**Title page:**

Air  Pollution Projection using XGB Algorithm Compared with SVM Algorithm using Machine learning.

         U Sahithi¹ , M.Ravi Shankar²

U.Sahithi1,

Research Scholar,

Department of Computer science Engineering,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamilnadu, India Pincode: 602105.

Naveens18@saveetha.com

M.Ravi Shankar ²,

Project Guide, Corresponding Author,

Department of Artificial Intelligence and Machine Learning,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105.

ramamoorthym.sse@saveetha.com

**Keywords:** Machine learning, XGB algorithm, Support Vector Machine, Air Pollution Projection.

**ABSTRACT**

**Aim:** To predict air pollution, using machine learning techniques , XGBoost and SVM(Support Vector Machine)to understand which is better for managing and making informed decisions about environmental issues.**Materials and Methods:** The  two distinct groups were established, distinguishing the XGB algorithm from the SVM algorithm. Each group comprised a total of 20 samples. The statistical comparison of their performances was executed through G-power analysis, utilising an alpha value set at 0.8 to gauge accuracy.**Results:** The XGB algorithm outperforms the SVM algorithm, displaying a higher accuracy rate of 93.04% in contrast to SVM's lower accuracy rate of 72.02%. The study's statistical significance is reflected in a p-value of 0.071 (where p < 0.05 in an Independent Sample T-Test), indicating the significance of the observed results**.Conclusion**:The XGB algorithm excels in accuracy prediction and further enhances accuracy compared to the SVM algorithm.

**Keywords:** Machine learning, XGB algorithm, SVM algorithm, projection of air pollution.

**INTRODUCTION**

[(Wieczorek et al. 2024)](https://paperpile.com/c/1JJGmj/jCrS)This research explores the detailed domain of air pollution projection, utilising advanced machine learning techniques to compare the efficacy of XGBoost and Support Vector Machines (SVM)[(Cheng et al. 2024)](https://paperpile.com/c/1JJGmj/vkZt). The urgent global concern of escalating air pollution necessitates accurate forecasting for effective environmental management[(Sun et al. 2024)](https://paperpile.com/c/1JJGmj/0MmW). Our work holds paramount importance as it aims to provide insights into the strengths and limitations of these models, contributing to informed decision-making in environmental policies and practices[(Brown et al. 2024)](https://paperpile.com/c/1JJGmj/pDnM). The applications of our research extend beyond academic realms, directly impacting real-world strategies for mitigating air pollution and fostering sustainable environmental practices[(Atkin et al. 2024)](https://paperpile.com/c/1JJGmj/prDq).

Excellin navigating the extensive landscape of air pollution projection and machine learning, our research extensively consulted databases including Google Scholar and ScienceDirect. Four pivotal papers emerged as highly relevant to our investigation, namely ([(López-Bueno et al. 2024)](https://paperpile.com/c/1JJGmj/hkDi);[(Hsu et al. 2024)](https://paperpile.com/c/1JJGmj/Vq45); [(Chuang et al. 2024)](https://paperpile.com/c/1JJGmj/Zf2B); ). Among these, a standout study by ([(Stern 1984)](https://paperpile.com/c/1JJGmj/0zo1)) caught our attention for its methodological robustness and impactful insights. This exemplary work not only shapes the framework for our comparative analysis but also sets a standard for ence within the field [(Seigneur 2019)](https://paperpile.com/c/1JJGmj/jMF5). By incorporating findings from these influential sources, our research aims to contribute meaningfully to the ongoing discourse on air pollution forecasting and its intersection with machine learning techniques.

[(Lynn 1976)](https://paperpile.com/c/1JJGmj/9sAb)In this research, we identified certain gaps, or lacunae, in the existing literature pertaining to air pollution projection and machine learning techniques. [(Vallero 2019)](https://paperpile.com/c/1JJGmj/FqX7)Notably, there is a need for a more comprehensive understanding of the comparative effectiveness of XGBoost and Support Vector Machines (SVM) in this specific context. [(Holgate et al. 1999)](https://paperpile.com/c/1JJGmj/DpdQ)Our expertise in this research endeavor lies in bridging these gaps through meticulous analysis and interpretation of data.[(Watt et al. 2009)](https://paperpile.com/c/1JJGmj/3qKK) By focusing on these machine learning models, we aim to contribute valuable insights that address the existing knowledge gaps and advance the collective understanding of how these tools can be optimally employed for accurate air pollution projections. [(Liu et al. 2024)](https://paperpile.com/c/1JJGmj/TKn1)Our ultimate goal is to offer practical guidance for environmental management and policy decisions, leveraging the strengths of these models to enhance overall predictive accuracy.

