iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.59394

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

Symptoms Based Disease Prediction System Using Machine Learning

Mr. N. Ashok[1], P. Imran Khan[2], M. N. V. Yasaswi[3], K. Surya Ratna[4], M. Gopi Krishna[5]
[1]*Assistant professor,* [2, 3, 4, 5] *Student, Dept. of Information Technology , Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt.,Andhra Pradesh.*

Abstract: ***The "Symptoms-Based Disease Prediction System using Machine Learning" project is a solution that intends to help users predict their potential diseases based on the symptoms they enter and other general information. This system uses several machine learning algorithms like Decision Tree, KNN, Naïve Bayes, and Random Forest, which are all implemented using the Python programming language. The project uses a dataset collected from hospitals, which allows the system to provide predictions with high accuracy, and ongoing development is in place to improve accuracy even further. The primary objective of this project is to address the challenges that the health industry faces, where people often ignore common symptoms, leading to more severe conditions due to factors such as laziness, time constraints, or reluctance to consult a doctor. Furthermore, this project intends to reduce the workload on doctors and provide a quick and accurate diagnosis using machine learning, thus supporting the health industry.***

Keywords***: Disease Prediction, Machine Learning Algorithms, Python Programming, Healthcare Industry, Symptoms based Prediction***

I. INTRODUCTION

The "Symptoms-Based Disease Prediction System using Machine Learning" is an innovative project that aims to help users predict their potential diseases based on their symptoms and other pertinent information. The system leverages several machine learning algorithms, including Decision Tree, KNN, Naïve Bayes, and Random Forest. All these algorithms are implemented using the popular Python programming language. The project relies on a dataset collected from reputable hospitals, which enables the system to provide highly accurate predictions. Ongoing development is in progress to enhance the system's accuracy even further.

The primary objective of this project is to address the challenges that the healthcare industry faces. Often, people ignore common symptoms, leading to more severe conditions, due to reasons such as laziness, time constraints, or reluctance to consult a doctor. The project aims to mitigate this issue by providing a quick and accurate diagnosis using machine learning. This, in turn, reduces the workload on doctors and supports the healthcare industry

To achieve its goal, the "Symptoms-Based Disease Prediction System using Machine Learning" project is working on developing multiple models to determine the most accurate one. The project collects and categorizes raw data based on symptoms using various machine learning algorithms. Traditionally, doctors have used trial-and-error methods to develop treatments, which has led to significant advancements. However, with the evolution of technology, particularly the high computing power enabling ML, new possibilities have emerged. ML algorithms can now uncover hidden patterns in data, providing valuable insights and enhancing disease detection and treatment project aims to provide an alternative health assessment method that does not require immediate medical consultation. By relying on Machine Learning algorithms and Python programming, the project leverages hospital datasets to make highly accurate predictions. The system is designed to address the need for early disease detection by providing a user-friendly prediction interface. By offering a convenient and accessible means of health check-up, this project contributes to the health industry. Moreover, it aims to bridge the gap for individuals who are hesitant to visit hospitals for minor health concerns.

II. LITERATURE SURVEY

1) Marshland, S. (2009) Machine Learnings an Algorithmic Perspectives. CRC Press, New Zealand, 6-7.Otoom et al., (2015) Effective Diagnosis and Monitoring of Heart Diseases. International Journal of Software Engineering and Its Application. 9, 143- 156.

2) Vembandasamy et al., (2015) Heart Disease Detection Using Naive Bayes Algorithms. IJISET-International Journal of Innovative Science, Engineering & Technology, 2, 441-444.

3) Tan et al., (2009) A Hybrid Evolutionary Algorithm for Attribute Selections in Data Mining. Journal of Expert Systems with Application, 36, 8616- https://doi.org/10.1016/j.eswa.2008.10.013.
4) Iyer, A., Jeyalatha, S. and Sumbaly, R. (2015) Diagnosis of Diabetes Using Classification Mining
5) Sarwar, A. and Sharma, V. (2012) Intelligent Naive Bayes Approaches to Diagnose Diabetes Type-2. Special Issues of International Journal of Computer Application (0975-8887) on Issues and Challenges in Networking, Intelligences and Computing Technologies-ICNICT 2012, 3, 14-16.
6) Otoom et al., (2015) Effective Diagnosis and Monitoring of Heart Diseases. International Journal of Software Engineering and Its Application. 9, 143-156.
7) Ephzibah, E.P. (2011) Cost Effective Approach on Feature Selection using Genetic Algorithm and Fuzzy Logics for Diabetes Diagnosis. International Journal on Soft Computing (IJSC), 2, 1-10. https://doi.org/10.5121/ijsc.2011.2101. ➢Vijayarani, S. and Dhayanand, S. (2015) Liver Diseases Prediction using SVM and Naive Bayes Algorithm. International Journals of Science, Engineering and Technology Researches (IJSETR), 4, 816-820.

