Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 1, pp. 398–414, DOI 10.12694/scpe.v26i1.3758

A NOVEL EFFECTIVE FORECASTING MODEL DEVELOPED USING ENSEMBLE MACHINE LEARNING FOR EARLY PROGNOSIS OF ASTHMA ATTACK AND RISK GRADE ANALYSIS

SUDHA YADAV[*], HARKESH SEHRAWAT[†], VIVEK JAGLAN[‡], SIMA SINGH[§], PRAVEEN KANTHA[¶], PARUL GOYAL[‖] AND SURJEET DALAL[**]

Abstract. Research curiosity enlarging the concern of clinician and researchers towards combination of medical science together with artificial intelligence to develop cost effective predictive model for asthma exacerbation. To accumulate the classification consequences, extensively known ensemble machine learning methods pivotal to artificial intelligence techniques are investigated and novel predictive model developed using catboost classifier that produced comparatively improved outcomes to predict the occurrence of asthma and asthma risk grade. Proposed model result is compared with other classifiers which are Support vector machine (SVM), K-Nearest neighbors (KNN), Logistic regression, Adaboost classifier, Gradient boosting classifier, Random forest, Decision tree. Model regulated classification accuracy as high as 93% with datasets selected for formation of early prognosis model of asthma disease by embracing only 20% of the features in the reduced feature set.

Key words: Asthma, Risk grade, Machine learning, Ensemble learning classifiers, Predictive model.

1. Introduction. Asthma is one of the most common severe chronic inflammatory disease of the airways that affect patient, family and healthcare system socially and financially both. Persons of all age weather children, adults and aged, influence from asthma. Asthma is persistent and its belongings are enduring. The clinical appearance of asthma is extremely diverse. Wheeze, shortness of breath, allergen, cough and chest tightness, comprises as basic indications of asthma. An asthmatic patient may endure with one or more combination of these signs, which may be irregular or persistent. Severe asthma attack may even leads to life-loss, accordingly instant medical contributions, either as an emergency department visit or admission to the hospital prerequisite. After consideration of all these circumstances, it is needed that prediction should be carried out at an early stage to overcome the likelihood of an asthma attack. Identification of high risk asthma patients in a timely manner and preventive involvements are the key points of advancing asthma care in the long–term.

Machine learning techniques occupy a diversity of probabilistic, statistics and optimization approaches to acquire from earlier experience and perceive valuable patterns from enormous, unstructured and complex datasets. Machine learning has the potential to develop more finely calibrated, personalized predictive probability scores for asthmatic patients than conventional statistical models. The use of machine learning methods is quickly increasing and deals with inventive methods for prognostic modeling to potentially permit a patient-centered

[*]Department of Computer Science & Engineering, Maharshi Dayanand University Rohtak Haryana, India (sudhayadav.91@gmail.com).
[†]Department of Computer Science and Engineering, Maharshi Dayanand University Rohtak Haryana, India (sehrawat_harkesh@mdu.ac.in).
[‡]Department of Computer Science and Engineering, Amity University Madhya Pradesh, Gwalior, India (jaglanvivek@gmail.com).
[§]Department of Planning and Architecture, Dada Lakshmi Chand State University of Performing and Visual Arts Rohtak, India (simasingh.2009@gmail.com).
[¶]Chitkara University School of Engineering and Technology, Chitkara University Himachal Pradesh, India (praveen.kantha@chitkarauniversity.edu.in).
[‖]Computer Science & Engineering Department, M. M. Engineering College, Maharishi Markandeshwar Deemed to be University, Mullana, Ambala – 133207, Haryana, India (parul.goyal@mmumullana.org).
[**]Department of Computer Science and Engineering, Amity University Madhya Pradesh, Gwalior, India (profsurjeetdalal@gmail.com)

approach. Machine learning techniques based models are appropriate to upgrade the detection, treatment, and management of asthma. One of the significant point regarding development and deployment of asthma related machine learning model is that entail large and assorted datasets, as well as attentive validation and assessment to confirm precision and consistency. These disease prediction models should always be applied in accumulation with clinical expertise but should not substitute the purpose of healthcare specialists in identifying and treating asthma or other disease. As per literature study no single machine learning algorithm is better than other in an unobserved situation. Traditional classifiers fail to accomplish better performance with different types of data available. To overcome this problem, we have implemented catboost classifier which is a boosting algorithm of ensemble machine learning to enhance the accuracy and other performance measures.

In this paper, a novel optimized prediction model were intended using ensemble learning that generalizes well with asthma prediction problems encompassing categorical data. The proposed work has been personalized to produce optimal performance with associated dataset that comprise independent attributes of categorical and/or numerical and a categorical binary dependent or target attribute. Performance comparison of the proposed model against other machine learning classifiers is done on the same dataset.

1.1. Research Contribution. There are the following research contributions as below:

- This paper optimised AdaBoost algorithm for Early Prognosis of Asthma Attack.
- This paper reduce the dangers produced with help of early Prognosis Of Asthma Attack .
- Recognizing and selecting relevant attributes increases the model's capacity to capture crucial patterns and correlations that improve predictive maintenance accuracy.
- The proposed method gains the accuracy to reduce level of medical errors.
- Adopting advanced analytics, machine learning, and optimization technologies improves industry efficiency and competitiveness.

1.2. Paper organization. The remainder of the article is structured as follows: A quick summary of the many literature evaluations already presented on the topic is provided in Section 2. The research approach is covered in Section 3. The research's findings are presented in Section 4. Potential applications are discussed in Section 5. The paper is ultimately concluded in Section 6.

2. Related Work. Asthma attack predictor system tied with smart mobile devices helpful for data collection developed by Tsang KCH [2022]. Machine learning coupled with smart devices enhances self-management by asthma prediction. Use of smart devices includes (smart-watch, smart inhaler and smart peak flow meter). Data was collected through questionnaires also. A two phase process in which first phase of manually data collection (a small team of persons were organized to collect data manually through questionnaires) and second phase comprises collection of data using smart devices.

Zhang [11] done analysis of PEF (Peak Expiratory Flow) along with asthma symptoms scores recorded by applicants for detection of asthma exacerbation on the basis of daily home monitoring. Data post processing done through normalization, standardization and smoothing filters. PCA (Principal Component Analysis) was used to diminish the huge amount of derived variables to a reduced amount of linearly independent components. Four different classifiers naïve Bayes, logistic regression, decision tree and perceptron algorithms were assessed. Stratified cross-validation method used for model accuracy evaluation. Outcome of proposed model is detection of exacerbation on the same day or up to coming three days in the future. Best performance measures values (sensitivity 90% and specificity 83%) attained by logistic regression model.

Machine learning model for asthmatic based on cough sound developed by Hwan [12]. Cough sounds of asthmatic children and healthy children were considered for development of classifier. Demographic detail, interval of cough, and previous data of respiratory status were considered. A dataset of cough sounds was prepared and randomly allocated to training and testing dataset. Mel-Frequency Cepstral Coefficients and Constant-Q Cepstral Coefficients, two audio features were removed. Classification model using Gaussian Mixture Model–Universal Background Model (GMM-UBM) developed. Predictive performance was also tested with the help of test set. Out of 1192 cough sounds, asthmatic cough sounds of 89 children and out of 1140 cough sounds, healthy sounds of 89 children were analyzed. This proposed model (audio-based classification) gave sensitivity 82.81% and specificity 84.76%.

