International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

# Early Chronic Pulmonary Disease Prediction System Using Machine Learning

***Nithya S[1],Risika V[2] , Muthukumar T[3]***
[1]*UG - Biomedical Engineering, Kongunadu College of Engineering and Technology, Trichy, Tamil Nadu – 621215, India*
[2] *UG - Biomedical Engineering, Kongunadu College of Engineering and Technology, Trichy, Tamil Nadu – 621215, India*
[3]*Assistant Professor - Biomedical Engineering, Kongunadu College of Engineering and Technology, Trichy,*
*Tamil Nadu -621215, India*

## Abstract
In this project motive is how to predict early stage in COPD disease. It know about chronic is a long term disease. It can be used to respiration sensor, gas sensor, smoke sensor, transformer, Arduino Uno, IoT module and LCD display. Smart breath analyzers are becoming into sophisticated diagnostic instruments. Respiration sensor is difficult to breath to be utilized. The gas sensor is detected to the gases in the environment. Arduino Uno is a controller it can be control with all the sensors. IoT module is collected all the sensor values. The LCD display shows to the sensor values. These gadgets allow for prompt medical intervention by continuously monitoring health conditions and forecasting when illness like Chronic Pulmonary Disease (COPD) will worsen. The system analyses patterns in the concentrations of gas, smoke and acetone from exhaled breath that is collected using a sampling bag and assessed using a gas-mixing bench using machine learning algorithms based on decision trees. The incorporation of Decision Tree algorithms facilities real-time pattern recognition and anomly detection, providing personalized health insights for high-risk patients. The models were trained on environmental and patient-specific datasets to enhance prediction accuracy. Prototypes incorporating a arduino uno were tested in diverse settings, including indoor and outdoor environments. The findings show that ML-driven smart breath analyzers enhance proactive healthcare and contribute to a more dependable and efficient telemedicine ecosystem by improving early COPD detection and monitoring.
**Keywords:** Machine learning, Smart breath analyzers, Detection tree algorithms, COPD prediction.

## 1. INTRODUCTION
Real-time clinical parameter monitoring, especially for the prevention and treatment of respiratory diseases like Chronic Obstructive Pulmonary Disease (COPD), has been made possible by the quick development of e-Health systems and smart sensor devices.
The respiratory rate is one of the most overlooked metrics in clinical practice, even though it is an essential vital sign for detecting and treating such illnesses. Early identification of exacerbations in COPD, sleep apnoea, cardiovascular illnesses, and metabolic disorders requires ongoing respiratory surveillance.

DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

When combined with Internet of Things (IoT) and machine learning (ML) technology, these devices offer continuous, real-time, non-invasive respiratory monitoring, which enhances patient care and diagnostic precision. The system can identify early warning indicators of COPD exacerbations and other respiratory difficulties by utilising sensor data and machine learning (ML)-driven pattern recognition. This enables prompt medical intervention. The system is further improved by the Internet of Medical Things (IoMT), which makes it possible for intelligent filtering, smooth data transfer, and individualised healthcare insights. The use of Decision Tree-based machine learning models in IoT-driven e-Health ecosystems is a major advancement in respiratory illness monitoring, especially considering the rising incidence of COPD and respiratory disorders linked to air pollution. This study examines the incorporation of machine learning (ML)-based smart breath analysers.

This accelerated loss of lung function may contribute to the development of COPD and consistent with this literature, though studies are not definitive, results suggest that outdoor air pollution exposure is linked to COPD incidence and prevalence. A recent meta-analysis showed that there was a trend towards increased prevalence of COPD, defined as chronic bronchitis or bronchitic symptoms, with higher PM exposure, but this association did not reach statistical significance. Long-term exposure to outdoor air pollution has also been linked to lung function decline and short-term changes in outdoor air pollutant concentrations have been associated with acute changes in lung function and increased respiratory symptoms among individuals with established COPD.

There have been studies linking secondhand smoke (SHS) exposure to the development of COPD, and among those with COPD there is evidence suggesting the SHS exposure contributes to worse quality of life, dyspnea and heightened risk for COPD exacerbation however, much less investigation exists exploring the role of other indoor pollutants to COPD morbidity.

## 2. LITERATURE SURVEY

They studied was designed to limit patient burden as much as possible. Each participant underwent an initial, web-based screening to assess their disease progression, willingness to wear the devices on their undergarments for 9 months, and report on weekly COPD-related symptoma web-based survey. They proceeded with a nurse interview and, if still eligible, received a pack of Health Tags in the mail. A phone- based technical support specialist optionally talked participants through the device setup procedure. Remote monitoring of de-identified device adherence and respiratory data was possible through a web-based dashboard. Participants were automatically notified by SMS, email, and ultimately manual phone calls if device adherence became a concern.

This existing system is an hybrid model which combine tree-based feature transformation with Bayesian non-parametric classification, to predict whether the patient should adopt NIPPV based on the their own physical condition. We delved into the feature importance and justified the rationality of using tree-based feature transformation. The proposed gaussian process classification (GPC) with gradient boosting decision tree (GBDT) feature transformation model has shown state-of-the-art results on both the NIPPV dataset and two simulated datasets with larger sample size. For critically ill COPD patients.

Identifying acute exacerbations in chronic obstructive pulmonary disease is of utmost importance for reducing the associated mortality and financial burden. In identification models for COPD and to compare the relative performance of different modelling paradigm to find the best model for this task.

In this existing system they constructed the prediction model for COPD severity using various machine learning techniques. By analyzing 36S samples of mild and severe COPD groups, we observed that the model using random forest performed the best (AUC =0.886) and Diffusing capacity of Lung CO, modified medical research council, and age were the most important features of the model.

In this existing system they constructed the prediction model for COPD severity using various machine learning techniques. By analyzing 36S samples of mild and severe COPD groups, we observed that the model using random forest performed the best (AUC =0.886) and Diffusing

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

capacity of Lung CO, modified medical research council, and age were the most important features of the model.
In this paper, Logistic and neural network models were established to predict individuals' risk estimates. Different threshold were set to explore the change of classification. For the prediction of COPD, it is reasonable to consider a range of acceptable risk thresholds to weigh the benefit of a true positive and the harm of a false positive. For individuals, continuous risk predictions are more useful than risk groups. The specificity and the positive predictive value should be given more priority in the choice of threshold for COPD prediction, also the benefits of true positives and the harms of false positives should be considered.
There search of emerging materials such as graphene monolayer and perovskite may revolutionize the field of point-of-care devices. These materials can boost the sensitivity and specificity of the detection, and therefore the detection can be performed in samples taken non-invasively, such saliva, and with less sample quantity. A graphene field effect transistor (GFET) coated with PEDOT:PSS and perovskite, bring advantages to the photo detection field, due to the unique proprieties of 2Dmaterials and the structure of perovskitek. This work presents a study of material characteristics comprising a GFET, with perspective to detect biomarkers of COPD.
Chronic obstructive pulmonary disease poses a large burden on health care current models for predicting severe COPD exacerbations lack accuracy, making it difficult to effectively target patients at high risk for preventive care management to reduce severe COPD exacerbations and improve outcomes.
Early diagnosis and intervention are crucial for reducing the burden of these conditions, yet traditional stethoscope-based auscultation remains subjective and dependent on clinician expertise.

## 3. PROPOSED SYSTEM

The proposed system integrates machine learning (ML) within a telemedicine-based IoT framework to create an advanced Smart Breath Analyzer for real- time health monitoring and disease prediction. It collects exhaled breath samples using a sampling bag, processes them via a gas-mixing bench, and uses sensors to measure concentrations of gases such as acetone, smoke, and other volatile organic compounds (VOCs).
A microcontroller processes the sensor data, which is then analyzed by ML algorithms to detect patterns and anomalies, providing early warnings for conditions like Chronic Obstructive Pulmonary Disease (COPD). The system employs an IoT module to transmit data to a cloud server, enabling real-time access and long-term storage for advanced analytics and trend monitoring. Patients and healthcare providers can interact with the system through a user- friendly telemedicine interface that displays results, alerts, and actionable recommendations.
Tested in diverse environments among high-risk COPD patients, the system demonstrates how ML enhances diagnostic accuracy and facilitates proactive healthcare. This approach combines real-time monitoring, personalized insights, and seamless telemedicine integration, contributing to smarter and more efficient healthcare systems.
The system is designed to be integrated into healthcare platforms, providing real-time analysis and risk assessment for physicians and patients. Wearable sensors can be incorporated to collect real-time physiological data, enhancing continuous monitoring. This machine learning based approach not only improves early detection but also aids in personalized treatment planning, ultimately reducing hospitalization rates and improving patients outcomes.
This proposed system shows to the early chronic pulmonary disease prediction system using machine learning.

DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

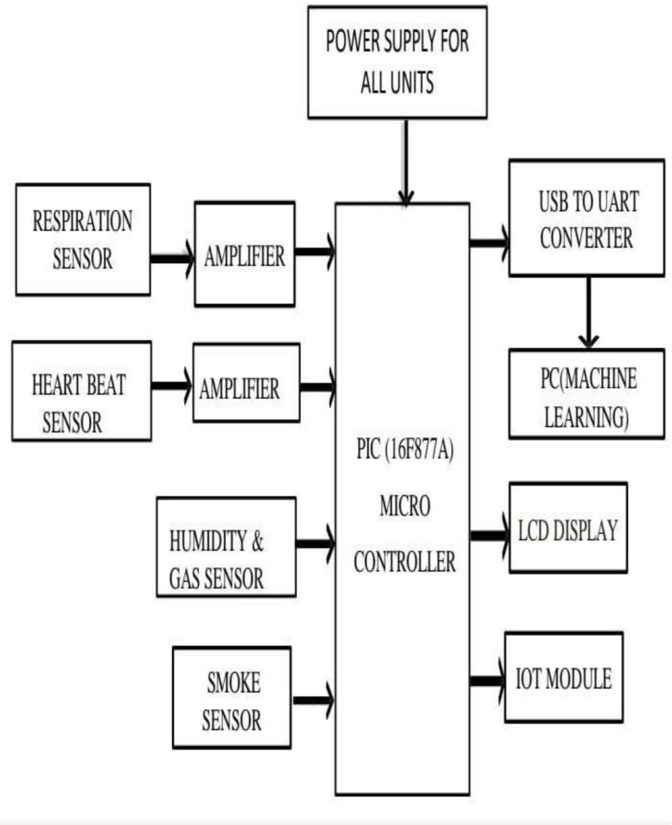


Fig 1. Block Diagram

The Respiration Sensor is used to monitor abdominal or thoracical breathing, in biofeedback applications such as stress management and relaxation training. Besides measuring breathing frequency, this sensor also gives you an indication of the relative depth of breathing. The first step in this process is breathing in air, or inhaling. The taking in of air rich in oxygen into the body is called inhalation and giving out of air rich in carbon dioxide from the body is called exhalation. The second step is gas exchange in the lungs where oxygen is diffused into the blood and the carbon dioxide diffuses out of the blood. The third process is cellular respiration, which produces the chemical energy that the cells in the body need, and carbon dioxide. Finally, the carbon dioxide from cellular respiration is breathed out of body from the lungs.
IR sensor for detecting the HEART BEAT. IR has less noise and ambient light than at normal optical wavelengths. The light is produced only when current passes through in the forward direction and block current in the reverse direction. Plethysmograph is an infrared photoelectric sensor used to record changes in pulsatile blood flow from the finger. The Plethysmograph operates by recording changes in blood volume as the arterial pulse expands and contracts the microvasculature.
Smoke detectors have been recognized as being very useful and desirable in providing an early warning of dangerous levels of smoke. A smoke detector can warn of a smoldering fire before any significant property damage occurs, and before the occupants of the structure find themselves in grave danger. When such detectors go into alarm, they usually generate an audible alarm indicating output. The alarm may be audible and/or visual within the monitored space, and may be electronically communicated to a remote monitoring site.
The Internet of things (IoT) is the network of physical devices, vehicles, home appliances and other

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

items embedded with electronics, software, sensors are uators, and connectivity which enables these objects to connect and exchange data. Each thing is uniquely identifiable through itsembedded computing system but is able to inter-operate within the existing Internet infrastructure.

Arduino / Genuino Uno is a microcontroller board based on the ATmega328P (datasheet). It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a 16 MHz quartz crystal, a USB connection, a power jack, an ICSP header and a reset button. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery to get started.UNO without worring too much about doing something wrong, worst case scenario you can replace the chip for a few dollars and start over again.

Uno board and version 1.0 of Arduino Software (IDE) were the reference versions of Arduino, now evolved to newer releases. The Uno board is the first in a series of USB Arduino boards, and the reference model for the Arduino platform; for an extensive list of current, past or outdated boards see the Arduino index of boards.

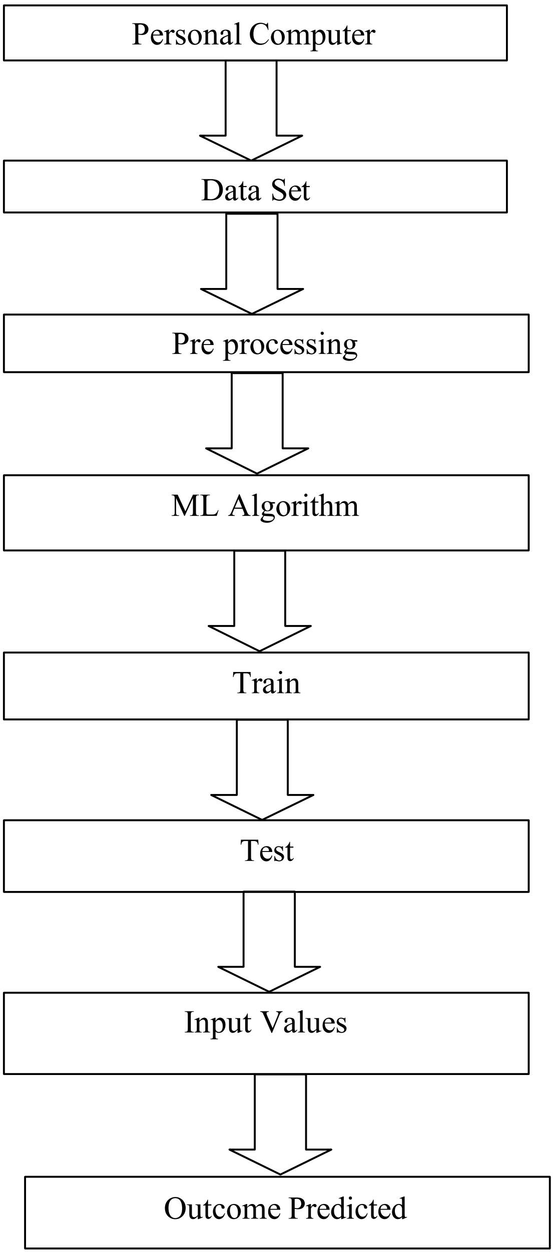


Fig 2.Machine learning work Flow

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

Python code typically has fewer lines than similar code. Python has quick development times, but its execution times are a little slower.
Python applications run more slowly when compared to fully compiling languages like C and C++.Naturally, with the current processing rates of computers, speed variations are typically only noticeable in benchmarking tests and not during actual operations. In most situations, Linux distributions and Mac OS X computers already come with Python installed.
Our Python NumPy Tutorial explains both the fundamental and more complex NumPy principles. Both professionals and beginners can benefit from our NumPy tutorial. Numeric Python, sometimes known as NumPy, is a Python library for computing and processing multidimensional and linear array elements.
It is a Python extension module that is primarily built in C. It has a number of routines that can quickly do calculations involving numbers. Many robust data structures, including multi-dimensional arrays and matrices, are available in NumPy. The best calculations for arrays and matrices are done using these data structures.

## 4. RESULTS AND DISCUSSION

The implementation and testing Smart Breathing Analyzers equipped with machine learning (ML) within a telemedicine-based IoT framework have demonstrated promising results in enhancing health monitoring and diagnostics. The prototypes, utilizing a Arduino Uno is a microcontroller, were tested across diverse environments, including indoor and outdoor settings, with a specific focus on high-risk COPD patients. The integration of ML algorithms significantly improved the system's ability to analyze gas, smoke, and acetone concentrations in exhaled breath, enabling accurate real-time pattern recognition and anomaly detection. The results highlighted the effectiveness of the system in identifying early warning signs of disease exacerbations, providing patients and healthcare providers with timely alerts and actionable insights. Furthermore, the incorporation of IoT allowed seamless data transmission and remote accessibility through a telemedicine platform, facilitating continuous monitoring and personalized care.

Table I: GOLD classification of Airflow Limitation Severity in COPD

| Gold Classification | Airflow Limitation Severity | $FEV_1$ Measurement |
|---|---|---|
| GOLD 1 | Mild | $FEV_1 \geq 80\%$ predicted |
| GOLD 2 | Moderate | $50\% \leq FEV_1 < 80\%$ predicted |
| GOLD 3 | Severe | $30\% \leq FEV_1 < 50\%$ predicted |
| GOLD 4 | Very severe | $FEV_1 < 30\%$ predicted |

GOLD stands for the Global Initiative for Chronic Obstructive Lung Disease. GOLD helps raise awareness of COPD and works with doctors and health experts to create better ways to prevent and treat this condition.
The original GOLD system used the term "stages" to refer to the different levels of COPD. Now they are called "grades".

Table II: pharmacologic Therapy for stable COPD as per Initiative for Obstructive lung disease

| Patient Group | Recommended First Choice | Alternative Choice | Other Possible Treatments |
|---|---|---|---|
| A | SA anticholinergic prn<br>OR<br>SA beta2-agonist prn | LA anticholinergic<br>OR<br>LA beta2-agonist<br>OR<br>SA beta2-agonist and SA anticholinergic | Theophylline |
| B | LA anticholinergic<br>OR<br>LA beta2-agonist | LA anticholinergic and LA beta2-agonist | SA beta2-agonist and/ or SA anticholinergic<br><br>Theophylline |
| C | ICS + LA beta2-agonist or LA anticholinergic | LA anticholinergic and LA beta2-agonist<br>OR<br>LA anticholinergic and PDE-4 inhibitor<br>OR<br>LA beta2-agonist and PDE-4 inhibitor | SA beta2-agonist and/ or SA anticholinergic<br><br>Theophylline |
| D | ICS + LA beta2-agonist and/ or LA anticholinergic | ICS + LA beta2-agonist and LA anticholinergic<br>OR<br>ICS + LA beta2-agonist and PDE-4 inhibitor<br>OR<br>LA anticholinergic and LA beta2-agonist<br>OR<br>LA anticholinergic and PDE-4 inhibitor | Carboxycysteine<br><br>N-acetylcysteine<br><br>SA beta2-agonist and/ or SA anticholinergic<br><br>Theophylline |

This table are considered for patients with severe COPD and frequent exacerbations despite standard therapy. Smoking cessation, pulmonary rehabilitation and vaccination are also components of COPD management.

Table III: Survival rates associated with pre existing COPD among elderly patients with non small cell lung cancer

| Group | ~ Percentage Alive | Median Survival |
|---|---|---|
| **Stage I NSCLC** | | |
| Non-COPD | 48.1% | 1,130 days |
| Emphysema | 38.3% | 811 days |
| Chronic bronchitis | 31.7% | 672 days |
| **Stage II NSCLC** | | |
| Non-COPD | 25.5% | 627days |
| Emphysema | 26.9% | 582 days |
| Chronic bronchitis | 17.3% | 445 days |
| **Stage III NSCLC** | | |
| Non-COPD | 10.6% | 255 days |
| Emphysema | 8.9% | 229 days |
| Chronic bronchitis | 7.2% | 222 days |
| **Stage IV NSCLC** | | |
| Non-COPD | 2.8% | 112 days |
| Emphysema | 2.4% | 110 days |
| Chronic bronchitis | 1.8% | 105 days |

COPD is an independent risk factor for decreased survival in NSCLC due to compromised pulmonary function and increased susceptibility to infections, respiratory failure and treatment complications.
Early stage NSCLC surgical resection may be curative, but COPD limits eligibility for lobectomy or pneumonectomy due to poor lung reserve.
COPD can lower the 5 years survival rate of NSCLC patients, especially in stages III and stages IV, where survival rates are already poor.
COPD worsens prognosis due to increased risk of respiratory decline and limited tolerance to chemotherapy or immunotherapy.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

Elderly patients with pre-existing chronic obstructive pulmonary disease who develop non small cell lung cancer have reduced survival rates due to compromised pulmonary function, increased risk of treatment related complications, and limited therapeutic options. The severity of COPD along with the advanced disease stages further decreasing survival outcomes.

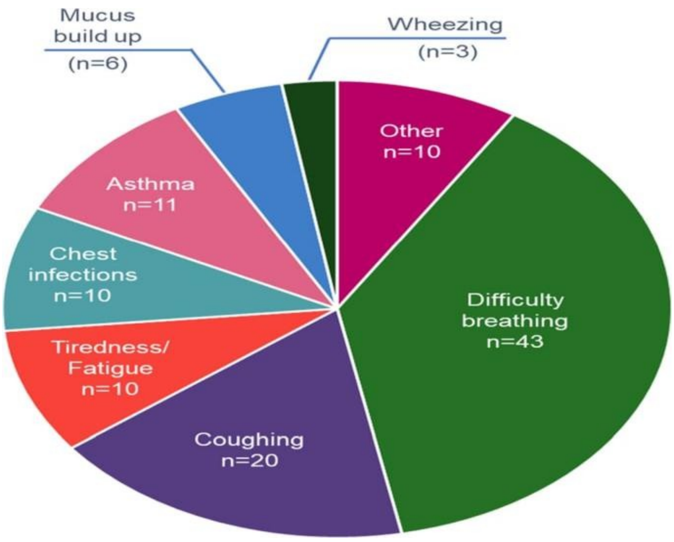


Fig 3.Symptoms Of COPD in Human Body Chronic Obstructive Pulmonary Disease is a progressive lung disease that primarily affects breathing.

Persistent cough often producing mucus especially in the morning. Difficulty breathing, especially during activity or exertion. Excess sputum, which may be clear, white, yellow or green. Whistling or squaky sound while breathing due to narrowed airways. Feeling of pressure or heaviness in the chest.

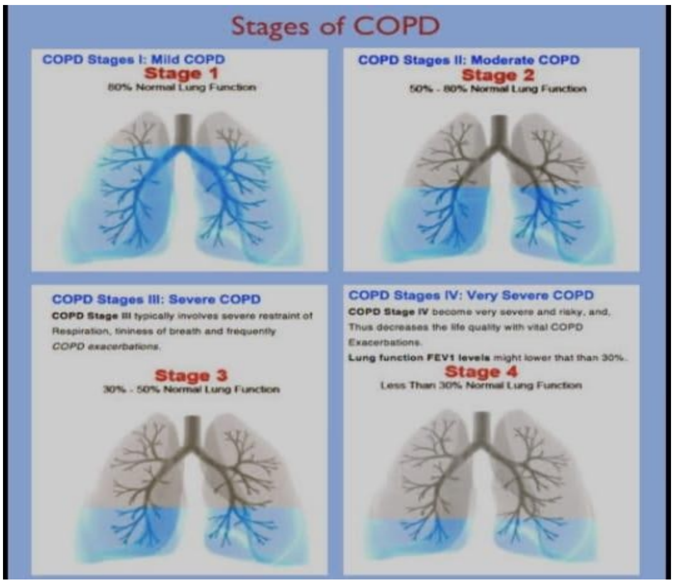


Fig 4. Stages Of COPD

Mild COPD is persistent but mild cough with mucus. Occasional shortness of breath, especially during physical activity. Lung function decline may not be noticeable. Regular exercise and healthy lifestyle choices.

Moderate COPD is increased breathlessness, especially during exertion. More frequent coughing and mucus production. Fatigue due to reduced oxygen levels. Some limitations in daily activities.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

Severe COPD significant difficulty in breathing, even with minimal exertion. Increased mucus production and chronic cough. Frequent exacerbations symptoms requiring medical intervention.
Vey Severe COPD extreme breathlessness, even at rest. Severe limitation in physical activity. Cyanosis means bluish lips and fingertips due to low oxygen.

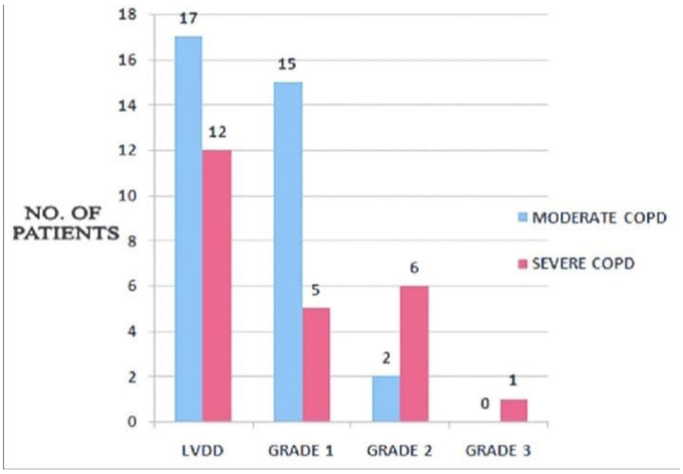

Fig 5.Correlated Graph of COPD

X-axis represents Time or disease stages. Y-axis represents key parameters lung function Decline, symptoms, Hospitalizations.
Negative correlation between FEV1% and COPD severity. Positive correlation between COPd severity and symptom burden.
Inverse correlation between exacerbations and FEV1 levels. Higher smoking leads to earlier onset and rapid progression.

## A. Future Scope

COPD is a progressive lung disease that often goes undiagnosed until symptoms become severe, and this delay in diagnosis can lead to irreversible damage to the lungs.By integrating machine learning algorithms, it becomes possible to detect early signs of COPD, allowing for earlier intervention and more effective management, ultimately improvingpatient outcomes. In the coming years, the scope of these prediction systems will likely expand as more data is collected from diverse sources such as electronic health records (EHRs), wearable devices, and imaging systems. ML algorithms, particularly supervised and unsupervised learning techniques, will be able to analyze complex patient data, including spirometry results, chest X- rays, CT scans, and genetic factors, enabling more accurate predictions of the onset and progression of COPD. This will allow clinicians to identify at-risk individuals even before symptoms become evident, providing a critical opportunity for early intervention.
In addition to improving diagnostic accuracy, machine learning can aid in the development of preventive measures and therapeutic strategies. ML systems can analyze historical patient data to predict how certain factors (e.g., smoking cessation, air quality) might influence the progression of the disease. These predictions can guide healthcare providers in recommending lifestyle changes or treatments that could delay or prevent the onset of severe COPD, contributing to improved quality of life and reduced healthcare costs**.** The integration of AI-powered tools in primary care settings is another key aspect of the future of COPD prediction. By making early warning systems available to general practitioners, ML models can empower doctors with decision-support tools to better assess the risks of their patients. This can lead to more proactive healthcare, where patients receive timely screenings and interventions**.** Future developments could involve refining model interpretability, incorporating additional data sources, and improving scalability across different healthcare environments.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.032 ISSN: 2581-4621

## 5. CONCLUSION

The early prediction system for chronic pulmonary disease (COPD) using machine learning demonstrates a significant advancement in the timely identification of the condition. By analyzing patient demographics, medical histories, and physiological data, machine learning models effectively predict the onset of COPD at early stages, enabling more proactive interventions. Algorithms such as Decision Trees, Support Vector Machines (SVM), and Random Forests show strong performance in terms of accuracy, providing valuable insights into patient health risks.

This system can play a critical role in healthcare by improving early diagnosis, reducing healthcare costs, and optimizing treatment strategies. With the integration of such technology, early detection of chronic pulmonary diseases becomes more efficient, enhancing the quality of patient care.

## 6. REFERENCE

1. HALPIN D M G, CELLI B R, CRINER G J, et al. The GOLD Summit on chronic obstructive pulmonary disease in low middle- income countries [J]. Int J Tuberc Lung Dis. 2019;23(11):1131–41.
2. HAIDER N S, SINGH B K, PERIYASAMY R, et al. Respiratory sound based classification of Chronic Obstructive Pulmonary Disease: a risk Stratification Approach in Machine Learning paradigm [J]. J Med Syst. 2019;43(8):255.
3. KANWADE A, BAIRAGI VK. Classification of COPD and normal lung airways using feature extraction of electromyographic signals [J]. J King Saud Univ-Comput Inform Sci. 2019;31(4):506–13.
4. FANG Y, WANG H, WANG L, et al.
Diagnosis of COPD based on a knowledge graph and Integrated Model [J]. IEEE Access. 2019;7:46004–13.
5. SORIANO JB, KENDRICK P J, PAULSON K R, et al. Prevalence and attributable health burden of chronic respiratory diseases, 1990– 2017: a systematic analysis for the global burden of Disease Study 2017 [J]. Lancet Respiratory Med. 2020;8(6):585–96. 6.
6. Roche, N. Systemic Medications in Chronic Obstructive Pulmonary Disease: Use and Outcomes. *Clin. Chest Med.* 2020, *41*, 485–494.
7. CHOI J Y, RHEE CK. Diagnosis and treatment of early chronic obstructive lung disease (COPD) [J]. J Clin Med. 2020;9(11):3426.
8. Chronic Obstructive Pulmonary Disease Group of Chinese Thoracic Society, Chronic Obstructive Pulmonary Disease Committee of Chinese Association of Chest Physician. Guidelines for the diagnosis and management of chronic obstructive pulmonary disease (revised version 2021) [J]. Chin J Tuberc Respir Dis. 2021;44(3):170–205.
9. JUNG T. VIJ N. Early diagnosis and real- time monitoring of Regional Lung function changes to Prevent Chronic Obstructive Pulmonary Disease progression to severe emphysema [J]. J Clin Med, 2021, 10(24).
10. FENG Y, WANG Y, ZENG C, et al. Artificial Intelligence and Machine Learning in Chronic Airway diseases: Focus on Asthma and Chronic Obstructive Pulmonary Disease [J]. Int J Med Sci. 2021;18(13):2871–89.
11. COATES J T, DE KONING C. Machine learning-driven critical care decision making [J].JR Soc Med. 2022;115(6):236–8.