III. METHODOLOGY

1) *Data Collection:* The project began by sourcing datasets from hospitals, containing essential information on patients' symptoms and diagnosed diseases. This step was critical to ensure the machine learning model had access to comprehensive and relevant data, necessary for accurate predictions.
2) *System Development:* Next, a Disease Prediction model was developed using multiple machine learning algorithms such as Decision Tree, KNN, Naïve Bayes, and Random Forest, in Python. The algorithms were chosen for their effectiveness in handling classification tasks, critical for disease prediction. Additionally, a user-friendly interface was created using Flask, HTML, and CSS to facilitate easy interaction with the system.
3) *Training the Model:* The collected datasets were then used to train the machine learning model, where the symptom-disease data was fed into the model to enable it to learn and make precise predictions based on the entered symptoms. This step was crucial in ensuring the model could accurately predict diseases by adhering to established medical guidelines, representing a significant step towards efficient, data-driven healthcare solutions in the modern era.

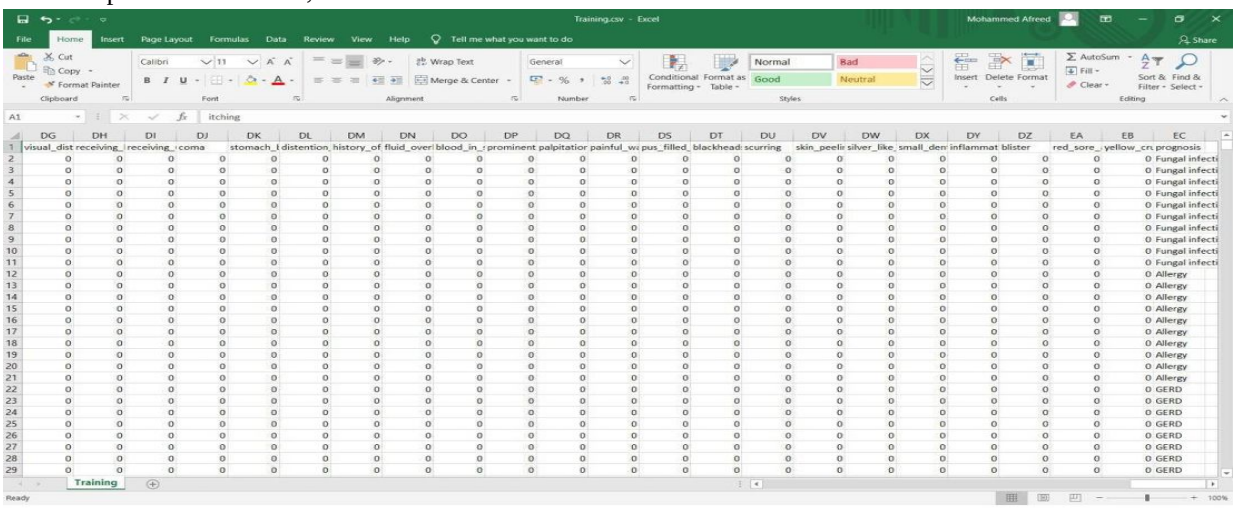


Fig. 1. Trained Data

4) *User Interaction:* To make the system accessible to users, a user-friendly interface was designed. This interface allows users to input their symptoms through a dropdown menu, making it easy for them to provide the necessary information for the disease prediction process.
5) *Prediction and Results:* Once the user inputs their symptoms, the system utilizes the trained machine learning model to predict the possible diseases based on the symptoms provided. The results are then presented to the user, along with a probability estimation for each predicted disease. This information can help users understand the likelihood of having a particular disease based on their symptoms, enabling them to take appropriate actions such as seeking medical advice or treatment.

