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. In many classification issues, the issue of class imbalance exists. Researchers constantly face the difficult issue of learning from unbalanced data, but often, they have produced possible solutions. In this study, our main goal is to deal with the imbalanced air quality data by classifying the data using the supervised machine learning (SVM) approach along with several kernels. The performance of these kernels has been evaluated using various statistical techniques. The results indicated the effectiveness of linear and radial kernels in addressing class imbalance issues and improving performance. As a result, accurate air quality classification using the SVM algorithm with a linear kernel will be beneficial for improving existing preventive policies and enhancing the capabilities of an effective emergency response in the worst pollution situations.

Keywords: Air Quality Index, Classification, Multi-class imbalance, Support Vector Machine, Accuracy measure

IIntroduction:

Climate change is no longer an anticipated future concern, but the world is now encountering the effects of climate change in its destructive incarnation. Every region of the world is experiencing climate emergencies due to the rising frequency of catastrophic climate events including droughts, floods, and forest fires. These changes appear to be caused by an increase in the per capita carbon footprint (Aydin and Turan, 2020). The effects of climate change are distributed globally, with some regions experiencing more severe effects than others. There is a clear connection between poor air quality and climate change. Air pollutants and greenhouse gases are often emitted at the same time, and many air pollutants have direct or indirect climate effects.

Air pollution occurs when air pollutant gaseous pollutants replace the function of air, causing disruptions in daily activities. Carbon emissions, which have led to a decline in air quality in large cities, are one of the main causes of climate change. Main cause being the burning of fossil fuels in the transport sector especially in big cities and industrial belts which has led to air quality degradation (Zhang et al., 2020). In countries like India, where nearly all of the major cities are densely populated and the air quality is worse due to the anthropological activities of the people, the problem of air pollution is more severe (Zheng et al., 2021).

Agriculture is also playing a defining role in enhancing this problem. The practice of burning crop residues in the field in India has caused the nation's capital and surrounding cities' air quality to deteriorate. Initiatives from both science and society are required to address this problem. The general population is now more aware of air pollution and its negative effects on human health, especially in large cities, and they wish to address this problem. After COVID caused a lockdown in India and the subsequent positive effects on air quality in major cities, the desire for pure air grew (Mahato and Ghosh (2020); Sharma et al., (2020); and Das et al., (2021)).

As a result of technical improvements and the use of autonomous devices in various fields, a vast amount of data about pollutant factors is generated. This data cannot be interpreted by the public for their uses; therefore, the concept of the Air Quality Index (AQI) was developed and is being used in almost all countries for measuring air quality (Ontario, 2013; Shenfeld, 1970). The Environmental Protection Agency (EPA) tracks the commonly known criteria pollutants, i.e., ground-level Ozone (O_3), Sulphur Dioxide (SO_2), Fine Particulate Matter (PM_{10} and $PM_{2.5}$), Carbon Monoxide (CO), Carbon Dioxide (CO_2), and Nitrogen Dioxide (NO_2), etc. (Xie and Deng, 2020; Dutta and Gupta, 2021). There are various air quality indices used by various nations, each of which corresponds to a specific standard (Rybarczyk and Zalakeviciute, 2018). reviewed a selection of the 46 most relevant journal papers and found more studies with O_3, NO_2, PM_{10}, and $PM_{2.5}$ and less on an overall AQI.

In India, many studies have been conducted to study the behavior of these pollutants (Mandal et al., (2020); Plocoste et al., (2020)). There are multiple ways in which this index is calculated, some agencies are using only five parameters for constructing the index, while others are using various parameters. Only PM_{10}, $PM_{2.5}$, SO_2, NO_2, CO, and O_3 are measured on a continuous basis at many air quality stations (including NH_3 at some stations); Pb, Benzene, Toulene, and Xylene are monitored using manual systems that do not give us continuous updates. Therefore, it is ideal to choose characteristics that are measured continuously to generate the AQI at short time intervals.

Currently, air quality monitoring is required to accurately anticipate pollutant concentrations and assess the level of air pollution. An accurate and effective modeling method or technique is required to categories air quality in order to improve existing monitoring efforts (Yang and Wang, 2017). One of the most important tasks is the classification of new items based on analogous cases. Machine learning algorithms like Decision Tree, Support Vector Machine (SVM), and Deep Neural Networks are frequently used in classification research for binary class balancing instances (Muhammad et al., 2015). In terms of risk reduction and generalization capacities, SVM models perform better than other machine learning models (Vapnik 1992; Suarez et al. 2011). This is supported by studies where the classification performance with high accuracy is above 90%, 93.2%, 97.01%, 98.5%, and 90%, such as (L. C. Wu (2018); Minarno et al., 2020; Mehdipour et al. 2018; and Sihag et al. 2019). SVM is not only known for its high accuracy but its applicability is also investigated as evidence of this method's superiority in various classification cases.

Many researchers had contributed their approaches, which were accurate to some extent (Iskandaryan D et al., 2020; Soh et al., 2019; Zhang Y et al., 2019). Most of the solutions were proposed for the binary class imbalance problems, however, they weren't appropriate for multi-class imbalance issues. The major issue with the class imbalance is that most data samples belong to one class while the remaining samples belong to different classes. The classification task becomes more complicated when one of the classes contains fewer instances than the other class (Menardi& Torelli, 2014). These limitations motivate us to deal with multi-class imbalance problems and also encourage us to give a possible contribution that will be able to solve the multi-class imbalance problem. In this study, the classification of air quality variables was done using the SVM approach, with various kernels that had been validated for their classification performance using different statistical measures.

I) Methods and Materials

a) 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.

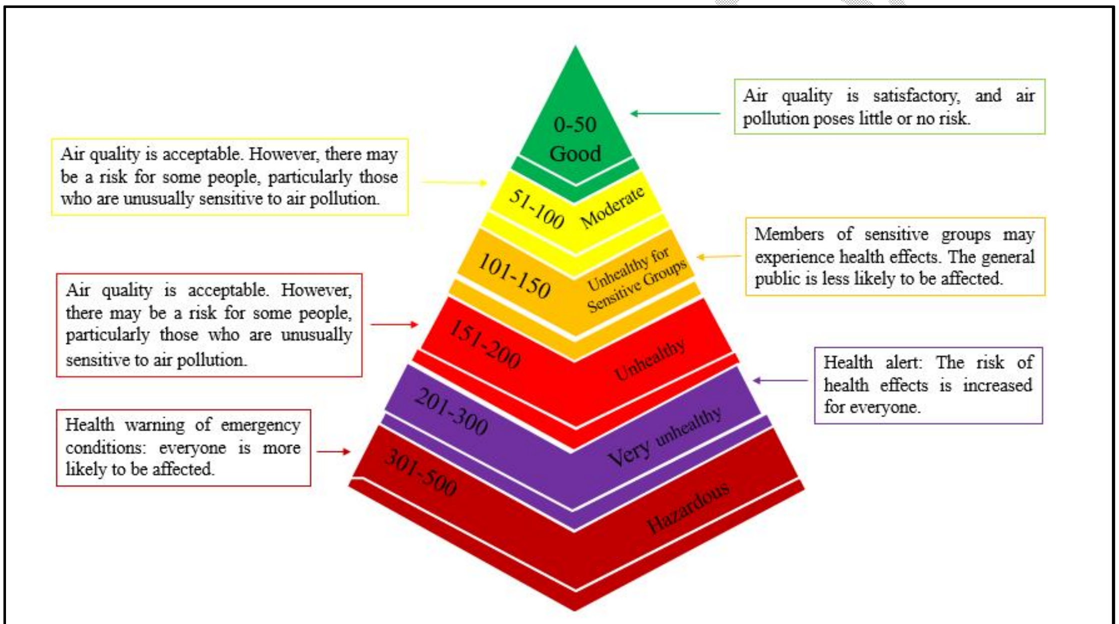

Figure 1. Class wise description of Air quality index

Support Vector Machine for Classification

Support Vector Machine is the supervised machine learning method that is more popular in Classification and Regression analysis. The basic idea of supervised learning is to model the unknown relationship between the known dependent and independent variables. It can achieve through the two phases. In the first phase, the training sample is used to learn about the relationship between the dependent and independent variables. In next phase, the identified relationship in the form of a model is used for classification and prediction purposes in the test sample. Vapnik and Chervonenkis (1964) 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 weight vector w is normal to the hyperplane, and defines its orientation, b is the bias term and T is the transpose. From Equation 1 a linear classifier (decision function) given by

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

Which classifying class C$_2$($y_i = +1$ if $w^T x + b \geq 0$)and class C$_1$($y_i = -1$ if $w^T x + b \leq 0$)patterns. Let $H_+(w^T x + b = +1)$ and $H_-(w^T x + b = -1)$ the shortest distance from the separating hyperplane to the closest class 2 (class 1) example, then the margin of the hyperplane is defined as the sum of $H_+$and H_-, i.e. $H_+ + H_-$. An optimal hyperplane for a linearly separable set of training data is defined as the linear decision function with maximal margin between the vectors of two classes, To find the optimal separating hyperplane, it is necessary to maximize the margin $H_+ + H_-$. Constructing the optimal hyperplane is therefore convex Quadratic Programming (QP) problem. We call this optimization problem is in the primal formulation. Therefore SVM formulations are done within a context of convex QP optimization problem. This is a constrained optimization problem in the primal weight space. Then Lagrange multipliers and Kursh-Kühn-Tucker (KKT) complimentary conditions are used to find the optimal solution. Under the condition for optimality, above QP problem is finally solved in the dual space of Lagrange function. Linear SVMs can very easily be generalized to include these nonlinear decision functions with the help of a new function which is called a kernel.

Kernel functions

Boser *et al.* (1992) 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.

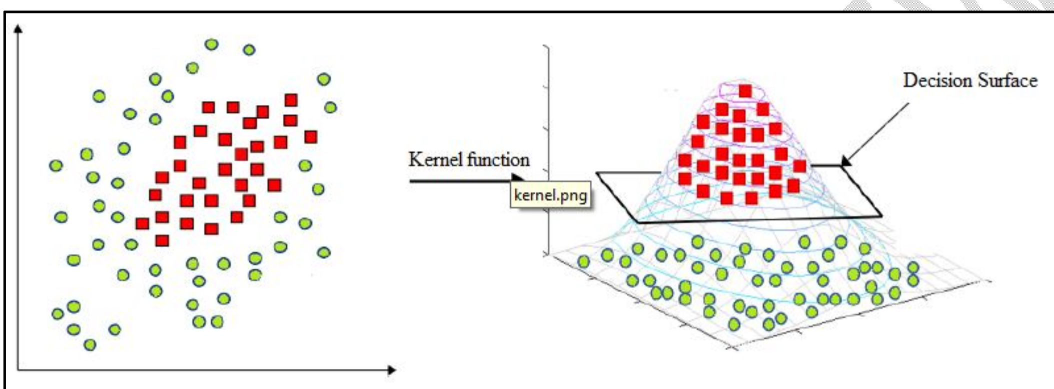

Figure 2. Transformation task using kernel

Choice of kernel is a tricky issue but there are certain guidelines which can be used. Usually RBF is a reasonable choice, because unlike the linear kernel it can model the nonlinear relationships between the variables also. It is also stated by Keerthi and Lin (2003) that a linear kernel is a special case of RBF. Further the number of hyper parameters is also less in case of RBF as compared to polynomial kernel and it has fewer numerical difficulties also. In some cases, RBF kernel is not suitable specially when the number of features is very large, in that case, linear kernel is more suitable (Hsu et al., 2016).

In multi-class classification, one against one and one against all, voting schemes are used Karatzoglou,Meyer, and Hornik,2006. In one against-one method which is also called pairwise classification (Knerr et al., 1990; Weston and Watkins, 1998; Crammer and Singer, 2000; Krebel, 1999), each classifier is trained on data from two classes resulting in ${}^{k}C_{2}$ classifiers. In the other method i.e one against all method each classifier is trained to separate one class form the rest resulting in k binary SVM classifiers, where k is the number of classes. The one against one classification has produced robust results when used with SVMs (Hsu and Lin, 2002), and therefore one against one classification was used in this study. The list of Kernels is mostly utilized for classification purpose in literature which is given in table 1.

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	$K(x, x_i) = tanh\,(\gamma x_i^T x + C)$

The comparison of kernel has done by shown steps in Fgure 3. For this purpose, SVM has performed using 'e-1071' package of R version 4.2.2.

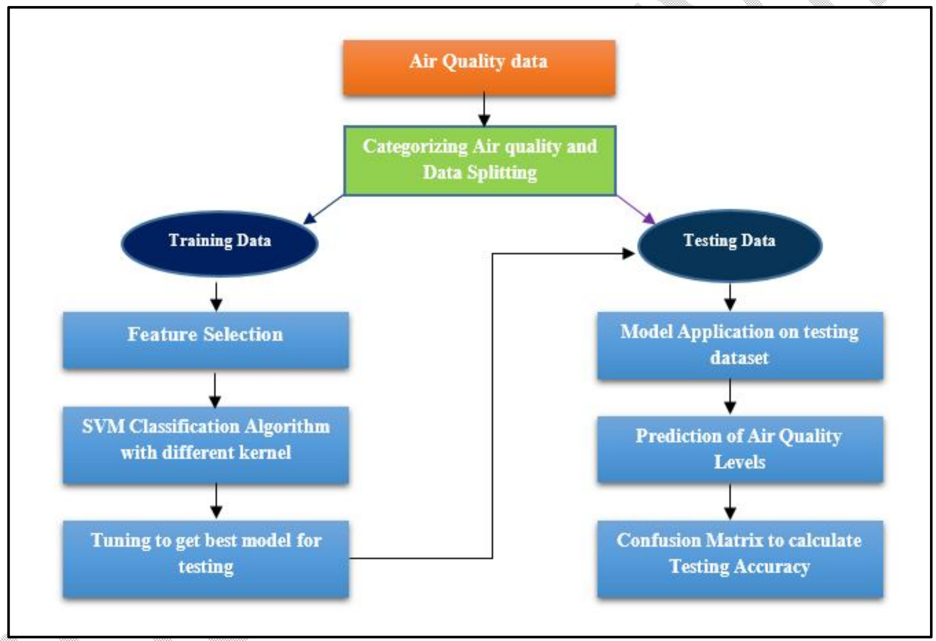

Figure 3. Flow chart of SVM Classification using different kernels

Statistical Evaluation criteria of Classification of Data

Confusion Matrix

Confusion matrix is defined as a matrix which is formed by the instances of predicted class as rows and instances of actual class as columns. It is the tabular way to visualize the performance of classification model. Each cell in a confusion matrix denotes the number of predictions by the model whether it classified the classes correctly or incorrectly. Figure 4 depicted the confusion matrix of 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 of Classification with reference to class b

Accuracy 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. 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 equation of accuracy measures is 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)	$\frac{FN}{FN + TN}$

*TP-True Positive class, TN-True Negative class, FP-False Positive class, FN-False Negative class

III 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

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 (Samsudin et al., 2021). 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

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.

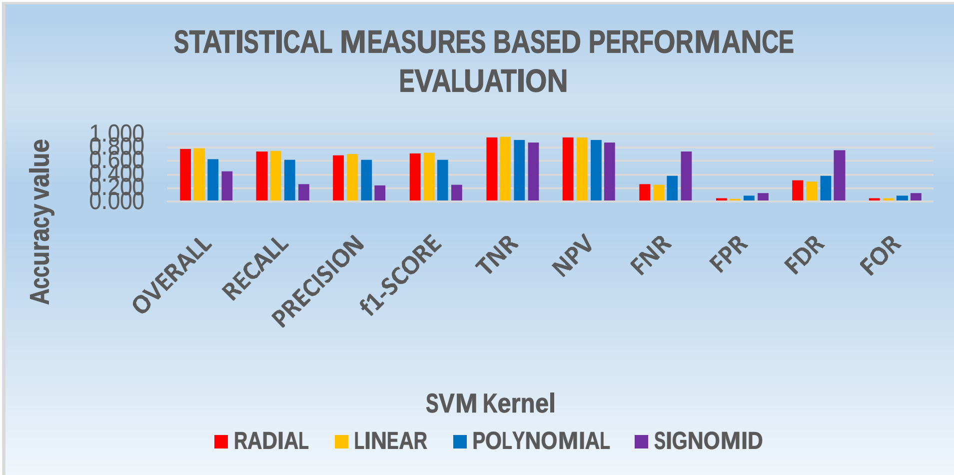

Figure 5. Accuracy measures of SVM classification with different kernels

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 Hatta et al., (2020) 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***

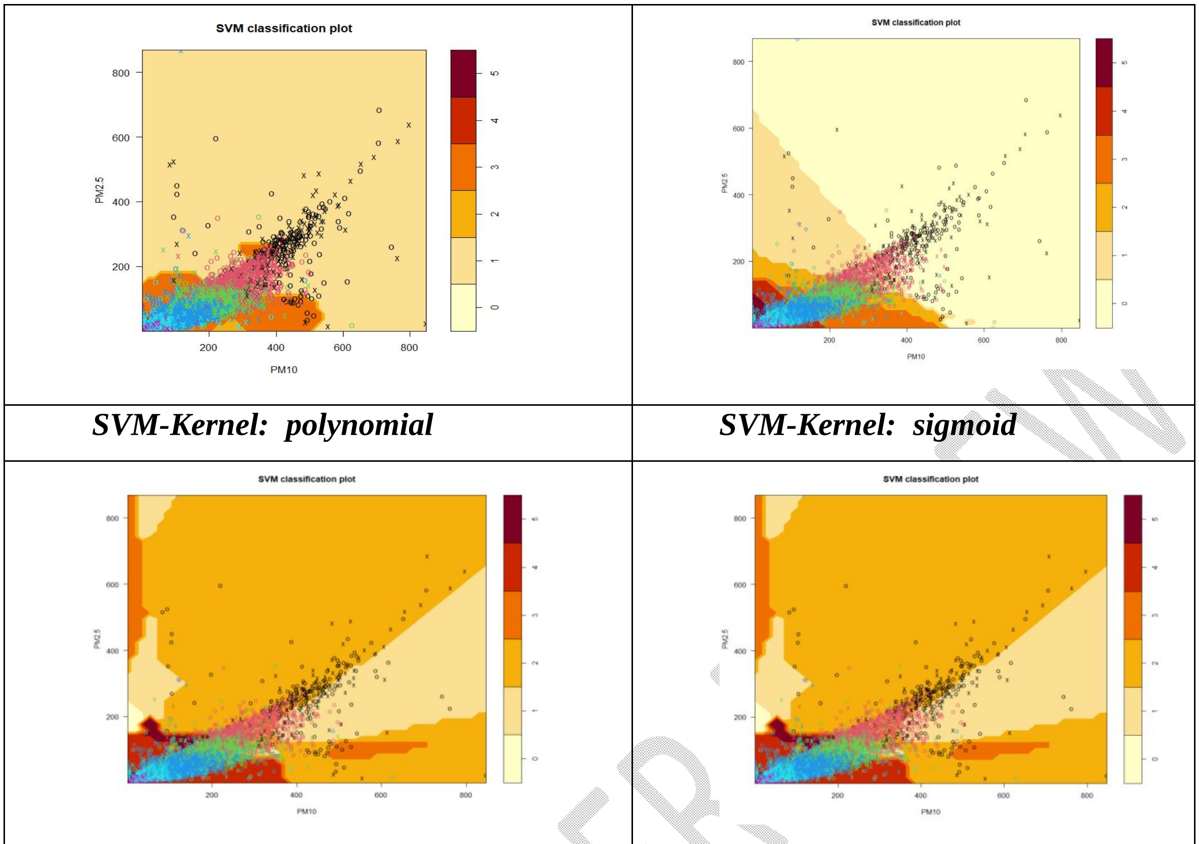

Figure 6. Performance of the classifiers for two variables $PM_{2.5}$ and PM_{10}

II) 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.

References:

Aydin, M. and Turan, Y. E. 2020. The influence of financial openness, trade openness, and energy intensity on ecological footprint: revisiting the environmental Kuznets curve hypothesis for BRICS countries.*Environ. Sci. Pollut. Control Ser.*, 27 (34), pp. 43233-43245.

Boser,B. E., Guyon,I. and Vapnik, V. 1992. A training algorithm for optimal margin classifiers. *In Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, ACM Press 144-152.

Boznar, M., Lesjak, M. and Mlakar, P. 1993. 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 Atmosphere* 27:221–230.

C. Cortes., and V. Vapnik. 1995.Support-vector network. *Machine Learning*, 20:273-297.

Caroline Mogollón-Sotelo, Alejandro Casallas, Sergio Vidal, Nathalia Celis, Camilo Ferro and Luis Belalcazar. 2021. A support vector machine model to forecast ground-level PM2.5 in a highly populated city with a complex terrain. *Air Quality, Atmosphere & Health*. 14:399–409.

Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin. 2016. A Practical Guide to Support Vector Classification. *Department of Computer Science. National Taiwan University, Taipe*i 106, Taiwan.

Comrie, A.C. 1997. Comparing neural networks and regression models for ozone forecasting. *J Air Waste Manage Assoc*, 47:653–663. Crammer, K. and Singer, Y. (2000). On the learnability and design of output codes for multiclass problems. *In: Proceedings of the 13th Annual Conference on Computational learning theory*. pp. 35–46.

Dutta, D. and Gupta, S. 2021. Rising trend of air pollution and its decadal consequences on meteorology and thermal comfort over gangetic West Bengal, India**.***Spatial Modeling and Assessment of Environmental Contaminants*, Springer, Cham ,pp. 689-720.

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

Hatta, N, M., Shah, Z, A., and Kasim, S. 2020. Evaluate the Performance of SVM Kernel Functions for Multiclass Cancer Classification. *International Journal of Data Science,* 1(1): 37-41.

Hsu, C.W. and Lin, C.J. 2002. A Comparison of Methods for Multi-class Support Vector Machines. *IEEE Transactions on Neural Networks*, 13, 1045–1052. URL

Karatzoglou, A. Meyer, D. and Hornik, K. 2006. Support Vector Machines in R.*Journal of Statistical Software*, 15(9).

Keerthi, S. S. and Lin, C.J. 2003. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Computation,* 15(7):1667-1689,

Knerr, S.Personnaz, L. and Dreyfus, G. 1990. Single-layer Learning Revisited: A Stepwise Procedure for Building and Training a Neural Network.*In J Fogelman* (ed.), Neurocomputing:Algorithms, Architectures and Applications, Springer-Verlag.

Krebel, U. 1999. Pairwise Classification and Support Vector Machines. *In B Sch¨olkopf, CJC Burges, AJ Smola (eds.),* Advances in Kernel Methods – Support Vector Learning, 255–268. MIT Press, Cambridge, MA.

Maciąg, P.S. Kasabov, N. Kryszkiewicz, M. and Bembenik, R. 2019. Air pollution prediction with clustering-based ensemble of evolving spiking neural networks and a case study for the London area. *Environ Model Softw*, 118:262–280.

Mahato, S. and Ghosh, K.G. 2020. Short-term exposure to ambient air quality of the most polluted Indian cities due to lockdown amid SARS-CoV-2. *Environ Res.*, 188: 109835

Mandal, S. Madhipatla, K.K. Guttikunda, S. Kloog, I. Prabhakaran, D. Schwartz, J.D. and Team G.H.I. 2020**.** Ensemble averaging based assessment of spatiotemporal variations in ambient $PM_{2.5}$ concentrations over Delhi, India, during 2010–2016. *Atmos. Environ.*, 224: 117309

Menardi, G. and Torelli, N. 2014. Training and assessing classifcation rules with imbalanced data. *Data Min Knowl Disc*, 28(1):92–122.

Mehdipour, V. Stevenson, D.S. Memarianfard, M. and Sihag, P. 2018. Comparing different methods for statistical modeling of particulate matter in Tehran, Iran. *Air Qual Atmos Health*, 11:1155–1165.

Ontario, 2013. A review of the Ontario air quality index and air quality health index system. *Air Resource Branch, Ontario Ministry of the Environment, Toronto, Ont., Canada.* ISBN978-1-4606-0936-1.

Plocoste, T., Calif, R., Euphrasie, Clotilde, L., and Brute, F.N. 2020.Investigation of local correlations between particulate matter (PM10) and air temperature in the Caribbean basin using Ensemble Empirical Mode Decomposition.*Atmospheric Pollution Research*, 11(10): 1692-1704, https://doi.org/10.1016/j.apr.2020.06.031

Rybarczyk, Y., and Zalakeviciute, R. 2018. Machine Learning Approaches for Outdoor Air Quality Modelling: A Systematic Review. *Appl. Sci*. 8: 2570.

Salazar-Ruiz, E. Ordieres. J.B. Vergara, E.P. and Capuz-Rizo, S.F. 2008. 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*, 23:1056–1069.

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

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

Sihag, P. Kumar, V. Afghan, F.R. Pandhiani, S.M. and Keshavarzi, A. 2019. Predictive modeling of PM2.5 using soft computing techniques: case study—Faridabad, Haryana, India. *Air Qual Atmos Health*, 12:1511–1520.

Suárez, A. García, P.J. Riesgo, P. Del Coz, J.J. and Iglesias-Rodríguez, F.J. 2011. Application of an SVM-based regression model to the air quality study at local scale in the Avilés urban area (Spain). *Math Comput Model*, 54:1453–1466.

Muhammad, S. Y., Makhtar, M., Rozaimee, A., Aziz, A. A., and Jamal, A. A 2015. Classification model for water quality using machine learning techniques. Int. J. Softw. Eng. its Appl., 9(6): 45–52.

Tao, Q. Liu, F. Li, Y. and Sidorov, D. 2019. Air pollution forecasting using a deep learning model based on 1D convnets and bidirectional GRU.*IEEE Access* 7:76690–76698.

Vapnik, V. 1992Principles of risk minimization for learning theory. In:D. S. Lippman, J. E.Moody, & D. S. Touretzky (Eds.), Advances in neural information processing systems (NIPS), pp 831–838

Vapnik.V. 1995. *The Nature of Statistical Learning Theory. Springer-Verlag, New York,NY.*

Weston, J. and Watkins, C. 1998.Multi-Class Support Vector Machines. *Technical Report CSD-TR-98-04*, Department of Computer Science, Royal Holloway, University of London, Egham, England, p. 10.

Wu, L. C. 2018. Detection of American Football Head Impacts Using Biomechanical Features and Support Vector Machine Classification. *Sci. Rep.*, 8(1): 1–14.

Xie X., Ai H. and Deng Z. 2020. Impacts of the scattered coal consumption on $PM_{2.5}$ pollution in China. *J . Clean. Prod.*, 245 , p. 118922.

Yang, Z., and Wang, J. 2017. A new air quality monitoring and early warning system: Air quality assessment and air pollutant concentration prediction. *Environ. Res.*, 158: 105–117.

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

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

Zhou Y., Chang F.J., Chang L.C, Kao I.F., Wang Y.S. and Kang C.C. 2018. Multi-output support vector machine for regional multi-step-ahead PM2.5 forecasting.*Sci Total Environ* 651:230–240.