Mindy et al. [16] also developed a pediatric asthma prediction model using novel ML algorithm, predictor pursuit (PP) and attempted to find out subclasses of pediatric asthma based on asthma treatment outcome status, to distinguish features associated with asthma control inside every determined pediatric phenotype and to envisage long-term asthma control among asthma-affected children. Attributes adherent to medication, age, sex, blood eosinophils, ethnicity, BMI, age of asthma onset, phenotypes, and severity were analyzed for estimation purpose. Four diverse phenotypes categories (A+/O−), (A−/O+), (A+/O+), (A−/O−) were discovered through pp model.

A proficient examination of asthma using competent machine learning algorithm done by soltani[13], on database of asthmatic (169) and non-asthmatic (85) patients visited two different hospitals. K-nearest neighbors, random forest and Support vector machine, Machine learning methods implemented on database after completion of preprocessing step. Data preprocessing completed by cross-fold da ta sampling and relief-F strategy method of data mining technique. Accuracy and specificity are two performance measures were evaluated for the proposed model and compared with previous research work.

Asthmatic patients required constant monitoring to keep records of their health condition as per Priya et al. (2021). A fog based healthcare system helpful to attain high-quality disease monitoring and control for asthmatic patients. In this paper, an IOT (Internet of Things)-based method is developed to assess severity condition of asthma in patient in a timely manner and help them to get rid-off from hospital admission. A model based on artificial neural network that can predict asthma attacks and notifies the relevant individuals, such as the patient and his or her family members. Notable precision achieved through this system was 86%.

Asthma exacerbation forecasting model using EHR (Electronic Health Record) designed by Martin [15].Structured data for gender, smoking status, race, age, environmental allergy testing, use of asthma medications, BMI status, and Asthma Control Test scores (ACT) were mined from a large repository of EHRs dataset. A subgroup of records of asthmatic patients with all prescribed asthmatic features used for primary analysis. Univariable and multivariable statistical analysis was finalized to recognize exacerbation factors. A risk forecasting model developed and verified centered on the multivariable analysis. A large dataset of 37,675 asthmatic patients was considered, out of which 1,787 records contained data of asthmatic patients and 979 of them experienced an .asthmatic attack. The AUC (area under the curve) performance measure value was 0. 67 in a collective derivation and substantiation cluster.

After considerations of 2-3 machine learning algorithms performance, Anne compare the performance of machine learning algorithm for asthma prediction and risk assessment purpose. Two ML methods (XGBoost, one class SVM) and logistic regression model provided their prediction result on the basis of asthma indicators. Best AUC result of XGBoost was 0.85 between the range (0.82–0.87) and 0.88 in the range of (0.86–0.90) for logistic regression.

Asthma prone area modeling using machine learning model done by Seyed Vahid by considering environmental and spatial factors. A spatial dataset of 872 asthma affected children locations was prepared. 13 environmental features (particulate matter (PM 10 and PM 2.5), rainfall, distance to parks and streets, temperature, pressure, humidity, wind speed, ozone (O3), carbon monoxide(CO), sulfur dioxide (SO2), and nitrogen dioxide (NO). Based on these environmental factors and spatial database, a random forest machine learning model was established to detect asthma prone areas.70% portion of dataset allocated to modeling and rest 30% portion of database allocated to validation process. Outcome of spatial correlation and random forest model presented that criteria of Particulate matter (PM 2.5 and PM 10) and distance to park and streets had major impression on asthma manifestation in study areas. The RF model accuracy measures represented AUC (area under the curve) 0.987 for training and 0.921 for testing data.

CAPP and CAPE model for childhood asthma prediction analyzed for early school age children's and preschool age children's. Kothawala[17] enforced seven machine learning methods to ascertain best predictive model and compare their performance with existing regression models. Risk category was also examined. (Recursive Feature Elimination) RFE recognized a novel optimal subgroup of features cooperative in prediction process of school-age asthma for each model. The best performance proved by Support Vector Machine (SVM) algorithms for both the CAPE (area under the receiver operating characteristic curve, AUC = 0.71) and CAPP (AUC = 0.82) models. Virtuous generalizability and excellent sensitivity demonstrated by both models in MAAS to predict a subgroup of persistent wheezes.

3. Material and Method.

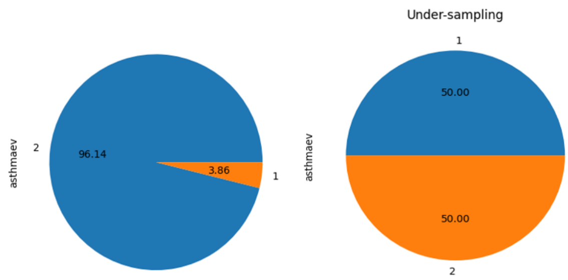
3.1. Dataset. We utilized a public dataset of asthmatic patients available on ISAAC website, a large repository of asthmatic patients data of different-different countries (state wise). This work consumed Indian dataset that contained 2961 no. of records primarily along with 58 attributes with respect to which different information is recorded.

Data elements used. Attributes contain information like demographic data (age, gender, date, country code, age group) and information which described indications of asthma such as frequency of recurrence in symptoms, speech, wheeze, wheeze12 (occurrence of wheezing in last 12 hours), sleep disturbance, shortness of breath, resulting from wheeze, nose irritation, rashes and presence of hay fever, breathing issues after exercise. Some attributes like age in groups, form version, date of information recorded, country code, age group and some variations of other attributes are carefully excluded because these attributes were not contributing in prediction process. Response values of all these attributes encoded in three different values 1, 2 and 9 where 1 signifies the presence of the appropriate attribute, 2 signifies the nonexistence (absence) of the same and 9 represents any other response value provided by respondents.

Outcome measure. The target attribute (asthma) was encoded as 9 (neither represents presence nor represents absence of asthma) for 112 samples out of 2981 records. Exclusion of all the 112 samples was done which resulted in a dataset with 2849 samples. Of these, only 110 records indicated presence of asthma and the rest were with non-asthma.

3.2. Dealing with class imbalance problem. Class imbalance problem is relatively natural here as the number of records with asthma positive (presence of asthma) was much fewer as compare to asthma negative (absence of asthma) records, which often worsens machine learning enactment. Intrinsically it is observed that one of the two classes is underrepresented. For that reason, class imbalance learning methods were investigated that operate by resampling the training data. These techniques increase or decrease the proportion of the training set that represents the minority or majority class respectively, targeting at creating models that are able to better recognize cases of the desirable outcome. Oversampling and undersampling are two primary substitutes to cope with class imbalance problem. Oversampling increase the quantity of samples in the minority class (class with lesser number of records) category, while under sampling deals with decrement in the quantity of samples representing the majority class (class with more number of records) of the training dataset. One key point is these techniques only applied to training dataset not on testing dataset from which absolute predictive model accuracy was determined. Outcome of both alternatives is a balanced class data that can be appropriately deployed as balanced input data for further modeling process. Fig. 1(i) represents 3.86 percentile of asthmatic records (indicated by 1) and 96.14 percentile of non-asthmatic (indicated by 2) out of total 2849 samples while (ii) represents equalization of asthmatic and non-asthmatic records. 110 samples of asthmatic with 110 randomly selected samples of non-asthmatic resulted by undersampling. With balanced input data predictive model can significantly produce consistent outcomes instead of treating the original class imbalanced data as the input. As prediction process here generally focused on minority class category (which implies asthmatic records) rather than the non-asthmatic records, under-sampling was opted to drop out the number of non-asthmatics records. Undersampling resulted in 110 randomly selected samples of non-asthmatic out of 2739 samples to create a balanced dataset of total 220 samples (110 of asthmatic and 110 of non-asthmatic). The authors describe the missing data identification and handling strategy here. The prescribed dataset included varying amounts of missing data, thus the authors utilized mean imputation method like k-nearest neighbors. The authors explained how they normalize characteristics before putting them into ensemble machine learning models. Fair comparisons and model training were achieved via z-score normalization. Incorporated are comprehensive instructions for data cleaning processes, such as methods for detecting and removing outliers. This guarantees that our forecasting model's performance and dependability are not compromised by any anomalies in the dataset used for training and evaluation. Fig. 3.1 Percentile distribution of asthmatic and non-asthmatic data before and after random sampling below.

3.3. Features Extraction. As a large quantity of overall elementary and derived variables is represented by dataset and it is uncertain which of them are utmost predictive of asthma, so features extraction techniques were investigated as attributes selection and reduction techniques. Valuable features contributed toward asthma



(i) (ii)

Fig. 3.1: Percentile distribution of asthmatic and non-asthmatic data before (i) and after random sampling (ii)

prediction were extracted through four diverse features extraction approaches which are correlation matrix represented by figure 3.2, chi-square, select k-best and information gain.

Correlation matrix: Represented in tabular form where each cell comprises the correlation coefficient which signifies the relationship between two variables. Correlation coefficient values lies between +1 to -1, +1 indicates strong relationship and -1 indicates weak relationship.

Chi-square method: Evaluates the dependency between outcome variables (target features) and rest of the other variables. Basically used to determine the features that are most strongly associated with the target variable.

Information gain: Primarily used for feature extraction by assessing the estimated amount of information gain for each variable in the context of the target variable.

Select k-best: It's a filter-based type method for feature selection, attribute extraction process execution takes place independently of any particular machine learning algorithm. Results evaluated from statistical tests like ANOVA F-test, mutual information score used internally to score and rank the features based on their relationship with the target variable.

Finally listed k- number of highest score features for final set of selected features. Fig. 3.2 shows the correlation coefficient of the attributes below.

Different features resulted from correlation matrix, chi-square method, information gain and select k-best method are shown by figure 3.3. A set of features listed by every feature extraction method based on their extraction procedure. Selected –features signified intersection of outcome of all four feature extraction methods that means those attributes were highly correlated with target variable (asthmaev). Mutual attributes were used for prediction purpose and rests were discarded as were not directly associated with the symptoms of asthma.

The extracted attributes subcategory is described by the subsequent attributes after applying features extraction techniques. Whezev (wheezing at any point of time in the past), Whez12 (wheezing issue occurred in the past 12 months), nwhez12 (wheezing issue occurred more than 4 in the past 12 months), cough12 (cough problem occurred during last 12 months), problem in breathing in the past 12 months, awake12 (sleep disturbance occurred due to wheeze for one or more nights in the past 12 months), shortness of breath, iactive12 (state of inactiveness in the past 12 months), eyes12 and pnose12 (watery eye and running nose symptoms in the past 12 months). It is evident from the fig. 4, that wheezing patterns and cough-related symptoms prioritized by the feature extraction technique.

3.4. Train-Test data split up using random sampling. Data samples were drawn randomly to create training and testing subsets. Size proportion of training and test subsets were chosen primarily describing the train-test split quantity, which was preferred as $80-20$ split portion that means 80% of the randomly selected data samples were allocated to training population while the rest 20% portion assigned to the test subset. The process of allocating samples to training and testing population was repeatedly occurred near about 10 times to make sure all the samples were placed in training and test populations at one or the other time. On the other hand with, stratified random sampling, the complete dataset was distributed among smaller clusters named as

Fig. 3.2: Correlation coefficient of attributes

index	Correlation matrix	value	Chi-square	value	Information gain	value	select_k_best	selected_features
0	speech12	0.158	whezev	8.542407e-05	whezev	0.048801	whezev	whezev
1	awake12	0.115	whez12	5.016579e-33	exwhez12	0.039042	whez12	whez12
2	nwhez12	0.109	nwhez12	1.052923e-14	cough12	0.029515	nwhez12	nwhez12
3	whez12	0.164	awake12	1.393268e-16	pnoseev	0.014302	awake12	speech12
4	whezev	0.106	speech12	3.094951e-25	awake12	0.011439	speech12	awake12
5	rashage	0.0829	exwhez12	9.100183e-03	nwhez12	0.011380	exwhez12	cough12
6	cough12	0.0811	cough12	2.476983e-04	speech12	0.009008	cough12	exwhez12
7	pnose12	0.812	pnoseev	3.983466e-03	whez12	0.008834	pnoseev	ieyes12
8	ieyes12	0.076	pnose12	1.372820e-10	hfeverev	0.008070	pnose1	iactive12
9	iactive12	0.068	ieyes12	2.378492e-10	rashev	0.007681	ieyes12	pnoseev
10	exwhez12	0.067	iactiv12	1.386638e-05	ieyes12	0.007429	iactiv12	pnose12
11	pnose12	0.046	hfevere	2.698542e-01	eczemaev	0.005134	rashage	NaN
12	hfeverev	0.018	rashev	3.983874e-01	rashage	0.003700	rclear12	NaN
13	rash12	0.039	rash12	4.639678e-02	rclear12	0.002214	eczemaev	NaN
14	rclear12	0.014	sitese	2.825219e-01	pnose12	0.002201	NaN	NaN
15	NaN	NaN	rashage	4.676054e-02	NaN	NaN	NaN	NaN
16	NaN	NaN	rclear1	1.216865e-02	NaN	NaN	NaN	NaN
17	NaN	NaN	rawake12	6.428554e-01	NaN	NaN	NaN	NaN
18	NaN	NaN	eczemaev	4.081560e-02	NaN	NaN	NaN	NaN

Fig. 3.3: Attributes selected from features selection methods

"strata". Samples constituted as a stratum were sharing some mutual characterization among attributes. Every individual strata were contributing towards formation of training and test datasets as samples were drawn proportionally from each and every distinct strata. The method was apparent as a more specific approach of implementing random sampling to test the performance of the models.

Fig. 3.3 shows the attributes selected from features selection methods.

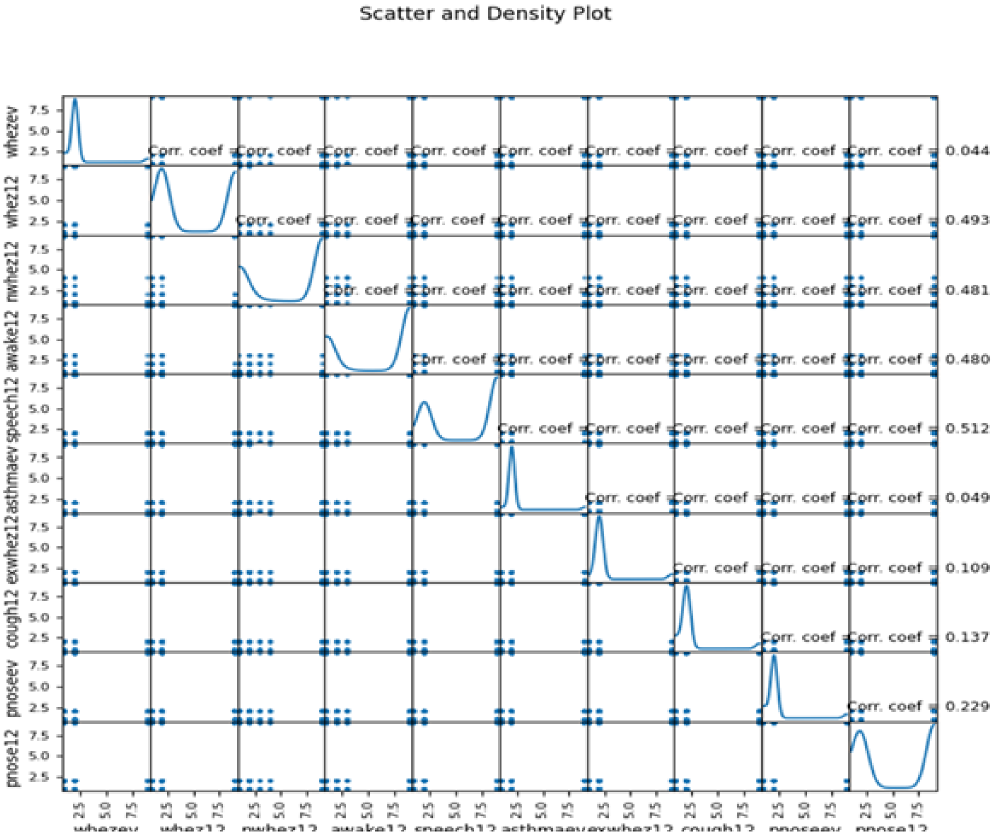


Fig. 3.4: Scattering points and density curve of input variables and target variable

3.5. Data analysis using scattering and density plot. This plot representing the strength of relationship among all the selected variables. Continuous lines indicates density distribution of one variables with their own respect. Scatter plot specified by dot points represents distribution of variables values with respect to other variables. As we can see that asthmaev is highly related to nwhez12 variable and least related to speech12. Variable nwhez12 is highly related to every other variables.

Fig. 3.4 depicts scattering points and density curve of input variables and target variable below.

4. Problem Formulation. Formulating the asthma detection and its risk grade prediction problem is a captious phase in developing a machine learning model for this job. Formulating a problem means describing the proposed predictive model's objectives, inputs and outputs (Kim et al., 2020; Lilhore, Dalal, et al., 2023).

4.1. Objective. The primary objective is to predict asthma and its risk grade or control level for an individual at a particular point of time. This early prediction can supportive for healthcare experts, patients, and caregivers to take conversant decisions concerning patient treatment and management. Collect asthma related data of numerous categories as per availability.

4.1.1. Inputs.

Patient Data. Collect comprehensive data about the patient, which may include:

- Demographic information (e.g., age, gender) if required.
- Allergy information, as any type of allergies like due to dust, due to food, due to smell can trigger asthma.
- Medical history including any past asthma diagnoses, asthma attack ever, comorbidities, family history and medications if patient taking currently or/and in the past. Documents reported symptoms, frequency,

and severity, as well as their medication usage (e.g., controller and rescue medications).
- Speech disturbance, sleep disturbance due to cough.
- Lifestyle factors, such as smoking status and drinking habit.

Environmental Factors. Comprise of environmental factors related data that may impact asthma, such as:
- Information about air quality (e.g., particulate matter and pollutant levels).
- Pollen and allergen factors like cough, cold and weather information (temperature, humidity) that can affect asthma symptoms.

Clinical data. Obtain clinical measurements relevant to asthma assessment, such as:
- Lung function measurements, including forced expiratory volume in one second (FEV1) and forced vital capacity (FVC).
- Peak expiratory flow rate (PEF) quantities, if available and Fractional exhaled nitric oxide (FeNO) levels to assess airway inflammation.

Wearable Sensor Data. If available, collect data from AI enabled wearable devices (e.g., smart watches or spirometers) that make available real-time information on patient physical activity, breathing rate, and other pertinent metrics.

4.1.2. Output. The crucial outcome is prediction of asthma and its risk grade or control level. Based on the explicit objectives of the predictive model, output is formulated in several ways:

Twofold Classification: Predict whether the patient is asthmatic or non-asthmatic based on training provided to the model.

Multiclass Classification: Predict asthma severity levels, such as High, medium, low based on no. of positive symptoms out of total no. of variables.

4.1.3. Problem Formulation. Formally, the authors have defined the asthma and risk grade prediction problem as follows:

Method: Ensemble machine learning (classification or regression).

Input Features: Patient data, clinical test results, symptom history, environmental data, and wearable sensor data.

Output: Predicted asthma and its risk grade or control level (binary classification, multiclass classification).

Objective: Develop a machine learning model to make precise predictions based on historical data and inputs features, permitting for early detection of worsening asthma and personalized management.

After completion of problem formulation, next steps are data preprocessing, feature selection, method selection, training, and assessment to develop a predictive model that can offer valuable observations into asthma severity or control for individual patients.

5. Proposed Methodology. Fig 5.1 depicts the Overall procedure of the proposed methodology below.

5.1. CatBoost Classifier. CatBoost, short for 'Categorical Boosting', an ensemble machine learning method uses decision trees for classification and regression. Breakdown of its name suggests two vital features, "cat" means it works with categorical data and "boost" indicates it uses gradient boosting process. Gradient boosting is a process involves construction of many decision trees iteratively. Result of each and every subsequent tree improves the result of their previous tree, leading to enhanced consequences. CatBoost also seems like advances in the original gradient boost method for a faster execution. Predicting asthma with catboost has the potential to support several new important expansions in both research and medical exercise. Machine learning's prognostic facilities may be used to improve personalize care for each patient. Through asthma severity prediction medical experts will be able to give patients the right dose of medication at an early stage and healthcare professionals can allocate necessary resources in a timely manner. So that the high risk patients receive prompt care, individuals who are predictable to have a more severe type of asthma may be given priority for more extensive interventions or examinations.

CatBoost overwhelms a restraint of other boosting methods which entail, typically, pre-processed data to convert categorical string variables into numerical values. This method can directly work with an assortment of categorical and non-categorical descriptive variables without preprocessing. CatBoost uses a method called ordered encoding to encode categorical features. Ordered encoding considers the target statistics from all the rows prior to a data point to calculate a value to replace the categorical feature. Ordered boosting, is a unique

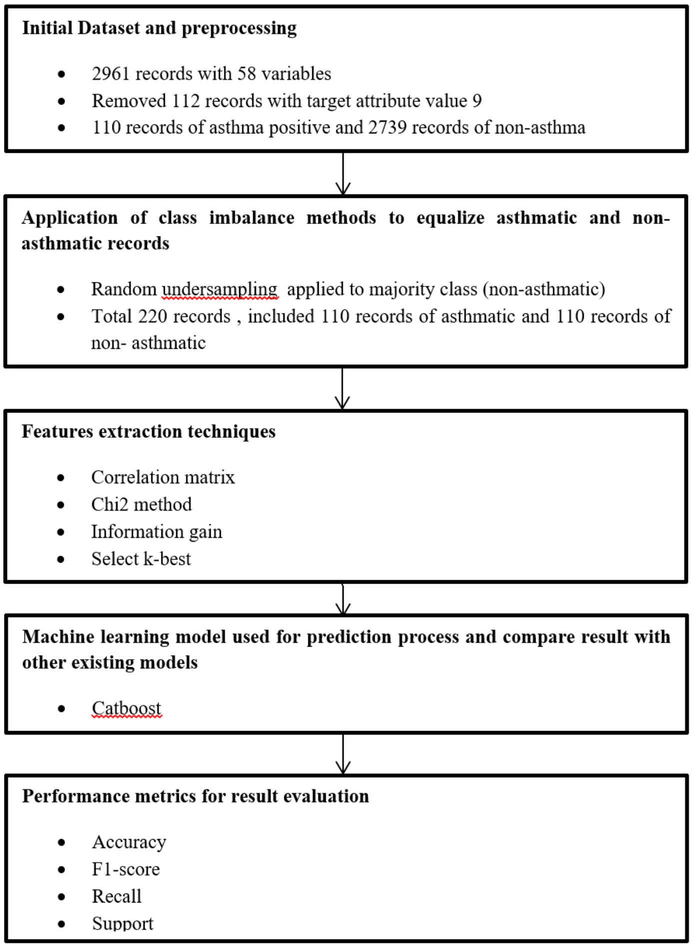


Fig. 5.1: Overall procedure of the proposed methodology

advancement of catboost classifier. It means catboost perform a random permutation over training dataset at each step of boosting to obtain non-shifted residuals by applying the current model to new training subset.

Advanced characteristics of catboost which make it better than other boosting algorithms:

Gradient Boosting: It's a prevailing ensemble learning procedure that syndicates weak prediction models, often decision trees, to produce a powerful predictive model. Main purpose of this is to add new models iteratively to the ensemble; and errors made by any previous models are corrected by the trained newly added model. CatBoost uses gradient boosting to enhance model accuracy.

Categorical Feature Support: CatBoost is an advanced algorithm to process categorical features flawlessly, ultimately develop time saving and effortless process of data preprocessing. For data with categorical features, attain improved accuracy as compared to another algorithm. Another gradient boosting algorithms required conversion of categorical data into numerical form through techniques like one-hot encoding to process that categorical data, but catboost classifier eradicates this need as it can directly work with categorical data without necessity of conversion.

Handling Missing Data Efficiently: CatBoost provides built-in support for missing data, dropping the steps related to preprocessing stage and make sure that model's performance don't obstruct by missing values.

Learning Rate: The learning rate attribute of catboost, monitors the step size at which the model learns throughout the boosting stage. CatBoost impulsively picks an ideal learning rate on the basis of dataset features, to stabilize the model's accuracy and learning speed.

Robust to Overfitting: CatBoost includes an assortment of methods, such as the execution of L2 regularization and a method known as ordered boosting, which contribute towards making it extremely resilient to overfitting and helps to control the complexity of the boosted trees. Ordered boosting, is a permutation-driven approach to the classical boosting method. It attains this by accumulating a regularization term to the loss function used throughout the training procedure.

Enhanced GPU Support: CatBoost employs GPU acceleration, provides faster training by controlling the corresponding processing power of graphics cards, making it ideal for big datasets.

User-Friendly Interface: CatBoost offers an in-built and simple API, making it easy to use for both learners and knowledgeable data scientists. For beginners, its user friendly nature confirms a faster learning curve.

Excellent Performance: CatBoost often require less parameter tuning as compared to other gradient boosting libraries to enhance accuracy, making it an attractive choice for real-world applications.

Symmetric Decision Tree: CatBoost uses symmetric decision trees, where the similar splitting condition is applied throughout complete level of the tree. Such trees are balanced; rare disposed to overfitting, and permits stepping up prediction significantly at testing time.

5.2. Mathematical depiction of CatBoost. In a training dataset which contains N no. of samples and M no. of features, where each sample is indicated as (xi, yi), as xi is a vector of M no. of features (or vector of input variables row wise) and yi is the corresponding target variable, CatBoost purposes to learn a function F(x) that predicts the target variable y.

F(x) represents that complete prediction function which catboost aims to learn. It takes an input vector x (input variables row wise) and predicts the corresponding target variable y.

$\sum$Mm=1 denotes the summation over the ensemble/group of decision trees. Range of summation is from 1 to M, where M indicates the total number of trees in the ensemble.

$\sum$Ni=1 signifies the summation over the training samples. Range of summation is from 1 to N, where N signifies the total number of training samples.

fm (xi) denotes the prediction of the m-th tree for the i-th training sample. Each and every tree in the ensemble/group makes their own predictions for every training sample and provides their contribution to the overall prediction.

The equation states that the overall prediction F(x) is obtained by adding up the initial guess F0(x) with the predictions of each tree Fm(xi) for every training sample. This summation is performed for all trees (m) and all training samples (i).

Steps involved into the Implementation of the proposed model are as follows:

Step1. Data collection and preprocessing: collect asthma related data and apply preprocessing (data sampling, scrutinizing) techniques.

Step2. Feature selection: four different feature selection techniques applied and then selected features which were common in all four techniques.

Step3. Dataset splitting: Dataset was partitioned into 2 subsets 1) training subset represented as (x_tr, y_tr) and testing subset represented as (x_tt,y_tt).

Step4. Create an instance of catboost classifier.

Step5. Set values of parameters on basis of that catboost classifier will be trained using (x_tr, y_tr)

- Iterations: This parameter indicates the number of boosting repetitions (decision trees) to be used during training.
- Depth: Defines the maximum depth level of the different decision trees in the ensemble/group, which directly affects model complexity. While smaller trees diminish overfitting but may supervise complex associations, deeper trees are further able to capture complete patterns but are also more disposed to overfitting.
- loss_function: The loss function used for training. For this proposed work loss-function value set as logloss.
- cat_features: An array of attributes representing list of features which are categorical.

	whezev	whez12	nwhez12	awake12	speech12	asthmaev	exwhez12	cough12	pnoseev	pnose12	ieyes12	iactiv12	no. of ones	Risk Grade
8	1	1	3	2	1	1	1	1	2	9	9	9	6	Moderate
32	2	2	2	1	1	1	1	2	1	1	1	2	7	High
39	1	1	2	2	2	1	2	1	2	2	2	1	5	Moderate
56	1	1	2	1	2	1	2	1	2	2	2	9	5	Moderate
80	2	9	9	9	9	1	2	2	2	9	9	9	1	Low
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2885	2	9	9	9	9	1	2	1	1	2	9	9	3	Low
2895	1	2	2	1	2	1	2	2	2	2	2	1	4	Moderate
2919	1	1	9	9	1	1	1	1	1	1	1	9	9	High
2947	1	1	9	9	1	1	1	1	1	1	1	9	9	High
2958	2	1	2	2	1	1	1	1	1	1	1	4	8	High

110 rows × 14 columns

Fig. 5.2: Asthmatic data records containing risk grade attribute

custom_metric: Performance metric used for evaluation.
random_seed: A numerical value to generate random number for producing the reproducible results.
verbose: Controls the amount of logging during training (higher values provide more detailed logging).

Step6. Catboost classifier iteratively builds up the ensemble of trees (trees of predefined depth) by reducing the loss function with the help of gradient descent. Every newly constructed tree will overcome the errors made by its predecessor and will contribute to make efficient prediction. This process will repeat until a preset number of trees have been added or a convergence condition has been encountered.

Step7. To make overall prediction, catboost summarize the prediction result obtain from ensemble of trees. This aggregated prediction outcome leads to highly precise and consistent models.

Step8. Calculate performance metrics of proposed model (Accuracy, F1-score, recall, support).

5.3. Risk Grade Prediction. Timely prediction of risk grade for asthmatic patients can support health experts to take concerning decision for better control and manage the situation. Risk grade consist of three different categories i.e. High, moderate and low. Figs. 5.2 and 5.3 represents dataset with risk grade attribute and distribution of all three categories of risk grade. Risk grade feature were added into dataset of asthmatic patients on the basis of three conditions:

If no. of positive features>=7 (out of 11) , Then risk grade==High
If no. of positive features>=4 && <=6, Then risk grade==Moderate
If no. of positive features<=3,Then risk grade==Low

Fig. 5.2 depicts asthmatic data records containing risk grade attribute.
Fig. 5.3 depicts distribution of risk grade categories.

6. Results and Discussion. The proposed model and other existing models i.e. support vector machine, k-nearest neighbors, linear regression, logistic regression, random forest, decision tree, Adaboost, gradient boosting were implemented on asthma dataset and their performance was evaluated using several metrics. Performance of proposed model compared with existing models performance. Cross validation and random sampling, two testing schemes used to authenticate the performance of the model.

6.1. Simulation result. In machine learning, feature importance is an essential point for deliberation. To better understand a dataset or to advances the prediction power of machine learning trained model, feature selection is executed to distinct the most significant features and parameters. CatBoost has an in-built feature to express the importance of each and every input variable along with their importance values. Table 6.1 illustrates importance value of every variable. Importance factor signifies contribution of every feature in prediction process.

Fig. 6.1 indicates an inclination from top to bottom and whezev is the highest significant feature for proposed prediction work; a falling slope from value 36-16 depicts cough12 is the second highest significant feature after that further features are progressing with small decrement in values. The authors validated our

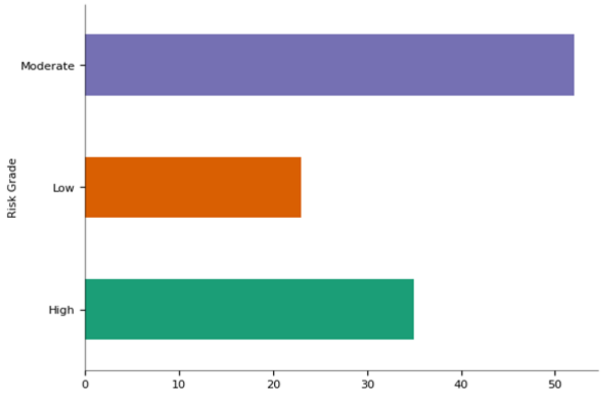


Fig. 5.3: Distribution of risk grade categories

Table 6.1: Feature importance with coefficient values

S.No	Features Id	Pmportance's
1	whezev	36.435764
2	cough12	16.850947
3	exwhez12	7.925729
4	pnose12	7.765365
5	ieyes12	5.770101
6	whez12	5.576501
7	iactiv12	4.817502
8	pnoseev	4.793485
9	awake12	4.184889
10	nwhez12	3.796222
11	speech12	2.083492

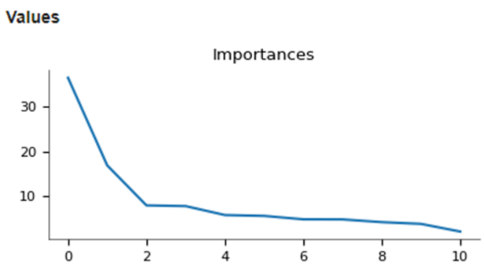


Fig. 6.1: Graphical representation of features importance

model using k-fold cross-validation. The authors did a 10-fold cross-validation on ten equal subsets of the dataset. Nine subgroups per fold were trained and the rest tested. Every subgroup was tested once during the ten-time method repeat. Averaged results were used to evaluate the model's efficacy.

Fig 6.2 presented that for improved decision-making, better model performance, and actionable insights, feature identification plays a critical role. Frequency distribution of features importance values for the proposed model, where X-axis indicates frequency counts and y-axis represents importance values of input variables. It represents frequency counts of importance values of input features in database.

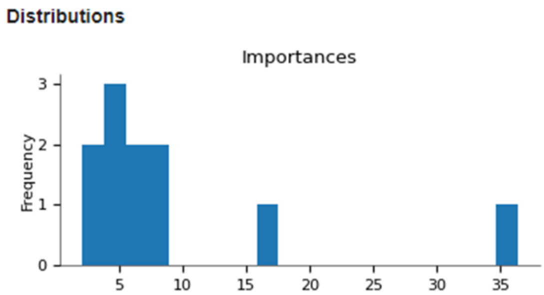


Fig. 6.2: Frequency distribution of features importance

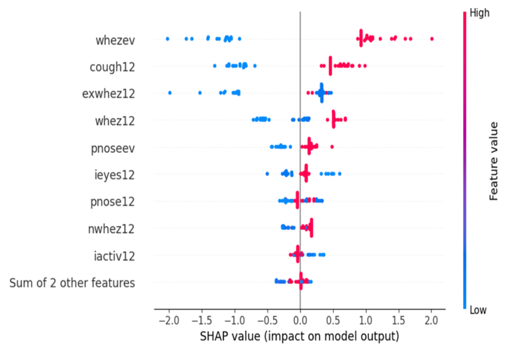


Fig. 6.3: Frequency distribution of features importance

Table 6.2: Comparative Table for results gained by Various model

Methods	Accuracy	Precision	Recall	F1-score
Logistic regression	47%	24.4%	47.7%	23.3%
RF	50%	25.5%	50%	33%
K-NN	77%	79.7%	77.3%	76.8%
SVM	84%	87.9%	84%	83.3%
Adaboost	84%	87.9%	83%	83.7%
Gradient Boosting	86%	86.9%	86.3%	86.3%
Proposed Method	98.91%	97.9%	98.2%	97.2%

6.2. SHAP. A Shapley additive explanation is basically used to illustrate the outcome of proposed prediction model. SHAP has various stimulating properties which permit it to be used on predictive model, to generate reliable descriptions and to handle complex model behaviors. Fig. 6.3 and 6.4 are presenting contribution of each input variables for model outcome using SHAP estimation. Fig 6.3 shows Variables contribution towards model performance. Fig 6.4 shows Values distribution of four features which are highly integrating in models prediction.

High precision, recall, f1-score and accuracy are always desirable and show better performance for a model. The proposed model has a higher indicator for all the performance assessing parameters, demonstrating the proposed model's strength over existing models. The proposed model exactly classifies the dataset into two categories: asthmatic and non-asthmatic. Table 6.2 shows the classification results of the proposed model.

Table 6.3 represents the ablation study of this work. Different types of situations have been involved in

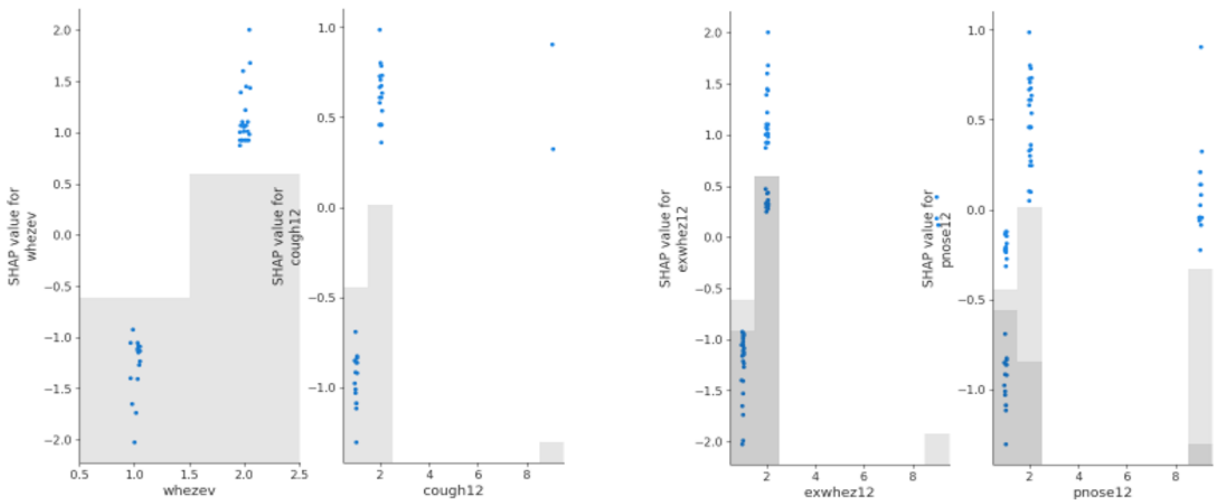


Fig. 6.4: Values distribution of four features which are highly integrating in models prediction

Table 6.3: Performance measures for Proposed work

Methods	Accuracy	Precision	Recall	F1-score
Existing CatBoost	0.8	0.78	0.85	0.89
Without Learning rate	0.78	0.73	0.81	0.83
Without Number of Trees)	0.75	0.79	0.85	0.80
Without Subsample)	0.65	0.93	0.87	0.69
Proposed Model with fine tuning	1.0	1.0	1.0	0.99

this study.

Several classifiers were evaluated for their accuracy, precision, recall, and F1-score. The evaluation of each classifier's performance in predicting asthma attacks was dependent on these essential criteria. The authors employed classifiers that exhibited unique error patterns and complementary strengths. Ensemble models benefit from diversity as it helps to mitigate overfitting and enhance generalization, resulting in increased robustness. We evaluated the processing efficiency of each classifier due to the need for swift predictions in medical applications. As a result of this need, the ensemble model demonstrated both accuracy and speed, making it suitable for practical use.

6.3. Discussion. We have analyzed that machine learning methods in combination with patient-reported asthma indication scores can predict asthma attack with good performance measures. In particular our proposed model, using catboost classifier together with intersection of all four feature extraction methods, achieved accuracy as high as 93% for asthma prediction and 100% accuracy for risk grade prediction. Class imbalance techniques played a vital role to better balance the asthmatic and non-asthmatic training data, allowing the proposed models to better discriminate minority cases. It was a necessary step due to the severe unbalancing in the original dataset (3.86% asthmatic cases, 98.14% non-asthmatic cases). There was possibility that most of the statistical and machine learning models would merely predict all cases as being in the majority class (ie. non-asthmatic cases) with a higher accuracy score near about 99%, but such a model would obviously not have medical efficacy. Consequently, implementations of class imbalance learning methods are essential to develop predictive models that give significant consequences.

Authors have extensively discussed our ensemble machine learning model's computing needs. These measures include examining inference and training times to identify issues as the dataset develops. The authors also discussed hardware acceleration and parallel processing to boost computer efficiency. We detailed the challenges of scaling our forecasting model to larger datasets. Memory consumption, model training duration, and real-time prediction efficiency are included. The authors also discuss model complexity-scalability compromises, emphasizing the need for more research to increase our approach's efficiency for larger applications. The authors' examples of computational complexity and scalability issues that potentially affect our forecasting model's actual deployment help contextualize these obstacles. Efficient computing resources are needed for timely and accurate asthma attack prognosis, including clinical deployment and integration with healthcare systems.

7. Conclusion and Future Scope. Machine Learning based catboost classifier outperforms the other classifiers as determined by a study of the numerous performance assessment strategies showing good consequences in terms of the main metrics specifically with an accuracy of 97.9% and F1-measure of 97.2% when subjected to performance evaluation using 10- fold cross validation. At the same time, the technique is seen to simplify well on any categorical data representing patient reported asthma response. A learning rate of 1.0 showed optimal consistent results with classification accuracy reaching as high as 93%. The method adopted for feature extraction (i.e. intersection of all four methods) has express good endeavors in selecting the most appropriate features with respect to the target variable asthmaev. Wheezing patterns and cough related symptoms are highly contributing features reserve place of one-third of the total number of features to create reduced feature set. Asthma risk grade prediction accuracy achieved through catboost classifier is 100%. Asthma is a progressive, long-lasting lung illness. Longitudinal predictions may be incorporated into future models to better understand how a disease progresses and how it responds to treatment. This proposed model prioritizes interoperability, ensuring smooth integration with existing EHR systems. This requires HL7 and FHIR data formats and protocols. Scalable architecture allows the model to manage large amounts of patient data and process it in real time, which is essential in clinical settings. Healthcare workers need an easy-to-use interface. Our method uses a clear GUI to display forecasts and risk grades. We recommend comprehensive medical staff training and technological support to enable successful integration. We're doing pilot tests with various healthcare institutions to verify the model's clinical efficacy. Our asthma episode and risk grade prediction method is accurate, which builds trust with healthcare providers. The model is reliable due to constant monitoring and updates. The strategy improves transparency and medical practitioner confidence by incorporating understanding into prediction.

CatBoost classifier is a new boosting algorithm which outperforms as compare to other boosting algorithm. Till now only limited studies have worked with catboost classifier. Early warnings delivered by models might be helpful to prevent or treat asthma exacerbations more effectively. Incorporation of few different modalities including genetic information, lung function testing and stress-related variables may also helpful to better elaborate asthma prediction. Research on techniques to advance the understandability and interpretability of models will remain to be important. A health expert or clinician must be first understood the model behavior and outcome produced by it before using it for their requirements or patients treatment. This proposed model can be implemented for another disease prediction work and can be worked with other dataset. Asthma-related metrics may be observed in real time with the assistance of incorporation with Internet of Things devices and wearable sensors.

Our model might be adapted to different patient demographics or healthcare settings with minimal retraining using transfer learning. We want to see how transfer learning can improve our forecasting model's generalizability and scalability. We may add biological traits and clinical data variables to our model to improve its predictive power. This includes investigating biomarkers, genetic data, and other physiological characteristics that may help diagnose asthma attacks and assess risk grades.

REFERENCES

[1] Siddiquee, J., Roy, A., Datta, A., Sarkar, P., Saha, S., & Biswas, S. S. (2016). Smart asthma attack prediction system using Internet of Things. *Proceedings of the 7th IEEE Annual Information Technology, Electronics and Mobile Communication Conference, IEEE IEMCON 2016*, 1–4. https://doi.org/10.1109/IEMCON.2016.7746252

[2] Achuth Rao, M. V., Kausthubha, N. K., Yadav, S., Gope, D., Krishnaswamy, U. M., & Ghosh, P. K. (2017). Automatic predic-

tion of spirometry readings from cough and wheeze for monitoring of asthma severity. *Proceedings of the 25th European Signal Processing Conference, EUSIPCO 2017*, 2017-January, 41–45. https://doi.org/10.23919/EUSIPCO.2017.8081165

[3] Do, Q. T., Doig, A. K., Son, T. C., & Chaudri, J. M. (2018). Personalized Prediction of Asthma Severity and Asthma Attack for a Personalized Treatment Regimen. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2018-July, 1–5. https://doi.org/10.1109/EMBC.2018.8513281

[4] Do, Q. T., Doig, A. K., Son, T. C., & Chaudri, J. M. (2018). Personalized Prediction of Asthma Severity and Asthma Attack for a Personalized Treatment Regimen. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2018-July, 1–5. https://doi.org/10.1109/EMBC.2018.8513281

[5] Luo, G., Stone, B. L., Fassl, B., Maloney, C. G., Gesteland, P. H., Yerram, S. R., & Nkoy, F. L. (2015). Predicting asthma control deterioration in children. *BMC Medical Informatics and Decision Making*, 15(1), 1-8.

[6] Gold, D. R., Damokosh, A. I., Dockery, D. W., & Berkey, C. S. (2003). Body-mass index as a predictor of incident asthma in a prospective cohort of children. *Pediatric Pulmonology*, 36(6), 514-521.

[7] Do, Q. T., Doig, A. K., & Son, T. C. (2019). Deep Q-learning for Predicting Asthma Attack with Considering Personalized Environmental Triggers' Risk Scores. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 562–565. https://doi.org/10.1109/EMBC.2019.8857172

[8] Kocsis, O., Lalos, A., Arvanitis, G., & Moustakas, K. (2019). Multi-model Short-term Prediction Schema for mHealth Empowering Asthma Self-management. *Electronic Notes in Theoretical Computer Science*, 343, 3–17. https://doi.org/10.1016/j.entcs.2019.04.007

[9] Hoq, M. N., Alam, R., & Amin, A. (2019). Prediction of possible asthma attack from air pollutants: Towards a high density air pollution map for smart cities to improve living. *Proceedings of the 2nd International Conference on Electrical, Computer and Communication Engineering, ECCE 2019*, 1–5. https://doi.org/10.1109/ECACE.2019.8679335

[10] Do, Q., Tran, S., & Doig, A. (2019). Reinforcement Learning Framework to Identify Cause of Diseases-Predicting Asthma Attack Case. *Proceedings of the 2019 IEEE International Conference on Big Data, Big Data 2019*, 4829–4838. https://doi.org/10.1109/BigData47090.2019.9006407

[11] Luo, J., & Long, Y. (2020). NTSHMDA: Prediction of Human Microbe-Disease Association Based on Random Walk by Integrating Network Topological Similarity. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 17(4), 1341–1351. https://doi.org/10.1109/TCBB.2018.2883041

[12] Priya, C. K., Sudhakar, M., Lingampalli, J., & Basha, C. Z. (2021). An Advanced Fog based Health Care System Using ANN for the prediction of Asthma. *Proceedings of the 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*, 1138–1145. https://doi.org/10.1109/ICCMC51019.2021.9418248

[13] Lisspers, K., Ställberg, B., Larsson, K., Janson, C., Müller, M., Łuczko, M., Bjerregaard, B. K., Bacher, G., Holzhauer, B., Goyal, P., & Johansson, G. (2021). Developing a short-term prediction model for asthma exacerbations from Swedish primary care patients' data using machine learning - Based on the ARCTIC study. *Respiratory Medicine*, 185(February). https://doi.org/10.1016/j.rmed.2021.106483

[14] Aditya Narayan, S., Nair, A. Y., & Veni, S. (2022). Determining the Effect of Correlation between Asthma/Gross Domestic Product and Air Pollution. *Proceedings of the 2022 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2022*, 44–48. https://doi.org/10.1109/WiSPNET54241.2022.9767145

[15] Tong, Y., Wang, Y., Zhang, Q., Zhang, Z., & Chen, G. (2022). A Reliability-constrained Association Rule Mining Method for Explaining Machine Learning Predictions on Continuity of Asthma Care. *Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2022*, 1219–1226. https://doi.org/10.1109/BIBM55620.2022.9995400

[16] Mahammad, A. B., & Kumar, R. (2022). Machine Learning Approach to Predict Asthma Prevalence with Decision Trees. *Proceedings of the International Conference on Technological Advancements in Computational Sciences, ICTACS 2022*, 263–267. https://doi.org/10.1109/ICTACS56270.2022.9988210

[17] Lilhore, U. K., Dalal, S., Faujdar, N., Margala, M., Chakrabarti, P., Chakrabarti, T., ... & Velmurugan, H. (2023). Hybrid CNN-LSTM model with efficient hyperparameter tuning for prediction of Parkinson's disease. *Scientific Reports*, 13(1), 14605.

[18] Kroes, J. A., Zielhuis, S. W., Van Roon, E. N., & Ten Brinke, A. (2020). Prediction of response to biological treatment with monoclonal antibodies in severe asthma. *Biochemical Pharmacology*, 179, 113978.

[19] Dalal, S., Lilhore, U. K., Simaiya, S., Jaglan, V., Mohan, A., Ahuja, S., ... & Chakrabarti, P. (2023). A precise coronary artery disease prediction using Boosted C5. 0 decision tree model. *Journal of Autonomous Intelligence*, 6(3).

[20] Saha, C., Riner, M. E., & Liu, G. (2005). Individual and neighborhood-level factors in predicting asthma. *Archives of Pediatrics & Adolescent Medicine*, 159(8), 759-763.

[21] Castro-Rodriguez, J. A., Cifuentes, L., & Martinez, F. D. (2019). Predicting asthma using clinical indexes. *Frontiers in Pediatrics*, 7, 320.

[22] Deshwal D, Sangwan P, Dahiya N, et al. COVID-19 Detection using Hybrid CNN-RNN Architecture with Transfer Learning from X-Rays. *Current Medical Imaging.* 2023 Aug. DOI: 10.2174/1573405620666230817092337. PMID: 37594157.

[23] Ram, S., Zhang, W., Williams, M., & Pengetnze, Y. (2015). Predicting asthma-related emergency department visits using big data. *IEEE Journal of Biomedical and Health Informatics*, 19(4), 1216-1223.

[24] Mrazek, D. A., Klinnert, M., Mrazek, P. J., Brower, A., McCormick, D., Rubin, B., ... & Jones, J. (1999). Prediction of early-onset asthma in genetically at-risk childre

[25] Monadi, M., Firouzjahi, A., Hosseini, A., Javadian, Y., Sharbatdaran, M., & Heidari, B. (2016). Serum C-reactive protein in asthma and its ability in predicting asthma control, a case-control study. *Caspian Journal of Internal Medicine*, 7(1), 37.

[26] Jaiswal, V., Saurabh, P., Lilhore, U. K., Pathak, M., Simaiya, S., & Dalal, S. (2023). A breast cancer risk predication and

classification model with ensemble learning and big data fusion. *Decision Analytics Journal*, 100298.
[27] Forno, E., & Celedón, J. C. (2019). Epigenomics and transcriptomics in the prediction and diagnosis of childhood asthma: are we there yet?. *Frontiers in Pediatrics*, 7, 115.
[28] Priya, C. K., Sudhakar, M., Lingampalli, J., & Basha, C. Z. (2021, April). An advanced fog based health care system using ann for the prediction of asthma. *Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1138-1145). IEEE.
[29] Kaan, A., Dimich-Ward, H., Manfreda, J., Becker, A., Watson, W., Ferguson, A., ... & Chan-Yeung, M. (2000). Cord blood IgE: its determinants and prediction of development of asthma and other allergic disorders at 12 months. *Annals of Allergy, Asthma & Immunology*, 84(1), 37-42.

Edited by: Manish Gupta
Special issue on: Recent Advancements in Machine Intelligence and Smart Systems
Received: Apr 17, 2024
Accepted: Jul 9, 2024