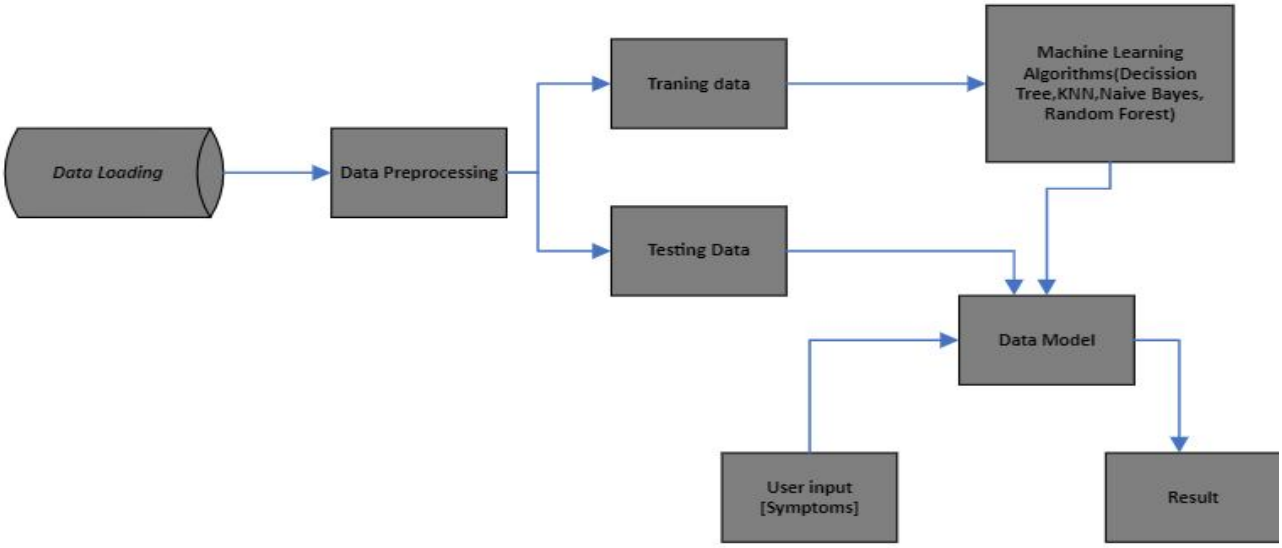


Fig. 2. Architecture

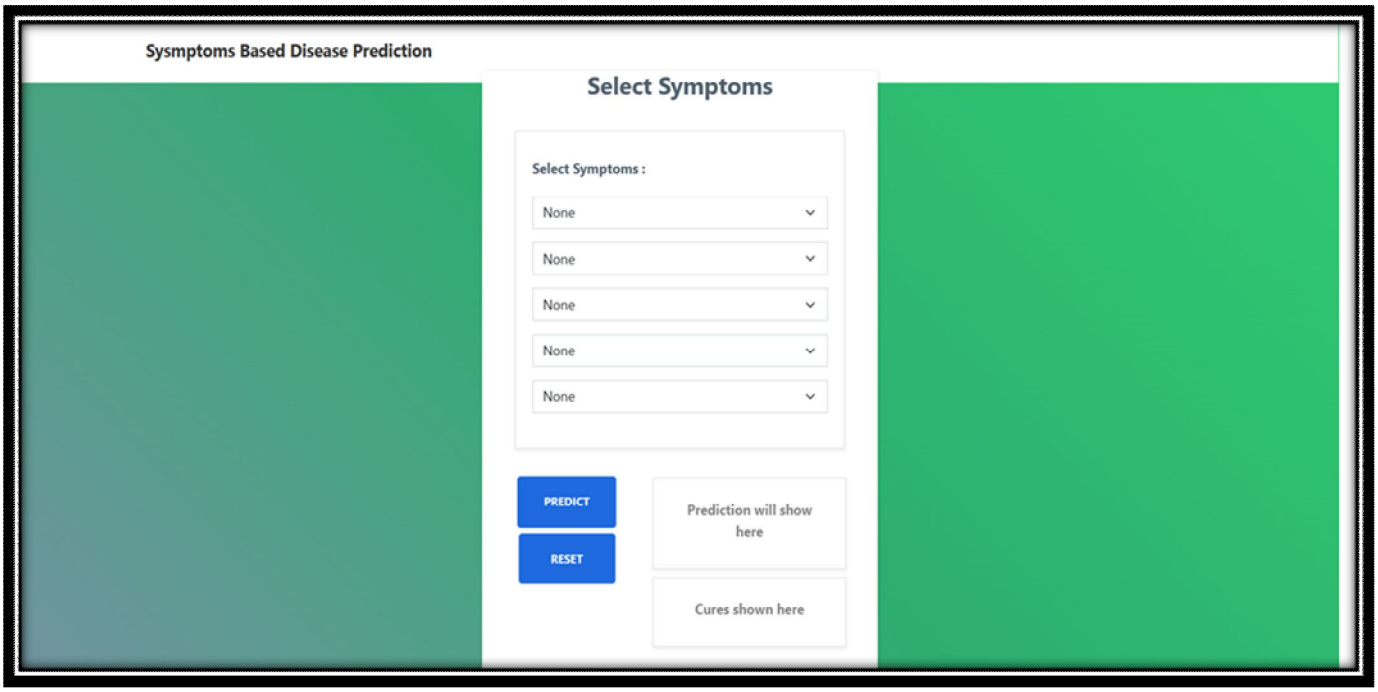


Fig. 3. Web Interface

IV. RESULTS AND DISCUSSIONS

Through rigorous experimentation, the project achieved a high accuracy rate in disease prediction. This was accomplished by combining structured and unstructured data, demonstrating enhanced accuracy in predictions. The system's ability to identify