iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VII **Month of publication:** July 2026

DOI: https://doi.org/10.22214/ijraset.2026.84307

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 14 Issue VII July 2026- Available at www.ijraset.com

Medicine Information and Advice System Using NLP and Decision Tree Algorithm

G Manoj Kumar[1], Bondada Sravani[2]
[1]*Assistant Professor,* [2] *MCA Final Semester, Master of Computer Applications, Sanketika Vidya Parishad Engineering College, Vishakhapatnam, Andhra Pradesh, India*

Abstract: ***The "Medicine Information and Advice System" is an application based on machine learning that aims to give reliable medicine information and individualized healthcare advice, enabling users to make better, informed choices more securely. It can handle both symptom-oriented and medicine-oriented questions through a simple interface, suiting all types of users. Using artificial intelligence, the system performs operations on user input either symptoms or medicine names to provide crucial health information. For medical questions, it provides comprehensive information such as, composition , product information, possible side effects, therapeutic class, uses. For symptom-based search, predictive modeling suggests possible diseases and recommends suitable medicines, diets, exercises, and precautions required. The system features voice command for hands-free use and is backed by a performance-tuned backend for rapid, reliable data access with ongoing learning from user behavior to enhance precision. Coupled with a simple, intuitive frontend and a large, robust medical database, the system improves health decision-making for patients and clinicians alike, enabling better, safer, and more tailored health outcomes.***

Keywords : Machine Learning, Medicine Information System, Disease Prediction, Symptom Analysis, Healthcare Recommendation, Drug Information, Artificial Intelligence (AI), Voice Assistance, Personalized Healthcare, Medical Database.

I. INTRODUCTION

The **Medicine Information and Advice System** is a machine learning-based healthcare application developed to provide accurate medicine information and personalized medical guidance. It helps users access reliable details about medicines, including their composition, uses, side effects, dosage, manufacturer, and available alternatives. The system also supports symptom-based queries by predicting possible diseases and recommending suitable medications, dietary suggestions, exercises, and precautions. It is designed to bridge the gap between complex medical information and the general public by presenting healthcare knowledge in a simple and understandable manner. Natural Language Processing (NLP) enables the system to interpret user queries efficiently and provide relevant responses. The application also includes voice command functionality, making it more accessible to users with different needs. A large medicine dataset and symptom–disease mapping dataset are used to train the machine learning models. Advanced preprocessing techniques improve the quality and accuracy of predictions. Classification algorithms analyze user inputs and generate meaningful healthcare recommendations. The system is capable of handling both medicine-oriented and symptom-oriented searches through a user-friendly interface. It reduces the time required to obtain reliable medicine information while minimizing the risk of medication errors. Healthcare professionals and patients can benefit from its intelligent recommendation capabilities. Continuous learning from user interactions further improves the system's performance over time. The proposed solution combines Artificial Intelligence, Machine Learning, and NLP to deliver fast, accurate, and personalized healthcare assistance. Overall, the Medicine Information and Advice System provides a smart, secure, and efficient platform for informed medical decision-making.

II. EXISTING SYSTEM

The existing medicine information systems mainly depend on manual searches, medical books, pharmacy websites, or consultations with healthcare professionals. These methods are often time-consuming and may not provide personalized recommendations to users. Most online medicine portals only display basic drug information such as usage, dosage, and side effects. They generally do not support intelligent symptom-based disease prediction or medicine recommendation. Existing systems also lack the ability to understand natural language queries effectively. Many platforms provide static information without considering the user's symptoms or medical context. Voice-based interaction is rarely available, making them less accessible for elderly or visually impaired users. Most systems do not combine medicine search and symptom analysis in a single application. They are unable to suggest suitable dietary plans, exercises, or preventive measures along with medicines.

The absence of machine learning limits their ability to improve recommendations over time. Many applications also fail to provide alternative medicines based on composition or therapeutic class. Some systems require users to have prior medical knowledge to understand the information. They may not support multilingual communication or personalized healthcare advice.

