ISSN: 2454-9940

# INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

https://zenodo.org/records/11545260

Vol 18, Issue 2, 2024

# "DRUG DISEASE PREDICTION USING MACHINE LEARNING"

**Syed Kamaal Uddin [1], Abdur Rahman Madani [2], Mohd Arbaaz Qureshi [3] Abdul Ahad Afroz [4]**

[1,2,3] B.E Student, Dept. of Information Technology, ISL Engineering College

[4]Assistant Professor (PhD), Dept. of Information Technology, ISL Engineering College

## ABSTRACT

***Data mining is sifting through large databases in search of meaningful trends and patterns. At this point in time, every single person on Earth relies on allopathic treatments and medicine. Data mining techniques are well-suited to medical databases because of the abundance of useful textual and visual information they include. To make sense of healthcare companies' mountain of unstructured data, data mining is essential. An API that can diagnose and prescribe medicine to users based on their symptoms is shown in this article. The API uses machine learning methods for this purpose. In order to identify the most particular condition that may be linked to symptoms, we use some intelligent data about the mining process. The patient could have little trouble recognising the diseases. With only a description of their symptoms, patients will be able to diagnose the problem, since the app will display potential disorders. When there is no nearby medical facility or when the patient is in dire straits, the system will become complacent. It is possible to do predictive analysis on the illness by taking into account various database attributes; this might result in pharmaceutical suggestions for the user. The trial's results could potentially inform future studies and healthcare technologies.***

## INTRODUCTION

Despite the technology's perceptual latency, it is nonetheless vital for numerous tasks, including x-ray imaging and surgical representation. With so many additional factors to consider, like the patient's medical history, their surrounding environment, their blood pressure, and the weather, the surgery is still highly dependent on the doctor's skill and understanding. Everyone thinks that a lot of aspects need to be included in order to grasp the full working process, but no model has succeeded in doing so thus far. To overcome this restriction, medical decision support systems are essential. Doctors may be led to the correct decision by the technology. A "medical decision support system" may refer to either the process of making an attempt at a diagnosis or the ultimate diagnosis that emerges from such an Endeavour. In the sense that it indicates a categorical kind of abnormality or a degree of abnormality on a continuum, the diagnostic opinion is important. Aside from medical considerations, non-medical factors that influence it include power ethics and financial incentives for the patient and practitioner. A story, a metaphor, a lengthy formulation, or a brief summary can all be possible formats. This code, which may be computer software or any other kind of communication, could trigger payment, prescription, notification, information, or advice. Signs of medical diagnostics include knowing what is considered normal and being able to assess the patient's current health. Automated decision support systems are rule-based systems that strive to automatically offer solutions to common management difficulties. Medical decision-making may be very specialist and difficult in some cases, such as when dealing with rare diseases or unusual conditions. Another possible reason for doctors making the wrong diagnosis is because they are nervous, tired, or just do not know enough. Traditional algorithms may sift through a mountain of hidden variables in addition to the patient's present health state, medical history, family medical history, and other pertinent information from their medical records. Differential diagnostic

techniques may be used to find the presence of an entity when there is more than one possible explanation, such in a case with many possible causes. The first step in using this method is to obtain or remove evidence that significantly lowers the probability of candidate conditions. Let me break it down for you: 1) the doctor starts by gathering all of the patient's information and making a list of all of the symptoms.2) It is essential that the doctor consider all potential causes of the patient's symptoms. Thirdly, place the most significant possible symptoms' causes at the head of the list. 4) The doctor should address acute health issues first, then look into and treat any underlying problems."Rule Out" in the sense of putting anything to the test or other scientific method. It will be required to rule out the possibility of such a diagnosis and, depending on the real condition, provide tests that should yield distinct results if no such diagnosis is available. Whether or whether this is feasible depends on the doctor's level of experience and education. This method is easy to implement. In order to find the most probable diseases using the K-Means approach, it is necessary to filter down the various parameters. This approach works better when dealing with several diseases. K-Mean is one of the most well-liked unsupervised learning algorithms when it comes to classification jobs. Finding the centres of the k clusters is the main objective. The different patient tests will serve as the basis for grouping. This method consistently produces the right diagnosis with fewer iterations since cluster boundaries are clear and do not overlap. Any user with an internet connection may access this system, which is made possible by its Service Oriented Architecture (SOA) architecture. The LAMSTAR Network can be utilised for weight computation, algorithm refinement, and overall speed testing, which yields better results.

## LITERATURE REVIEW

**"A preeclampsia diagnosis approach using Bayesian networks"**

Hypertension is the main reason why mothers die during childbirth. Preeclampsia may develop at any point in a woman's pregnancy. Finding those who are more likely to have preeclampsia is the first step in reducing the severity of the illness and its effects. A thorough examination is necessary to assist decision-makers in many medical circumstances requiring massive volumes of data. Compiling all available data and extracting useful insights from it is made feasible with the aid of sophisticated decision support systems. In the face of uncertainty, Bayesian networks provide models that facilitate data collection. Using Bayesian networks, this study recommends developing a system to assist prenatal care doctors in making rational decisions when diagnosing preeclampsia. Qualitative and quantitative modelling, as well as network construction, is other subjects addressed. The primary result of this research is the implementation of a Bayesian network that may aid decision makers in the care of pregnant women during times of uncertainty.

**"Recent developments in data mining applications and techniques"**

Many new subfields of data mining were developed and extensively researched between 2013 and 2015. Topics covered include healthcare, education, data streams, the internet, and big data. In the first part of the essay, the study's methodology and basic data mining processes and approaches are presented. Next, for each field of study, we discuss the methods, approaches, applications, and challenges revealed by the relevant surveys and research. The essay concludes with a review of the challenges encountered by researchers throughout data mining studies and recommendations for further study to address any gaps in our understanding.

**"MobDBTest: A machine learning based system for predicting diabetes risk using mobile devices"**

It is quite probable that diabetes mellitus (DM) has epidemic proportions in India. The enormous quantity of illness and destruction caused by diabetes and its potential complications has put a tremendous strain on health care systems and families worldwide. Worryingly, diabetes is being diagnosed at younger and younger ages throughout the country, and it is now recognized to be linked to several complications. The rising incidence of diabetes in India is due to a number of causes, one of which is the fact that people's lifestyles have changed as a result of their migration from rural to urban areas. Because people don't know enough about diabetes, many people die too soon. A competence that increases diabetes awareness may therefore have an effect on the Indian populace. To combat the widespread ignorance about diabetes, this research proposes an android app-based strategy. By using state-of-the-art machine learning methods, the programmed is able to predict the user's

blood sugar level. At the same time, the gadget provides data and suggestions pertaining to diabetes. In this comparative study, four distinct ML approaches were tested simultaneously. The greatest performance was achieved by the Decision Tree (DT) classifier out of the four ML approaches. In light of this, the diabetes prediction app's backend is built using a DT classifier utilising real-world datasets obtained from a world-class hospital in Chhattisgarh, India.

**"Advance in Intelligent Systems and computing"**

The "Advances in Intelligent Systems and Computing" journal publishes articles on a wide range of AI-related subjects, including theory, applications, and design techniques. Among the several disciplines included are engineering, economics, e-commerce, healthcare, the environment, life science, computer science, and information and communication technology. Computational intelligence, soft computing (including neural networks and fuzzy systems), evolutionary computing (and the fusion of these paradigms), social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, vision and perception, systems based on DNA and the immune system, systems that self-organize and adapt, e-learning and teaching, systems that are centred around humans and their needs, recommender systems, intelligent control, robotics and mechatronics (including human-machine teaming), data analysis, knowledge management, intelligent agents, intelligent decision make, The majority of the articles published under the title "Advances in Intelligent Systems and Computing" are proceedings from conferences, symposiums, and congresses. They review the most recent theoretical and practical developments in the field. The novels in the series are famous for their fast publication and widespread dissemination. The results of studies may now be more widely and rapidly communicated.

## PROBLEM SYSTEM:

Machine learning algorithms provide a game-changing opportunity for advancements in medication development and disease prediction. These algorithms supplement conventional approaches by incorporating insights derived from data. Predicting the effectiveness of treatments and their side effects is one of its main uses. Clinical data, genetic variants, and molecular structures may be analyzed using machine learning (ML) approaches to uncover drug chemistry correlations with treatment results. This shortens the time it takes to bring new drugs to market and helps the pharmaceutical business save money.

## PROPOSED SYSTEM:

To implement the previous strategy, two conditions must be met: first, the ability to accurately identify cars on roads and second, the ability to monitor their velocities. In this case, we may use OpenCV, which makes use of the Haar cascade, to train our system to identify the object—the automobile. We estimate the velocities of the autos sensed on the highways using a python script and a Haar cascade that we built. The real-time application of this project is showing great promise because of its low development costs, fast processing speeds, and simplicity of implementation. Furthermore, the system might be integrated into modelling programmers to track the speeds of vehicles. Improving its ability to recognize different kinds of vehicles and identify those who ignore traffic signals is within the realm of possibility. Improving the project may be achieved, for example, by expanding the haar cascade to detect more vehicles on the road. Better algorithmic identification of these vehicles will allow for a more streamlined search procedure.

**Advantage:**

- More Accuracy.
- More Efficiency.

## SYSTEM REQUIREMENTS HARDWARE REQUIREMENTS:

Since they may serve as the basis for an implementation contract, the hardware requirements should provide a thorough and uniform description of the whole system. As a starting point for the system design, they are used by software engineers. The systems' functionality, not their planned implementation, is shown.

## OPERATING SYSTEM:

Windows: 7 or newer

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

https://zenodo.org/records/11545260

Vol 18, Issue 2, 2024

- Processor - Pentium –IV
- Speed - 1.1 GHz
- RAM - 256 MB(min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse

## SOFTWARE REQUIREMENTS

You may find the system requirements in the software requirements document. It needs to have a description as well as a set of requirements. It gives forth general aims for the system to fulfill rather than detailing particular techniques. It could be useful for estimating costs, arranging team activities, carrying out jobs, and evaluating progress throughout development.

- Operating System - Windows7/8/10/11
- Programming Language - Python

## SYSTEM DESIGN

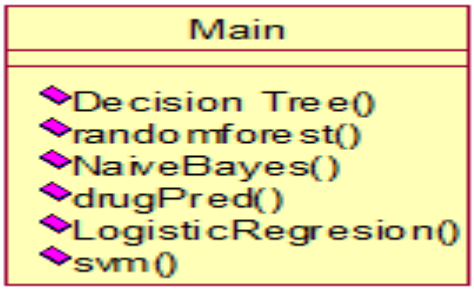
**UML Diagram:**

**Class Diagram:**

The class diagram is the main building block of Design method based on objects. Both low-level modeling—which entails translating models into code—and high-level conceptual modelling of the application's architecture are its primary purposes. Class diagrams are often used in data modelling. Class diagrams display the main objects and their interactions with the programmed, as well as the classes that need coding. A three-part box represents each category in the figure:

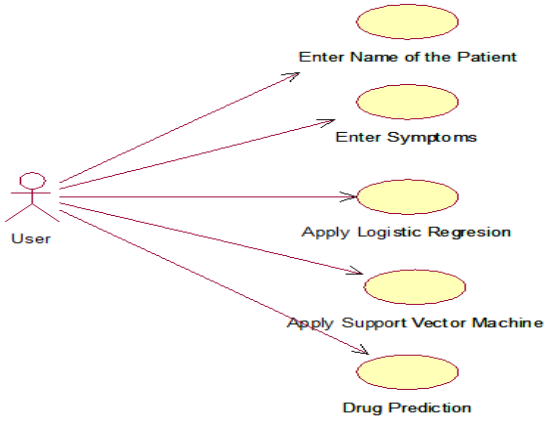
The upper part holds the name of the class

- The middle part contains the attributes of the class
- The bottom part gives the methods or operations the class can take or undertake.



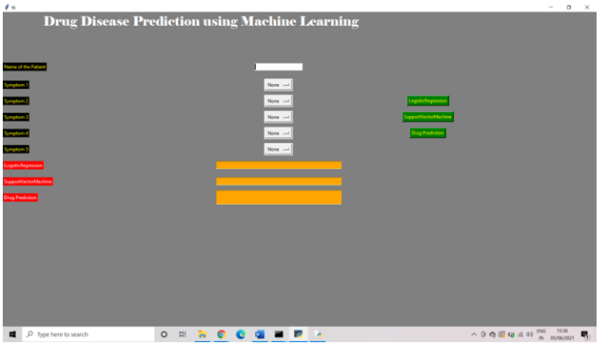

**Use case Diagram:**

In its most basic form, a use case diagram is a graphical depiction of the steps a user takes to complete a use case inside a system. A use case diagram may show how various users interact with a system and the different sorts of users that utilize it. Along with the textual use case, this diagram type is often utilised, and other diagram kinds are frequently included as well.




## OUTPUTSSCREENSHOTS

Click on run.bat file in your project directory




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.

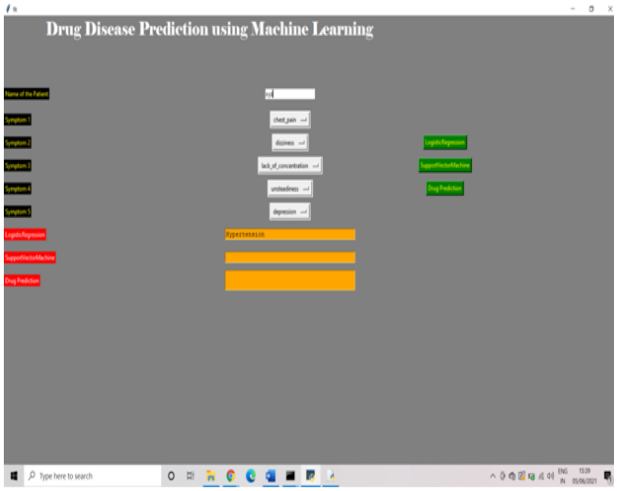
INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

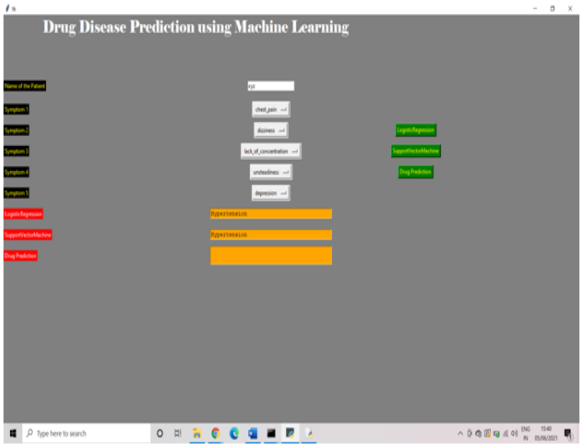
www.ijasem.org

https://zenodo.org/records/11545260

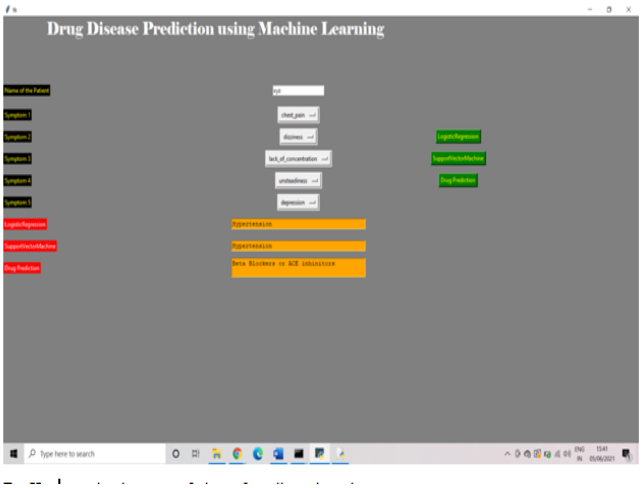
Vol 18, Issue 2, 2024




From the above figure for given symptoms it predicted Hypertensio by using Logistic regression Now test for SVM also.




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




For Hypertension it suggest 2 drugs for relieve the pain

**Symptoms for Diabetes:**

- Blurred &Distored Vision,
- Obesity,
- Excessive hunger
- Polyuria
- Increased appetite
- Symptoms for Hypertension:
- Lack of concentration, Unsteadiness,
- Diziness, chest pain,

## CONCLUSION AND FUTURE ENHANCEMENTS

Finally, a critical new area in healthcare is emerging at the crossroads of illness research and medication development. The discovery of new medicines, the knowledge of disease causes, and the personalization of therapeutic methods are all being expedited by the continuous developments in technology, particularly data analytics and machine learning. Nevertheless, there are still obstacles to overcome, such as the need of thorough validation, ethical concerns, and fair access to new medicines. To overcome these obstacles and fully use drug development in the fight against illnesses, multidisciplinary teamwork, rigorous scientific investigation, and a focus on the patient will be crucial in the future.

The area of illness diagnostics and drug development has recently seen the emergence of machine learning as a potent tool. It is able to sift through mountains of data in search of promising medication candidates or illness outcomes predictions thanks to its complex algorithms. Researchers have shown that deep learning, support vector machines, and random

forests are some of the most effective machine learning approaches for improving illness diagnosis accuracy and streamlining the medication development process. But there are still obstacles, such as issues with data quality, model interpretability, and ethical concerns. To sum up, machine learning has enormous potential to transform the pharmaceutical and medical diagnostic industries, but it will need more study and cooperation to reach its full potential.

## REFERENCES

1. Mário W. L. Moreira, Joel J. P. C. Rodrigues, Antonio M. B. Oliveira, Ronaldo F. Ramos, Kashif Saleem, "A preeclampsia diagnosis approach using Bayesian networks", Communications (ICC) 2016 IEEE International Conference on, pp. 1-5, 2016, ISSN 1938-1883.

2. Ameera M. Almasoud, Hend S. Al-Khalifa, Abdulmalik Al-Salman, "Recent developments in data mining applications and techniques", Digital Information Management (ICDIM) 2015 Tenth International Conference on, pp. 36-42, 2015.

3. K Sowjanya, Ayush Singhal, Chaitali Choudhary, "MobDBTest: A machine learning based system for predicting diabetes risk using mobile devices", Advance Computing Conference (IACC) 2015 IEEE International, pp. 397-402, 2015.

4. Bharathan Venkatesh, Danasingh Asir Antony Gnana Singh, Epiphany JebamalarLeavline, Advance in Intelligent Systems and Computing, vol. 517, pp. 633, 2017, ISSN 2194-5357, ISBN 978- 981-10-3173-1.

5. Felix Gräßer, Surya Kallumadi, Hagen Malberg, and Sebastian Zaunseder. 2018. Aspect-Based Sentiment Analysis of Drug Reviews Applying Cross-Domain and Cross-Data Learning. In Proceedings of the 2018 International Conference on Digital Health (DH '18). ACM, New York, NY, USA, 121-125.

6. Effieient Algorithms to find Frequent Itemset Using Data MiningSagar Bhise1, Prof. Sweta Kale International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056.

7. Text Data Mining of Care Life Log by the Level of Care Required Using KeyGraph Muneo Kushima, Kenji Araki, Tomoyoshi Yamazaki, Sanae Araki, Taisuke Ogawa, Noboru Sonehara Proceedings of the International MultiConference of Engineers and Computer Scientists 2017 Vol I, IMECS 2017, March 15 - 17, 2017, Hong Kong.

8. Chen, W., Guo, H., Zhang, F., Pu, X., and Liu, X. (2012). Mining Schema Matching Between Heterogeneous Databases. In Consumer Electronics, Communications and Networks (CECNet), 2012 2nd International Conference on, pages 1128–1131. IEEE.

9. Han, J. and Gao, J. (2009). Research Challenges for Data Mining in Science and Engineering. Next Generation of Data Mining, pages 1– 18.

10. H. Abe, S. Hirano and S. Tsumoto, “Evaluation Temporal Models Based on Association Mining From Medical Documents,” Japan Association for Medical Informatics, vol. 27, no. 1, pp. 33-38, 2007.

11. Sun, Y., Han, J., Yan, X., and Yu, P. S. (2012). Mining Knowledge from Interconnected Data: A Heterogeneous Information Network Analysis Approach. Proceedings of the VLDB Endowment.

12. M. Akhil jabbar& Dr. Priti Chandrab “Heart Disease Prediction System using Associative Classification and Genetic Algorithm” International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies-ICECIT, 2012.

13. Nikhil N. Salvithal“ Appraisal Management System using Data mining “International Journal of Computer Applications (0975 – 8887) Volume 135 – No.12, February 2016.

14. Tanvi Sharma, Anand Sharma & Vibhakar Mansotra “Performance Analysis of Data Mining Classification Techniques on Public Health Care Data” International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 6, June 2016.

15. V.Krishnaiah , Dr.G.Narsimha, Dr.N.Subhash Chandra “Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques”(IJCSIT)International Journal of

Computer Science and Information Technologies, Vol. 4 (1) , 2013, 39 - 45.

16. Jaimini Majali, Rishikesh & Niranjan, Vinamra Phatak "Data Mining Techniques For Diagnosis And Prognosis Of Cancer" International Journal of Advanced Research in Computer and CommunicationEngineering Vol. 4, Issue 3, March 2015.

17. Data Mining Techniques to Predict Diabetes Influenced Kidney Disease Swaroopa Shastri, Surekha, SaritaInternational Journal of Scientific Research in Computer Science, Engineering and Information Technology © 2017 IJSRCSEIT | Volume 2 | Issue 4 | ISSN : 2456-3307.

18. DeFronzo Ralph. From the triumvirate to the ominous octet: a new paradigm for the treatment of type 2 diabetes mellitus. Diabetes. 2009;58:773-95.