diseases early potentially prevents severe health complications and reduces the immediate need for medical consultation. These promising results highlight the potential to democratize health information access, empowering users to proactively manage their health by identifying issues at an early stage.

Fig. 3. Select Symptoms

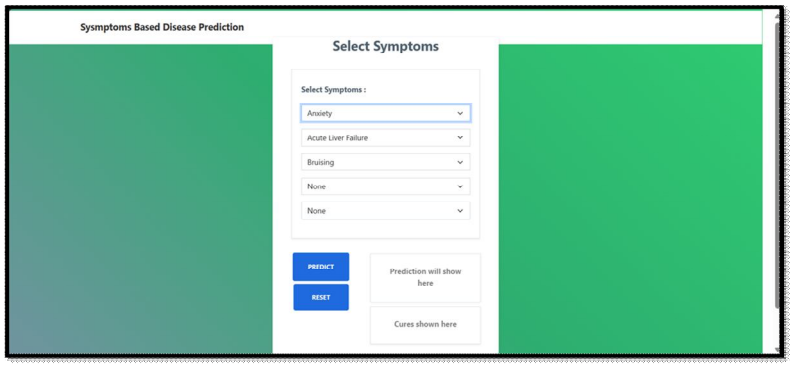
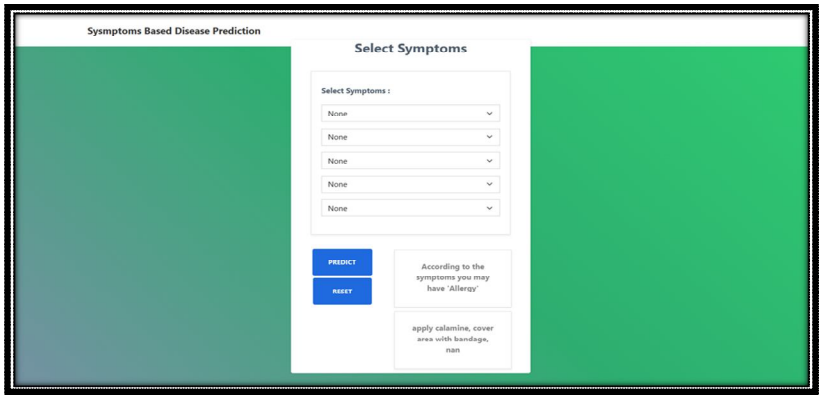


Fig. 4. Disease Prediction



V. FUTURE WORK

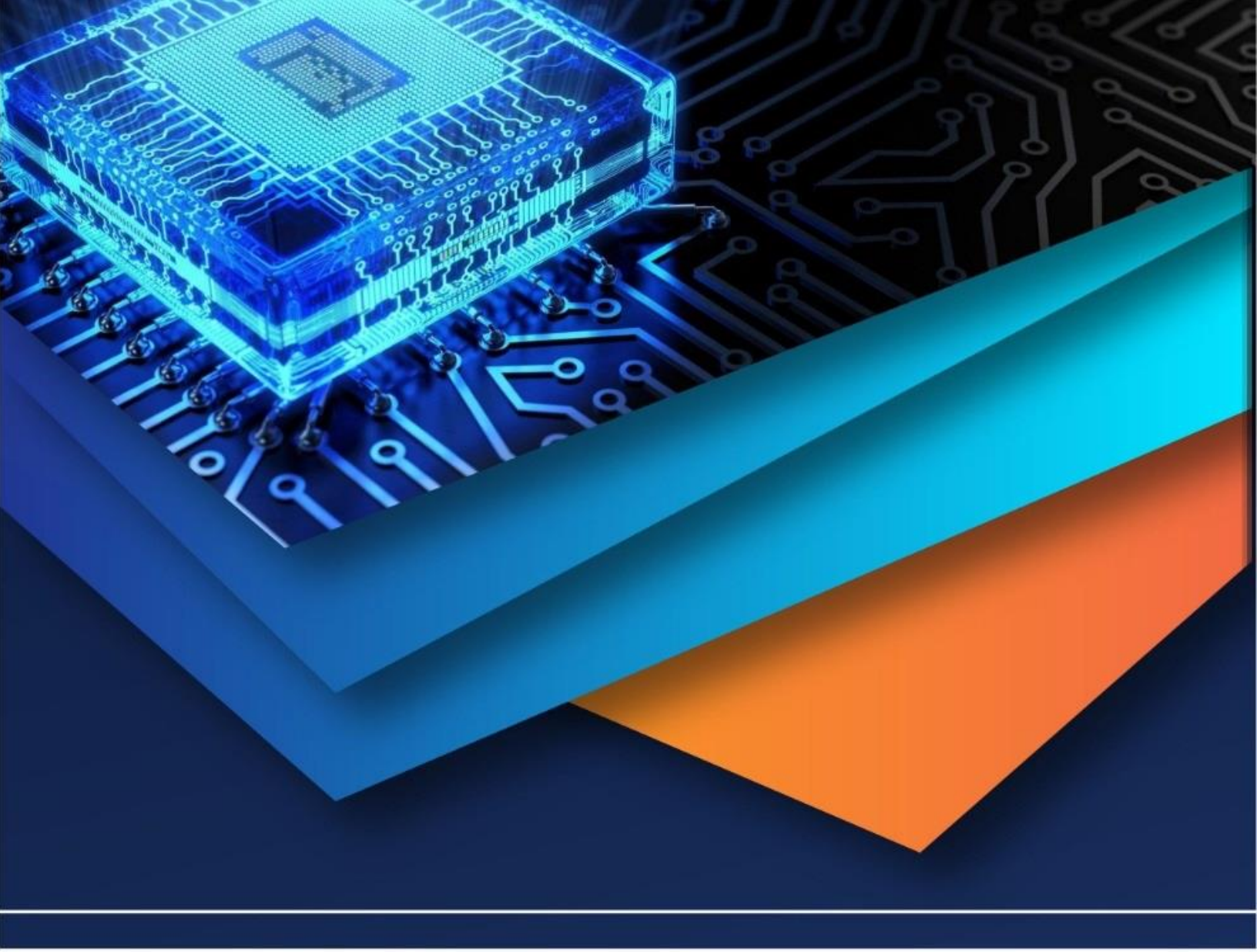
This project aims to enhance user interaction by developing a more interactive user interface. It will include features for creating backups of user data and ensuring the security and availability of important information. Additionally, the system will be designed to be mobile-friendly, enabling users to access it conveniently on their smartphones. The interface will provide more comprehensive details on diseases and include information on the latest diseases, ensuring that users have access to up-to-date and relevant information.

VI. CONCLUSION

This project on disease prediction using machine learning is highly beneficial for both individuals and the healthcare sector. It allows users to predict diseases based on their symptoms and general information without visiting a hospital or clinic. For the healthcare industry, this system can significantly reduce the workload of doctors by providing quick and accurate disease predictions. By identifying diseases early, it helps prevent them from becoming fatal and causing further complications. Overall, this project has the potential to improve healthcare efficiency and patient outcomes.

REFERENCES

[1] Gulia, A et al., (2014) Liver Patients Classification Using Intelligent Technique. (IJCSIT) International Journal of Computer Science and Information Technology, 5, 5110-5115.
[2] Rajeswari, P. and Reena,G.S. (2019) Analysis of Liver Disorders Using Data Mining Algorithms. Global Journal of Computer Science and Technology, 10, 48-52.
[3] Tarmizi, N.D.A et al., (2018) Malaysia Dengue Outbreaks Detection Using Data Mining Model. Journals of Next Generation Information Technology (JNIT), 4, 96- 107.
[4] Fathima, A.S. and Manimeglai, D. (2016) Predictive Analysis for the Arbovirus Dengue using SVM Classifications. International Journals of Engineeringand Technology, 2, 521-527.
[5] Rambhajani, M et al., (2015) A Survey on Implementations of Machine LearningTechnique for Dermatology Disease Classifications. International Journal of Advance in Engineering & Technology, 8, 194-195.
[6] Otoom et al., (2015) Effective Diagnosis and Monitoring of Heart Diseases. International Journal of Software Engineering and Its Application. 9, 143-156.
[7] Reena,G.S. (2019) Analysis of Liver Disorders Using Data Mining Algorithms. Global Journal of Computer Science and Technology, 10, 48-52.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY