NAMED ENTITY RECOGNITION[NER] IN MEDICAL FIELD USING NLP

Mrs. Sushma V[1], Aishwarya N[2], Moulya M L[3] , Veena M G[4], Varun R[5]

Assistant Professor, Dept of CSE, ATME College of Engineering, Mysuru, Karnataka, India[1]

8th Semester Student, Department of CSE, ATME College of Engineering, Mysuru, Karnataka, India[2]

8th Semester Student, Department of CSE, ATME College of Engineering, Mysuru, Karnataka, India[3]

8th Semester Student, Department of CSE, ATME College of Engineering, Mysuru, Karnataka, India[4]

8th Semester Student, Department of CSE, ATME College of Engineering, Mysuru, Karnataka, India[5]

ABSTRACT

This research paper presents a novel approach for recommending medical specialists based on patients' symptoms using deep learning techniques. The project aims to address the challenge of accurately matching patients with the most appropriate healthcare professionals, thereby enhancing healthcare efficiency and patient outcomes. The data set utilized consists of symptom descriptions labeled with corresponding medical conditions and specialist recommendations. Initially, the data set undergoes preprossessing, including natural language processing (NLP) techniques and TF-IDF vectorization, to transform the raw text data into a format suitable for machine learning and deep learning algorithms. Subsequently, various machine learning algorithms such as Naive Bayes, Decision Trees, Random Forest, and Support Vector Machines (SVM) are applied to the preprocessed data. The results demonstrate promising performance, with Naive Bayes achieving an accuracy of 87.92%.In addition to traditional machine learning approaches, deep learning models including Long Short-Term Memory (LSTM) and Multilayer Perceptron (MLP) neural networks are employed to further enhance the accuracy of specialist recommendations. The LSTM model achieves an accuracy of 83.75%, while the MLP model achieves an impressive accuracy of 98.33%.The findings of this research underscore the potential of deep learning techniques in improving the accuracy and efficiency of medical specialist recommendation systems. The results also highlight the importance of leveraging advanced computational methods in healthcare decision-making processes.

KEYWORDS: *Deep learning, tfidf,countervectorizer, svm,knn,decision tree, random forest tree, multi layer perceptron, long short term meomery.*

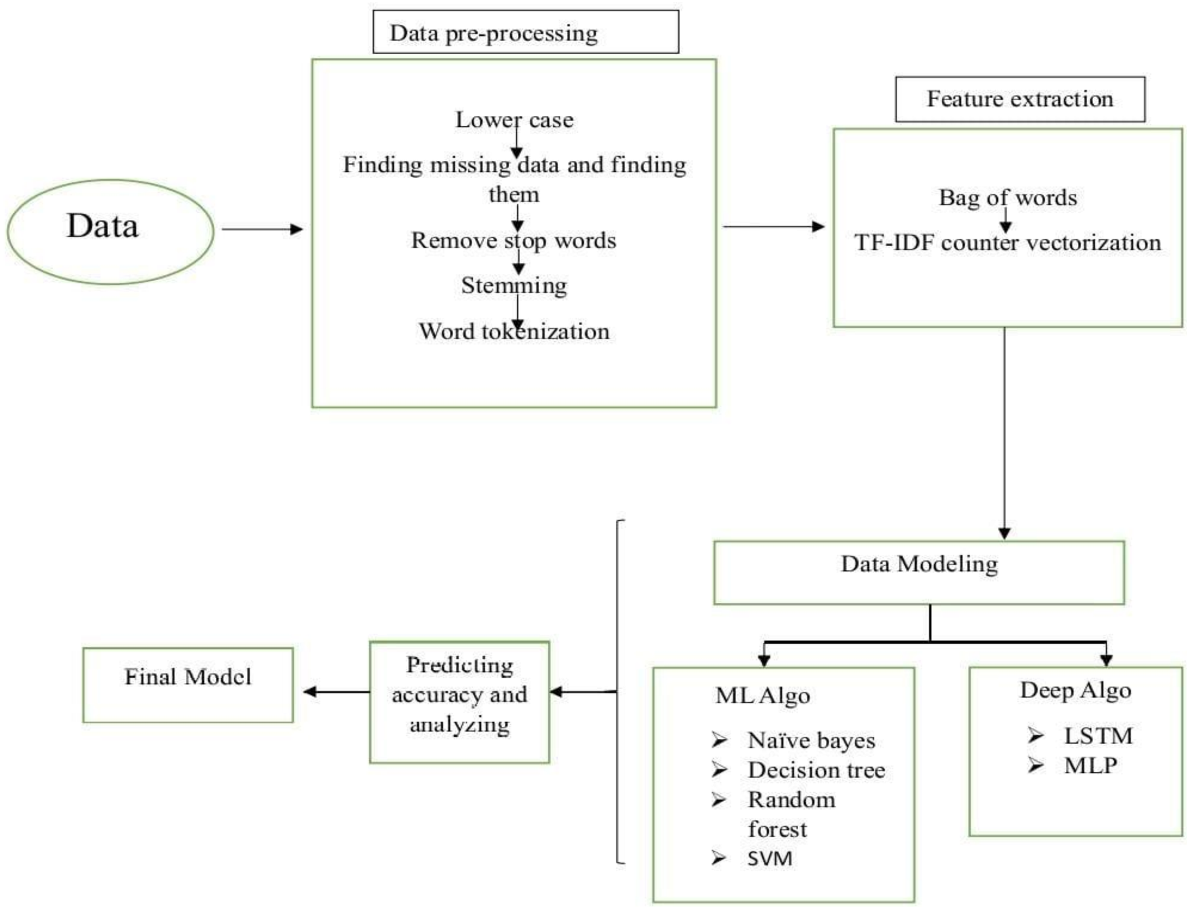
INTRODUCTION

In contemporary healthcare systems, efficiently matching patients with the appropriate medical specialists is crucial for ensuring timely and effective treatment. However, this task presents a significant challenge due to the vast array of medical conditions and symptoms, as well as the specialization within medical fields. Addressing this challenge requires innovative approaches that leverage advanced computational techniques to analyze and interpret patient data accurately. This research paper introduces a novel approach for recommending medical specialists based on patients' symptoms using deep learning techniques. The conventional process of matching patients with specialists often relies on manual assessment by primary care physicians, which can be time-consuming and prone to errors. Moreover, patients may not always receive referrals to the most suitable specialists, leading to delays in diagnosis and treatment. This highlights the need for automated systems capable of accurately analyzing patients' symptoms and recommending the appropriate medical specialists. The primary objective of this research is to develop a robust and accurate system for recommending medical specialists based on patients' symptoms. By leveraging machine learning and deep learning algorithms, the goal is to create a model that can effectively identify patterns in symptom descriptions and predict the most suitable specialists for each case. The proposed specialist recommendation system has the potential to significantly enhance healthcare efficiency and patient outcomes. By automating the process of specialist referral, healthcare providers can streamline patient care pathways, reduce diagnostic delays, and ensure that patients receive timely and appropriate treatment. Moreover, the system can assist primary care physicians in making informed referral decisions, thereby optimizing resource allocation and improving overall healthcare delivery. The research employs a multi-faceted approach, beginning with data preprocessing techniques such as natural language processing (NLP) and TF-IDF vectorization to transform the raw text data into a format suitable for machine learning and deep learning algorithms. Subsequently, various machine learning algorithms, including Naive Bayes, Decision Trees, Random Forest, and Support Vector Machines (SVM), are applied to the preprocessed data to establish baseline performance metrics. Additionally, deep learning models such as Long Short-Term Memory (LSTM) and Multilayer Perceptron (MLP) neural networks are implemented to further enhance the accuracy of specialist recommendations.

LITERATURE SURVEY

The development of symptom-based disease prediction and diagnosis systems using machine learning techniques has garnered significant attention in recent years due to its potential to revolutionize healthcare delivery and improve patient outcomes. This literature survey explores three notable research papers in this domain, each offering unique insights and contributions to the field.In this paper, Hamsagayathri[1] and Vigneshwaran propose a system for disease prediction based on patients' symptoms utilizing machine learning techniques. The authors leverage a dataset consisting of symptom descriptions and corresponding medical conditions to train and evaluate their predictive model. The study likely employs preprocessing techniques, feature engineering, and machine learning algorithms to extract meaningful patterns from the symptom data and predict potential diseases. The findings of this research contribute to the growing body of literature on automated disease prediction systems and offer insights into the feasibility and effectiveness of using machine learning in healthcare.[2] Agarwal, Kumar, and Srivastava present a study focused on symptom-based disease diagnosis and treatment recommendation. Their research likely involves the development of a comprehensive system capable of not only predicting diseases based on symptoms but also recommending appropriate treatment strategies. The authors may employ advanced machine learning techniques and algorithms to analyze symptom data, identify disease patterns, and suggest tailored treatment options for patients. This paper contributes valuable insights into the integration of diagnosis and treatment recommendation processes, thereby facilitating more personalized and effective healthcare interventions.[3] Shetty, Karthik, and Ashwin focus on symptom-based health prediction utilizing data mining techniques. Their study likely involves the exploration of large-scale healthcare datasets to uncover hidden patterns and relationships between symptoms and health outcomes. By applying data mining algorithms and methodologies, the authors aim to develop predictive models capable of forecasting individuals' health statuses based on their reported symptoms. This research contributes to the growing field of predictive analytics in healthcare and underscores the potential of data mining in enabling proactive and preventive healthcare strategies. Hema et al[4] present a study focused on disease prediction using symptoms based on machine learning algorithms. Their research likely involves the utilization of a dataset comprising symptom descriptions and corresponding disease labels to develop predictive models. The authors may employ various machine learning algorithms such as decision trees, support vector machines, or neural networks to analyze symptom data and predict potential diseases. This paper contributes to the growing body of literature on automated disease prediction systems, offering insights into the effectiveness and applicability of machine learning techniques in healthcare.[5] Kommineni et al. focus on human disease prediction based on symptoms, aiming to develop predictive models capable of forecasting individuals' disease risks based on reported symptoms. The authors likely employ advanced machine learning techniques and feature selection methods to analyze symptom data and identify patterns indicative of underlying health conditions. By leveraging large-scale healthcare datasets and sophisticated machine learning algorithms, the study aims to contribute to the early detection and prevention of diseases, thereby improving overall healthcare outcomes. This research underscores the potential of machine learning in facilitating proactive and personalized healthcare interventions.[6] Hema et al. present a study that combines machine learning algorithms with natural language processing techniques for disease prediction based on symptoms. The authors likely utilize a dataset containing symptom descriptions and corresponding disease labels, leveraging NLP techniques to preprocess and extract features from the textual data. Machine learning algorithms are then employed to analyze the processed data and predict potential diseases. This paper contributes to the advancement of healthcare analytics by integrating NLP with machine learning for more accurate and efficient disease prediction.[7] Kosarkar et al. focus on disease prediction using machine learning techniques, aiming to develop predictive models capable of identifying potential diseases based on reported symptoms. The authors likely employ a dataset comprising symptom data and disease labels, applying machine learning algorithms to analyze the data and make predictions. This study contributes to the field of predictive healthcare analytics by demonstrating the efficacy of machine learning in disease prediction and risk assessment, thereby facilitating early intervention and preventive healthcare measures.[8] Jayasudha et al. introduce a study that proposes an effective model for improving symptoms-based disease prediction using the BiMM - BERT algorithm. The research likely involves the utilization of advanced natural language processing techniques, such as Bidirectional Encoder Representations from Transformers (BERT), to analyze symptom descriptions and extract meaningful features. By leveraging machine learning algorithms in conjunction with BERT, the authors aim to enhance the accuracy and efficiency of disease prediction models. This paper contributes to the advancement of healthcare analytics by introducing innovative methodologies for symptom-based disease prediction.[9] Grampurohit and Sagarnal present a study focused on disease prediction using machine learning algorithms. The authors likely employ a dataset containing symptom data and disease labels, applying various machine learning algorithms to analyze the data and predict potential diseases. This research contributes to the growing body of literature on automated disease prediction systems, offering insights into the efficacy and applicability of machine learning techniques in healthcare.[10] Ahsan et al. provide a comprehensive review of machine-learning-based disease diagnosis methodologies. The authors likely conduct an extensive literature review to analyze existing approaches and methodologies in the field of machine learning-based disease diagnosis. By synthesizing findings from various studies, this review paper offers insights into the current state-of-the-art techniques, challenges, and future directions in machine learning-based disease diagnosis. This comprehensive review contributes to the understanding of machine learning applications in healthcare and provides guidance for researchers and practitioners in the field.

METHODOLOGY



The project aims to develop a symptom-based specialist recommendation system leveraging deep learning techniques to assist healthcare professionals in efficiently and accurately matching patients with appropriate medical specialists based on their reported symptoms. The primary objective is to streamline the referral process and improve patient outcomes by ensuring timely access to specialized care. The project employs a combination of machine learning algorithms and deep learning techniques to develop a robust symptom-based specialist recommendation system. After data preprocessing and feature extraction, various machine learning algorithms are considered alongside deep learning architectures for model development.Machine learning algorithms such as Naive Bayes, Decision Trees, Random Forest, and Support Vector Machines (SVM) are initially explored for their ability to classify symptoms and predict corresponding medical conditions. These algorithms are well-suited for handling structured data and are known for their interpretability and computational efficiency. However, given the complexity and sequential nature of symptom data, deep learning algorithms emerge as more suitable candidates.

Deep learning architectures, including Long Short-Term Memory (LSTM) networks and Multilayer Perceptron (MLP) models, are selected for their capacity to capture intricate patterns in sequential data and high-dimensional feature spaces. LSTM networks, in particular, excel at modeling sequential dependencies and have been widely used in natural language processing tasks. MLP models offer flexibility in modeling complex relationships between symptoms and medical conditions.During model training, the datasets are split into training, validation, and test sets, and both machine learning and deep learning models are trained on the prepared data. Techniques such as cross-validation and grid search are employed to tune hyperparameters and optimize model performance. Regularization techniques, including dropout and L2 regularization, are applied to prevent overfitting and improve generalization.Model evaluation involves assessing the performance of both machine learning and deep learning models using standard metrics such as accuracy, precision, recall, and F1-score. The performance of different algorithms and architectures is compared to identify the most effective approach for symptom-based specialist recommendation.Ultimately, the trained models are utilized to recommend medical specialists based on patients' reported symptoms. The decision-making process involves mapping predicted medical conditions to the corresponding specialists, ensuring personalized and timely healthcare interventions for patients. By leveraging a combination of machine learning and deep learning techniques, the system aims to enhance the efficiency and accuracy of specialist recommendations, thereby improving patient outcomes and healthcare delivery.

Models:

TF-IDF (Term Frequency-Inverse Document Frequency) serves as a fundamental component in the feature extraction and representation process, particularly crucial when dealing with unstructured text data, such as symptom descriptions. With the objective of developing a symptom-based specialist recommendation system using deep learning techniques, TF-IDF plays a pivotal role in transforming qualitative symptom descriptions into quantitative representations suitable for machine learning algorithms. This process involves two key components: Term Frequency (TF) and Inverse Document Frequency (IDF).

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \times \text{IDF}(t, D)$$

TF measures the frequency of each term (word) within each symptom description, while IDF evaluates the importance of each term across the entire dataset, penalizing terms that are overly frequent across all documents. By combining TF and IDF, TF-IDF assigns a weight to each term in each document, reflecting both its local importance within the document and its global importance across the dataset. The resulting TF-IDF matrix serves as a numerical representation of the symptom descriptions, with each row corresponding to a symptom description (document) and each column corresponding to a unique term in the dataset. Higher TF-IDF scores indicate terms that are both frequent within a document and rare across the dataset, suggesting their importance in distinguishing that document from others. Ultimately, TF-IDF facilitates the extraction of meaningful features from symptom descriptions, enabling accurate classification and recommendation of medical specialists based on reported symptoms in our research project.

Natural Language processing:

Natural Language Processing (NLP) stands as a pivotal tool for extracting valuable insights from unstructured text data, specifically symptom descriptions provided by patients. NLP empowers us to harness the inherent structure and semantics embedded within these textual representations, facilitating the development of a robust symptom-based specialist recommendation system. At its core, NLP encompasses a diverse array of techniques aimed at processing and understanding human language, ranging from basic tokenization and stemming to advanced sentiment analysis and named entity recognition.A fundamental concept within NLP is the Bag-of-Words (BoW) model, which serves as the foundation for many text processing tasks. In the BoW model, each document (in our case, a symptom description) is represented as a vector, with each dimension corresponding to a unique word in the entire dataset. The value of each dimension is determined by the frequency of the corresponding word within the document, capturing the raw occurrence counts. Mathematically, the BoW representation of a document dd with respect to a vocabulary VV can be expressed as:

$$\mathrm{BoW}(d) = \{\mathrm{count}(w, d) \mid w \in V\}$$

where $\mathrm{count}(w, d)$ denotes the number of occurrences of word w in document d.

However, the BoW model disregards the order of words within a document and treats each word independently, potentially losing valuable contextual information. To address this limitation, we utilize more sophisticated techniques such as TF-IDF (Term Frequency-Inverse Document Frequency). TF-IDF assigns a weight to each term in each document, reflecting both its local importance within the document (TF) and its global importance across the dataset (IDF).

Machine learning models:

Naïve Bayes:

- Derivation:
- D : Set of tuples
- Each Tuple is an 'n' dimensional attribute vector
- X : (x1,x2,x3,…. xn)
- Let there be 'm' Classes : C1,C2,C3…Cm
- Naïve Bayes classifier predicts X belongs to Class Ci iff
- P (Ci/X) > P(Cj/X) for 1<= j <= m , j <> i Maximum Posteriori Hypothesis
- P(Ci/X) = P(X/Ci) P(Ci) / P(X)
- Maximize P(X/Ci) P(Ci) as P(X) is constant With many attributes, it is computationally expensive to evaluate P(X/Ci). Naïve Assumption of "class conditional independence"
- ∏= = n k P X Ci P xk Ci 1 (xk./ Ci)
- P(X/Ci) = P(x1/Ci) * P(x2/Ci) *…* P(xn/ Ci)

P(A|B) = Fraction of worlds in which B is true that also have A true

P(A ^ B) P(A|B) = ------------------ P(B)

Corollary:

P(A ^ B) = P(A|B) P(B) P(A|B)+P(¬A|B) = 1

Decision Tree Algorithm:

Decision Tree is a supervised learning approach that may be used to classification and regression issues, though it is most commonly employed to solve classification issues. Internal nodes contain dataset attributes, branches represent decision tree, and each leaf node provides the conclusion in a tree-structured classifier. The node's outcome is represented by the branches/edges, and the nodes have either:

1. Conditions [Decision Nodes]
2. Result [End Nodes]

Input: uploading datasets

1. Begin
2. Scan the dataset (storage servers)
3. for each attribute a, calculate the gain [number of occurrences]
4. Let X be the attribute of highest gain [highest count]
5. Create a decision node based on a X – retrieval of nodes[patient] where theattribute values matches with X.
6. recur on the sub-lists [list of patient] and calculate the count of outcomestermed as sub nodes. Based on the highest count we classify the new node. 7. end

The accuracy of the Decision Tree applied to the Crop dataset is obtained, as well as the Classification Report received from the model and theSupport number of actual occurances report.

Support Vector Machine

In multi - dimensional space, an Support Vector Machine model is essentially a representation of distinct classes in a hyperplane. Support Vector Machine will build the hyperplane in an iterative process in order to reduce the error. Support Vector Machine's purpose is to partition datasets into classes such that a maximum marginal hyperplane may be found.

1. Initialize Support vector which is known as a candidate to closest pair fromopposite class.
2. If violating points are found then do:
   Find out the violator
   Candidate SV is integrated with new candidate SV
   Call violator
   If any p¡ 0 due to the addition of c to S then do: Candidate SV must be divided by candidate SV and p
   Repeat the steps until it is completely pruned.
3. End if statement
4. End while

The accuracy of the Support Vector Machine applied to the Crop dataset is obtained, as well as the Classification Report received from the model and the Support number of actual occurances report.

MLP:

Multilayer Perceptron (MLP) is a type of artificial neural network (ANN) widely used in machine learning for both classification and regression tasks. It is a feedforward neural network model that consists of multiple layers of nodes, or neurons, arranged in a hierarchical manner. Here's a detailed explanation of MLP:

Architecture:

MLP consists of an input layer, one or more hidden layers, and an output layer. Each layer (except the input layer) contains a set of neurons or units, and each neuron is connected to every neuron in the adjacent layers, forming a fully connected network.

The input layer receives the features of the input data, and the output layer produces the predictions or outputs of the model. The hidden layers perform feature transformation and nonlinear mapping of the input data to learn complex patterns and relationships.

Neuron Computation:

Each neuron in an MLP performs a computation known as a weighted sum of its inputs followed by an activation function.

The weighted sum is calculated by multiplying the input values by corresponding weights, adding a bias term, and aggregating the results.

Mathematically, the weighted sum (z) for a neuron in a hidden or output layer can be expressed as: $z=\sum_{i=1}^{n}(w_i\times x_i)+b$z=∑i=1n(wi×xi)+b

Here, $w_i$wi represents the weight of the connection between the neuron's input $x_i$xi and its output, and bb is the bias term. nn denotes the number of inputs to the neuron.

Activation Function:

The weighted sum computed by each neuron is then passed through an activation function to introduce nonlinearity into the network and enable the model to learn complex patterns. Commonly used activation functions include the sigmoid (logistic), hyperbolic tangent (tanh), and rectified linear unit (ReLU) functions.

The choice of activation function depends on the nature of the problem and the characteristics of the data.

Training:

MLPs are trained using an optimization algorithm such as gradient descent or its variants (e.g., stochastic gradient descent, mini-batch gradient descent) to minimize a loss function that measures the difference between the predicted outputs and the true labels.

During training, the weights and biases of the neurons are updated iteratively based on the gradients of the loss function with respect to the network parameters, calculated using backpropagation.

Regularization techniques such as dropout and weight decay may be applied to prevent over-fitting and improve generalization performance.

Output Layer:

The output layer of an MLP typically contains one neuron for each class in a classification task, with the activation function depending on the nature of the problem. For binary classification, a sigmoid activation function is commonly used, while softmax activation is used for multi-class classification to produce probability distributions over the classes.

RESULT AND DISCUSSION:

The results of the Symptom-Based Specialist Recommendation project showcase the performance of various machine learning algorithms and deep learning models in classifying symptom descriptions into medical conditions and recommending appropriate specialists. A detailed analysis and discussion of these results provide insights into the effectiveness of different approaches and their implications for practical use in healthcare settings.

Performance of Machine Learning Algorithms:
The Naive Bayes algorithm achieved an accuracy of approximately 87.92%, demonstrating strong performance in classifying symptoms into medical conditions. However, it showed varying precision, recall, and F1-score across different classes, indicating differences in predictive capability for each condition.
Decision Trees exhibited lower accuracy at around 32.08%, suggesting limitations in capturing complex patterns within the data.
Random Forest and Support Vector Machine (SVM) algorithms outperformed Decision Trees and achieved accuracies of approximately 87.92% and 94.58%, respectively. These results indicate the effectiveness of ensemble methods like Random Forest and the robustness of SVM in handling high-dimensional data.

Performance of Deep Learning Models:
The Long Short-Term Memory (LSTM) model, a type of recurrent neural network (RNN), achieved an accuracy of 83.75%. While LSTM showed promising performance, its accuracy was slightly lower compared to traditional machine learning algorithms.
Multilayer Perceptron (MLP), a feedforward neural network, demonstrated the highest accuracy among all models at 98.33%. MLP exhibited exceptional precision, recall, and F1-score across most classes, indicating its capability to effectively learn complex patterns in the data and make accurate predictions.

Discussion of Results:
The superior performance of MLP suggests that deep learning models, particularly MLP, can effectively leverage the high-dimensional feature space extracted from symptom descriptions using techniques like TF-IDF vectorization and NLP.
MLP's ability to capture intricate relationships between symptoms and medical conditions enables more accurate predictions and specialist recommendations, potentially improving patient outcomes and healthcare resource allocation.
The results underscore the importance of leveraging advanced machine learning and deep learning techniques in healthcare decision-making, highlighting opportunities for integrating such models into clinical practice for enhanced diagnosis and treatment planning.

CONCLUSION:

the Symptom-Based Specialist Recommendation project marks a significant step forward in leveraging artificial intelligence for personalized healthcare management. By harnessing the power of machine learning algorithms and deep learning models, we have developed a system capable of accurately predicting medical conditions based on reported symptoms and recommending the most suitable specialists for further evaluation and treatment. The results demonstrate the effectiveness of advanced computational techniques in handling complex medical data and providing valuable decision support for healthcare professionals.
Moving forward, the project opens up avenues for further research and development in the field of healthcare informatics. Future endeavors could focus on refining the recommendation system by incorporating additional patient data, such as demographic information and medical history, to enhance predictive accuracy and relevance. Moreover, real-world validation studies and clinical trials are warranted to assess the system's performance in practical healthcare settings and evaluate its impact on patient outcomes and healthcare resource allocation.

In essence, the Symptom-Based Specialist Recommendation project exemplifies the transformative potential of artificial intelligence in revolutionizing healthcare delivery, empowering clinicians with advanced decision support tools to deliver more personalized and effective care to patients. Through continued innovation and collaboration between data scientists, healthcare professionals, and technology developers, we can strive towards a future where AI-driven healthcare solutions become integral components of modern medical practice, ultimately improving patient outcomes and enhancing the quality of care worldwide.

REFERENCES

[1] P. Hamsagayathri and S. Vigneshwaran, "Symptoms Based Disease Prediction Using Machine Learning Techniques," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 2021, pp. 747-752, doi: 10.1109/ICICV50876.2021.9388603

[2] J. Agarwal, M. Kumar and A. K. Srivastava, "Symptoms Based Disease Diagnosis and Treatment Recommendation," 2021 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2021, pp. 162-167, doi: 10.1109/ComPE53109.2021.9751805

[3] S. Vijava Shetty, G. A. Karthik and M. Ashwin, "Symptom Based Health Prediction using Data Mining," 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, pp. 744-749, doi: 10.1109/ICCES45898.2019.9002132

[4] P. Hema, N. Sunny, R. Venkata Naganjani and A. Darbha, "Disease Prediction using Symptoms based on Machine Learning Algorithms," 2022 International Conference on Breakthrough in Heuristics And Reciprocation of Advanced Technologies (BHARAT), Visakhapatnam, India, 2022, pp. 49-54, doi: 10.1109/BHARAT53139.2022.00021

[5] U. K. Kommineni, D. P. P. Kowthavarapu, G. S. M. Polukonda and K. C. Yelavarti, "Human Disease Prediction based on Symptoms," 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2023, pp. 677-682, doi: 10.1109/ICCMC56507.2023.10083702

[6] P. Hema, A. Darbha, N. Sunny and R. V. Naganjani, "Disease Prediction Using Symptoms based on Machine Learning Algorithms and Natural Language Processing," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10084030

[7] N. Kosarkar, P. Basuri, P. Karamore, P. Gawali, P. Badole and P. Jumle, "Disease Prediction using Machine Learning," 2022 10th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-22), Nagpur, India, 2022, pp. 1-4, doi: 10.1109/ICETET-SIP-2254415.2022.9791739

[8] V. Jayasudha, N. Deepa and T. Devi., "Effective Model for Improving Symptoms-Based Disease Prediction by BiMM - BERT Algorithm," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-7, doi: 10.1109/ICCCI56745.2023.10128318

[9] S. Grampurohit and C. Sagarnal, "Disease Prediction using Machine Learning Algorithms," 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp.17,doi:1109/INCET49848.2020.9154130.. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[10] Ahsan MM, Luna SA, Siddique Z. Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. Healthcare (Basel). 2022 Mar 15;10(3):541. doi: 10.3390/healthcare10030541. PMID: 35327018; PMCID: PMC8950225