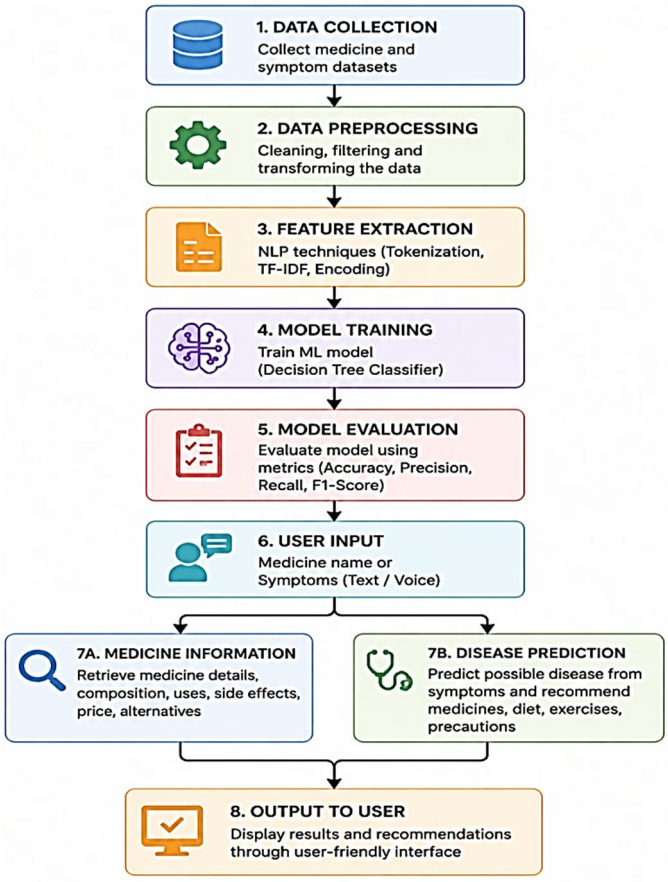


Figure1 : Flow Chart

A. Challenges

The existing medicine information systems face several challenges that reduce their effectiveness and reliability. Most systems provide only basic medicine details without offering personalized healthcare recommendations. They cannot accurately predict diseases based on user symptoms using intelligent algorithms. Many applications rely on manual keyword searches, making information retrieval slow and inefficient. The lack of Natural Language Processing (NLP) limits their ability to understand user queries in simple language. Existing platforms usually do not support voice-based interaction, reducing accessibility for elderly and visually impaired users. Many systems fail to suggest suitable alternative medicines when a specific drug is unavailable. Drug interaction analysis is often limited, increasing the risk of medication errors. Most applications do not provide integrated advice on diet, exercise, and preventive measures. The quality of recommendations depends heavily on incomplete or outdated medical databases. Many healthcare platforms cannot continuously learn from user feedback to improve prediction accuracy. Handling ambiguous or multiple symptoms remains a difficult task for traditional systems. Scalability becomes a challenge when processing large medicine datasets and numerous user requests. Ensuring data privacy and protecting sensitive medical information is another major concern. Achieving high prediction accuracy while minimizing false recommendations is also challenging. Integration of machine learning models with real-time healthcare data requires significant computational resources. Maintaining updated medicine information and treatment guidelines is a continuous challenge. These limitations highlight the need for an intelligent AI and Machine Learning-based Medicine Information and Advice System capable of delivering accurate, fast, and personalized healthcare assistance.

III. PROPOSED SYSTEM

The proposed Medicine Information and Advice System is an AI and Machine Learning-based application developed to provide accurate medicine information and personalized healthcare guidance. It supports both medicine-oriented and symptom-based queries through a simple and user-friendly interface. The system retrieves detailed medicine information, including composition, dosage, side effects, manufacturer, and available alternatives. For symptom-based inputs, it predicts possible diseases using machine learning classification algorithms. It also recommends suitable medicines, dietary plans, exercises, and necessary precautions. Natural Language Processing (NLP) is used to understand and process user queries efficiently. Voice command functionality allows users to interact with the system easily and improves accessibility. A large medicine database ensures accurate and reliable recommendations. Advanced preprocessing techniques improve prediction accuracy and system performance. The application continuously learns from user interactions to enhance future recommendations. It minimizes medication errors by providing trustworthy medical information and alternative medicines. The system delivers quick responses with high accuracy and better user experience. It serves as a smart healthcare assistant for both patients and healthcare professionals. Overall, the proposed system provides fast, secure, and intelligent medical advice for informed healthcare decision-making.