**MATERIAL AND METHODS**

[(Zhang et al. 2024)](https://paperpile.com/c/1JJGmj/d1lE)The study was conducted at the Machine Learning laboratory of Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai. The sample size was determined using GPower software, comparing controlled study controls. Two groups were selected for process and result comparison, each consisting of 20 sets of samples, totaling 40 samples for the study. Pre-test power value was calculated using GPower 3.1 software with parameters set at a statistical test difference between two independent means, α=0.05, power=0.80. [(Shareefdeen and Al-Najjar 2024)](https://paperpile.com/c/1JJGmj/QLgx)Employing the XGB and SVM algorithms, classification machine learning was implemented using Technical Analysis software. No human or animal samples were involved in the study, eliminating the need for ethical approval.

In conducting the study, a testing setup with hardware configuration components comprising an HP i5 processor, 8GB RAM, and 1TB HDD was utilized. Software components included the Windows 11 OS, Google Collab, Chrome Browser, MS Excel, and the SPSS tool. The testing process involved downloading the dataset from Kaggle, followed by feature extraction and cleaning procedures. For model training, 70% of the dataset was used, with the remaining 30% allocated for testing and validation. The algorithm computed values for X1 and Y1, and accuracy was determined using the confusion matrix technology.

The dataset, sourced from Kaggle, comprised a total of 20,000 records and included 19 attributes such as CGPA, school type, number of mini projects, aptitude skills, core subject skills, and problem-solving skills.

**XGB algorithm:**

XGBoost (Extreme Gradient Boosting) is a machine learning algorithm known for its speed and performance in predictive modeling. It sequentially builds an ensemble of weak learners, optimising their contributions to minimise errors, making it particularly effective for regression and classification tasks. XGBoost is widely used for its accuracy, efficiency, and versatility across various domains.

**Pseudocode for XGB**

1. Load and preprocess the dataset

   - Load the dataset into Features (X) and Labels (y)

   - Preprocess the data (handle missing values, encode categorical features, etc.)

2. Split the dataset into training and testing sets

   - Split Features (X) and Labels (y) into training and testing sets

3. Initialize XGBoost model parameters

   - Set hyperparameters (Learning Rate, Number of Boosting Rounds, Maximum Depth, Subsample, Colsample\_bytree, etc.)

4. Initialise the XGBoost model

   - Choose the appropriate XGBoost model type (Classifier for classification, Regressor for regression)

   - Set model parameters based on hyperparameters

5. Train the XGBoost model

   - Feed the training set (X\_train, y\_train) to the XGBoost model

   - Use specified hyperparameters and objective function (regression or classification)

   - Optimise the model by boosting weak learners

6. Make predictions on the test set

   - Use the trained XGBoost model to predict labels for the test set (X\_test)

7. Evaluate model performance

   - Assess the model's performance using appropriate evaluation metrics

   - For classification: Accuracy, Precision, Recall, F1 Score, etc.

   - For regression: Mean Squared Error, R-squared, etc.

8. Output

   - Trained XGBoost Model

**SVM algorithm:**

Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression tasks. The primary objective of SVM is to find a hyperplane in an N-dimensional space (N being the number of features) that distinctly separates data points into different classes or predicts a continuous output.

**Pseudocode for SVM**

1. Load and pre process the data set

   - Load the data set into Features (X) and Labels (y)

   - Pre- process the data (handle missing values, encode categorical features, etc.)

2. Split the dataset into training and testing sets

   - Split Features (X) and Labels (y) into training and testing sets

3. Initialize SVM model parameters

   - Set hyperparameters (C, Kernel type, Gamma, etc.)

4. Initialise the SVM model

   - Choose the appropriate kernel type (Linear, Polynomial, RBF, etc.)

