A Novel EHR Based Extraction and Association Techniques to Identify Prevalent Medical Conditions.

Manjesh B N[1]
Research scholar
Jain university

Dr. Raja Praveen N[2]
Associate professor
Jain university

Abstract :

Due to the state of the environment and human lifestyles today, people suffer from a variety of ailments. To stop such diseases from becoming extremely severe, it is crucial to recognise and anticipate them in their earliest stages. For the most part, diagnosing diseases accurately by hand is challenging for doctors. Identifying and predicting patients with more prevalent chronic illnesses is the aim of this article. This might be done by employing a state-of-the-art Ml_methods to make sure that this categorisation accurately identifies people with chronic conditions. Disease forecasting is a difficult endeavour as well. In order to forecast diseases, data mining is essential. Predicting mortality from chronic diseases sooner allows for disease prevention, which is the only approach to combat this problem. A patient has a requirement for such a model, and machine learning is highly advised in this situation. For a doctor, however, it is impossible to make a precise forecast based just on symptoms. The most difficult duty is making a diagnosis of an illness. Data mining is crucial for diagnosing the sickness and addressing this issue.

Index Terms: Chronic Diseases, Machine Learning, Diseases Prediction and accuracy, Deep learning,CNN

Introduction:

The study of computer systems that learn from data and experience is known as machine learning. ML can be classified as either unsupervised (which deals with clustering of various groups for a specific function) or supervised (which uses output variables that are predicted from input factors). With the use of machine learning (ML), complicated models can be determined and medical knowledge extracted, presenting new concepts to experts. In clinical practice, machine learning (ML) predictive models can emphasize and enhance rules in the selection of specific patient treatment options. These are also capable of independently diagnosing various disorders in accordance with clinical guidelines. Long-lasting and challenging to treat, some illnesses can only be prevented if they can be identified sooner.

A lot of data generated by the healthcare sector has been used to enhance machine learning in making predictions and choices. We streamline machine learning methods for accurate chronic disease outbreak prediction. Only a small portion of disease prediction using ML approaches has been explored in various research. By using machine learning techniques, we offer a unique method that tries to uncover meaningful features with the goal of increasing the accuracy of disease prediction. To increase the learning process' accuracy, numerous such algorithms are used. Then, it can be tested using the datasets that are currently accessible. On an average more than 70% of people expenditure is on chronic diseases treatment.

2.REVIEW CRITERIA
[1]. Chronic Diseases Prediction Using Machine Learning (Shweta Agarwal1 , Dr.Chander Prabha2 , Dr.Meenu Gupta):

Chronic diseases has been a major setback for welness of human life. According to WHO majority of the percent of deaths in the world are direct or indirect effects of chronic diseases on human body . New developments in ML are widening the window for medical decision support systems.The advance of Machine Learning has opened the pathway for predicting these diseases.
The accuracy of various machine learning (ML) strategies employed by other researchers to diagnose chronic diseases is discussed in this work. Such a disease prediction algorithm makes it possible to diagnose illness sufferers based just on their symptoms. In this paper,they emphasise the usage of various multiple algorithms than single algorithm for better accuracy.

[2]. Estimation of Chronic Illness Severity Based on Machine Learning Methods.(Yue Chang1 and Xudong Chen):

In this paper the author uses logistic regression model for prediction purpose. The logistic function should be chosen first. Getting the logistic regression's prediction function is the second step. Finding the best answer for the classification task is the third phase.
To forecast the severity of a victim's chronic disease, this study presented a model of sickness severity-levels.
In the beginning, we compared seven distinct classification models and ten relative characteristics based on the initial dataset.
The XGBoost and LightGBM models offer the best performance, according to the comparison. After that, the feature effects on chronic diseases were examined via feature selection, which involved choosing the most pertinent features.

[3]. Comparative analysis of machine learning techniques based on chronic kidney disease dataset(Vineeta Gulati1 , Neeraj Raheja):

This study compares and contrasts the efficacy of existing methods for the diagnosis and prognosis of chronic renal disease using data from numerous studies. Additionally, by utilizing a variety of machine learning approaches. It is found that employing real-world datasets and accounting for all relevant features results in an accurate machine learning prediction of the presence of chronic renal disease.

This research offers a analysis of ml methods based on a dataset for chronic kidney disease with various features taken from either an existing database or a real-world database.

[4]: Machine-Learning-Based Disease Diagnosis:
(Md Manjurul Ahsan, Shahana Akter Luna, Zahed Siddique):

The sickness or disorders that account for a person's symptoms and indicators are identified by a medical diagnosis. Humans are fallible, thu that a victim could experience overdiagnosis more frequently. misdiagnosis will result in issues like unneeded treatment, harming people's health and the economy.

In this publication, There is also discussion of various further ML-based disease diagnosis methods. Our main conclusion shows that ML is the most widely used method among researchers due to its outstanding effectiveness in building a robust model.

[5]: Prediction of Chronic Kidney Disease Using Machine Learning Algorithm(Siddheshwar Tekale1 , Pranjal Shingavi2 , Sukanya Wandhekar3 , Ankit Chatorikar4):

In this paper, we looked at various Ml algorithms. In order to estimate the accuracy of several Ml techniques like Decision Tree and SVM, we analysed 14 distinct CKD patient-related characteristics. The accuracy of the decision tree and SVM algorithms, respectively, is 91.75% and 96.75%, respectively, according to the results analysis. The prediction process takes less time with this technology, which is an advantage.

It will enable medical professionals to diagnose more people more quickly and to begin treating patients with CKD earlier. The researchers' aim in this case is to create a model that can predict if CKD is present or not, and if so, how severe it is, with the greatest degree of accuracy
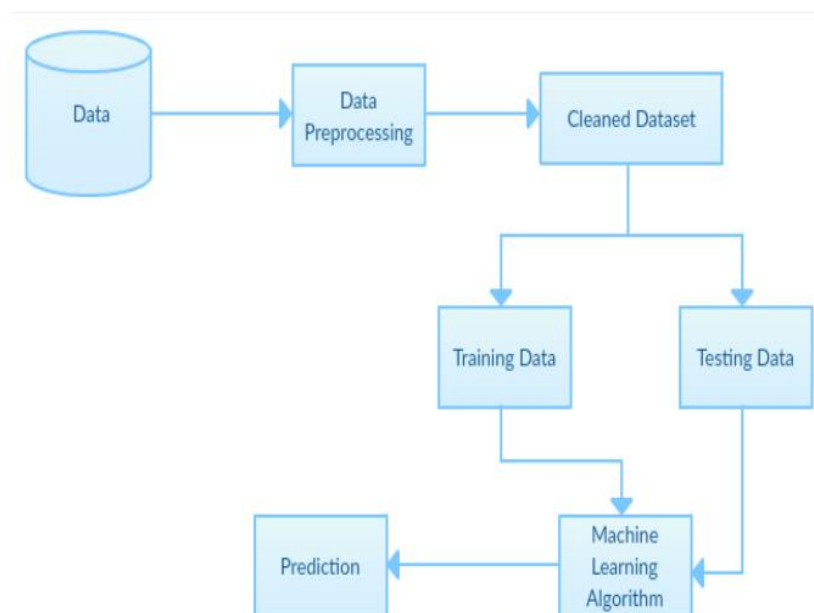
Fig1: Methodology used for prediction

[6]: Comparing different supervised machine learning algorithms for disease prediction
(Shahadat Uddin, Arif Khan, Md Ekramul Hossain, Mohammad Ali Moni):

In this paper they have considered publictions for research that used more than one supervised machine learning method in order to avoid the possibility of selection bias. In distinct research contexts, the same supervised learning algorithm can produce various outcomes.
In order to forecast diseases, this research aimed to compare the abilities of various supervised machine learning algorithms. Despite changes in effectiveness and frequency, the results highlight the promise of these families of algorithms for disease prediction.

[7]: A novel early diagnostic framework for chronic diseases with class imbalance
(XiaohanYuan, Shuyu Chen*, Chuan Sun & LuYuwen):

In this paper, we looked into a general learning algorithm built on PNNs for the early detection of chronic diseases. To contrast the NLPNN algorithm with five cutting-edge baseline algorithms, According to the findings of the experiments, NLPNN has the highest accuracy on the nine datasets for chronic diseases. The generalisation recall, specificity value of this model have all reached 1.000 for the early diagnosis of chronic renal illness .

Acoording to authors –
By improving the recall value of the model, researchers hope to increase the likelihood that sick cases will be correctly diagnosed.
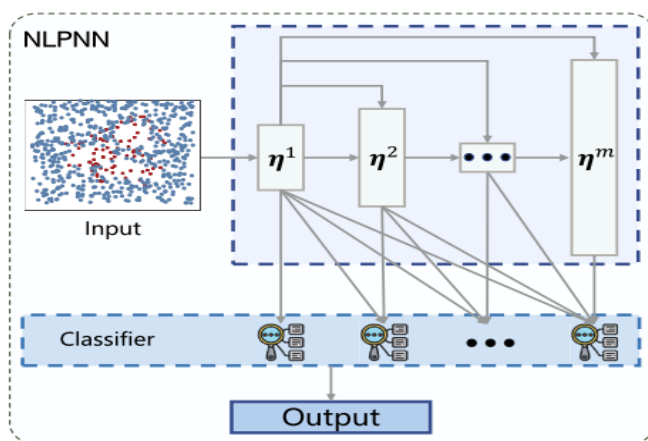
ig2:Flowchart of proposed algorithm

[8]: Coronavirus disease (COVID-19) cases analysis using machine-learning applications
(Ameer Sardar Kwekha-Rashid1 · Heamn N. Abduljabbar2,3 · Bilal Alhayani):

In this paper,the authors have analyzed the recent pandamic disease covid-19 and its effects in a holistic manner.
The publications employed in this study's examination of machine learning applications in COVID-19 disease for diverse objectives were the main focus. Our final words are that ML applications demonstrated promising outcomes with high sensitivity, accuracy, and specificity utilising various models and algorithms. The research findings generally showed that supervised learning is more accurate in identifying COVID19 cases than unsupervised learning, which was only slightly more accurate (7.16%).
The most popular machine learning approach used for manufacturing lines is supervised learning.
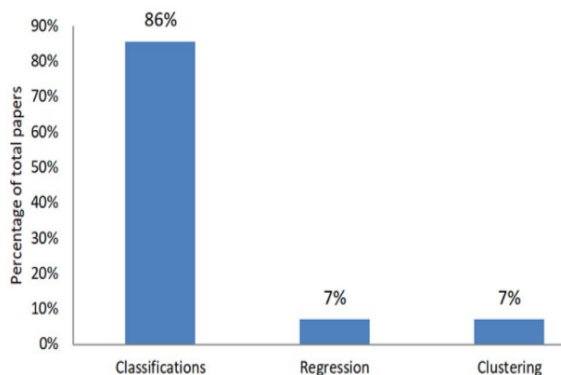

Fig3: Depiction of ML tasks distributed

[9]: Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosi
(Gopi Battineni, Getu Gamo Sagaro, Nalini Chinatalapudi, Nalini Chinatalapudi):

A significant amount of the cost of healthcare worldwide is caused by chronic illnesses (CDs). These disorders require lifetime treatment for patients.
The current study assessed the research pertaining to the diagnosis of chronic illnesses. However, recent researchers suggest that few ML models are hindered by growing included datasets with harmful material, which can have serious repercussions.

On the other hand, limits in diagnosis, according to authors, may cause life-threatening attacks and can even be a factor in fatalities. Contrarily, a false diagnosis causes doubt about the use of machine learning, which may cause policymakers to steer clear of the application of predictive models. Reviews of prediction models can therefore offer support for the great approaches that are suggested for CDs diagnosis.

[10]: Smart healthcare disease diagnosis and patient management: Innovation, improvement and skill development
(Arkadip Ray, Avijit Kumar Chaudhuri0:

The current studies of novel predictive models built on ML techniques show potential for scientific study. Medical personnel need precise forecasts of the outcomes of the numerous ailments that patients experience. Additionally, timing is an important factor that influences clinical decisions for accurate predictions. Regarding technique, methods, and performance, the writers have examined a large number of works in this field.
According to a review of similar studies, even when the same dataset is used, the forecast produced by current forecasting models varies as per this paper. As new methods need to be enhanced, predictive models are also crucial.
This paper suggests that new ML alogorithms have different degrees of accuracy for the same dataset and a new ML
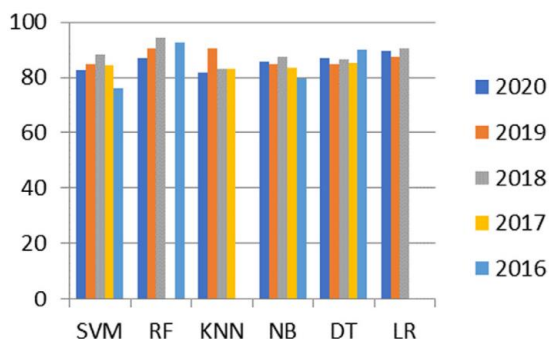algorithm can be worked upon for better results.

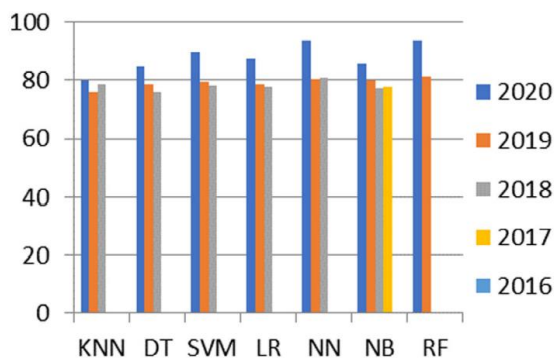**Fig 4: Accuracy of heart disease prediction   of ML models**

Fig 5: Accuracy of diabetes prediction of various Ml models

3. **GAP IDENTIFICATION**

TABLE I. Comparative Study

Sl No.	Author(s)	Year	Approach	Gap identification
1	Shweta Agarwal1 Dr.Chander Prabha2 , Dr.Meenu Gupta3	2021	Have analyzed various Ml models to predict chronic diseases and accuracy of the predict.	Authentic medical reports like can be used for better aresults of the model .The classifiaction of structure and unstructured data can be improved.
2	Yue Chang1 Xudong Chen 2	2021	Different ML models are used to predict severity-level.	Because there is a dearth of relevant medical background information that makes evaluation of disease worthy of medical relevance, there should be more scope for severity .
3	Vineeta Gulati1 Neeraj Raheja2	2021	Kideny disease prediction based on GFR value	Real life dataset with all possible attributes would yield more results.
4	Md Manjurul Ahsan1 Shahana Akter Luna2 Zahed Siddique3	2022	ML based disease dignosis approach	The review focuses on disease classifications, datasets, ML and DL algorithms, and evaluation measures. This work does not analyse the recommended ML procedure to the extent that reference material does, despite the fact that it is exhaustively examined in that literature.Bibilometric analysis and model building with Deep learning algorithms
5	Siddheshwar Tekale1 Pranjal Shingavi2 Sukanya Wandhekar3 Ankit Chatorikar4	2018	Usage of ML algorithms like decision tree,Svm to predict CKD patients severity	Strength of data is not very high and there are missing values which is an hindrance for accuracy
6	Shahadat Uddin1 Arif Khan2 Md Ekramul Hossain3 Mohammad Ali Moni4	2019	Research on optimal ML algorithm for forseeing ailments	There should be a clear distinguis between clinical data and research scope which affects the model accuracy.
7	XiaohanYuan1 Shuyu Chen2 Chuan Sun3 & LuYuwen4	2022	Predict chronic diseases at an aerly stage with class imbalance	There is still scope for clarity regarding choosing of hidden features of PNN
8	Ameer Sardar Kwekha-Rashid1 Heamn N. Abduljabbar2 Bilal Alhayani3	2021	This study proposes the use of machine learning (ML) predictive models	There is a scope to use real time hospiotal dataset for better clarity on accuracy of ML models
9	Gopi Battineni1 Getu Gamo Sagaro2 Nalini Chinatalapudi3 Francesco Amenta4	2020	In the diagnosis of chronic diseases, this study discusses the use of machine learning (ML) predictive models	The expansion of electronic data access and the establishment of new medical discoveries create new opportunities for results.These models are created to highlight the obligation of high-quality victim care and lower the cost expenses.
10	Arkadip Ray1 Avijit Kumar Chaudhuri2	2021	This study shows that even using the same dataset, newly developed ML techniques have varying degrees of accuracy.	Balancing the imbalanced medical set could be improved further for better accuracy of ML models for the same dataset.

4. CONCLUSION:

In this present-day world with the advancement of cut-edge technologies like Ml_methods, there is a huge scope for the prediction of chronic disease which is widespread according to many surveys. However, there needs to be selective usage of models based on its accuracy levels. There should be proper classification of datasets. This accurate prediction helps misdiagnosis of doctors in medicine field.

REFERENCES

[1] Ahsan, M.M.; Luna, S.A.; Siddique, Z. Machine-LearningBased Disease Diagnosis: A Comprehensive Review. Healthcare 2022, 10, 541..

[2] Yue Chang1 , Xudong Chen 2,Estimation of Chronic Illness Severity Based on Machine Learning Method Hindawi Wireless Communications and Mobile Computing Volume 2021, Article ID 1999284
.

[3Vineeta Gulati1,Neeraj Raheja2 Comparative analysis of machine learning techniques based on chronic kidney disease dataset ICETCE 2021 1131 (2021) 012010
.

[4] Ahsan, M.M.; Luna, S.A.; Siddique, Z. Machine-LearningBased Disease Diagnosis: A Comprehensive Review. Healthcare 2022, 10, 541.
.

[5] D. Nagavci, M. Hamiti, and B. Selimi, "Review of prediction of disease trends using big data analytics," Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 8, pp. 46–50, 2018, doi: 10.14569/ijacsa.2018.090807.

[6] M. E. Farooqui and D. J. Ahmad, "a Detailed Review on Disease Prediction Models That Uses Machine Learning," Int. J. Innov. Res. Comput. Sci. Technol., vol. 8, no. 4, pp. 326–330, 2020, doi: 10.21276/ijircst.2020.8.4.14.

[7] M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C. H. Youn, "Wearable 2.0: Enabling Human-Cloud Integration in Next Generation Healthcare Systems," IEEE Commun. Mag., vol. 55, no. 1, pp. 54–61, 2017, doi: 10.1109/MCOM.2017.1600410CM.

[8] P. Shimpi, "A Machine Learning Approach for the," vol. 6, no. Iccmc, pp. 603–607, 2017. [9] G. V. Gayathri and S. C. Satapathy, "A Survey on Techniques for Prediction of Asthma," Smart Innov. Syst. Technol., vol. 159, pp. 751–758, 2020, doi: 10.1007/978- 981- 13-9282-5_72.

[9] Battineni G, Sagaro GG, Chintalapudi N, Amenta F. The Benefits of Telemedicine in Personalized Prevention of Cardiovascular Diseases (CVD): A Systematic Review. J Pers Med. 2021 Jul 14;11(7):658. doi: 10.3390/jpm11070658. PMID: 34357125; PMCID: PMC8304370.

[10] Chaudhuri, Avijit & Ray, Arkadip & Das, Anirban & Chakrabarti, Prasun & Banerjee, Dilip. (2020). Early Detection of Cardiovascular Disease in Patients with Chronic Kidney Disease using Data Mining Techniques. ASIAN JOURNAL OF CONVERGENCE IN TECHNOLOGY. 6. 10.33130/AJCT.2020v06i03.011.