IV. ALGORITHM

A. Data Collection and Preprocessing

The first stage of the Medicine Information and Advice System involves collecting a comprehensive dataset containing medicine details and symptom–disease mappings. The medicine dataset includes medicine names, salt composition, manufacturer, dosage, therapeutic class, side effects, price, and alternative medicines. The symptom dataset contains disease names along with their associated symptoms, medicines, dietary suggestions, exercises, and precautions. After collecting the data, preprocessing is performed to improve its quality. Missing values are handled, duplicate records are removed, and inconsistent entries are corrected. Text fields are converted into lowercase, unwanted symbols are removed, and Natural Language Processing (NLP) techniques such as tokenization and TF-IDF vectorization are applied. Categorical attributes are encoded into numerical values so that machine learning algorithms can process them efficiently. The cleaned dataset is then prepared for model training and testing.

B. Model Training

Once preprocessing is completed, the dataset is divided into training and testing datasets using an 80:20 ratio. The training data is used to build machine learning models capable of predicting diseases from symptoms and retrieving medicine information accurately. Decision Tree is selected as the primary classification algorithm because of its simplicity, interpretability, and efficient classification performance. During training, the algorithm learns relationships between symptoms and diseases by constructing decision nodes based on important features. The trained model also learns medicine-related information such as drug composition, side effects, and alternatives. Hyperparameter tuning and validation techniques are applied to improve prediction accuracy and reduce overfitting. The trained model is then saved for deployment.

C. Prediction and Recommendation

After training, the model is integrated into the application. Users can provide either a medicine name or a list of symptoms through text or voice input. If a medicine name is entered, the system retrieves detailed information such as composition, dosage, manufacturer, side effects, price, and available alternatives from the database. If symptoms are provided, the trained Decision Tree classifier predicts the most probable disease. Based on the prediction, the system recommends suitable medicines along with dietary advice, exercise suggestions, and preventive precautions. NLP techniques help the system understand user queries even when different words or sentence structures are used. The recommendation process is completed within a few seconds, providing users with fast and accurate healthcare guidance.

D. Model Evaluation and Performance

The final stage focuses on evaluating the performance of the machine learning model. The trained model is tested using unseen data to measure its prediction capability. Standard evaluation metrics such as Accuracy, Precision, Recall, and F1-Score are calculated to determine the effectiveness of the system. Confusion matrices are analyzed to identify incorrect predictions and improve model performance. User feedback is also collected to continuously refine the recommendation process. The system updates its knowledge base periodically with new medicine information and treatment guidelines to maintain reliability. Overall, the combination of Decision Tree, Natural Language Processing, and a large healthcare dataset enables the Medicine Information and Advice System to deliver accurate, secure, and personalized medical recommendations while improving user experience and healthcare decision-making.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

V. TECHNIQUES

A. Machine Learning (ML)

Machine Learning is the core technology used in the Medicine Information and Advice System. It enables the application to learn patterns from large medical datasets and make intelligent predictions without being explicitly programmed for every situation. The system is trained using datasets containing medicine information, symptoms, diseases, and treatment recommendations. Classification algorithms are applied to predict possible diseases based on user symptoms and to retrieve relevant medicine details. Machine Learning improves the accuracy of recommendations by learning from historical medical data and user interactions. It also supports continuous model improvement through retraining with updated datasets. By automating prediction and recommendation processes, Machine Learning helps users receive fast, reliable, and personalized healthcare guidance.

