Disease And Drug Prediction Using Different Machine Learning Algorithms

[1] Archana Ponnam, [2] MD Sirajuddin

[12]Vaageswari College of Engineering, Karimnagar, Telangana, India

[1] archanaponnam98@gmail.com [2] siraj569@hotmail.com

Abstract_ In order to predict diseases from patients' or users' symptom reports, we use a system called Disease Prediction Using Machine Learning. A user inputs their symptoms into the system, and the system returns a probability of disease. The disease prediction is performed using the supervised machine learning algorithm, Naive Bayes classifier. The Nave Bayes algorithm determines the disease's likelihood. Accurate analysis of medical data aids in early disease detection and patient care, which is especially important as the volume of available biomedical and healthcare data continues to grow. Diseases such as diabetes, malaria, jaundice, dengue, and tuberculosis can be predicted with the help of linear regression and decision trees.

Indexed Terms- Diseases, Drugs, Machine Learning, Prediction, And Random Forest.

1.INTRDOUCTION

The croakers' still require technology in every imaginable way since sophisticated computing appeared, such as surgical representation and x-ray photography, but the technology has kept in the background. The system still needs the croaker's expertise and experience for medical records, weather reports, atmospheric conditions, blood pressure, and other reasons. No model has properly anatomized the many factors needed to understand the entire working process. Medical decision support tools are key to overcoming this gap. This approach helps croakers choose wisely. "Medical decision support system" involves both diagnosing a patient's disease and the resulting expert opinion. Subjective judgement specifies a specific category or range of anomalies. Case or croaker is a

narrative about power ethics and financial incentives. It may be a tale or gimmick. It could be a computer-like type of communication that initiates compensation, ritual, announcement, news, or advise. A doctor's suggestion will examine the patient's health and normalcy. Automated decision support systems utilise rules to answer business questions.

In unusual or essential situations, making a medical decision might be complex. Misdiagnosis owing to weariness or lack of understanding about possible causes is inevitable. The usual method may consider prior conditions, family history, case history, and retired factors. Differentiating opinion types reveal the seeker druthers' globe of several probable druthers. In this system, seeker situations must be reduced to a low probability, hence elimination or information transfer is needed. It's four-fold 1) The croaker collects symptom data. Second, the croaker should list all symptom triggers. Third, the croaker should rank the causes from most to least hazardous. Fourth, the croaker should treat the most significant causes first. " Discard," to eliminate a hypothesis by testing or other science. If no other opinions are the same, we can test that view and exclude it if it's wrong. This uses the croaker's background and experience. This system's implementation is easy.

The K-Means method narrows probable outcomes to the most likely. This method is great for organising several variables. K-Mean can address the clustering problem. Locating k cluster centres is crucial. Case diagnostics will be used as clustering attributes. This algorithm reduces duplicates and gives more precise results for every viewpoint because cluster boundaries are clearly defined and don't overlap. The system uses Service Oriented Architecture (SOA), making it accessible to anyone with an internet connection. The LAMSTAR Network can be used to calculate weight, improving the algorithm's finesse, test speed, and result quality.

2.LITERATURE SURVEY

More than one-third of the global population has been diagnosed with diabetes, making it one of the major causes of death. It also plays a role in the development of many other diseases and ailments, such as heart failure, blindness, urinary tract infections, and more. After seeing a specialist, patients in this situation must travel to a diagnostic centre to collect their test results. The reason being the ongoing time and monetary investment required. However, thanks to the advancement of Machine Learning methods, we are able to actively seek a solution to the issue at hand; at present, we possess a highly developed system based on information processing that can determine with high accuracy whether or not a given patient suffers from polygenic disease. The ability to foresee the spread of an illness is useful because it enables medical professionals to prepare for a potential surge in patient demand. information extracted from a sea of diabetes records. A more accurate way of predicting a patient's risk of acquiring diabetes could be developed with the help of this analysis. Models are developed using methods such as decision trees, ANNs, I Bayes, and support vector machines. There is an 85% success rate with the Decision Tree model, 77% with the I Bayes model, and 77.3% with the Support Vector Machine model. A great degree of accuracy was achieved using these procedures, as shown by the findings.

2, With the advancement of information technology, assistive technologies have made significant progress in a variety of areas. However, these developments have resulted in a far higher and more complex data volume to manage and process than ever before. In addition, the information created from a wide range of devices in a relatively short time period might, to an exaggerated degree, be viewed as a massive data problem, due to the properties of those data, specifically, that they endure in diverse formats and formed swiftly. In order to improve the service and environment for patients, this study offers a cyber-

physical system for case centric aid operations and service (Health-CPS) that is built on pall and large information analytics at the DT tree's nodes. There are three tiers to this approach: a data-centric service layer, a knowledge management layer that facilitates distributed storage and parallel computation, and a knowledge gathering layer that uses a uniform common standard. This study shows that the extensive use of cloud and big data technology in the humanitarian sector has allowed for the creation of a plethora of useful applications and services. Without any context or background information, SVM produces very good results. Even semi- and unstructured data, such as text, pictures, and trees, can be successfully processed using the SVM algorithm. The SVM algorithm has the limitation that optimal classification results depend on the right choice of numerous important parameters. Decision trees provide easy-to-understand rules that anyone can follow. Instability is a hallmark of decision trees; even a small change to the ideal decision tree's data structure can have far-reaching consequences. When they're wrong, it's usually by a small margin. Certainly, Mr. Bayes It is trustworthy because it avoids the trouble of estimating probabilities for the cases of missing data. depends on how one processes data coming in. Growing the size of the dataset used for training can induce bias. ANN is easy to apply and produces reliable projections. Not ideal for using complicated models on massive data sets. call for a great deal of thought and analysis. Having abnormally high levels of sugar in the blood is the root cause of diabetes. It is currently one of the worst diseases in the world. The entire over, people are suffering from this devastating disease. Heart disease, paralysis, kidney failure, eye disease, and blindness are all common outcomes of uncontrolled diabetes. Multiple automated systems for diagnosing diabetes have been planned and detailed. Diabetic patient diagnosis has historically necessitated more time and money. However, with the development of machine learning,

we now have the chance to find a solution to this pressing issue. As a result, we've developed a system to determine whether or not a patient has diabetes. This study's primary focus is on creating a web-based app that makes use of the superior prediction accuracy of a powerful machine learning algorithm. Because of its potential for diagnostic-based prediction of diabetes onset, the Pima Indian dataset was used as a benchmark in this study. Because Artificial Neural Networks (ANNs) have proven to be so effective at increasing precision, we have developed a novel, interactive web application for diabetes prediction..

3.1DATASET

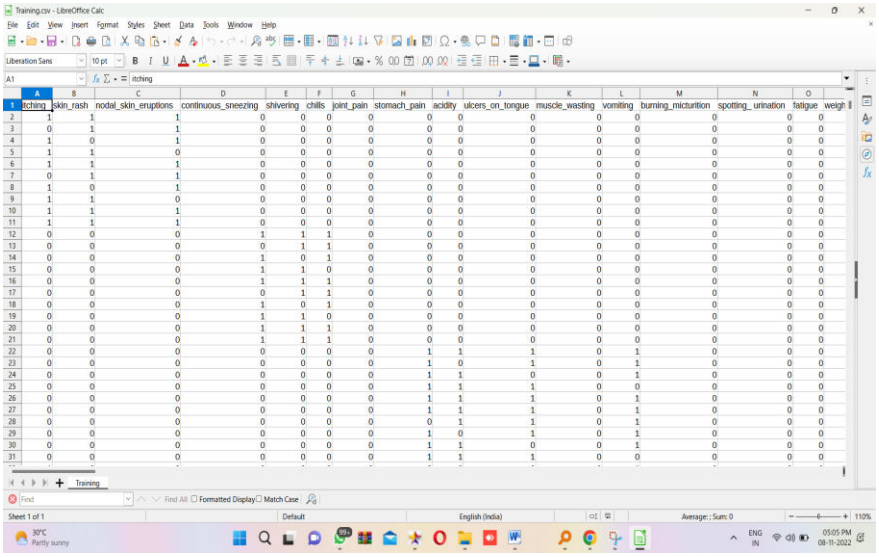

Fig 1:Dataset

4.RESULTS AND DISCUSSION

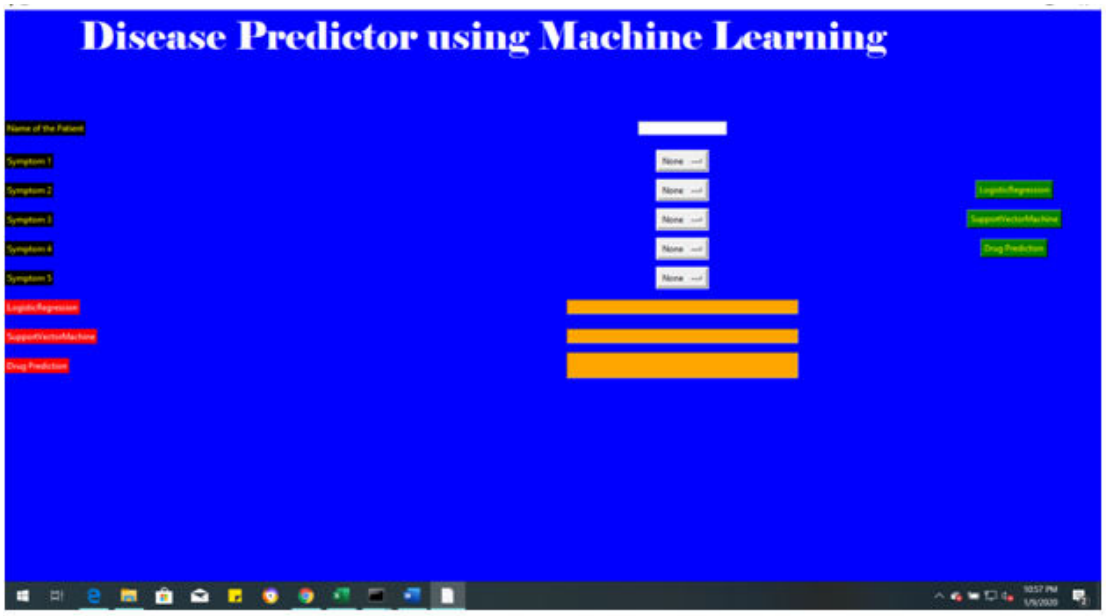

Fig 2:Enter the Name of patient and enter the symptoms of the patient to prediction the disease. And then click on algorithm from which you want to predict.

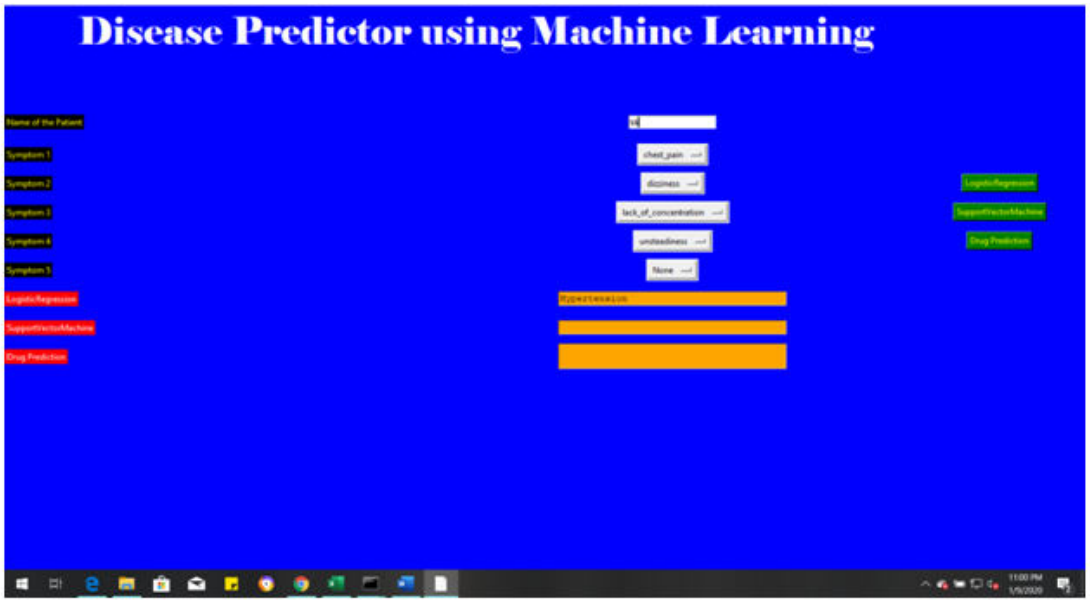

Fig 3:From the above figure for given symptoms it predicted Hypertensio by using Logistic regression

Now test for SVM also.

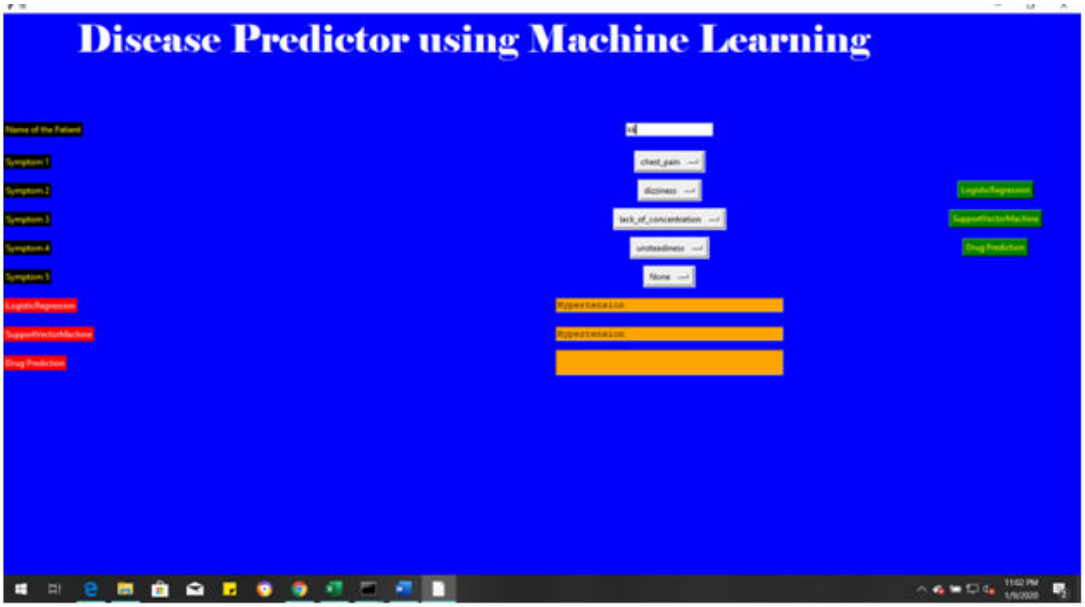

Fig 4:For SVM also for the given symptoms it predicted Hypertension. Now predict the Drug for the disease.

5.CONCLUSION

This paper discusses an algorithm that predicts a disease from its symptoms. The user chooses any five of the available symptoms from a drop-down menu, and an algorithm uses these to predict the condition. This method can also recommend medications that are frequently prescribed for a specific condition. The major objective is to identify the disease early on and anticipate it. Doctors can also utilise this approach to forecast diseases without being confused. This system can aid medical professionals.

BIBLIOGRAPHY

1. Brown, A.S., Patel, C.J.: A standard database for drug repositioning. Scientic

Data 4, 170029 (2017)

2. Callahan, A., Cruz-Toledo, J., Ansell, P., Dumontier, M.: Bio2rdf release 2: improved coverage, interoperability and provenance of life science linked data. In:

Extended Semantic Web Conference.pp. 200{212. Springer (2013)

3. Campillos, M., Kuhn, M., Gavin, A.C., Jensen, L.J., Bork, P.: Drug target identi-cation using side-e ect similarity. Science 321(5886), 263{266 (2008)

4. Chiang, A.P., Butte, A.J.: Systematic evaluation of drug{disease relationships to

identify leads for novel drug uses. Clinical Pharmacology & Therapeutics 86(5), 507{510 (2009)

5. Gottlieb, A., Stein, G.Y., Ruppin, E., Sharan, R.: Predict: a method for inferring

novel drug indications with application to personalized medicine. Molecular

systems biology 7(1), 496 (2011)

6. Guney, E.: Reproducible drug repurposing: When similarity does not suce. In:

PACIFIC SYMPOSIUM ON BIOCOMPUTING 2017. pp. 132{143 (2017)

7. Hay, P.J., Claudino, A.M.: Bulimia nervosa: online interventions. BMJ clinical

evidence 2015 (2015)

8. Hu, G., Agarwal, P.: Human disease-drug network based on genomic expression

proles. PloS one 4(8), e6536 (2009)

9. Kuhn, M., Letunic, I., Jensen, L.J., Bork, P.: The sider database of drugs and side

E ects. Nucleic acids research 44(D1), D1075{D1079 (2015)

10. Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C.,Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., et al.: The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. Science 313(5795), 1929{1935 (2006) 1