Predictive Risk-Assessment System (PRAS) Platform for Healthcare Analytics and Visualization

G.JayaLakshmi, Dr.Sangeetha Yalamanchili

Abstract— Predictive modeling assumption and implementation has established accomplishment in online business intelligence and is at present affecting from theory into practical implementation in healthcare. In the advancement of expansion for health care applications, clinical decision frameworks (CDFs) were intended to advise practitioners of probable on contraindication. CDFs are proposed to progress care and decrease outlay. The framework is implanted into online healthcare record software. The idea was to develop a predictive model for in sequence systems to recognize patients at improved hazard for chronic diseases and to vigilant medicinal professionals to take necessary precautionary actions. And the level of healthcare data has grown extremely in recent years, raising the need to present the data in a way that are more comprehensible and perceptive.

Keywords: healthcare; machinelearning; predictive analytics; EHRs; Big Data

1. INTRODUCTION

In this digital era, healthcare industry is being distorted by the progressions in machine learning (or deep learning) and artificial intelligence (AI). By binding enormous big data, ML is being used in healthcare to present superior patient care and has resulted in improved better outcomes. Analysts are predicting that by the year 2018, about 30% of the healthcare providers will make use of cognitive analytics to better evaluate the patient data. These days, machine learning supports personalized care through enhanced diagnostics and predictive and prescriptive healthcare advices.

By using Big Data technology next to with machine learning and artificial intelligence, healthcare companies can build accurate decisions, significantly get better operating efficiencies, and reduce unwanted costs. With their enhanced efficiencies, healthcare companies can now save lives. Organizations will be capable to verify which patient is at a higher risk of contracting a particular disease. In addition, post-discharge outcomes can too be kept under control and the amount of re-admissions can be reduced considerably. Furthermore, diagnoses will take less time and patients will be able to know instantly what they are suffering from and take action for treatment.

'Big Data' is the huge volume of data that is exponentially growing. Data analytics embraces extracting the functional information from the data by building every possible relation amongst various data[2]. Big data has mainly three characteristics: volume, velocity and variety. Big data analytics helps in resulting solutions for problems like cost reduction, lowering the hazard in decision making and time-saving. Machine learning in data analytics helps to manage big data better and used to guide machines by feeding datasets and making algorithms that facilitate machines in decision making and problem solving. The intention of machine learning is to achieve as several positive outcomes with increasingly accurate predictions. Both big data and machine learning technologies can be employed in computers to facilitate smart intelligent decisions.

A. Transformation of Healthcare by Big Data and Machine Learning

The health sector has huge data volumes from health records, genomic data, health scans, smartphone apps, and wearable's. As a result, healthcare organizations are progressively finding ways to constantly produce data-driven insights through the implementation of machine learning which delivers solutions for processing data, developing and deploying algorithms to generate insights to help solve ongoing healthcare challenges [2].

In machine learning, a computer system uses algorithms to discover patterns in data without specific commands from humans. The sophisticated algorithms are used to understand data (from "training set") with the use of classifiers (features or attributes for classification of subjects) to make predictions (from "test set"). There are supervised and unsupervised techniques. Popular concepts of machine learning consist of supervised techniques such as hidden Markov models, naive Bayesian classifiers ,neural networks, and support vector machines. These methods receive raw data and use analgorithm to predict the outcome depending upon a prior training set of data.

Revised Version Manuscript Received on February 14, 2019.

G.JayaLakshmi, Assistant Professor:Dept. of Information Technology, V R Siddhartha Engineering College, Vijayawada, AndhraPradesh, India (Email: jaya1123@gmail.com)

Dr.Sangeetha Yalamanchili, Associate Professor: Dept. of Information Technology, V R Siddhartha Engineering College, Vijayawada, AndhraPradesh, India (Email: sangeetha18.yalamanchili@gmail.com)

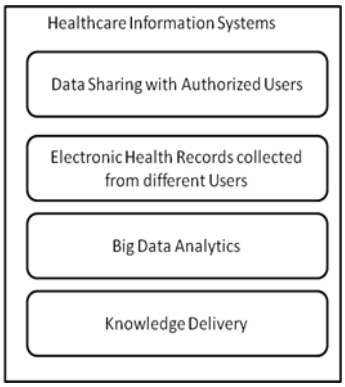

Fig. 1. Knowledge discovery in Healthcare systems

II. RELATED WORK

"No free lunch" problem in machine learning declares that no exclusive algorithm facilitates finest for any application [3]. As effect, as people devise algorithm for elegant healthcare function, the analysis becomes monotonous as one ought to go through several algorithms and select the one with better performance. It is imperative to only try proper algorithms for the application. Meant for example, clustering algorithms seem unsuitable in resolving classification problems. Deep learning has achieved recognition in recent years, attributable to its advantage in resolving difficult problem. On the other hand, it is reminded that deep learning is not always the suitable or necessary for all the applications. It depends on the difficulty of the problem, accessible quantity of data, computing power and preparation time [4].

Sequence sensitive assessments express that house wealth is connected with an increased probability of avoiding transfer processes and imminent tolerant and costly familiar and classified professionals prior to access a open clinic [5]. A variety of analytical paths that be in the patient-centric healthcare scheme from the outlook of different users [6]. Healthcare experts prioritize private and organizational put on as well as justifying individual and executive loss above privacy of EMRs. Healthcare experts sturdily disagreed that they would hold back information to harden gains or lessen losses. [7]. Childhood physical action (PA) is vital for health athwart the lifetime. Time pre-schoolers splurge outside house, which is connected with more PA, is possible subjective by parents' observation of locality familiar social organize related to pre-schoolers PA, clear as the motivation of neighbors to intercede to make certain social order and a secure the public atmosphere for adolescent children's lively play[8].

III. PROPOSED WORK

Acquiring the data is one significant feature of coming up with solutions. Data attainment in the healthcare area is extremely difficult and compounded by vast amounts of data.

PRAS is a scalable and supple framework for predictive analytics and treatment. It provides a data form and typical running time that supports a toolbox of hazard assessment mechanism that can recognize high risk patients. The capability of this system is to allow entity workings to be written once and used frequently within diverse clinical settings, risk assessments calculation, analytics design to deploy specific use cases with no varying the platform and supports for active patient data to present practical value to outsized organizations

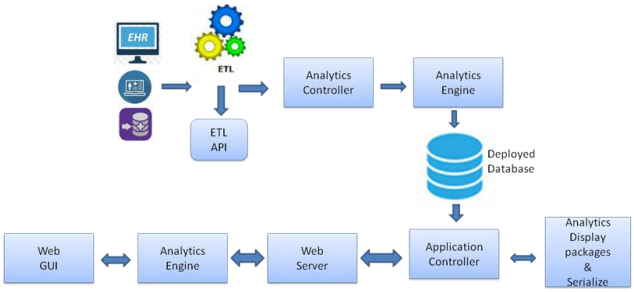

Fig.2. PRAS flexible platform to expose risk assessment.

The heart of the system consists of three databases; the two databases supply as the link connecting the two main computing parts of analytics processing and an analytics subsystem to analyze to uncover risk assessment. Exhaustive patient data includes analyzes, labs, prescriptions, and procedures and ingested by PRAS into deployed database and staged where it is practiced by analytics to predict high risk in patients. After this method completion, the data will be available to patients. The database devise consent to original patient data to be loaded and evaluated incrementally without influencing the value of examine for users who contact patient data. The third database used to store analytics results by creating case file to high risk patients by analytics. The case file links together all risks associated with a given patient and contains metadata. The data file connects collectively all risks connected with a known patient and have metadata and a historical data. The file in addition records the history of a patient's risk assessments to ensure that data is matched as new data arrives and is evaluated by the classification.

A model is defined for evaluation process consequent with the predictive analytic model .The primary step was to discover the most widespread disease hazard factors. These hazard factors must be legitimate and consistent according to subsistence as searchable ordered data in the electronic healthcare records. The scheme would then relate a reason procedure to determine an association of risk factors and start dates to discover patients at augmented risk for acquiring the chronic disease. Medical professionals with the system would get an electronic observant to timely an experimental decision for involvements.

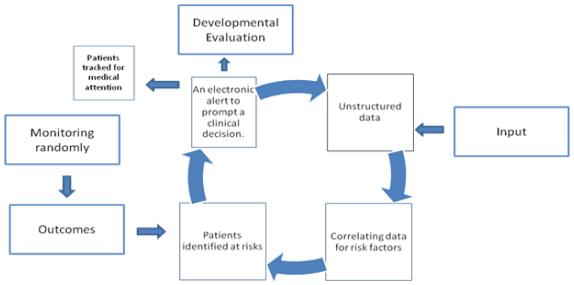

Fig.3. Predictive analytics for the avoidance of chronic disease.

The idea of analytics in a healthcare organization is to add insights to improve a procedure to address a subject such as, improving clinical value and patient security or improving the physical condition of a particular patient. Organizations normally make one of two analytic reporting models. One is a centralized model, where the analytics cluster is its own entity, independent of any particular cluster. The second is a decentralized model where the analysts effort directly for the dissimilar groups or departments. Data visualization gives data in a visual circumstance to facilitate people understand the data's importance. The procedure of visualization makes large volumes of data to get at the real meaning of that data and express key insights.

IV. RESULTS AND OBSERVATIONS

Initially, the framework that the end user visits is as shown below fig 4. The python programming is used to design the framework.

The python packages installed are

```
from tkinter import *
import numpy as np
import pandas as pd
# from gui_stuff import *
```

The dataset for one of the chronic disease is:

l1=['back_pain','constipation','abdominal_pain','diarrhoea',

l1=['back_pain','constipation','abdominal_pain','diarrhoea','mild_fever','yellow_urine',
'yellowing_of_eyes','acute_liver_failure','fluid_overload','swelling_of_stomach',
'swelled_lymph_nodes','malaise','blurred_and_distorted_vision','phlegm','throat_irritation',
'redness_of_eyes','sinus_pressure','runny_nose','congestion','chest_pain','weakness_in_limbs',
'fast_heart_rate','pain_during_bowel_movements','pain_in_anal_region','bloody_stool',
'irritation_in_anus','neck_pain','dizziness','cramps','bruising','obesity','swollen_legs',
'swollen_blood_vessels','puffy_face_and_eyes','enlarged_thyroid','brittle_nails',
'swollen_extremeties','excessive_hunger','extra_marital_contacts','drying_and_tingling_lips',
'slurred_speech','knee_pain','hip_joint_pain','muscle_weakness','stiff_neck','swelling_joints',
'movement_stiffness','spinning_movements','loss_of_balance','unsteadiness',
'weakness_of_one_body_side','loss_of_smell','bladder_discomfort','foul_smell_of urine',
'continuous_feel_of_urine','passage_of_gases','internal_itching','toxic_look_(typhos)',
'depression','irritability','muscle_pain','altered_sensorium','red_spots_over_body','belly_pain',
'abnormal_menstruation','dischromic _patches','watering_from_eyes','increased_appetite','polyuria','family_history','mucoid_sputum',
'rusty_sputum','lack_of_concentration','visual_disturbances','receiving_blood_transfusion',
'receiving_unsterile_injections','coma','stomach_bleeding','distention_of_abdomen',
'history_of_alcohol_consumption','fluid_overload','blood_in_sputum','prominent_veins_on_calf',
'palpitations','painful_walking','pus_filled_pimples','blackheads','scurring','skin_peeling',
'silver_like_dusting','small_dents_in_nails','inflammatory_nails','blister','red_sore_around_nose',
'yellow_crust_ooze']

And the diseases can be:

disease=['Fungal infection','Allergy','GERD','Chronic cholestasis','Drug Reaction',
'Peptic ulcer diseae','AIDS','Diabetes','Gastroenteritis','Bronchial Asthma','Hypertension',
' Migraine','Cervical spondylosis',
'Paralysis (brain hemorrhage)','Jaundice','Malaria','Chicken pox','Dengue','Typhoid','hepatitis A',
'Hepatitis B','Hepatitis C','Hepatitis D','Hepatitis E','Alcoholic hepatitis','Tuberculosis',
'Common Cold','Pneumonia','Dimorphic hemmorhoids(piles)',
'Heartattack','Varicoseveins','Hypothyroidism','Hyperthyroidism','Hypoglycemia','Osteoarthristis',
'Arthritis','(vertigo) Paroymsal Positional Vertigo','Acne','Urinary tract infection','Psoriasis',
'Impetigo']

The diseases can be predicted using the designed framework by choosing the different classification techniques as below.

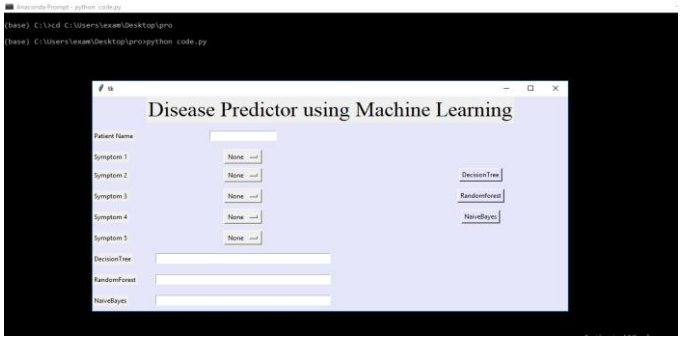

Fig.4. Predictive analytics for the avoidance of chronic disease

It consists of the following entries: Patient Name, Symptom1, Symptom2, Symptom3, Symptom4, and Symptom5. These are the fields that the user should field.

The user selects his/her symptoms from the provided list as shown below:

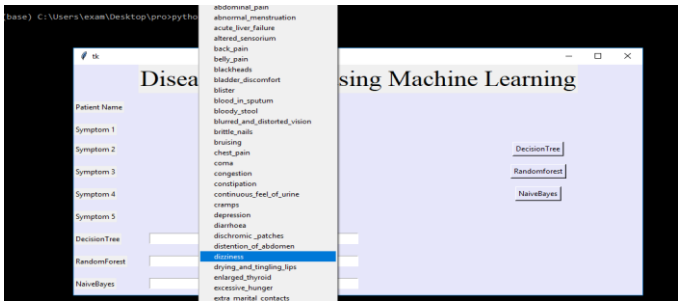

Fig.5. Predictive analytics for the avoidance of chronic disease

Once the user completes entering his/her symptoms (fig:4.1©), he must click on the right side buttons, which represents results of (i) Decision Tree, (II) Random Forest, (iii) NaiveBayes.

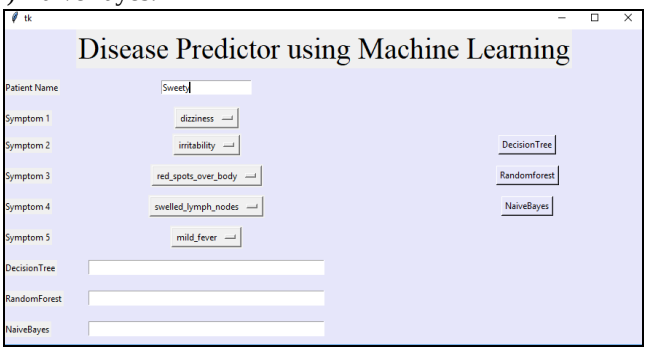

Fig.5. Selection of algorithms for disease prediction

The classification algorithms are selected and the diseases are predicted according to the options selected in the framework.

The accuracy and likelihood are visualized using visualization tools.The pictorial representation of the results are given below

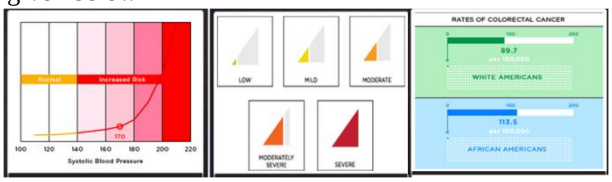

Fig.6. Healthcare data visualization for classification of risks and differences in likelihood

CONNCLUSION

More than people, machines can infer data better than human mind. Computers and robots will most likely never replace doctors and nurses, but changing the way health institutions are providing healthcare. With the capability to predict what and how insistent a person's condition might be, and to recognize which treatments may or may not work well, machine learning has become an essential and ultimately crucial part of medical care. This work was projected to here the idea of predictive analytics to be worn by clinicians for prevention of chronic disease. The conclusion for this work was planned to predict a typical quantity of time that a patient will acquire a diagnosis. Predictive analytics are before now invaluable tools in the novel healthcare delivery models. As the data improves and novel algorithms are planned, their value will amplify further, but only if they're connected to workflow computerization solutions that make their imminent actionable.

Many healthcare organizations effort with the requirement to admittance and facility to exploit data composed during non-integrated conventional systems for data mining ,knowledge discovery and assessment making. For thriving healthcare organizations, it is significant to allow personnel and management for deliberate assessment making through data warehousing based on significant opinion. In this paper, the common architectural approaches for predictive analytics solution have been summarized and its main development components have been introduced in regulate to give a elevated level analytics with reporting regarding healthcare informatics.

REFERENCES

1. Álvaro BrandónHernández,María S.Perez, SmratiGupta, VictorMuntés-Mulero. Using machine learning to optimize parallelism in big data applications. Future Generation Computer Systems,Volume 86, September 2018, Pages 1076-1092
2. Gandomi A., Haider M.Beyond the hype: Big data concepts, methods, and analytics.International Journal of Information Management, 35 (2) (2015), pp. 137-144
3. Wolpert, D.H. The Supervised Learning No-Free Lunch; Springer: London, UK, 2002.
4. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444
5. Marco J. Haenssgen, Proochista Ariana,Healthcare access: A sequence-sensitive approach,SSM - Population Health,Volume 3,2017,Pages 37-47,ISSN 2352-8273
6. Venketesh Palanisamy, Ramkumar Thirunavukarasu,Implications of big data analytics in developing healthcare frameworks – A review,Journal of King Saud University - Computer and Information Sciences,2017,ISSN 1319-1578.
7. Michael Lapke, Christopher Garcia, David Henderson,The disconnect between healthcare provider tasks and privacy requirements,Health Policy and Technology,Volume 6, Issue 1,2017,Pages 12-19,ISSN 2211-8837.
8. Ester Cerin, Yi Nam Suen, Anthony Barnett, Wendy Y.J. Huang, Robin R. Mellecker,Validity of a scale of neighbourhood informal social control relevant to pre-schoolers' physical activity: A cross-sectional study,SSM - Population Health,Volume 3,2017,Pages 57-65,ISSN 2352-8273.
9. Min Cheni, Yixue Haol, Kai Hwang, Lu Wangi and Lin Wang, "Disease Prediction by Machine Learning Over Big Data From Healthcare Communities,"IEEE2017
10. DP. Siriyasatien, S. Cahdsuthi, K. Jampachaisri, and K. Kesorn,"Dengue Epidemics Prediction: A survey of the State-of-the-Art based on Data Science Processes." IEEE 2018.