B. Natural Language Processing (NLP)

Natural Language Processing (NLP) enables the system to understand and process user queries written or spoken in natural language. Users can enter medicine names or describe their symptoms in simple sentences without following a fixed format. NLP techniques such as tokenization, stop-word removal, text normalization, and TF-IDF feature extraction convert textual information into machine-readable data. These processed features are then used by the machine learning model for accurate prediction and information retrieval. NLP also improves search efficiency by recognizing similar medical terms and handling spelling variations. This technique makes the system more interactive, user-friendly, and capable of providing meaningful healthcare responses.

C. Decision Tree Classification

The Decision Tree algorithm is used as the primary classification technique for predicting diseases based on user symptoms. It works by dividing the dataset into multiple decision nodes according to important symptom features. Each branch represents a possible decision, while the leaf node represents the final disease prediction. Decision Trees are simple to understand, easy to implement, and provide high classification accuracy for structured medical datasets. They require minimal data preprocessing and can effectively handle both numerical and categorical data. The trained Decision Tree model helps the system quickly identify possible diseases and recommend suitable medicines, precautions, dietary plans, and exercises for users.

D. Voice Command and Recommendation System

The system incorporates voice command functionality to improve accessibility and ease of use. Users can interact with the application using speech instead of typing, making it useful for elderly and visually impaired individuals. Speech input is converted into text using speech recognition techniques and then processed through NLP and machine learning models. After analyzing the input, the recommendation engine retrieves detailed medicine information or predicts possible diseases based on symptoms. It provides medicine suggestions, dosage details, side effects, alternative medicines, diet plans, exercise recommendations, and preventive measures. This intelligent recommendation system delivers quick, accurate, and personalized healthcare assistance, enhancing the overall user experience.

VI. METHODOLOGY

The methodology of the Medicine Information and Advice System begins with collecting medicine and symptom datasets from reliable healthcare sources. The collected data is preprocessed by removing missing values, duplicate records, and unnecessary information.

Natural Language Processing (NLP) techniques are applied to convert user text into machine-readable features. The processed dataset is divided into training and testing datasets for machine learning. A Decision Tree classification algorithm is trained to predict diseases based on user symptoms. The trained model is integrated with a medicine database to retrieve detailed drug information and recommendations.

Users can provide medicine names or symptoms through text or voice input. The system analyzes the input and generates accurate medicine details, disease predictions, and healthcare advice. Model performance is evaluated using accuracy, precision, recall, and F1-score metrics. Finally, the system delivers fast, reliable, and personalized healthcare recommendations through a simple user-friendly interface.

