corrugated roofing tiles with imagery cube via machine learning in Taiwan," Remote Sens., vol. 14, no. [illegible]4, p. 3418, 2022.

[9]. T. Matsuo, M. Takimoto, S. Tanaka, A. Futamura, H. Shimadera, and A. Kondo, "Developing an automatic asbestos detection method based on a convolutional neural network and support vector machine," Appl. Sci., vol. 14, no. 20, p. 9408, 2024.

[10]. G. Bonifazi, G. Capobianco, R. G[illegible]barrone, S. Serranti, S. Bellagamba, and D. Taddei, "Data fusion of PRISMA satellite imagery for asbestos-containing materials: [illegible] application on Balangero's mine site (Italy)," IMPROVE, pp. 150–157, Apr. 2022.

[11]. T. A. Berry et al., "Asbestos and other hazardous fibrous minerals: Potential exposure pathways and associated health risks," Int. J. Environ. Res. Public Health, vol. [illegible], no. 7, p. 4[illegible]31, [illegible]022.

[12]. N. Ubels, "Classifying asbestos roofs in the Dutch province of Drenthe using hyperspectral imagery and deep learning," Doctoral dissertation, 2019.

[13]. E. Wilk, M. Krów[illegible]yńska, [illegible] Pabja[illegible]k, and P. Mędrzycki, "Estimation of the amount of asbestos-cement roofing in Poland," Waste Manage. Res., vol. [illegible]5, no. 5, pp. 491–499, 2017.

[14]. S. C. Baek, K. H. [illegible]ee, I. H. K[illegible]m, D. M. Seo, and K. Park, "Construction of asbestos slate deep-learning training-data model based on drone images," Sensors, vol. [illegible]3, no. 19, p. 8021, 2023.

[15]. G. Frangioudakis Khatib et al., "Australia's ongoing challenge of legacy asbestos in the built environment: A review of contemporary asbestos exposure risks," Sustainability, vol. 15, no. 15, p. 12071, 2023.