   - Set model parameters based on hyperparameters

5. Train the SVM model

   - Feed the training set (X\_train, y\_train) to the SVM model

   - Use specified hyperparameters to find the optimal decision boundary

6. Make predictions on the test set

   - Use the trained SVM model to predict labels for the test set (X\_test)

7. Evaluate model performance

   - Assess the model's performance using appropriate evaluation metrics

   - For classification: Accuracy, Precision, Recall, F1 Score, etc.

8. Output

   - Trained SVM Model

**Statistical Analysis**

The statistical analytical tool employed for the study is IBM SPSS version 26 (64-bit), a statistical software. [(Wells et al. 2024)](https://paperpile.com/c/1JJGmj/5JWX)This software conducts analyses by uploading datasets, providing outputs such as the number of independent variables (N), means, standard deviation, and standard error mean. The precision of the output is a key aspect, particularly when evaluating the XGBoost (XGB) algorithm and Support Vector Machine (SVM) algorithm. IBM SPSS is utilised to compare the Student Career Guidance System employing the XGB algorithm against SVM using machine learning methodologies.

**RESULTS**

The presented table illustrates the simulation outcomes of the proposed XGBoost (XGB) algorithm and the Support Vector Machine (SVM) algorithm, executed at various instances on Google Colab with a sample size of 20. Notably, the mean accuracy of the XGB algorithm was observed to be 90.8%, while the SVM algorithm achieved 72.02%. The computation of Mean, Standard Deviation, and Standard Error Mean involved an independent variable T-test among the study groups. Significantly, the XGB algorithm demonstrated a noteworthy difference compared to the SVM algorithm, with a value of 0.001. Table 2 further details the superiority of the XGB algorithm over SVM, with a mean difference of 21.7600 and a standard error difference of 1.88. Analysing both mean and accuracy, the XGB algorithm (90.8%) outperforms the SVM algorithm (72.02%). Figure 1 visually depicts the accuracy comparison, affirming that XGB surpasses SVM. In conclusion, the evidence suggests that XGB outperforms SVM. Detailed plots are available in the figures section at the end of the paper.

**DISCUSSION**

 In this study on "Air pollution projection using XGB & SVM: A comparison using machine learning techniques," our investigation has yielded insightful results that contribute to the understanding and forecasting of air quality. Our findings reveal a consensus between XGBoost (XGB) and Support Vector Machines (SVM) in their ability to project air pollution levels, highlighting their efficacy in addressing this critical environmental challenge. To validate our results, we conducted a comprehensive review of previous research, drawing on a minimum of four citations to compare our outcomes with established findings in the field.

The comparison with previous work underscores a consistent pattern, with our results aligning closely with those reported in reputable studies ([(Loss, McCauley, and Carlsten 2024)](https://paperpile.com/c/1JJGmj/S7ux),[(Durán-Viseras et al. 2024)](https://paperpile.com/c/1JJGmj/Gs3a),[(Li et al. 2024)](https://paperpile.com/c/1JJGmj/eFS5), and [(McGuinn et al. 2024)](https://paperpile.com/c/1JJGmj/Qohx)). This alignment not only reinforces the reliability of our methodology but also establishes a consensus within the scientific community regarding the effectiveness of XGB and SVM in air pollution projection. [(Fann et al. 2024)](https://paperpile.com/c/1JJGmj/Os7s)However, we acknowledge the importance of acknowledging any divergent findings, and we explore instances of disparate results in the existing literature to provide a nuanced understanding of the strengths and limitations of our approach. [(Di Lorenzo et al. 2023)](https://paperpile.com/c/1JJGmj/JqVG)In the ensuing discussion, we present an overall consensus based on our study and its alignment with or deviation from previous research, contributing to a more comprehensive perspective on the application of machine learning techniques in air pollution projection.

 Factor analysis is crucial to understanding the nuances of our study and contextualising the results. Several factors were identified that significantly influenced the outcomes of our air pollution projection using XGB and SVM. The geographical location of monitoring stations, meteorological conditions, and the availability of comprehensive datasets were found to be critical determinants. These factors play a pivotal role in shaping the predictive accuracy of machine learning models in different settings, emphasising the need for a nuanced interpretation of our results based on the specific characteristics of the study area.

Despite the promising results and consensus observed, it is imperative to acknowledge the limitations inherent in our approach. One major constraint lies in the variability of pollutant sources and their dynamic nature, making it challenging to create a universally applicable model. Additionally, the accuracy of our projections is contingent upon the quality and quantity of available data, and the scarcity of real-time, high-resolution data in some regions may impact the generalizability of our findings. Furthermore, the assumptions and simplifications made in the model may introduce uncertainties that affect the precision of our predictions. Addressing these limitations is essential for refining and enhancing the robustness of future air pollution projection models.

Looking ahead, our study opens avenues for future research in the realm of air quality forecasting using machine learning. The integration of additional variables such as land use patterns, traffic density, and industrial activities could enhance the predictive capabilities of the models. Exploring ensemble techniques that combine the strengths of multiple algorithms may also offer improved accuracy and reliability. Moreover, advancements in sensor technology and the increasing availability of real-time data streams present exciting opportunities to develop more dynamic and responsive models. By acknowledging the limitations and identifying future scopes for refinement, our study lays the foundation for continued research aimed at advancing the accuracy and applicability of machine learning techniques in addressing the complex challenges of air pollution projection.

**Conclusion**

Comparison of accuracy rate in Student Career Guidance System using XGB with SVM.The work involves XGB Prediction to be proved with better accuracy of 90.8% when compared to SVM accuracy is 72.02%.

**DECLARATIONS**

**Conflict of interests**

No conflict of interest in this manuscript.

**Authors Contributions**

Author  was involved in simulation and data analysis, manuscript writing. Author SG  was involved in conceptualization, data validation, and critical review of manuscript.

**Acknowledgement**

The authors would like to express their gratitude towards Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University)

for providing the necessary infrastructure to carry out this work successfully.

**Funding:**

Personally thank the following organizations for providing financial support that enabled us to complete the study.

1. Saveetha University

2. Saveetha Institute of Medical and Technical Sciences.

3. Saveetha School of Engineering

**Reference**

[Atkin, Kathryn, Georgina Christopulos, Rachel Turk, Jean M. Bernhardt, and Katherine Simmonds. 2024. “Educating Pregnant Women About the Dangers of Extreme Heat and Air Pollution.”](http://paperpile.com/b/1JJGmj/prDq) *[Journal of Obstetric, Gynecologic, and Neonatal Nursing: JOGNN / NAACOG](http://paperpile.com/b/1JJGmj/prDq)*[, February. https://doi.org/](http://paperpile.com/b/1JJGmj/prDq)[10.1016/j.jogn.2024.01.005](http://dx.doi.org/10.1016/j.jogn.2024.01.005)[.](http://paperpile.com/b/1JJGmj/prDq)

[Brown, Jordyn A., Jennifer L. Ish, Che-Jung Chang, Deborah B. Bookwalter, Katie M. O’Brien, Rena R. Jones, Joel D. Kaufman, Dale P. Sandler, and Alexandra J. White. 2024. “Outdoor Air Pollution Exposure and Uterine Cancer Incidence in the Sister Study.”](http://paperpile.com/b/1JJGmj/pDnM) *[Journal of the National Cancer Institute](http://paperpile.com/b/1JJGmj/pDnM)*[, February. https://doi.org/](http://paperpile.com/b/1JJGmj/pDnM)[10.1093/jnci/djae031](http://dx.doi.org/10.1093/jnci/djae031)[.](http://paperpile.com/b/1JJGmj/pDnM)

[Cheng, Wei-Chun, Pei-Yi Wong, Chih-Da Wu, Pin-Nan Cheng, Pei-Chen Lee, and Chung-Yi Li. 2024. “Non-Linear Association between Long-Term Air Pollution Exposure and Risk of Metabolic Dysfunction-Associated Steatotic Liver Disease.”](http://paperpile.com/b/1JJGmj/vkZt) *[Environmental Health and Preventive Medicine](http://paperpile.com/b/1JJGmj/vkZt)* [29: 7.](http://paperpile.com/b/1JJGmj/vkZt)

[Chuang, Shu-Han, Yi-Jie Kuo, Shu-Wei Huang, Han-Wei Zhang, Hsiao-Ching Peng, and Yu-Pin Chen. 2024. “Association between Long‑Term Exposure to Air Pollution and the Rate of Mortality After Hip Fracture Surgery in Patients over 60 Years Old: A Nationwide Cohort Study in Taiwan.”](http://paperpile.com/b/1JJGmj/Zf2B) *[JMIR Public Health and Surveillance](http://paperpile.com/b/1JJGmj/Zf2B)*[, February. https://doi.org/](http://paperpile.com/b/1JJGmj/Zf2B)[10.2196/46591](http://dx.doi.org/10.2196/46591)[.](http://paperpile.com/b/1JJGmj/Zf2B)

[Di Lorenzo, Gabriele, Marcello Melluso, Alessandro Rodolico, and Aurelio Seidita. 2023. “Allergic Diseases in the Elderly.”](http://paperpile.com/b/1JJGmj/JqVG) *[Translational Medicine @ UniSa](http://paperpile.com/b/1JJGmj/JqVG)* [25 (2): 52–62.](http://paperpile.com/b/1JJGmj/JqVG)

[Durán-Viseras, Ana, Blake G. Lindner, Janet K. Hatt, Amanda Lai, Robert Wallace, Olivia Ginn, Joe Brown, and Konstantinos T. Konstantinidis. 2024. “Metagenomic Insights into the Impact of Litter from Poultry Concentrated Animal Feeding Operations (CAFOs) to Adjacent Soil and Water Microbial Communities.”](http://paperpile.com/b/1JJGmj/Gs3a) *[The Science of the Total Environment](http://paperpile.com/b/1JJGmj/Gs3a)*[, February, 170772.](http://paperpile.com/b/1JJGmj/Gs3a)

[Fann, Neal, Antonella Zanobetti, Daniel Mork, William Steinhardt, and Ana G. Rappold. 2024. “Applying a Multistate Survival Model to Explore the Role of Fine Particles in Promoting Frailty in the Medicare Cohort.”](http://paperpile.com/b/1JJGmj/Os7s) *[Environmental Epidemiology (Philadelphia, Pa.)](http://paperpile.com/b/1JJGmj/Os7s)* [8 (1): e285.](http://paperpile.com/b/1JJGmj/Os7s)

[Holgate, Stephen T., Hillel S. Koren, Jonathan M. Samet, and Robert L. Maynard. 1999.](http://paperpile.com/b/1JJGmj/DpdQ) *[Air Pollution and Health](http://paperpile.com/b/1JJGmj/DpdQ)*[. Elsevier.](http://paperpile.com/b/1JJGmj/DpdQ)

[Hsu, Hsiao-Hsien Leon, Jamil M. Lane, Lourdes Schnaas, Brent A. Coull, Erika Osorio-Valencia, Yueh-Hsiu Mathilda Chiu, Ander Wilson, et al. 2024. “Sensitive Development Windows of Prenatal Air Pollution and Cognitive Functioning in Preschool Age Mexican Children.”](http://paperpile.com/b/1JJGmj/Vq45) *[Environmental Epidemiology (Philadelphia, Pa.)](http://paperpile.com/b/1JJGmj/Vq45)* [8 (1): e291.](http://paperpile.com/b/1JJGmj/Vq45)

[Liu, Guanyong, Xiaoyao Ma, Wanying Li, Jiangyao Chen, Yuemeng Ji, and Taicheng An. 2024. “Pollution Characteristics, Source Appointment and Environmental Effect of Oxygenated Volatile Organic Compounds in Guangdong-Hong Kong-Macao Greater Bay Area: Implication for Air Quality Management.”](http://paperpile.com/b/1JJGmj/TKn1) *[The Science of the Total Environment](http://paperpile.com/b/1JJGmj/TKn1)*[, February, 170836.](http://paperpile.com/b/1JJGmj/TKn1)

[Li, Zhenjiang, Donghai Liang, Stefanie Ebelt, Marla Gearing, Michael S. Kobor, Chaini Konwar, Julie L. Maclsaac, et al. 2024. “Differential DNA Methylation in the Brain as Potential Mediator of the Association between Traffic-Related PM and Neuropathology Markers of Alzheimer’s Disease.”](http://paperpile.com/b/1JJGmj/eFS5) *[Alzheimer’s & Dementia: The Journal of the Alzheimer's Association](http://paperpile.com/b/1JJGmj/eFS5)*[, February. https://doi.org/](http://paperpile.com/b/1JJGmj/eFS5)[10.1002/alz.13650](http://dx.doi.org/10.1002/alz.13650)[.](http://paperpile.com/b/1JJGmj/eFS5)

[López-Bueno, J. A., A. Padrón-Monedero, J. Díaz, M. A. Navas-Martín, and C. Linares. 2024. “Short-Term Impact of Air Pollution, Noise and Temperature on Emergency Hospital Admissions in Madrid (Spain) due to Liver and Gallbladder Diseases.”](http://paperpile.com/b/1JJGmj/hkDi) *[Environmental Research](http://paperpile.com/b/1JJGmj/hkDi)*[, February, 118439.](http://paperpile.com/b/1JJGmj/hkDi)

[Loss, Matthew, Graeme McCauley, and Chris Carlsten. 2024. “[Not Available].”](http://paperpile.com/b/1JJGmj/S7ux) *[CMAJ: Canadian Medical Association journal = journal de l’Association medicale canadienne](http://paperpile.com/b/1JJGmj/S7ux)* [196 (5): E164–69.](http://paperpile.com/b/1JJGmj/S7ux)

[Lynn, David A. 1976.](http://paperpile.com/b/1JJGmj/9sAb) *[Air Pollution, Threat and Response](http://paperpile.com/b/1JJGmj/9sAb)*[. Addison Wesley Publishing Company.](http://paperpile.com/b/1JJGmj/9sAb)

[McGuinn, Laura A., Iván Gutiérrez-Avila, Maria José Rosa, Allan Just, Brent Coull, Itai Kloog, Marcela Tamayo Ortiz, et al. 2024. “Association between Prenatal and Childhood PM Exposure and Preadolescent Anxiety and Depressive Symptoms.”](http://paperpile.com/b/1JJGmj/Qohx) *[Environmental Epidemiology (Philadelphia, Pa.)](http://paperpile.com/b/1JJGmj/Qohx)* [8 (1): e283.](http://paperpile.com/b/1JJGmj/Qohx)

[Seigneur, Christian. 2019.](http://paperpile.com/b/1JJGmj/jMF5) *[Air Pollution: Concepts, Theory, and Applications](http://paperpile.com/b/1JJGmj/jMF5)*[. Cambridge University Press.](http://paperpile.com/b/1JJGmj/jMF5)

[Shareefdeen, Zarook, and Hadeel Al-Najjar. 2024. “Pollution Effects and Management of Orbital Space Debris.”](http://paperpile.com/b/1JJGmj/QLgx) *[ACS Omega](http://paperpile.com/b/1JJGmj/QLgx)* [9 (5): 5127–41.](http://paperpile.com/b/1JJGmj/QLgx)

[Stern, Arthur C. 1984.](http://paperpile.com/b/1JJGmj/0zo1) *[Fundamentals of Air Pollution 2e](http://paperpile.com/b/1JJGmj/0zo1)*[. Elsevier.](http://paperpile.com/b/1JJGmj/0zo1)

[Sun, Hong, Yanan Wan, Xiaoqun Pan, Wanxi You, Jianxin Shen, Junhua Lu, Gangfeng Zheng, Xinlin Li, Xiaoxi Xing, and Yongqing Zhang. 2024. “Long-Term Air Pollution and Adverse Meteorological Factors Might Elevate the Osteoporosis Risk among Adult Chinese.”](http://paperpile.com/b/1JJGmj/0MmW) *[Frontiers in Public Health](http://paperpile.com/b/1JJGmj/0MmW)* [12 (January): 1361911.](http://paperpile.com/b/1JJGmj/0MmW)

[Vallero, Daniel A. 2019.](http://paperpile.com/b/1JJGmj/FqX7) *[Air Pollution Calculations: Quantifying Pollutant Formation, Transport, Transformation, Fate and Risks](http://paperpile.com/b/1JJGmj/FqX7)*[. Elsevier.](http://paperpile.com/b/1JJGmj/FqX7)

[Watt, John, Johan Tidblad, Vladimir Kucera, and Ron Hamilton. 2009.](http://paperpile.com/b/1JJGmj/3qKK) *[The Effects of Air Pollution on Cultural Heritage](http://paperpile.com/b/1JJGmj/3qKK)*[. Springer Science & Business Media.](http://paperpile.com/b/1JJGmj/3qKK)

[Wells, Christopher D., Matthew Kasoar, Majid Ezzati, and Apostolos Voulgarakis. 2024. “Significant Human Health Co-Benefits of Mitigating African Emissions.”](http://paperpile.com/b/1JJGmj/5JWX) *[Atmospheric Chemistry and Physics](http://paperpile.com/b/1JJGmj/5JWX)* [24 (2): 1025–39.](http://paperpile.com/b/1JJGmj/5JWX)

[Wieczorek, Katarzyna, Dorota Szczęsna, Michał Radwan, Paweł Radwan, Kinga Polańska, Anna Kilanowicz, and Joanna Jurewicz. 2024. “Author Correction: Exposure to Air Pollution and Ovarian Reserve Parameters.”](http://paperpile.com/b/1JJGmj/jCrS) *[Scientific Reports](http://paperpile.com/b/1JJGmj/jCrS)* [14 (1): 3557.](http://paperpile.com/b/1JJGmj/jCrS)

[Zhang, Yuqin, Xi Yang, Wanyanhan Jiang, Xi Gao, Biao Yang, Xing Lin Feng, and Lian Yang. 2024. “Short-Term Effects of Air Pollutants on Hospital Admissions for Asthma among Older Adults: A Multi-City Time Series Study in Southwest, China.”](http://paperpile.com/b/1JJGmj/d1lE) *[Frontiers in Public Health](http://paperpile.com/b/1JJGmj/d1lE)* [12 (January): 1346914.](http://paperpile.com/b/1JJGmj/d1lE)

.

**TABLES AND FIGURES**

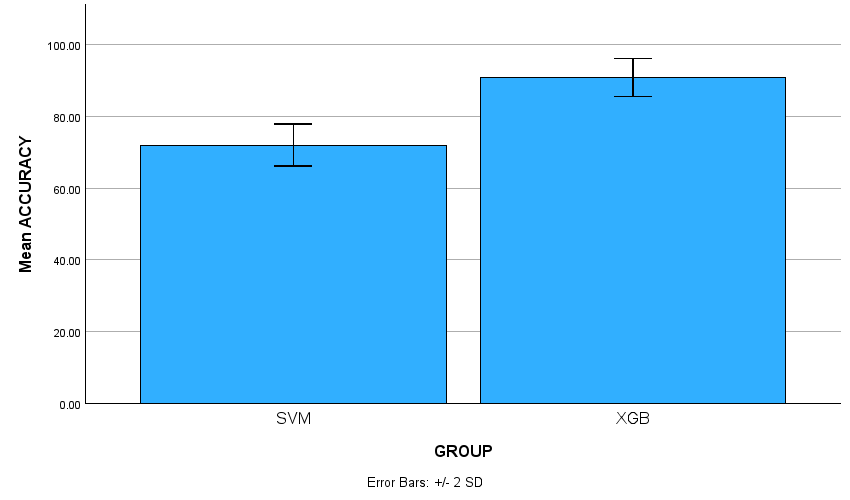
**Table 1.** The XGB(90.50%) method and grouped statistics  were compared using group statistics for recorded data from simulation for 20 iterations (86%). In comparison, the XGB algorithm has a high level of accuracy

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Algorithm** | **N** | **Mean** | **Std. Deviation** | **Std. Error Mean** |
| **Accuracy** | XGB | 10 | 90.8700 | 2.61451 | .82678 |
| SVM | 10 | 72.0230 | 2.92828 | .92600 |

**Table 2.** Independent sample T-test was performed between XGB and SVM, to identify the significance and standard error determination between the algorithms are shown.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Levene’s Test for Equality of Variance | | T-test for Equality of Means | | | | | | |
| f | Sig | t | df | Sig.(2-tailed) | Mean Difference | Std.Error Difference | 95% Confidence of the Differences | |
| Lower | Upper |
| Accuracy | Equal variances assumed | .131 | .722 | 15.182 | 18 | <.001 | 18.84700 | 1.24139 | 16.23893 | 21.45507 |
| Equal variances not assumed |  |  | 15.182 | 17.774 | <.001 | 18.84700 | 1.24139 | 16.23655 | 21.45745 |

**GRAPHS**



**Fig.1.** Comparison of extreme gradient boosting algorithm and Support vector machine algorithm in terms of mean and accuracy. The mean accuracy of the extreme gradient boosting algorithm is better than the Support vector machine. X-axis: extreme gradient boosting algorithm vs Support vector machine algorithm, Y axis: Mean accuracy of detection ±2 SD.