AI Powered Asthma Prediction Towards Treatment Formulation: An Android App Approach

MD MEHEDI HASAN

19 May 2022

Department of Information and Communication Engineering
Noakhali Science And Technology University

Abstract

Asthma is a disease which attacks the lungs and that affects people of all ages. Asthma prediction is crucial since many individuals already have asthma and increasing asthma patients is continuous. Machine learning (ML) has been demonstrated to help individuals make judgments and predictions based on vast amounts of data. Because Android applications are widely available, it will be highly beneficial to individuals if they can receive therapy through a simple app. In this study, the machine learning approach is utilized to determine whether or not a person is affected by asthma. Besides, an android application is being created to give therapy based on machine learning predictions. To collect data, we enlisted the help of 4,500 people. We collect information on 23 asthma-related characteristics. We utilized eight robust machine learning algorithms to analyze this dataset. We found that the Decision tree classifier had the best performance, out of the eight algorithms, with an accuracy of 87integrate machine learning with an Android application. We accomplished asthma therapy using an Android application developed in Java and running on the Android Studio platform.

Keyword

Artificial intelligence; asthma prediction; machine learning; android application.

1 Introduction

Asthma is a lung condition that affects children and adults alike. The air passages in the lungs constrict due to inflammation and stiffness of the muscles around the small airways. In 2019, asthma afflicted an estimated 262 million individuals, resulting in 461000 deaths [1]. The most prevalent chronic illness among children is asthma. Although it affects people of all ages, asthma is more common in children and is the most common cause of chronic airway disease. As patients become older, it becomes increasingly challenging to identify [2] adult-onset asthma from other conditions such as chronic obstructive pulmonary disease (COPD) or Asthma-COPD overlap syndrome (ACOS), leading to frequent misdiagnosis or underdiagnosed [3]. Furthermore, asthma imposes a high cost on the social, economic, and health-care systems of many countries. Inhaled medication can help individuals with asthma control their symptoms and live every day, active life. Eliminating asthma triggers may also help to reduce asthma symptoms. Due to fast population expansion, automatic disease identification has become a critical topic in medical science. Coronavirus (COVID-19) has spread extensively and has become one of the most severe and acute infections in recent years. This paper [4] describes a deep learning strategy for diagnosing COVID-19 from X-ray pictures that combines a convolutional neural network (CNN) and a long short-term memory (LSTM). In [5], the author gave an overview of newly created systems based on deep learning techniques that used several medical imaging modalities such as CT and X-ray. In recent years, a slew of Asthma prediction techniques has been developed and published. Tong et al. [6] developed an automated technique for producing explanations for rules for the predictions of any machine learning model based on imbalanced tabular data and for proposing tailored actions without compromising prediction precision. Our technology has worked fine to explain our Intermountain Healthcare model forecasts [7]. However, the technique has a drawback in that it is not used in reality. The disadvantage of this paradigm is that it is only implemented for children, but our model is for individuals of all ages. In [8], the author has created prediction algorithms to assess the health data for children with Asthma. These prediction models are developed using machine learning classifiers, including linear regression, decision trees, the Random Forest, KNN [9] and methods of the Naive Bayes [10,11]. They got a good accuracy [12] for random forest classifiers and it's around 91machine learning for Asthma risk prediction (ML) [14]. The author created a model for prediction but did not provide any treatment for the patient, which is something we did in our model. The complete technology is used via Internet-of-Things resources on a mobile-health (m-health) phone (IoT). This investigation shows a connection between the indoor and outdoor particulate matter (PM) and the PEFR. The mining of data comprising patients' past health information has utilized an efficient technique for Asthma illness prediction [15]. These studies utilize five methods of categorization for machines called Naive Bayes, J48, Random Forest and Random Tree. The authors of [16] have created a breast cancer prediction model that employs 10-fold cross validation to achieve a reliable result. Wisconsin breast cancer diagnosis data collection is collected from the UCI machine learning repository for breast cancer. Another prediction model for breast cancer is also being developed in [17]. In [18], a new mathematical model for breast cancer was created. The difference between those three papers and our research is that they focused on breast cancer while we focused on asthma. In [19], the authors propose an approach for employing a deep neural network to diagnose diabetes by training its features in a five-fold and ten-fold cross validation fashion. They only utilised one model in this study, but we used eight in ours. Using Random Forests (RFs) and Artificial Neural Networks, this research [20] proposes an expert system for the classification of liver disorders (ANNs). The approaches employ 10-fold cross validation to train the input features, while we used 23. Among all of these disorders, heart disease is the most frequent, and its effects are far more deadly than those of any other disease. The author of this paper [21] compared a lot of computational intelligence algorithms for coronary artery heart disease prediction, but they don't provide any practical applications like Android apps. In [22] the 2012–2014 real-world Asthmatic patients were trained in a big town in China on prediction models including logistic regression, random forest, supporting vector machines, regression classification tree and backpropagation neural network [23]. The goal of the author was to create models that used machine learning to forecast the probability of exacerbations. The goal of the author [24] was to create models that used machine learning to forecast the probability of exacerbations. AUPRC = 0.007 (950.0002) was provided for the model validated on the test data, suggesting that it may be unable to estimate an incident of Asthma exacerbation for the near future by historical clinical data only. Machine learning is often recognized as the most effective method for predicting illnesses such as heart disease and asthma. For asthma illness prediction, we employ 23 characteristics. Compared to previous approaches in works like [15] the findings demonstrate a higher degree of performance. We utilize eight machine learning algorithms, and the decision tree achieved the highest accuracy (87from that, the number of Android users is growing by the day. To carry out the therapy, we use the Java programming language to develop an Android application in which we recommended necessary lab tests and medications for Asthma patients. In recent years, inhaler therapy has been a popular therapeutic option in the medical sector. Followings are the concrete contents of contributions in this paper: 1. Using machine learning techniques, improve the accuracy of asthma illness forecasts. We used 23 asthma risk factor characteristics and consulted with a renowned asthma doctor to complete this assignment. All of the characteristics are significantly linked to asthma. We discovered that the Decision Tree algorithm has the highest accuracy, which is 87 percent. 2. Provide proper treatment for asthma patients. After we've trained all of the asthma characteristics, our model will tell us whether or not this person has asthma. If the model finds a positive output, it indicates he or she has asthma, and the model will provide treatment; otherwise, the model will notify that he or she is asthma-free. 3. Developed an Android app for identifying asthma patients and administering medication. For asthma patients, this application contains 23 fields that are utilized to identify the patient. After entering all of his information, the app will notify him/her if he/she has asthma or not. This application also makes appropriate recommendations based on the output. The remainder of the article is arranged as follows, with Section 2 delving into the methodology. We went through all of the preprocessing, ML algorithms, and Java applications. Section 3 goes through the results and discussions. We have demonstrated all of the expected results and

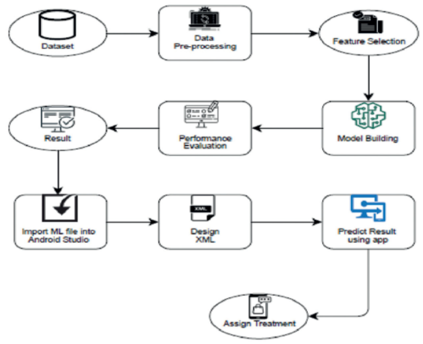

Figure 1: Experiment workflow

the development of the Android application in this section. Section 4 describes the work that will be done in the future. A conclusion is included in Section 5.

2 Methodology

In this work, we perform the job in two modules. The data analysis and prediction are accomplished by machine learning, whereas the treatment is performed by an android application developed in Java. We utilized Google Colab to do Asthma data pre-processing and prediction [25]. We use five key steps to obtain the forecast: data sources, label verification, data prepossessing, training data using ML algorithms, and prediction. The data quality is maintained by using a large number of samples with a wide range of variance. To get the prediction in machine learning, we have utilized eight robust ML algorithms and made a comparison between those algorithms. After that, we have utilized Android Application 9.0 to development an application, and the languages used are Java and XML. In this case, Java is used for coding, while an XML file is used for the design section. Fig. 1 describes how the machine learning models and android application in the forecast node have been implemented. During training, the selection of features is employed to remove duplicates that do not significantly contribute to the prediction result and enhance overall fidelity to the model. The imputation will initially be done to adapt the missing data, followed by the feature scaling to standardize the dataset's value range. Finally, the binary classifiers are fitted to the data during the k-fold cross validation phase with all the samples for training, validation, and testing to provide a more robust classification [13] following that, the machine learning file is loaded into Android Studio. We generated a file using Tensorflow after receiving predictions from ML models, and this file was imported into Android Studio. Based on the ML algorithm, this smartphone app attempts to determine whether or not this person has Asthma. Finally, this application provides the appropriate therapy based on this forecast.

2.1 Machine Learning

For each iteration m, the performance of each model built based on 22 features and ML methods employed is recorded. We have assessed our dataset using eight ML algorithms to achieve a better outcome. Sub-Section (1) covers data sources, Sub-Section (2) examines the verification of the label, Sub-Section (3) provides predisposition of information, Sub-Section (4) describes the implementation of ML algorithms, sub-Section (5) compare the performance between those implemented algorithms and finally best outcome is shown as result and this result is transfer to android studio.

2.1.1 Data Sources

We consult with Asthma professionals before collecting data. The doctor explained to us the characteristics that cause asthma and we had taken a short note from his description. Finally, we gathered 23 features from him. We built a Google form based on those features. For data collection, we used social media and went to the doctor's chamber. To gather data via social media, we shared this Google form with an open public group, posted it on our timeline, emailed it to our known contacts and we got a good response from there. Beside this, we spent many days physically collecting data from the doctor's chamber to take the data from asthma patents and who has symptoms of asthma. Some prior studies only collected data from children [9], but we gathered data from people of all ages in our study. We don't have any age restrictions. Finally, we collected around 4500 data points from social media and the visiting doctor chamber. Tab. 1 shows all the features based on which we collect the data.

2.1.2 Data Pre-Processing

Following the gathering of various records, data on asthma illness is pre-processed [13]. The collection comprises about four thousand five hundred entries, many of which have missing values. We utilized a mathematical model, mean, and median to fill in the missing value. **Mean:**In mathematics and statistics, the idea of mean is crucial. The arithmetic average of all terms is the most frequent formulation for the mean of a statistical distribution containing a discrete random variable. In a group of numbers, the mean is the average or most frequent value which is expressed by the Eq. (1)

$$Mean = [\frac{\sum_{i=1}^{n} X_i}{n}] \tag{1}$$

Median: Whether the number of terms in a discrete random variable distribution is even or odd affects the median. If the total number of observations (n) is odd, the median can be represented by Eq. (2).

$$Median = [\frac{n+1}{2}]obsservation \tag{2}$$

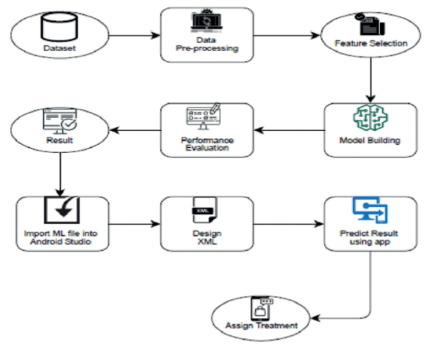

Figure 2: Android apps development steps of asthma prediction

2.2 Android App Development

The Platform utilized in the development of this application is 9.0, and the coding language is Java. This application's design is finished in XML. To complete the coding and design we have used Android Studio. Fig. 2 explain the development process of this application.

Table 3: Eight algorithm and metrics score

Algorithm	Accuracy	F1 score	Recall	Precision
Random Forest Classifier	0.80	0.73	0.72	0.76
K-Nearest Neighbor	0.81	0.72	0.71	0.78
Bernoulli Naive Bayes	0.75	0.71	0.72	0.70
MLP Classifier	0.81	0.75	0.76	0.77
Adaboost Classifier	0.80	0.76	0.76	0.76
Decision Tree	0.87	0.81	0.80	0.83
SVM Classifier	0.80	0.73	0.72	0.76
Impact Learning	0.85	0.78	0.78	0.81

Reference

WHO, Asthma, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/asthma.

M. A. Awal, M. S. Hossain, K. Debjit, N. Ahmed, R. D. N. Nemri et al., "An early detection of asthma using bomla detector," IEEE Access, vol. 9, pp. 58403–58420, 2021.

O. Enilari and S. Sinha, "The global impact of asthma in adult populations," Annals of Global Health, vol. 85, no. 1, pp. 622, 2019.

M. Z. Islam, M. M. Islam and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," Informatics in Medicine Unlocked, vol. 20, pp. 100412, 2020. [5] M. M. Islam,

F. Karray, R. Alhajj and J. Zeng, “A review on deep learning techniques for the diagnosis of novel
coronavirus (COVID-19),” IEEE Access, vol. 9, pp. 30551–30572, 2021. [6] Y. Tong, A. I. Messinger and G. Luo, “Testing the generalizability of an automated method for explaining machine learning predictions on asthma patients asthma hospital visits to an academic healthcare system,” IEEE Access, vol. 8, pp. 195971–195979, 2020.