



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 14, Issue 8, August 2025**

# **Performance Evaluation of SVM and RF Models in COPD Severity Classification**

**Meghana K R, Prashanth Ankalkoti**

Student, Department of MCA, JNN College of Engineering, Shivamogga, Karnataka, India

Assistant Professor, Department of MCA, JNN College of Engineering, Shivamogga, Karnataka, India

**ABSTRACT:** COPD is a progressive respiratory disorder that makes breathing increasingly difficult as the condition advances. It is commonly associated with factors such as cigarette smoking, industrial exposure, and polluted environments. The present work focuses on classifying the severity of COPD into four categories: mild, moderate, severe, and very severe. For this purpose, machine learning techniques Support Vector Machine and Random Forest are employed using a publicly available dataset obtained from Kaggle. A comparative study of these two models is carried out to evaluate their performance in terms of accuracy and reliability. The outcome of this research provides insights into the more effective algorithm that can support healthcare practitioners in the timely assessment of COPD.

**KEYWORDS:** Chronic Obstructive Pulmonary Disease, Support Vector Machine, Random Forest, Machine Learning, Comparative Analysis

## **I. INTRODUCTION**

Chronic Obstructive Pulmonary Disease is a widespread lung disorder that reduces airflow and causes persistent breathing difficulties. The risk is particularly high among individuals exposed to harmful industrial gases, dust particles, or polluted surroundings, as well as habitual smokers. Since COPD is a progressive disease, early and accurate diagnosis is crucial to prevent complications and improve quality of life. However, current diagnostic methods largely rely on manual interpretation of spirometry results by medical professionals, which can lead to inconsistent evaluations, late identification, and inappropriate treatment decisions. The shortage of lung specialists in many healthcare settings further adds to the challenge of timely diagnosis. To address these gaps, the present study explores the development of an intelligent system that automates the analysis of breathing test results. Specifically, Support Vector Machine (SVM) and Random Forest (RF) algorithms are applied to classify COPD severity into different stages. By assessing their performance on parameters such as accuracy, sensitivity, and robustness, the study aims to determine the most effective model that can serve as a reliable tool for assisting clinicians in decision-making..

## **II. LITERATURE SURVEY**

Sandeep Bodduluri et al.[1] This research paper examines how deep neural networks and machine learning models are employ to detect COPD through spirometry data. I also see the advancements in the abilities of DNNs and ML models identify this condition. This study used both the CT scan and Spirometer data for the analysis. This paper uses the models FCN and random forest classifier for the identification of COPD. In the end of the analysis the Flow- volume curve gave the best result in identify the COPD.Ching-Hsiung Lin et al. [2] This research paper uses an app-based spirometer called as Spirobank which is a portable app. The dataset uses the CAT questionnaires score and readings from the standard diagnostic spirometer data. This paper used the logistic regression to the analysis. This also used to detect the risk factor associated with COPD. This work is done to verify the app ‘Spirobank’ works effectively for the identification of COPD. Kofi Odame et.al [3] This paper researches the severity of the COPD patients. The dataset collected is the tidal breathing parameters which is from the wearable device. This input is given to the ML algorithm which is proposed for the process the input for classify the severity of the COPD. This research gave 94.6% accuracy. Jorge L.M. Amaral et al. [4] In this paper it researches on the automatic identification of the COPD disease. The dataset collected in this paper is the spirometry data which contains parameters that are necessary for the identification. They used the ML models like SVM, DT, and RF. For robust model performance they used the cross- validation techniques. This result predicts the severity of COPD. Yinhe Feng et al. [5] This paper researches focus on the diagnosis and monitor of Asthma and COPD using AI and ML technologies. This focuses on the areas like screening, diagnosis. The



datasets are trained on clinical, spirometer dataset and from some generic information. This help as the warning system to the patients. This helps the patients to decrease the hospital visit. Sudipto Bhattacharjee et.al [6] The study developed machine learning models including SVM, Random Forest, Naive Bayes, and MLP to classify obstructive and non-obstructive pulmonary diseases using spirometry data. The MLP model achieved the highest accuracy of 83.7% and was also validated on a blind dataset. Their model was deployed as a web application to aid early diagnosis for clinicians and patients. Leeran Talker et.al [7] This paper presents the comparison of the COPD classification using the different ML and DL methods. Thia is done over the different datasets like bio samples and all which are evaluated by the models like SVM and CNN and so on. Many models showed the high accuracy. Ramadoss Ramalingam et al. [8] This paper done the work by using the Fractional dynamics for predicting the COPD. This used the 2 datasets both datasets contain the Phycological signals. For the feature extraction they used the couple matrix. They worked on the FDDL, LSTM and CNN. Their model got 98% of accuracy. Chenzhong Yin et al. [9] This paper researches and proposed the algorithms like SVM, RF, and Naive Bayes and MLP which are the Machine learning models for the obstructive and non-obstructive COPD classification. This uses 1163 patients spirometry data as the dataset. Along these algorithms the MLP gave the high accuracy of about 93%. Taloba, A. I., et.al [10] The paper worked on proposing the ML framework for COPD prediction early using the Spirometer data. MFCCs are used for the feature extraction in this data. The classification is done by Hybrid SVM and KNN model. This approach led to achieve 94-98.9% accuracy in early prediction of COPD.

### III. METHODOLOGY

In this detection of the severity level of the COPD, SVM and RF algorithms are used for the prediction. The algorithms are chosen to compare the performance for this severity level prediction in COPD disease.

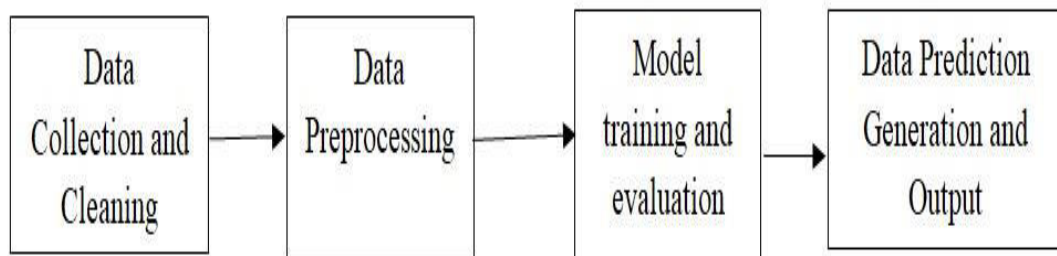


Fig 1: Methodology of proposed system

The following steps are involved in this project,

- Data collection and cleaning:  
Collection of data (spirometry data) from kaggle and cleaning by retaining the required data.
- Data preprocessing:  
Handling missing values, mapping the vales to the numerical form and encoding the numerical format for ML algorithms SVM and RF.
- Model training and Evaluation:  
Splitting the data for training and testing and initializing the Machine Learning models for prediction.
- Prediction generation and Output:  
prediction of severity by the models trained and saving the data frame in a CSV file.
- Tools and libraries used for this are python 3.8, PyCharm, pandas, numpy, matplotlib, and seaborn and so on.

#### IV. EXPERIMENTAL RESULTS

The study on the performance evaluation of SVM and RF on the COPD severity level prediction shows the random forest model gives the more accuracy than the Support vector machine model. The SVM gave 94.29% of accuracy and the RF gave the accuracy of 77.14% of accuracy. By this we can conclude that RF model predicts the Severity of the COPD stage better than the SVM model.

Author	Method used	Efficiency
Sudipto Bhattacharjee et.al [6]	SVM, RF Naive Bayes and MLP	83.7%
Chenzhong Yin et al [9]	SVM, Naive Bayes and MLP	93%
Taloba, A. I., et.al	Hybrid model	94-98.9%
Our system	SVM and RF	94.29%

Table 4.1: Results of other existing system

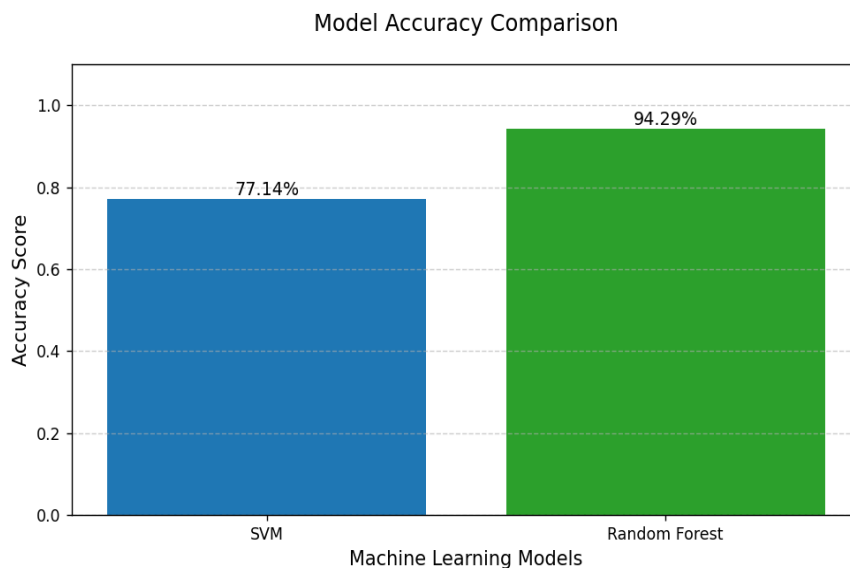


Fig 2: Model accuracy graph

#### V. CONCLUSION

The final conclusion of the paper is that the used algorithms of ML Support Vector Machine and Random Forest are correctly detecting the COPD severity using the Spirometry data. On processing the code for this objective, the Random Forest outperformed better than the Support vector Machine algorithm. The Random Forest accuracy for this is 94.29%. In further enhancement the web application or mobile app for this may help patients to get the severity of COPD

**REFERENCES**

1. Sandeep Bodduluri, Purushotham V. Bangalore, Surya P. Bhatt “Deep neural network analyses of spirometry for structural phenotyping of chronic obstructive pulmonary disease” JCI Insight.;5(13): e132781. 2020
2. Ching-Hsiung Lin et al. “Novel App-Based Portable Spirometer for the Early Detection of COPD” by Diagnostics, 11, 785. 2021
3. Kofi Odame et.al “Inferring COPD severity using tidal breathing” arXiv:2011.07684v3 [eess.SP] 24 Apr 2021
4. Jorge L.M. Amaral et al. “Machine learning algorithms and forced oscillation measurements applied to the automatic identification of chronic obstructive pulmonary disease” computer methods and programs in biomedicine 105 :183-193 2021
5. Yinhe Feng et al. “Artificial Intelligence and Machine Learning in Chronic Airway Diseases: Focus on Asthma and Chronic Obstructive Pulmonary Disease” International Journal of Medical Sciences; 18(13): 2871-2889. 2021
6. Sudipto Bhattacharjee et.al “Classification of the obstructive and non-obstructive pulmonary diseases on the basis of spirometry using ML techniques” Journal of Computational Science Volume 63, September, 101768 2022
7. Leeran Talker et.al “Machine diagnosis of COPD using novel fast response using Capnometer” Respiratory Research 24:150 2023
8. Ramadoss Ramalingam et al. “A comparative analysis of chronic obstructive pulmonary disease using machine learning and deep learning” International Journal of Electrical and Computer Engineering (IJECE) Vol. 13, No. 1,2023
9. Chenzhong Yin et al. “Fractional dynamics foster deep learning of COPD stage prediction” arXiv:2303.07537v1 [cs.LG] 13 Mar 2023
10. Taloba, A. I., et.al “Detecting respiratory diseases using machine learning-based pattern recognition on spirometry data” Alexandria Engineering Journal 113; 44–59, 2025



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details