International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

E-health Chain and Anticipation of Future Disease

Rohit Dhonde[1], Pradnesh Khedekar[2], Pradeep Kshirsagar[3] and Prof. Manasi Kulkarni[4]

[1,2,3]Member, Pillai College of Engineering, New Panvel, Maharashtra – 410206, India
[4]Guide, Pillai College of Engineering, New Panvel, Maharashtra – 410206, India

---

Abstract - *E-Health Chain & Anticipating Future Diseases is a system which aims at maintaining Electronic Health Records (EHRs) in a more efficient way as compared to traditional way of storing and maintaining paper based health records. The Digital prescription module can be used by patients to buy medicines from pharmaceutical stores by just providing a unique ID of the patient which helps pharmacists to access the latest prescribed medicines. Electronic Health Records (EHRs) allows doctors to access a patient's health records simply from one electronic file, doctors can browse, check results as they are entered, together with image files like X-rays even from remote hospitals. In an emergency scenario, a doctor can use a patient's Unique ID code to browse time-critical data, like allergies, blood groups, recent treatments. In situations of emergency, the historical data will assist the doctors to take effective actions and safeguard the life of the patients. Tab reminder alert helps patients to take medicines on time as prescribed by doctor. System uses machine learning algorithms to predict and examine future diseases which helps patients to take preventive measures.*

Key Words: **E-Health records(EHR), E-healthchain, disease predictor, Random Forest, E-prescription, NLP, E-ambulance**

1. INTRODUCTION

E-Health Chain & Anticipating Future Diseases is a system which aims at maintaining Electronic Health Records (EHRs) in a more efficient way as compared to traditional way of storing and maintaining paper based health records. Digital prescription modules can be used by patients to buy medicines from pharmaceutical stores by just providing a unique ID of the patient which helps pharmacists to access the latest prescribed medicines. Electronic Health Records (EHRs) allows doctors to access a patient's health records easily from a single electronic file, doctors can read test results as they are entered, including image files such as X-rays even from remote hospitals. E-Ambulance is a quick-response solution that can detect and place a phone call for the ambulance and send the ambulance to the required destination. In an emergency situation, a doctor can use a patient's Unique code to read time-critical data, such as,recent treatments, allergies, and blood type. In situations of emergency, the historical data will assist the doctors to take effective actions and safeguard the life of the patients. Tab reminder alert helps patients to take medicines on time as prescribed by doctor. System uses machine learning algorithms to predict and examine future diseases which helps patients to take preventive measures.

2. LITERATURE SURVEY

Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization[1]. The aim of this work was to compare algorithms with all different performance measures using machine learning. All data was pre-processed and used to take a glance at the prediction. every worked higher in some things and worse in others. K-Nearest Neighbour K-NN, and Random Forest RF and Artificial Neural Network MLP are the models apparently to work best inside the knowledge set used in this study. Experimental results show that the improvement hybrid approach can increase the predictive accuracy of medical data sets. The projected ways that are compared to supervised algorithms based on existing approximate sets and classification accuracy measurements are used to measure the performance of the proposed approaches. Therefore, the analysis section clearly incontestable the effectiveness of hybrid PSO and ACO approaches to malady diagnosing compared to different existing approaches. The projected optimized model by FCBF, PSO and ACO succeed an accuracy score of 99.65% with KNN and 99.6% with RF.

Liver disease prediction by using different decision tree techniques[2]. The study used some decision tree formula like J48, LMT, Random Forest, Random tree, REPTree, decision Stump and Hoeffding Tree to predict the disease at an earlier stage. These formulas provide varied results supported Accuracy, Mean Absolute Error, Precision, Recall, kappa statistics and Runtime. These techniques were evaluated and their performance was compared. From the analysis, Decision Stump outperforms well than different algorithms and its achieved accuracy is 70.67%. The performance measure used for comparison are listed within the table (Table 2) the application of decision tree in predicting disease can benefit in managing the health of people. However, within the future, we are going to collect the very recent data.

Disease prediction using machine learning[3]. A machine learning and new multimodal disease risk prediction algorithmic rule supported the convolutional neural network (CNN-MDRP) using structured and unstructured data. By giving the input of symptoms we'll get correct disease prediction as output, which can facilitate us perceive the level of disease risk prediction. This technique leads to low time

consumption and minimal value possible for illness prediction.

Prediction of probability of disease based on symptoms using machine learning algorithms[4]. A machine learning and new convolutional neural network based multimodal disease risk prediction (CNN-MDRP) algorithm using structured and unstructured data from hospital for effective prediction of diseases. Existing work is not focused on both data types in the area of healthcare. Compared to several typical prediction algorithms, the proposed algorithm accuracy prediction reaches 94.8% than that of the CNN-based unimodal disease risk prediction (CNN-UDRP) algorithm.

Performance Analysis of Machine Learning Algorithms on Diabetes Dataset using Big Data Analytics[5]. Today's world people are more involved in their hectic schedules by not taking care of their health, that ends up in chronic issues like diabetics. in this paper, the author tries to provide a comprehensive comparative study on different machine learning algorithms. This comparative study is done based on totally different metrics like Accuracy, Kappa, Precision, Recall, Sensitivity and Specificity. The achieved results show that RF formula is predicting hectic a lot of properly and accurately.

Disease Prediction by Machine Learning Over Big Data From Healthcare Communities[6]. A new convolutional neural network based multimodal disease risk prediction (CNN-MDRP) algorithm using structured and unstructured data from hospital. To the best of our knowledge, none of the existing work focused on both data types in the area of medical big data analytics. Compared to several typical prediction algorithms, the prediction accuracy of our proposed algorithm reaches 94.8% with a convergence speed which is faster than that of the CNN-based unimodal disease risk prediction (CNN-UDRP) algorithm.

Predictive Analytics for Chronic Kidney Disease Using Machine Learning Techniques[7]. The predictive models by using machine learning methods including K-nearest neighbors (KNN), support vector machine (SVM), logistic regression (LR), and decision tree classifiers to predict chronic kidney disease.

Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease[8]. Data mining techniques for various analysis of medical data may be a smart methodology. The performance of decision tree methodology was found to be 91 accurate compared to naive bayes methodology. Classification algorithmic rule on diabetes dataset performance was obtained as 94% Specificity and 95% Sensitivity. We additionally found that mining helps to retrieve correlations from attributes that aren't direct indicators of the category that we tend to try to predict. It is more working on enhancing the performance of prediction system accuracy in neural networks and bunch algorithmic rule knowledge analysis.

Data Analysis on Health Management Systems for Improving Doctor's Advice on Patients[9]. Focusing on improving doctors' advice for patients via the data of physical examinations, and helps patients get to know their physical condition as accurately as possible. On basic data collecting, we consider that the exchange of data with EMR is an advisable way to let doctors obtain more information about a patient's medical history. As an assisted diagnostic method, the relationships about some unknown abnormal items with special diseases can be discovered by analyzing the data based on HMS.

Machine learning applications in cancer prognosis and prediction[10]. The concepts of ML while we outlined their application in cancer prediction/prognosis. Most of the studies that have been proposed in the last few years and focus on the development of predictive models using supervised ML methods and classification algorithms aiming to predict valid disease outcomes. Based on the analysis of their results, it is evident that the integration of multidimensional heterogeneous data, combined with the application of different techniques for feature selection and classification can provide promising tools for inference in the cancer domain.

3. PROPOSED SYSTEM

Existing System Architecture

The existing manual method of maintaining a patient record, maintaining doctor's data, day to day activities and request is hard and thus a system or application which might complete these tasks in an exceedingly straightforward to use is what we are able to deliver through this application.

A health system consists of all organizations, folks and actions whose primary intent is to market, restore or maintain health. This includes efforts to influence determinants of health in addition as additional direct health-improving activities. A health system is thus over the pyramid of publically closely-held facilities that deliver personal health services. It includes, as an example, a mother caring for a sick kid at home; personal providers; behaviour modification programmes; vector-control campaigns; insurance organizations; activity health and safety legislation. It includes inter-sectoral action by health workers, as an example, encouraging the ministry of education to market feminine education, a documented determinant of higher health.

Proposed System Architecture

E-Health Chain & Anticipating Future Diseases is a system which aims at maintaining Electronic Health Records (EHRs) in a more efficient way as compared to traditional way of storing and maintaining paper based health records. The

Digital prescription module can be used by patients to buy medicines from pharmaceutical stores by just providing a unique ID of the patient which helps pharmacists to access the latest prescribed medicines. Electronic Health Records (EHRs) allows doctors to access a patient's health records easily from a single electronic file, doctors can read test results as they are entered, including image files such as X-rays even from remote hospitals. E-Ambulance may be a quick-response answer which will notice and place a phone call for the ambulance and quickly send the emergency ambulance to the required destination. In an emergency scenario, a doctor will use a patient's unique ID code to scan time-critical info, like allergies,blood type, and up to date treatments. In situations of emergency, the historical data will assist the doctors to take effective actions and safeguard the life of the patients. A system in which it will store all the patient's historical data and the data will be used to predict and examine the future diseases using Machine Learning Algorithm.If user's symptoms don't specifically match any sickness within the info, then it's shows the diseases user might most likely have based on his/her symptoms..This can be used by Doctor to see Users historical data as well as by the Pharmacist to provide medicine prescribed by Doctor's. User can login using Unique ID which can be used by Doctors to Access their previous records.Electronic Health Records (EHRs) allows doctors to access a patient's health records simply from one electronic file, doctors will browse take a look at results as they're entered, together with image files like X-rays even from remote hospitals.In E-ambulance it can used to enter symptoms during the time when the patient is in the ambulance.

Patient Registration: If Patient could be a new user he can enter his personal details and he can use user Id and password through that he will login to the system.

Patient Login:If Patient already has an existing account then he/she will log into the system.

View Details: Patient and Doctor each will read their entered details. Doctor can also read Patients details and patients can read solely doctors very little information.

Diseases Prediction: Patient can specify the symptoms caused because of his unhealthiness. System will ask certain questions regarding his illness and system predict. The disease based on the symptoms specified by the patient and system will also suggest doctor based on the disease.

Doctor: Doctor will record new data.

Pharmacist: Will provide medicine to the Patient which is prescribed by the Doctor.

Role Based Access Control: Role-based access control (RBAC) is a method of restricting network access based on the roles of individual users within an enterprise. RBAC lets Doctor's have access rights only to the information they need to do their jobs(eg. Patient and Disease Profile) and prevents them from accessing information that doesn't pertain to them.

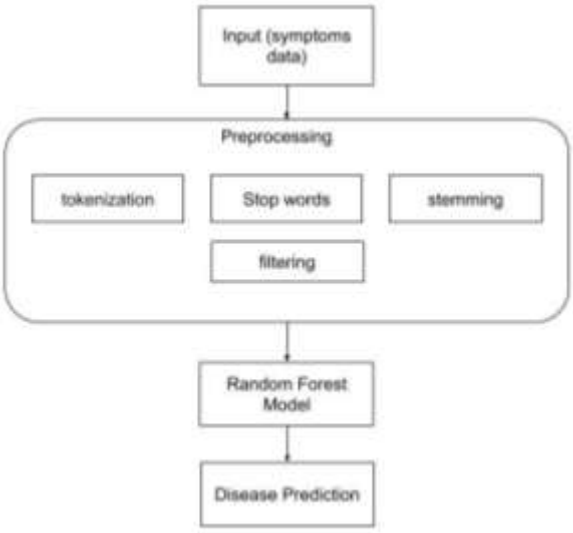

Figure1: Proposed system architecture

Preprocessing of data:

1. **Punctuation Removal**: The punctuation marks are removed from the text because they add no meaning to the data thus of no use.
2. **Blank/White space Removal**: To remove leading and ending spaces, you can use the *strip()* function. It is helping to reduce the memory uses and increase the efficiency of the model.
3. **Stemming/Lemmatization:** The aim of lemmatization, like stemming, is to reduce inflectional forms to a common base form. As opposition stemming, lemmatization doesn't merely lop off inflections. Instead it uses lexical information bases to urge the right base types of words.
4. **Stop-Word Removal:** The removal of Stopwords (such as "is","the".etc) is called Stop Word Removal. Stopwords add very little meaning so if removed the database space is saved and processing speed improves.

Disease Prediction Algorithm

Random Forest:

Random forests or random decision forests or associate ensemble learning methodology for classification, regression and different tasks that operates by constructing a large number of decision trees at training time and outputting the class that's the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees habit of overfitting to their training set. In decision prediction the Random

Forest algorithm is used to process the symptoms as input and gives diseases as output. Based on the probability of predicted diseases top 4 diseases with brief information are shown as output.

1) Preliminaries: decision tree learning
2) Bagging

The training algorithm for random forests applies the general technique of bootstrap aggregating, or bagging, to tree learners. Given a training set $X = x_1, ..., x_n$ with responses $Y = y_1, ..., y_n$, bagging repeatedly (B times) selects a random sample with replacement of the training set and fits trees to these samples:

For $b = 1, ..., B$:

1. Sample, with replacement, n training examples from X, Y; call these X_b, Y_b.
2. Train a classification or regression tree f_b on X_b, Y_b.

Accuracy Parameter:

1. Recall

Recall actually calculates how many of the Actual Positives our model captures through labeling it as Positive (True Positive).Recall shall be the model metric we use to select our best model when there is a high cost associated with False Negative.
Recall = (True Positive)/(True Positive+False Negative)

2. Precision

Precision talks about how precise/accurate your model is out of those predicted positive, how many of them are actual positive. Precision is a good measure to determine when the costs of False Positive is high.
Precision = (True Positive)/(True Positive+False Positive)

3. Confusion Matrix

A Confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known.

REFERENCES

[1] Machine learning applications in cancer prognosis and prediction.

[2] Predictive Analytics for Chronic Kidney Disease Using Machine Learning Techniques

[3] Performance Analysis of Machine Learning Algorithms on Diabetes Dataset using Big Data

[4] Disease Prediction by Machine Learning Over Big Data From Healthcare Communities.

[5] Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease.

[6] PREDICTION OF PROBABILITY OF DISEASE BASED ON SYMPTOMS USING MACHINE LEARNING ALGORITHM.

[7] Data Analysis on Health Management Systems for Improving Doctor's Advice on Patients.

[8] Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization.

[9] DISEASE PREDICTION BY USING MACHINE LEARNING.

[10] LIVER DISEASE PREDICTION BY USING DIFFERENT DECISION TREE TECHNIQUES.