**Disease Prediction using ML**

**ABSTRACT:**

With big data growth in biomedical and healthcare communities, accurate analysis of medical data benefits early disease detection, patient care, and community services. However, the analysis accuracy is reduced when the quality of medical data is incomplete. Moreover, different regions exhibit unique characteristics of certain regional diseases, which may weaken the prediction of disease outbreaks. In this paper, we streamline machine learning algorithms for effective prediction of chronic disease outbreak in disease-frequent communities. We experiment the modified prediction models over real-life hospital data collected from different parts of county in 2013-2015. To overcome the difficulty of incomplete data, we use a latent factor model to reconstruct the missing data. We experiment on a regional chronic disease of cerebral infarction. In this Machine Learning project is used to predict the disease based on the symptoms given by the user.It predicts using three different machine learning algorithms. So,the output is accurate .It uses tkinter for GUI. In this we will analyze data using three algorithms

**INTRODUCTION:**

At present, when one suffers from particular disease, then the person has to visit to doctor which is time consuming and costly too. Also if the user is out of reach of doctor and hospitals it may be difficult for the user as the disease can not be identified. So, if the above process can be completed using a automated program which can save time as well as money, it could be easier to the patient which can make the process easier. There are other Heart related Disease Prediction System using data mining techniques that analyzes the risk level of the patient. Disease Predictor is a web based application that predicts the disease of the user with respect to the symptoms given by the user. Disease Prediction system has data sets collected from different health related sites. With the help of Disease Predictor the user will be able to know the probability of the disease with the given symptoms.

As the use of internet is growing every day, people are always curious to know different new things. People always try to refer to the internet if any problem arises. People have access to internet than hospitals and doctors. People do not have immediate option when they suffer with particular disease. So, this system can be helpful to the people as they have access to internet 24 hours.

**Existing system:**

Different rules are made to get the best outcome. In this research age , sex, smoking, overweight, alcohol intake, blood sugar, hear rate, blood pressure are the parameters used for making the decisions. Risk level for different parameters are stored with their id’s ranging. ID lesser than of 1 of weight contains the normal level of prediction and higher ID other than 1 comprise the higher risk levels .K-means clustering technique is used to study the pattern in the dataset. The algorithm clusters informations into k groups. Each point in the dataset is assigned to the closed cluster. Each cluster center is recomputed as the average of the points in that cluster.

**Disadvantages:**

In most of the existing system specific disease related prediction is done but no based on symptoms.

There are many tools related to disease prediction. But particularly heart related diseases have been analyzed and risk level is generated. But generally there are no such tools that are used for prediction of general diseases. So Disease Predictor helps for the prediction of the general diseases.

**Proposed System:**

Disease Prediction has been already implemented using different techniques like Neural Network, decision tree and Naïve Byes algorithm. Particularly heart related disease is mostly analyzed. From the analysis it was found that Naïve Bayes is more accurate than other techniques. So, Disease Predictor also uses Naïve Bayes for the prediction of different diseases.

**Advantages:**

* Using this method we can analyze prediction based on three algorithms and Decision tree Randomforest and NaiveBayes.
* Disease Prediction is done by implementing the Naïve Bayes Classifier. Naïve Bayes Classifier calculates the probability of the disease. Therefore, average prediction accuracy probability 90% is obtained.

**SYSTEM REQUIREMENTS:**

**HARDWARE REQUIREMENTS:**

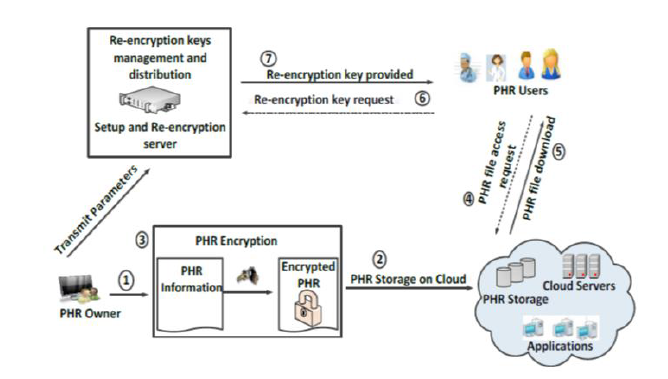
* System : Pentium Dual Core.
* Hard Disk : 120 GB.
* Monitor : 15’’ LED
* Input Devices : Keyboard, Mouse
* Ram : 1GB.

**SOFTWARE REQUIREMENTS:**

* Operating system : Windows 7.
* Coding Language : python
* Tool : anaconda, flask framework
* Database : MYSQL

**SYSTEM DESIGN**

**ARCHITECTURE :**

****

**REFENRENCES:**

* K. Gai, M. Qiu, Z. Xiong, and M. Liu, “Privacy-preserving multi-channel communication in Edge-of-Things,” Future Generation Computer Systems, 85, 2018, pp. 190-200.
* K. Gai, M. Qiu, and X. Sun, “A survey on FinTech,” *Journal of Network and Computer Applications*, 2017, pp. 1-12.
* Abbas, K. Bilal, L. Zhang, and S. U. Khan, “A cloud based health insurance plan recommendation system: A user cen-tered approach, “*Future Generation Computer Systems*, vols. 43-44, pp. 99-109, 2015.
* N. Khan, ML M. Kiah, S. A. Madani, M. Ali, and S. Sham-shirband, “Incremental proxy re-encryption scheme for mo-bile cloud computing environment,”*The Journal of Supercom-puting*,Vol. 68, No. 2, 2014, pp. 624-651.
* R. Wu, G.-J. Ahn, and H. Hu, “Secure sharing of electronic health records in clouds,” In *8th IEEE International Conference on Collaborative Computing: Networking, Applications and Work-*
* A. Abbas andS. U. Khan, “A Review on the State-of-the-Art Privacy Preserving Approaches in E-Health Clouds,” *IEEE Journal of Biomedical and Health Informatics,*vol. 18, no. 4, pp. 1431-1441, 2014.
* [7] M. H. Au, T. H. Yuen, J. K. Liu, W. Susilo, X. Huang, Y. Xiang, and Z. L. Jiang, “A general framework for secure sharing of personal health records in cloud system,” Journal of Comput-er and System Sciences, vol. 90, pp, 46-62, 2017.
* [8] J. Li, “Electronic personal health records and the question of privacy,” *Computers*, 2013, DOI: 10.1109/MC.2013.225.
* [9] D. C. Kaelber, A. K. Jha, D. Johnston, B. Middleton, and D. W. Bates, “A research agenda for personal health records (PHRs),” *Journal of the American Medical Informatics Association*, vol.15, no. 6, 2008, pp. 729-736.
* [10] S. Yu, C. Wang, K. Ren, and W. Lou, “Achieving secure, scala-ble and fine-grained data access control in cloud computing,” in *Proceedings of the IEEE INFOCOM*, March 2010, pp. 1-9.
* [11] S. Kamara and K. Lauter, “Cryptographic cloud storage,” *Financial Cryptography and Data Security*, vol. 6054, pp. 136–149, 2010.
* T. S. Chen, C. H. Liu, T. L. Chen, C. S. Chen, J. G. Bau, and T.C. Lin, “Secure Dynamic access control scheme of PHR in cloud computing,” *Journal of Medical Systems*, vol. 36, no. 6, pp. 4005–4020, 2012.
* [13] K. Gai, M. Qiu, “Blend arithmetic operations on tensor-based fully homomorphic encryption over real numbers,” *IEEE Transactions on Industrial Informatics*, 2017, DOI: 10.1109/TII.2017.2780885..
* [14] M. Li, S. Yu, Y. Zheng, K. Ren, andW. Lou, “Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption,” *IEEE Transactions on Parallel and Distributed Systems*, 2013, vol. 24, no. 1, pp. 131–143.
* [15] “Health Insurance Portability and Accountability,” http://www.hhs.gov/ocr/privacy/hipaa/administrative/privacyrule/, accessed on October 20, 2014.
* [16] Z. Xiao and Y. Xiao, “Security and privacy in cloud compu-ting,” *IEEE Communications Surveys and Tutorials*, vol. 15, no. 2, pp. 1–17, Jul. 2012.