A. Input

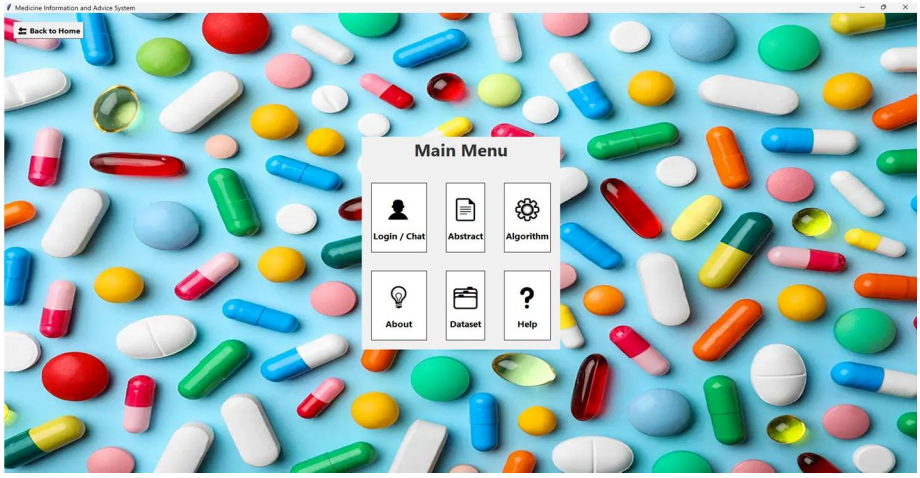


Figure 2: Info Screen Input

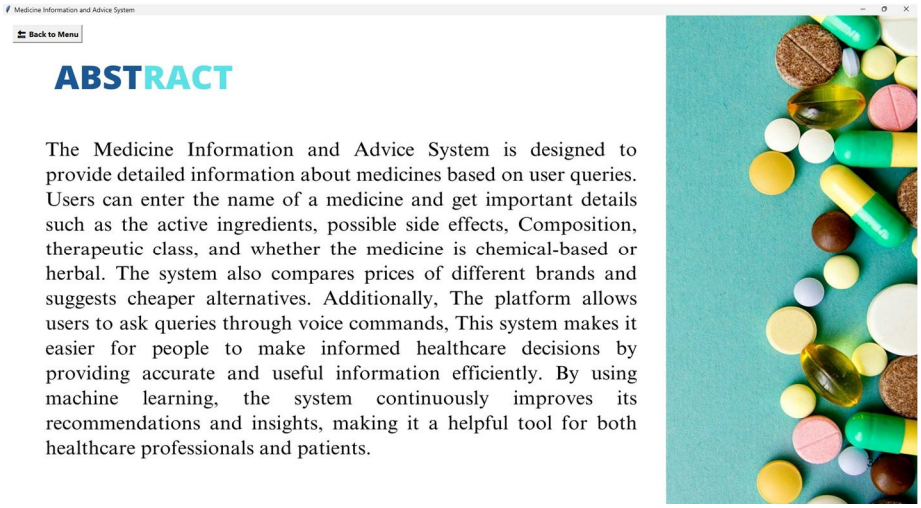


Figure 3:Abstract Screen

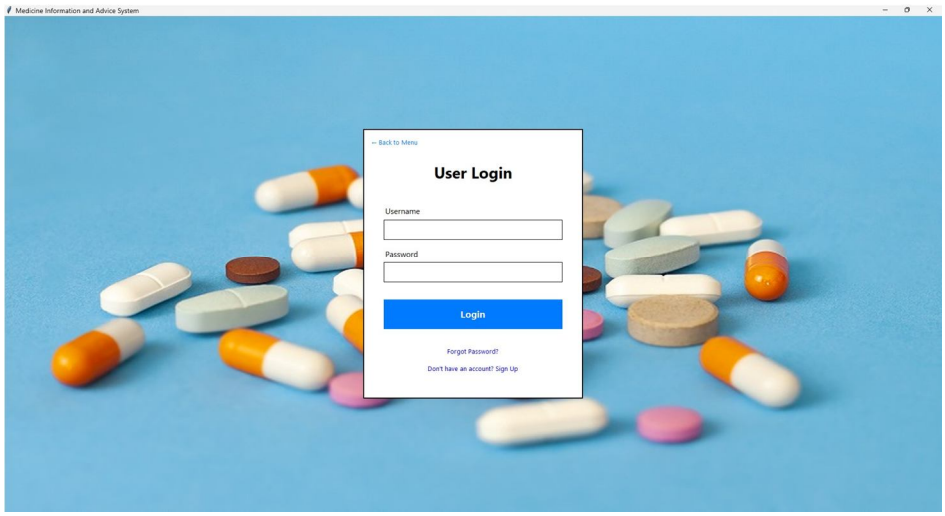


Figure 4: Login Screen

B. Output

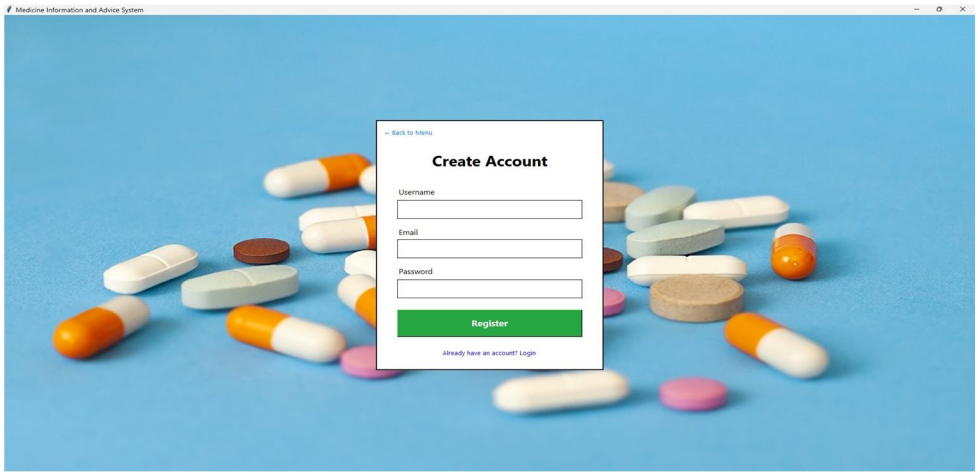


Figure 5: Register Screen

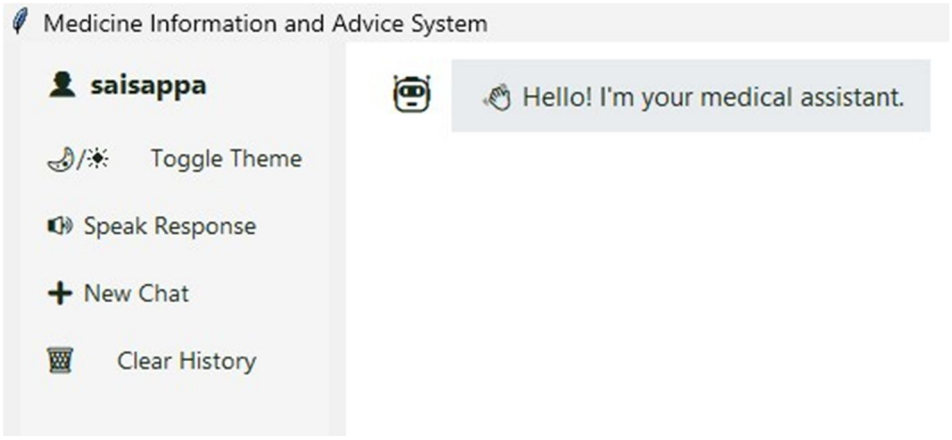


Figure 6: Chat Interface Screen

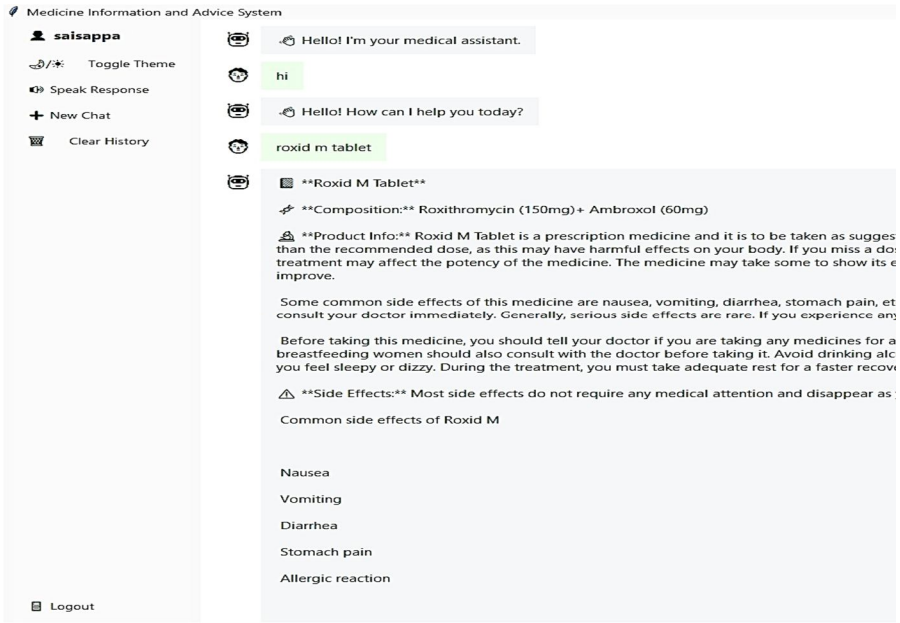


Figure 7: Alternative medicine info Screen

VII. CONCLUSION

The AI-Powered Student Placement Prediction System Using Machine Learning is an intelligent web-based application developed to simplify and improve the student placement process. The system uses the Logistic Regression Machine Learning algorithm to analyze important student attributes such as CGPA, academic percentages, aptitude score, communication skills, technical skills, certifications, internship experience, and project details to predict placement opportunities accurately. By automating the prediction process, the system reduces manual effort, minimizes human errors, and supports data-driven decision-making for educational institutions.

The application provides a secure and user-friendly platform where students can register, manage their profiles, generate placement predictions, and download detailed PDF reports. In addition to predicting placement status, the system also provides placement probability, confidence score, and personalized recommendations that help students identify areas for improvement and enhance their employability. These recommendations encourage students to strengthen their technical knowledge, communication skills, and practical experience before participating in campus recruitment.

The integration of Python, Flask, Scikit-Learn, SQLite, HTML, CSS, and Bootstrap ensures efficient system performance, secure data management, and an interactive user interface. The proposed system can assist placement officers in monitoring student performance, identifying students who require additional training, and improving overall placement management. Overall, the project demonstrates the practical application of Artificial Intelligence and Machine Learning in the education sector by providing an accurate, reliable, and scalable solution for predicting student placement outcomes and supporting better career planning for students.

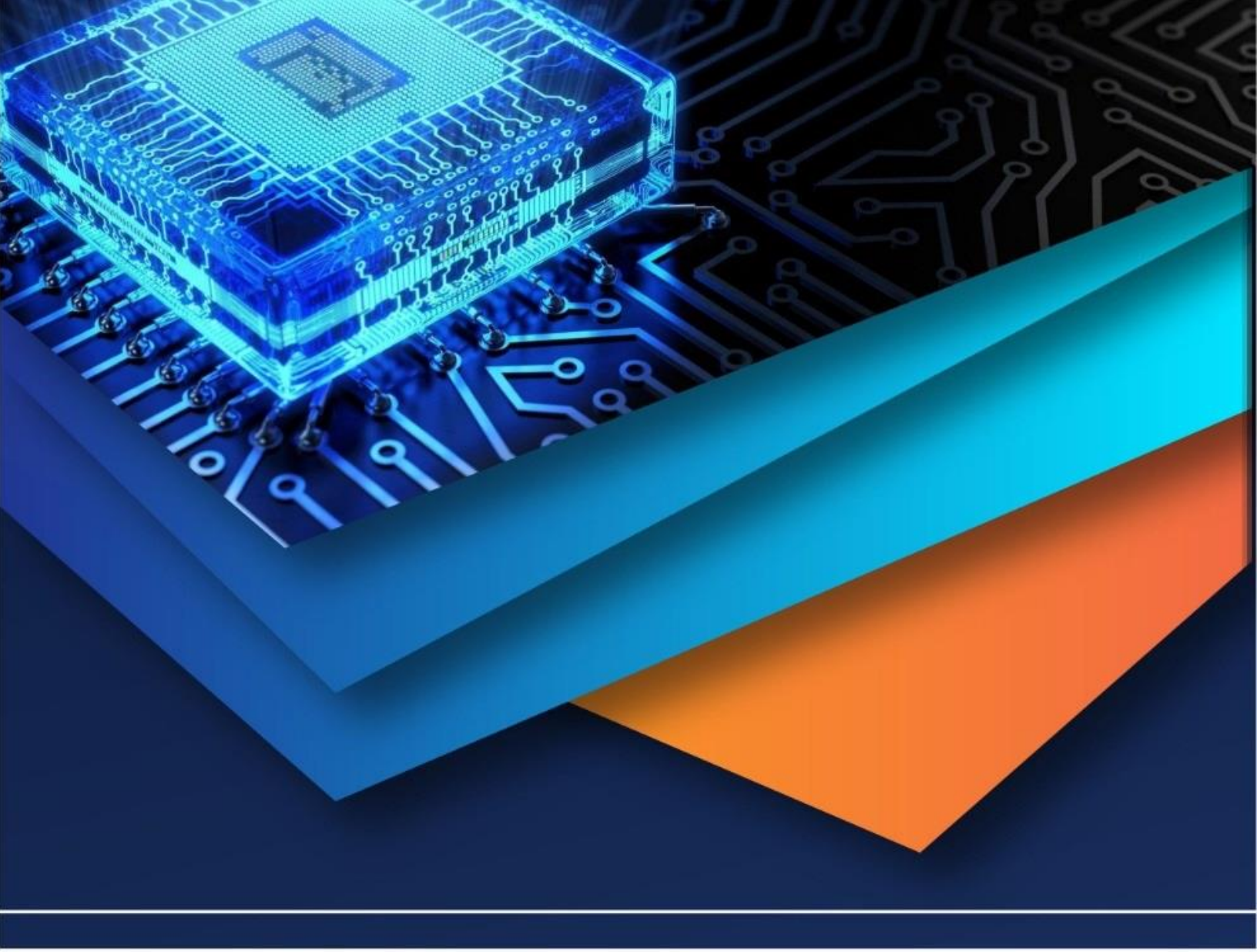
REFERENCES

[1] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. https://www.deeplearningbook.org
[2] Mitchell, T. M. (1997). Machine Learning. McGraw-Hill Education. https://www.cs.cmu.edu/~tom/mlbook.html
[3] Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer. https://link.springer.com/book/10.1007/978-0-387-45528-0
[4] Pedregosa, F., et al. (2011). Scikit-Learn: Machine Learning in Python. Journal of Machine Learning Research. https://scikit-learn.org
[5] Python Software Foundation. (2024). Python Documentation. https://docs.python.org/3/
[6] Flask Development Team. (2024). Flask Documentation. https://flask.palletsprojects.com
[7] SQLite Consortium. (2024). SQLite Documentation. https://www.sqlite.org/docs.html
[8] McKinney, W. (2022). Pandas Documentation. https://pandas.pydata.org/docs/
[9] Harris, C. R., et al. (2020). NumPy Documentation. https://numpy.org/doc/
[10] Bootstrap Team. (2024). Bootstrap Documentation. https://getbootstrap.com/docs/
[11] Chollet, F. (2021). Deep Learning with Python (2nd ed.). Manning Publications. https://www.manning.com/books/deep-learning-with-python-second-edition
[12] Géron, A. (2022). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (3rd ed.). O'Reilly Media. https://www.oreilly.com/library/view/hands-on-machine-learning/
[13] Brownlee, J. (2019). Machine Learning Algorithms for Classification. Machine Learning Mastery. https://machinelearningmastery.com
[14] IEEE Xplore Digital Library. (2024). Research Papers on Student Placement Prediction Using Machine Learning. https://ieeexplore.ieee.org
[15] Springer Nature. (2024). Research Publications on Artificial Intelligence and Machine Learning. https://link.springer.com
[16] Elsevier. (2024). ScienceDirect: Machine Learning and Educational Data Mining Research. https://www.sciencedirect.com
[17]

BIBLOGRAPHY

G.Manoj Kumar Working As An Assistant Professor In Masters Of Computer Applications (MCA) Sanketika Vidya Parishad Engineering College, Which Is Accredited With An ‘A’ Grade By NAAC, Affiliated To Andhra University. Completed Post Graduation In Andhra University College Of Engineering (AUCE). With Accredited By NAAC With His Areas Of Interest In Java,Python,DBMS,Web Technology, OS.

Bondada sravani Is Currently Pursuing Her Final Semester Of Master Of Computer Applications (MCA) At Sanketika Vidya Parishad Engineering College, Which Is Accredited With An ‘A’ Grade By NAAC, Affiliated To Andhra University, And Approved By AICTE. With A Keen Interest In Machine Learning And Artificial Intelligence, She Has Undertaken Her Postgraduate Project Titled “**MEDICINE INFORMATION AND ADVICE SYSTEM USING NLP AND DECISION TREE ALGORITHM**” The Project Has Been Successfully Carried Out Under The Guidance Of Mr. G. Manoj Kumar, Assistant Professor, SVPEC.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY