Statistical Evaluation of the Performance of SVM Kernels for Air Quality Classification: case study on India

ABSTRACT

Air pollution has now become a burning problem in the developing world, particularly in countries like India. Deteriorating air quality has led to a widespread increase in diseases related to the lungs in the population, leading to a stupendous increase in the economic burden on the general populace. Agriculture has also led to an increase in air pollution due to the stubble burning activity of the farmers. Every year in the months of winter, air pollution halts all the social and economic activities of major Indian metro cities. Air quality classification has emerged as one of the most significant research and modeling issues as a result of the significant increase in air pollution. Only if the data are properly classified will it be feasible to reduce the impact of air pollution on human health. Accurate air quality classification has become crucial for addressing these problems. However, classifying air quality data is challenging due to class imbalances. In this study, we tackle this issue by using the Support Vector Machine (SVM) approach with various kernels. Statistical evaluation showed that the linear kernel performed best with an accuracy of 0.78, followed closely by the radial kernel. These results suggest that SVM with a linear kernel can improve air quality classification, aiding in better preventive measures and emergency responses during severe pollution events.

Keywords: *Accuracy measure*; *Air Quality Index;Kernels; Multi-class imbalance; Support Vector Machine.*

1.INTRODUCTION

Air quality degradation has become a critical concern, particularly in the context of climate change and urbanization. The increasing frequency of extreme weather events such as droughts, floods, and forest fires is directly linked to rising carbon emissions, a primary driver of climate change [1]. Air pollution, resulting from the emission of gaseous pollutants like carbon dioxide (CO2), is a significant contributor to this environmental crisis [9]. In densely populated nations like India, air pollution is aggravated by anthropogenic activities, including the burning of fossil fuels in the transportation sector and agricultural practices like crop residue burning [38-39]. The resulting deterioration of air quality has led to widespread concern, with both the public and policymakers seeking effective solutions to monitor and mitigate pollution levels.The COVID-19 lockdown in India, which led to temporary improvements in air quality, has heightened public awareness and the desire for cleaner air [10,18,26].

One of the key tools used to measure air quality is the Air Quality Index (AQI), which simplifies complex data about various pollutants into a single value for public understanding [22,27]. However, the accurate classification and prediction of air quality require advanced techniques, especially given the growing volume of environmental data [17,19]. This has led to the application of machine learning models, such as

Support Vector Machines (SVM), for classifying air quality parameters. SVM has gained attention for its ability to handle complex classification tasks, but its performance depends on the choice of kernel functions, which can vary significantly in different contexts. This study focuses on the statistical evaluation of the performance of various SVM kernels in classifying air quality in India, where pollution levels are a major public health and environmental concern.

2.LITERATURE REVIEW

Researchers have turned to machine learning algorithms such as Decision Trees, Support Vector Machines (SVM), and Deep Neural Networks to classify air quality variables [3,5,7,15,30-32]. SVM models have been shown to outperform others machine learning techniques in terms of accuracy and risk reduction, with studies reporting high classification performance [21,28-29,35]. While several researchers have proposed solutions for binary class imbalance problems, these approaches have limitations when addressing multi-class imbalance issues [20-25].

2.1 Problem Statement / Objective of the study

The key problem addressed in this study is the multi-class imbalance in air quality classification. While binary class imbalance has been explored extensively, handling multi-class imbalance remains a challenge, especially when one class has significantly fewer instances than others [20]. The study aims to address this gap by applying Support Vector Machines (SVM) with various kernels, validating their performance using statistical measures to classify air quality variables accurately. This research seeks to contribute to resolving the multi-class imbalance issue in air quality data classification.

3.Methods and Materials

3.1 Data

The data set used for the study of air quality classification for major cities in India which was collected on a daily basis from the Central Pollution Control Board (CPCB). Air pollution levels are obtained through actual observation and published by the National Air Quality Index-CPCB which categorizes levels and air quality features into six air pollution levels. The features selected for the study include PM 2.5, PM10, NO_2, NH_3, CO, SO2, and O_3. The different classes of AQI is given in figure 1 which explains the AQI range, label, levels of class and its health impact.

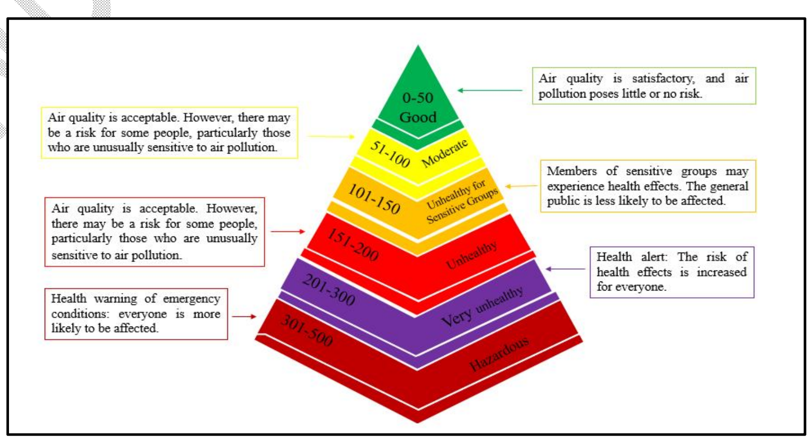


Figure 1. Class wise description of Air quality index

3.2 Support Vector Machine for Classification

Support Vector Machine (SVM) is a popular supervised machine learning method used for classification and regression analysis [4]. It models the relationship between dependent and independent variables in two phases: first, learning the relationship using a training sample, and then using the model for classification and prediction on a test sample. [33] proposed an original linear SVM formulation for separable data. For a given binary classification problem, the objective is to estimate functions f with parameter vector θ such that $\{f(x;\theta): R^n \rightarrow \{-1, +1\}\}$ using a finite set of training data. Let the training data set consist of $\{x_i, y_i\}_{i=1}^{N}$ with input patterns $x_i \in R^n$ and their respective class labels $y \in \{-1, +1\}$. When the training data is linearly separable, a separating hyperplane (a hyperplane that separates the positive from the negative examples) of the form

$$w^T x + b = 0 \qquad (1)$$

can be fitted to correctly classify training patterns, where w^T (the transpose of the weight vector) is normal to the hyperplane and defines its orientation, while b is the bias term that adjusts the position of the hyperplane. From Equation (1), the linear classifier or decision function used to classify the data is expressed as:

$$y(x) = sign(\, w^T x + b) \qquad (2)$$

Which classifying class C2 is defined as($y_i = +1$ if $w^T x + b \geq 0$)and class C_1($y_i = -1$ if $w^T x + b \leq 0$). The hyperplanes $H_+(w^T x + b = +1)$ and $H_-(w^T x + b = -1)$ represent the shortest distances from the separating hyperplane to the closest points of classes C2 and C1. The margin is the sum of these distances, H_+ and H_-, and the goal is to maximize this margin. Finding the optimal hyperplane for a linearly separable dataset involves solving a convex Quadratic Programming (QP) problem, which can be done using Lagrange multipliers and Karush-Kuhn-Tucker (KKT) conditions [16]. The solution is found in the dual space, and nonlinear decision functions can be handled using kernel functions.

3.2.1 Kernel functions

[2] introduced the kernel trick to accomplish generalization of nonlinear SVM. A nonlinear transformation can be done on the set of input vectors to a higher dimensional space (where the dot product is defined) which helps to facilitate the classification of data in linear separation. Figure 2 shows how kernel project the data from original space to transform high dimensional space. To transform the nonlinearly separable data into linearly separable data, the data are mapped in the form $\varphi(x_i): R^n \rightarrow R^{n_{\mathcal{H}}}$ using a nonlinear function $\varphi(x_i)$ into a higher dimensional feature space, which is also a Hilbert space of finite or infinite dimension. Then Lagrange multipliers and Kursh-Kühn-Tucker (KKT) complimentary conditions are used to find the optimal solution.

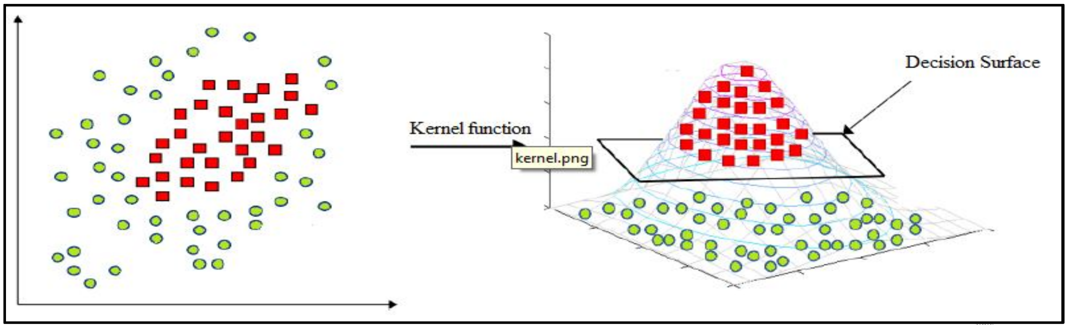


Figure 2. Transformation task using kernel

The choice of an appropriate kernel for SVM models can be challenging, but certain guidelines can help. The Radial Basis Function (RBF) kernel is often a good choice because, unlike the linear kernel, it can model nonlinear relationships between variables. As noted by [14], the linear kernel is a special case of the RBF kernel. Moreover, the RBF kernel involves fewer hyper parameters compared to the polynomial kernel and presents fewer numerical issues. However, when dealing with a large number of features, the linear kernel may be more suitable, as noted by [6].

In multi-class classification, two common strategies are "one-against-one" and "one-against-all" voting schemes [13]. The one-against-one method has shown to provide robust results when combined with SVMs [34,8,12], and therefore, it was the method used in this study.A list of commonly used kernels for classification is provided in Table 1. The steps used for kernel comparison are illustrated in Figure 3. The SVM implementation in this study was performed using the ‘e1071’ package in R version 4.2.2.

Table 1. Some typical choices of kernel function

Kernel Type	Expression
Linear	$K(x, x_i) = x_i^T x$
Polynomial	$K(x, x_i) = (x_i^T x + k)^d$
Radial Basis	$K(x, x_i) = exp\{-\|x - x_i\|^2/2\sigma^2\}$
Sigmoid	$K(x, x_i) = tanh\,(\gamma x_i^T x + C)$

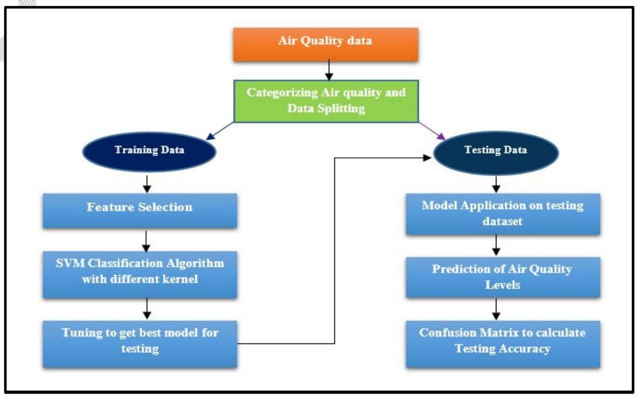


Figure 3. Steps used for kernel comparison

3.3 Statistical Evaluation criteria of Classification of Data

3.3.1 Confusion Matrix

A confusion matrix is a table that displays the performance of a classification model, with predicted classes as rows and actual classes as columns. Each cell shows the number of correct or incorrect predictions made by the model. True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are used to represent the outcomes. Figure 4 illustrates the confusion matrix for multiclass classification.

		ACTUAL					
	Classes	a	b	c	d	e	f
PREDICTED	a	TN	FN	TN	TN	TN	TN
	b	FP	TP	FP	FP	FP	FP
	c	TN	FP	TN	TN	TN	TN
	d	TN	FP	TN	TN	TN	TN
	e	TN	FP	TN	TN	TN	TN
	f	TN	FP	TN	TN	TN	TN

Figure 4. Confusion matrix for multiclass classification.

3.3.2Accuracy measures

In this section, the different type of accuracy measures used to evaluate the performance of SVM classification with different kernels. These measures provide a single value to compare the different methods and identify the more effective one which has been calculated from the confusion matrix [37]. In this study, ten statistical measures such as overall accuracy, precision, recall, F1 score, TNR, NPV, FNR, FDR, and FOR have been considered for evaluation purposes. The formulas of accuracy measures are given in Table 2.

Table 2. The formula for accuracy measures

S.no	Accuracy measures	Formula
1	Overall accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
2	Precision	$\frac{TP}{TP + FP}$
3	Recall	$\frac{TP}{TP + FN}$
4	F1 score	$\frac{2 \times Precision \times Recall}{Precision + Recall}$
5	True Negative Rate (TNR)	$\frac{TN}{TN + FP}$
6	Negative Predictive Value (NPV)	$\frac{TN}{TN + FN}$
7	False Negative Rate (FNR)	$\frac{FN}{FN + TP}$
8	False Positive Rate (FPR)	$\frac{FP}{FP + TN}$

9	False Discovery Rate (FDR)	$\frac{FP}{FP + TP}$
10	False Omission Rate (FOR)	$\frac{FN}{FN + TN}$

4. Results and Discussion

The air quality data is enriched with more features which helps to perform the better classification task. The total number of samples collected for analysis is 29532 from different cities of India. Among them, 14402 samples and 7 attributes have been used for classification purpose after removing the missing values. The attributes of the air quality data include pollutants such as $PM_{2.5}$, PM_{10}, NO_2, NH_3, CO, SO_2 and O_3. Table 3 shows the various features of the data with the help of several parameters, such as name of the parameter with its mean values, median values, measuring unit, standard deviation, and actual and prescribed range of variables.

Table 3. The statistical description of the various parameters

	Mean	Median	Units	Standard Deviation	Actual Range		Prescribed Range	
					Min	Max	Min	Max
$PM_{2.5}$	56.64	42.54	µg/m^3	51.57	0.16	868.66	0.00	60.00
PM_{10}	115.84	94.44	µg/m^3	87.65	0.18	847.41	0.00	100.00
NO_2	28.10	23.64	µg/m^3	20.66	0.01	162.50	0.00	200.00
NH_3	20.78	15.48	µg/m^3	17.35	0.06	207.14	0.00	200.00
CO	1.04	0.82	µg/m^3	1.17	0.00	30.44	0.00	4.00
SO_2	11.50	9.54	µg/m^3	8.53	0.01	113.82	0.00	80.00
O_3	34.96	31.63	µg/m^3	21.84	0.01	257.73	0.00	18.00

The Air Quality Index have divided into six classes such as severe (0), very poor (1), poor (2), moderately polluted (3), satisfactory (4) and good (5) according to the range from 0 to more than 400. Table 4 represents the data in class wise which contains class-wise distribution and the range of six parameters for severe class to good class. It confirms that the presence of imbalance occurred in our data.

Table 4. The concentration ranges of various pollutants for each AQI Class

Name of Class	Samples in each class	$PM_{2.5}$	PM_{10}	NO_2	NH_3	CO	SO_2	O_3
Severe (0)	285	15-685	83-847	0-163	5-112	0-15	6-79	5-258
Very Poor(1)	859	45-350	92-503	0-150	1-167	0-30	3-92	0-200
Poor (2)	1358	17-354	31-627	0-130	1-106	0-16	1-104	0-162
Moderate (3)	5323	7- 869	0-486	0-108	0-207	0-11	0-114	0-172
Satisfactory (4)	5515	3-153	0-388	0-92	0-196	0-6	0-74	0-97
Good (5)	1062	0-97	0-195	0-80	0-83	0-4	0-45	0-45

4.1 Classification Model: Comparison between different kernels

In SVM classification, Identification of best kernel is the important criteria which helps to utilize the features of parameter in proper way. The primary aim of this study to find the appropriate kernel for SVM classification model. The kernels which have been used in this tasks are Radial, Linear, Polynomial and Sigmoid. The performance of kernels is compared by performance validation measures such as precision and accuracy. For validation purpose, the data have divided into training set and test set. Among 14402 samples, the training set had 9602 samples and the 4800 samples was in test set.

As given in Table 5, Four types of kernel were used to classify the data and compare the prediction abilities of the SVM's. The number of support vectors were varied for the SVM classification with different kernels. The number of support vector has given the idea about the data which approach or far away from hyperplane during classification. The underfitting or overfitting of SVM model depends on the number of support vectors. The kernel of the SVM gives best classification which will have medium number of support vectors [36]. From Table 5, the maximum and minimum number of support vectors are 6018 and 4880 respectively. The medium number of support vectors are present in Linear kernel (5187) and Radial kernel (5028). Similarly, the highest number of support vectors was available in class poor and the lowest was in class good.

In multivariate classification, Confusion matrix is the best way to visualize the actual and predicted values of each class. The diagonal of the confusion explains the exact match of actual and predicted and non-diagonal values shows incorrect prediction. The confusion matrix of the different kernels is given in Table 6 which represents the performance of classification task.

Table 5. Parameter Summary of the Classifiers

	SVM-Kernel: Radial	**SVM-Kernel: Linear**	**SVM-Kernel: Polynomial**	**SVM-Kernel: Sigmoid**
Number of Support Vectors	5028	5187	4880	6018
Number of Classes	6	6	6	6
Support vectors class wise				
Severe(0)	1507	1734	1577	1753
Very Poor(1)	396	568	429	693
Poor(2)	1751	1673	1721	2019
Moderate(3)	723	758	692	825
Satisfactory(4)	472	350	360	536
Good(5)	179	104	101	192

4.2 Performance evaluation

The SVM classification result of all kernels has been evaluated in the form of accuracy measures such as overall accuracy, precision, recall, F1 score, TNR, NPV, FNR, FPR, FDR and FOR. Overall accuracy, Precision and recall are the measure of quantity which decides if the classification method provided correct results. F1 score is the harmonic mean of precision and recall. TNR is helped to utilized to quantify the specificity. The classification will consider as the best which classification provides large value of five measures. The remaining five measures such as NPV, FNR, FPR, FDR and FOR calculated the instance of mismatched or false classification. Therefore, the lower values of these five measures represent a good classification algorithm. The results of various measures are calculated from the confusion matrix which is provided in Table 7.

Table 6. Confusion matrix between the actual and predicted values of AQI

Radial Kernel						
True / Pred	**0**	**1**	**2**	**3**	**4**	**5**
0	52	12	1	0	0	0
1	23	207	49	6	1	0
2	2	56	249	102	1	1
3	16	21	125	1382	198	2
4	0	0	3	260	1610	120
5	0	0	0	3	85	213

Linear Kernel						
True / Pred	**0**	**1**	**2**	**3**	**4**	**5**
0	60	17	2	0	0	0
1	31	210	34	0	0	0
2	2	64	269	67	1	0
3	0	5	118	1444	215	3
4	0	0	4	241	1593	152
5	0	0	0	1	86	181

Polynomial Kernel						
True / Pred	**0**	**1**	**2**	**3**	**4**	**5**
0	49	47	2	0	0	0
1	37	151	23	1	0	0
2	7	91	343	326	1	0
3	0	7	49	1180	768	1
4	0	0	9	241	1061	133
5	0	0	1	5	65	202

Sigmoid Kernel						
True / Pred	**0**	**1**	**2**	**3**	**4**	**5**
0	0	3	33	72	6	0
1	3	22	37	48	3	0
2	11	72	75	106	8	0
3	61	162	262	925	455	119
4	18	37	19	466	1114	194
5	0	0	1	137	309	23

Table 7. Performance evaluation of SVM classification based on different kernels

KERNAL	**Overall**	**Precision**	**Recall**	**F1**	**TNR**	**NPV**	**FNR**	**FPR**	**FDR**	**FOR**
Radial	0.77	0.74	0.69	0.71	0.95	0.95	0.26	0.05	0.31	0.05
Linear	0.78	0.75	0.70	0.72	0.95	0.95	0.25	0.05	0.30	0.05
Polynomial	0.62	0.62	0.61	0.62	0.91	0.91	0.38	0.09	0.39	0.09
Sigmoid	0.45	0.26	0.24	0.25	0.87	0.87	0.74	0.13	0.76	0.13

*TNR – True Negative Rate, NPV–Negative Predictive Value, FNR-False Negative Rate, FPR-False Positive Rate, FDR-False Discovery Rate, FOR-False Omission Rate

From this table, the observed range of overall accuracy is 0.45 to 0.78 which is lesser than expectation. Because the classification of air quality index has performed using seven variables instead of eleven variables namely PM $_{2.5,}$ PM_{10}, NO, NO_2, NH3, CO, SO2, O3, Benzene, Toulene, Xylene. Table 7 showed that the highest overall accuracy is observed in Linear kernel (0.78) followed by Radial kernel (0.77). Linear kernel had the highest value of Precision, Recall and F1 score followed by Radial kernel. Based on TNR values, Linear kernel and Radial kernel showed equal performance.

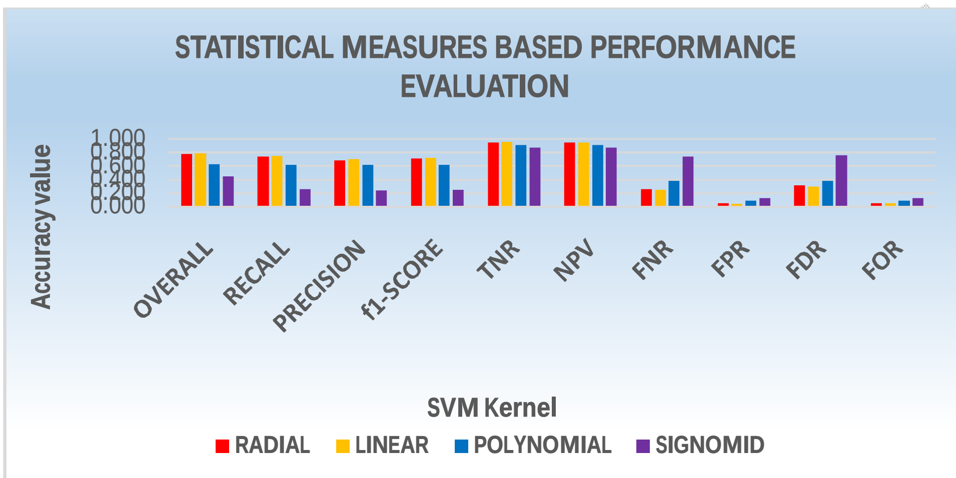
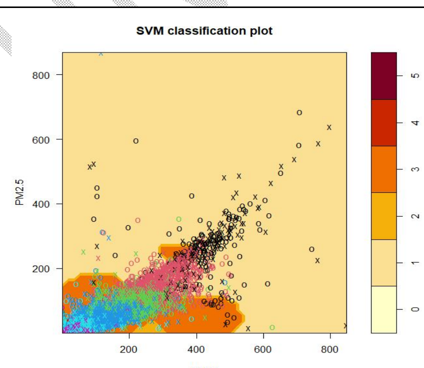
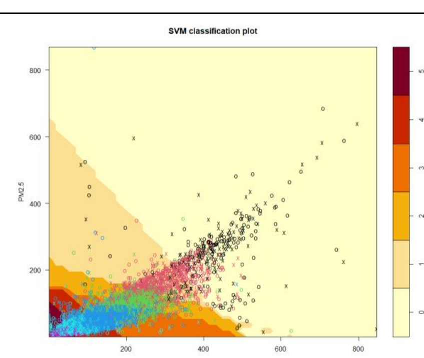


Figure 5. Accuracy measures of SVM classification with different kernels

The lowest value of error measures like FNR, FDR revealed that the linear kernel is better than the radial kernel. The values of NPR, FPR, FOR are equal in both linear and radial kernel. Finally, all measures showed that this air quality index data are suited more to Linear and Radial kernel function compared to Polynomial and sigmoid kernel function (Figure 5). Among these both, Linear is chosen to be the best kernel function. The similar result was reported by [11] for multiclass cancer classification. The graphical representation of the Support Vector Machines (Figure 6) with respect to only two attributes PM_{10} and $PM_{2.5}$, clearly indicates that the performance of the linear kernel is best as there are clear-cut boundaries of the different classes.

SVM-Kernel: radial	***SVM-Kernel: linear***
	

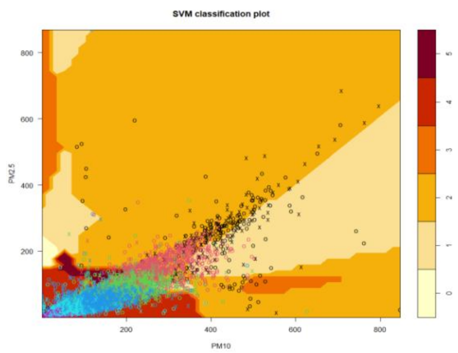
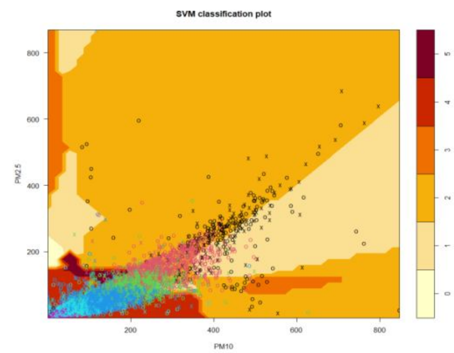
SVM-Kernel: polynomial	***SVM-Kernel: sigmoid***
	

Figure 6. Performance of the classifiers for two variables $PM_{2.5}$ and PM_{10}

5.Conclusions

Classification of Air quality is an important aspect of pollution studies. The complex relationship between the air pollutants and AQI can be automated with the use of Support Vector Machines. Therefore, this study in which different SVM kernels have been tested to access the classification capabilities of SVM becomes pertinent for the planners especially in city areas for round the clock monitoring of the air pollution. The use of Support vector machines can enhance the reach of automated machines to even remote areas. It was found that even when some variables which cannot be measured on a continuous scale are left out and only variables which are measured automatically are used the accuracy of the classification is 78 percent. Among the SVM kernels the linear kernel and the Radial kernel outperformed the other kernels viz polynomial and the sigmoid. The study can further be extended with new data sets to explore the possibilities of using Support Vector machines for automatic forecasting of the AQI.

Statements and Declarations

Data Availability Statement- The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Disclaimer (Artificial intelligence): Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

6.References

1. Aydin M, Turan YE. The influence of financial openness, trade openness, and energy intensity on ecological footprint: revisiting the environmental Kuznets curve hypothesis for BRICS countries. Environ SciPollut Control Ser. 2020;27(34):43233-45.

2. Boser BE, Guyon I, Vapnik V. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. ACM Press; 1992. p. 144-52.

3. Boznar M, Lesjak M, Mlakar P. A neural network-based method for short-term predictions of ambient SO2 concentrations in highly polluted industrial areas of complex terrain. Atmos Environ Part B Urban Atmos. 1993; 27:221-30.

4. Cortes C, Vapnik V. Support-vector network. Mach Learn. 1995; 20:273-97.

5. Mogollon-Sotelo C, Casallas A, Vidal S, Celis N, Ferro C, Belalcazar L. A support vector machine model to forecast ground-level PM2.5 in a highly populated city with a complex terrain. Air QualAtmos Health. 2021; 14(1):399-409.

6. Hsu CW, Chang CC, Lin CJ. A practical guide to support vector classification. Department of Computer Science, National Taiwan University, Taipei 106, Taiwan; 2016.

7. Comrie AC. Comparing neural networks and regression models for ozone forecasting. J Air Waste Manag Assoc. 1997;47(2):653-63.

8. Crammer K, Singer Y. On the learnability and design of output codes for multiclass problems. In: Proceedings of the 13th Annual Conference on Computational Learning Theory. 2000. p. 35-46.

9. Dutta D, Gupta S. Rising trend of air pollution and its decadal consequences on meteorology and thermal comfort over Gangetic West Bengal, India. In: Spatial Modeling and Assessment of Environmental Contaminants. Springer, Cham; 2021. p. 689-720.

10. Das P, Mandal I, Debanshi S, Mahato S, Talukdar S, Giri B, Pal S. Short term unwinding lockdown effects on air pollution. J Clean Prod. 2021; p. 126514.

11. Hatta NM, Shah ZA, Kasim S. Evaluate the performance of SVM kernel functions for multiclass cancer classification. Int J Data Sci. 2020;1(1):37-41.

12. Hsu CW, Lin CJ. A comparison of methods for multi-class support vector machines. IEEE Trans Neural Netw. 2002; 13(1):1045-1052.

13. Karatzoglou A, Meyer D, Hornik K. Support vector machines in R. J Stat Softw. 2006;15(9).

14. Keerthi SS, Lin CJ. Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Comput. 2003;15(7):1667-1689.

15. Knerr S, Personnaz L, Dreyfus G. Single-layer learning revisited: a stepwise procedure for building and training a neural network. In: Fogelman J, editor. Neurocomputing: Algorithms, Architectures and Applications. Springer-Verlag; 1990.

16. Krebel U. Pairwise classification and support vector machines. In: Schölkopf B, Burges CJC, Smola AJ, editors. Advances in Kernel Methods – Support Vector Learning. MIT Press; 1999. p. 255-268.

17. Maciąg PS, Kasabov N, Kryszkiewicz M, Bembenik R. Air pollution prediction with clustering-based ensemble of evolving spiking neural networks and a case study for the London area. Environ Model Softw. 2019;11(8):262-280.

18. Mahato S, Ghosh KG. Short-term exposure to ambient air quality of the most polluted Indian cities due to lockdown amid SARS-CoV-2. Environ Res. 2020;18(8):109835.

19. Mandal S, Madhipatla KK, Guttikunda S, Kloog I, Prabhakaran D, Schwartz JD, et al. Ensemble averaging based assessment of spatiotemporal variations in ambient PM2.5 concentrations over Delhi, India, during 2010–2016. Atmos Environ. 2020; 224:117309.

20. Menardi G, Torelli N. Training and assessing classification rules with imbalanced data. Data Min Knowl Disc. 2014;28(1):92-122.

21. Mehdipour V, Stevenson DS, Memarianfard M, Sihag P. Comparing different methods for statistical modeling of particulate matter in Tehran, Iran. Air QualAtmos Health. 2018;11(1):1155-65.

22. Ontario. A review of the Ontario air quality index and air quality health index system. Air Res Branch, Ontario Ministry of the Environment. 2013. ISBN 978-1-4606-0936-1.

23. Plocoste T, Calif R, Euphrasie C, Brute FN. Investigation of local correlations between particulate matter (PM10) and air temperature in the Caribbean basin using ensemble empirical mode decomposition. AtmosPollut Res. 2020;11(10):1692-704. https://doi.org/10.1016/j.apr.2020.06.031

24. Rybarczyk Y, Zalakeviciute R. Machine learning approaches for outdoor air quality modelling: a systematic review. Appl Sci. 2018; 8:2570.

25. Salazar-Ruiz E, Ordieres JB, Vergara EP, Capuz-Rizo SF. Development and comparative analysis of tropospheric ozone prediction models using linear and artificial intelligence-based models in Mexicali, Baja California (Mexico) and Calexico, California (US). Environ Model Softw. 2008;23(2):1056-69.

26. Sharma S, Zhang M, Gao J, Zhang H, Kota SH. Effect of restricted emissions during COVID-19 on air quality in India. Sci Total Environ. 2020; 728:138878. https://doi.org/10.1016/j.scitotenv.2020.138878

27. Shenfeld L. Note on Ontario's air pollution index and alert system. J Air Pollut Control Assoc. 1970;20(9):622-25.

28. Sihag P, Kumar V, Afghan FR, Pandhiani SM, Keshavarzi A. Predictive modeling of PM2.5 using soft computing techniques: case study—Faridabad, Haryana, India. Air QualAtmos Health. 2019;12(2):1511-20.

29. Suárez A, García PJ, Riesgo P, Del Coz JJ, Iglesias-Rodríguez FJ. Application of an SVM-based regression model to the air quality study at local scale in the Avilés urban area (Spain). Math Comput Model. 2011;54(1):1453-66.

30. Muhammad SY, Makhtar M, Rozaimee A, Aziz AA, Jamal AA. Classification model for water quality using machine learning techniques. Int J SoftwEng Its Appl. 2015;9(6):45-52.

31. Tao Q, Liu F, Li Y, Sidorov D. Air pollution forecasting using a deep learning model based on 1D convnets and bidirectional GRU. IEEE Access. 2019;7(2):76690-98.

32. Vapnik V. Principles of risk minimization for learning theory. In: Lippman DS, Moody JE, Touretzky DS, editors. Advances in neural information processing systems (NIPS). 1992. p. 831-38.

33. Vapnik V. The nature of statistical learning theory. New York (NY): Springer-Verlag; 1995.

34. Weston J, Watkins C. Multi-class support vector machines. Technical report CSD-TR-98-04. Egham (England): Department of Computer Science, Royal Holloway, University of London; 1998. p. 10.

35. Wu LC. Detection of American football head impacts using biomechanical features and support vector machine classification. Sci Rep. 2018;8(1):1-14.

36. Xie X, Ai H, Deng Z. Impacts of the scattered coal consumption on PM2.5 pollutions in China. J Clean Prod. 2020; 245:118922.

37. Yang Z, Wang J. A new air quality monitoring and early warning system: air quality assessment and air pollutant concentration prediction. Environ Res. 2017;158(1):105-17.

38. Zhang Z, Zeng Y, Zheng N, Luo L, Xiao H. Fossil fuel-related emissions were the major source of NH3 pollution in urban cities of northern China in the autumn of 2017. Environ Pollut. 2020; 256:113428.

39. Zheng Y, Huang L, Zhai J. Divergent trends of urban thermal environmental characteristics in China. J Clean Prod. 2021; 287:125053.