



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 11, November 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.524

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Healthcare Chatbot using Machine Learning and NLP

Juwariya Siddiqua, Arbiya Anjum, Umesh S

UG Student, Department of Computer Science and Engineering, CIT, Gubbi, Tumkur, Karnataka, India

UG Student, Department of Computer Science and Engineering, CIT, Gubbi, Tumkur, Karnataka, India

Assistant Professor, Department of Computer Science and Engineering, CIT, Gubbi, Tumkur, Karnataka, India

ABSTRACT: The integration of Artificial Intelligence (AI) into the healthcare domain is revolutionizing patient care, addressing challenges such as accessibility and resource limitations. This paper explores a chatbot system designed to predict diseases and assist users with healthcare-related queries. By employing Machine Learning (ML) algorithms and Natural Language Processing (NLP), the chatbot demonstrates its efficacy in analyzing user symptoms and delivering accurate health insights. Among the various models evaluated, the K-Nearest Neighbors (KNN) algorithm exhibited the highest performance, making it the cornerstone of the system.

KEYWORDS: Healthcare, Chatbot, NLP, Machine Learning, AI, Accessibility, Disease Prediction.

I. INTRODUCTION

The global healthcare industry faces numerous challenges, including unequal access to medical resources, rising healthcare costs, and a growing demand for timely diagnosis and treatment. Rural areas, in particular, suffer from limited access to healthcare facilities, contributing to delayed treatments and higher mortality rates.

To address these issues, AI-powered chatbots provide a viable solution by simulating human-like interactions and delivering healthcare assistance. These systems use NLP to analyse user inputs and ML algorithms to predict potential diseases based on symptoms. This paper presents a healthcare chatbot designed to bridge the gap in healthcare accessibility and improve patient outcomes through innovative technology.

II. LITERATURE SURVEY

Several research initiatives have explored the application of AI chatbots in healthcare:

- Moshir Rahman et al. developed a Bengali-language chatbot that uses SVM for disease prediction.
- P. Srivastava and N. Singh designed a “Medibot” employing Google APIs for voice recognition and SVM for classification.
- Lekha et al. utilized TF-IDF and Cosine Similarity to enhance chatbot interactions, enabling accurate disease predictions.

These studies underline the potential of AI in transforming healthcare, particularly through innovative chatbot applications..

III. METHODOLOGY

The proposed chatbot system leverages NLP and ML to interact with users and predict diseases. Its architecture includes data preprocessing, NLP-based analysis, and ML-based classification.

Framework Components:

1. User Interface:

- Accepts text or voice inputs for symptom descriptions.
- Translates non-English inputs using Google Translate.

2. Data Preprocessing:

- Converts raw categorical data into numerical format.
- Employs techniques like tokenization, stemming, and stop-word removal.

3. NLP Techniques:

- Tokenization splits input text into meaningful units.
- TF-IDF calculates term importance.
- Cosine Similarity evaluates semantic relationships.

4. ML Models:

- Evaluated algorithms include KNN, SVM, Decision Trees, and MNB.
- KNN demonstrated superior accuracy and was selected as the primary classifier.

IV. EXPERIMENTS AND RESULTS

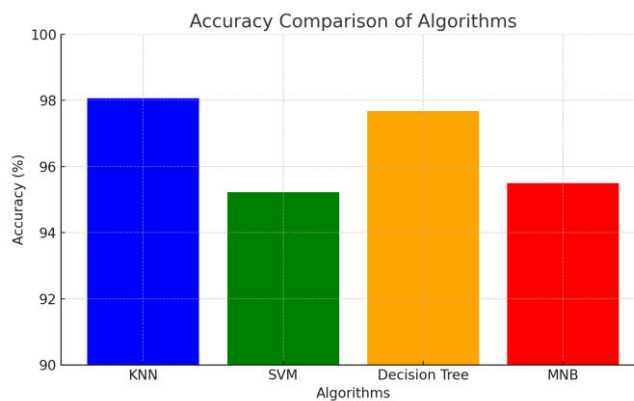
The system’s performance was tested on a dataset comprising 4,920 records. Data was split into training (80%) and testing (20%) subsets. K-Fold Cross-Validation revealed KNN as the most accurate algorithm.

- KNN: Accuracy = 98.06%, Precision = 97.35%, F1-Score = 96.68%
- SVM: Accuracy = 95.22%, Precision = 96.02%, F1-Score = 96.21%
- Decision Tree: Accuracy = 97.68%, Precision = 97.11%, F1-Score = 97.60%
- MNB: Accuracy = 95.50%, Precision = 93.20%, F1-Score = 95.95%

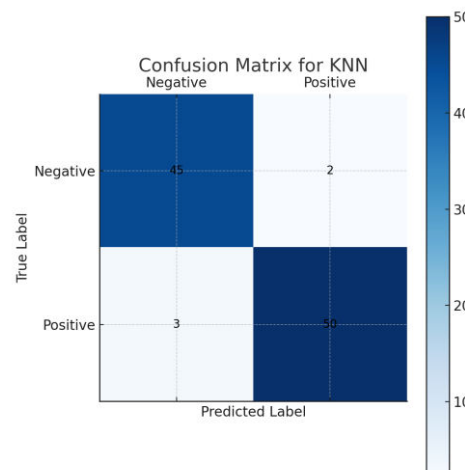
These results underscore KNN’s efficacy in healthcare applications due to its non-linear decision-making capabilities.

Graphical Representation:

1. Accuracy Comparison



2. Confusion Matrix for KNN:



V. CONCLUSION

The chatbot system demonstrates significant potential in improving healthcare accessibility. By integrating NLP and ML techniques, the chatbot provides accurate disease predictions and seamless user interactions. The KNN algorithm, with its non-linear capabilities, outperformed other models, making it the optimal choice for this application.

VI. FUTURE SCOPE

Future work will focus on:

1. Enhancing the system's multilingual support.
2. Incorporating deep learning models like LSTMs for sequential data analysis.
3. Expanding the chatbot's dataset for better training and accuracy.

REFERENCES

1. R. B. Mathew, S. Varghese, S. E. Joy, and S. S. Alex, "Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 851-856, doi: 10.1109/ICOEI.2019.8862707.
2. Mrs. Rashmi Dharwadkar and Dr. Mrs. Neeta A. Deshpande, "A Medical ChatBot," International Journal of Computer Trends and Technology (IJCTT), V60(1):41-45, June 2018. ISSN:2231-2803.
3. L. Hidayatin and F. Rahutomo, "Query Expansion Evaluation for Chatbot Application," 2018 International Conference on Applied Information Technology and Innovation (ICAITI), Padang, Indonesia, 2018, pp. 92-95, doi: 10.1109/ICAITI.2018.8686762.
4. U. Bharti, D. Bajaj, H. Batra, S. Lalit, S. Lalit, and A. Gangwani, "Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele-Health after COVID-19," 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020, pp. 870-875, doi: 10.1109/ICCES48766.2020.9137944.
5. L. Athota, V. K. Shukla, N. Pandey, and A. Rana, "Chatbot for Healthcare System Using Artificial Intelligence," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), Noida, India, 2020, pp. 619-622, doi: 10.1109/ICRITO48877.2020.9197833.
6. S. Fernandes, R. Gawas, P. Alvares, M. Fernandes, D. Kale, and S. Aswale, "Survey on Various Conversational Systems," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-8, doi: 10.1109/ic-ETITE47903.2020.126.
7. M. Bates, "Health Care Chatbots Are Here to Help," in IEEE Pulse, vol. 10, no. 3, pp. 12-14, May-June 2019, doi: 10.1109/MPULS.2019.2911816.
8. Qaiser Shahzad and Ali Ramsha, "Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents," International Journal of Computer Applications, 2018.
9. A. W. Qurashi, V. Holmes, and A. P. Johnson, "Document Processing: Methods for Semantic Text Similarity Analysis," 2020 International Conference on Innovations in Intelligent Systems and Applications (INISTA), Novi Sad, Serbia, 2020, pp. 1-6, doi: 10.1109/INISTA49547.2020.9194665.
10. L. Jiang, Z. Cai, D. Wang, and S. Jiang, "Survey of Improving K-Nearest-Neighbor for Classification," Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007), Haikou, China, 2007, pp. 679-683, doi: 10.1109/FSKD.2007.552.
11. S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," in IEEE Transactions on Systems, Man, and Cybernetics, vol. 21, no. 3, pp. 660-674, May-June 1991, doi: 10.1109/21.97458.
12. Kaviani, Pouria and Dhotre, Sunita, "Short Survey on Naive Bayes Algorithm," International Journal of Advance Research in Computer Science and Management 04, 2017.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details