

AN INTUITIVE APPROACH ON LSTM-HYBRID MODEL FOR MULTI-DISEASE PREDICTION AND CHATBOT RECOMMENDATION SYSTEM

¹Savithri, ²Dr. V. Goutham

¹Research Scholar, Department of Computer Science, Bharatiya Engineering Science & Technology Innovation University (BESTIU), Gownivaripalli, Gorantla, Andhra Pradesh
and

Asst. Professor, St. Francis College for Women, Begumpet, Hyderabad,

²Professor in Computer Science and Engineering Department,
Nalla Narasimha Reddy Group of Institutions, Hyderabad, Telangana

ABSTRACT:

In the realm of healthcare, accurate disease classification is essential for effective diagnostics and treatment planning. Traditional methods often struggle with the complexity and variability of medical data. This paper introduces a novel approach by combining user query responses with a hybrid Long Short-Term Memory (LSTM) model that integrates Convolutional Neural Networks (CNNs) and a chatbot-based recommendation system. User-generated queries are processed using natural language processing (NLP) techniques and then analyzed by the hybrid LSTM-CNN model, which captures both sequential and contextual dependencies to enhance classification accuracy. The CNN enhances feature extraction, while the LSTM manages long-term dependencies, improving the precision and recall of the disease classification. The chatbot not only interacts with users to provide tailored recommendations and information but also aids in gathering contextual data that feeds into the classification model. Evaluation on a diverse dataset shows that this combined approach significantly outperforms traditional methods, offering a scalable, context-aware solution for multi-disease classification and demonstrating substantial potential for advancing healthcare outcomes through sophisticated machine learning techniques and interactive systems.

INTRODUCTION:

Accurate disease classification is fundamental to effective healthcare, guiding diagnostic and treatment decisions. Traditional methods, while critical, often struggle with the complexity and variability of medical data. With the advent of machine learning (ML), artificial intelligence (AI), and deep learning (DL), novel approaches have emerged to address these challenges, offering the potential for enhanced diagnostic accuracy and efficiency. This paper introduces a hybrid Long Short-Term Memory (LSTM) model, integrating LSTM with convolutional neural networks (CNNs), to improve multi-disease classification using data from the UCI Machine Learning Repository. This approach aims to leverage both the sequential and feature extraction capabilities of these advanced techniques to provide more accurate disease diagnoses.

Current Techniques in Disease Classification

Traditional disease classification relies heavily on clinical expertise and manual interpretation of symptoms, medical history, and diagnostic tests. These methods are essential but often limited by their inability to handle complex, high-dimensional data. Machine learning models such as decision trees and random forests offer more sophisticated classification by creating a series of decision rules or aggregating multiple decision trees. Support Vector Machines

(SVMs) are also used, finding optimal hyperplanes to separate classes, but they can struggle with noisy or overlapping data. In contrast, artificial intelligence approaches like rule-based systems apply predefined rules but lack adaptability, while basic neural networks model non-linear relationships but may not handle sequential data effectively.

Advancements with Deep Learning Models

Deep learning has significantly advanced disease classification capabilities. Convolutional Neural Networks (CNNs) excel at extracting features from complex data, particularly in image-based tasks, but are less effective for sequential data. Recurrent Neural Networks (RNNs), including their advanced variant LSTMs, are designed to manage sequential data and long-term dependencies, making them suitable for tasks involving temporal relationships. However, standard RNNs can struggle with very long sequences due to issues like vanishing gradients. Combining CNNs and LSTMs in a hybrid model aims to overcome these limitations by integrating feature extraction and sequential learning capabilities.

Hybrid LSTM Model for Disease Classification

Our proposed hybrid LSTM model integrates CNNs and LSTMs to address the challenges of multi-disease classification effectively. Initially, user query responses, rich in contextual and symptomatic information, are processed and structured using natural language processing (NLP). The CNN component extracts relevant features from these queries, identifying local patterns and important features. The LSTM component then captures the temporal dependencies and contextual nuances within the sequential data. By combining these strengths, the hybrid model enhances the ability to make precise classifications based on complex and dynamic input data. This integration allows the model to leverage both local feature patterns and long-term temporal information, resulting in a more robust and accurate classification system. The UCI Machine Learning Repository offers a diverse range of datasets suitable for evaluating disease classification models.

Concept:

Applying our hybrid LSTM model to these datasets demonstrates its effectiveness in improving classification accuracy. The model outperforms traditional methods and standalone ML or DL approaches by effectively handling both sequential and complex features. This hybrid approach addresses the limitations of existing techniques, providing a scalable and adaptable framework for healthcare analytics. The success of our hybrid LSTM model highlights its potential in advancing disease classification methodologies, offering a promising solution to the challenges faced in healthcare diagnostics and underscoring the value of integrating advanced machine learning techniques for improved healthcare outcomes.

Problem statement:

Accurate disease classification is a cornerstone of effective healthcare, yet traditional diagnostic methods frequently encounter challenges when faced with the complexity and variability of medical data. Conventional approaches rely heavily on manual interpretation of clinical symptoms, medical history, and diagnostic tests, which are essential but often insufficient for handling high-dimensional and complex data. Traditional methods such as decision trees, random forests, and support vector machines (SVMs) offer certain advantages but struggle with noisy data, overlapping classes, and the inability to capture intricate temporal or contextual relationships in user queries. Moreover, basic neural networks and rule-based AI systems, while useful, either lack the ability to handle sequential dependencies or fail to adapt dynamically to complex data patterns.

In response to these limitations, there is a pressing need for more sophisticated models that can effectively manage the complexity and temporal aspects of medical data. The integration of advanced machine learning and deep learning techniques promises to address these issues

by enhancing feature extraction and classification accuracy. This paper introduces a hybrid Long Short-Term Memory (LSTM) model combined with Convolutional Neural Networks (CNNs) to overcome the shortcomings of traditional methods. By utilizing user query responses, which provide rich contextual and symptomatic information, our approach aims to deliver a more robust and precise disease classification system. This method addresses the need for a scalable and adaptable framework that can dynamically handle complex and variable medical data.

Objectives

1. **Develop and validate a hybrid LSTM-CNN model** for multi-disease classification that leverages user query responses to improve diagnostic accuracy.
2. **Demonstrate the effectiveness of the hybrid model** on datasets from the UCI Machine Learning Repository, showcasing its ability to handle sequential and complex features better than traditional methods.

Overview.

This paper presents a novel approach to multi-disease classification by introducing a hybrid Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) model that utilizes user query responses. The methodology involves processing these responses using natural language processing (NLP) to structure the data, which is then analysed by the hybrid model to capture both temporal dependencies and local feature patterns. The paper evaluates this model using diverse datasets from the UCI Machine Learning Repository, demonstrating its superior performance in classification accuracy compared to traditional methods and standalone machine learning or deep learning models. The results highlight the model's effectiveness in handling complex, sequential, and contextual data, offering a promising solution for more precise and adaptable disease classification in healthcare.

RELATED WORK:

Sánchez-Adame et al. [1] establish essential heuristics for evaluating chatbot effectiveness, providing a structured approach that can enhance usability and performance, though empirical validation in diverse contexts remains necessary. Kuhail et al. [2] explore the dynamics of personality alignment between chatbots and users, revealing that higher personality congruence leads to improved user satisfaction, thereby advocating for a psychological perspective in chatbot design, while calling for exploration of various personality types across different demographics. Medeiros et al. [3] address the role of chatbots in mental health applications, demonstrating that supportive chatbots can significantly reduce users' perceived stress levels, although future studies should investigate the long-term effects and specific supportive strategies. Benaddi et al. [4] provide a systematic review of chatbots in tourism, classifying different types and assessing their impact, while emphasizing the need for empirical studies to explore real-world implementation challenges. Daniel et al. [5] introduce the Xatkit framework for low-code chatbot development, empowering non-technical users; however, further exploration of scalability and performance is warranted. Ren et al. [6] highlight the utility of the SOCIO chatbot in UML modeling, presenting a practical example in educational settings but suggesting additional comparisons with other modeling tools. Park et al. [7] investigate prompt engineering's role in enhancing generative AI performance, offering valuable strategies for developing effective conversational agents, while also identifying the need for broader applications across diverse scenarios. Santos et al. [8] propose a conversation-driven approach for managing chatbot interactions, enhancing user engagement, and knowledge management, but future research should consider user behavior variations in different contexts. Ren et al. [9] emphasize the importance of usability evaluation in chatbot research,

presenting findings from a secondary study that validates usability metrics, though longitudinal studies are needed to track usability over time. Gunnam et al. [10] assess cloud-based heterogeneous chatbot systems, identifying key performance metrics and challenges, while advocating for detailed analysis of deployment scenarios. Wu et al. [11] discuss the integration of hybrid model chatbots in K-12 education, outlining advantages and constraints, and suggesting further exploration of user experiences in various educational settings. Abdellatif et al. [12] provide a comparative analysis of NLP platforms for chatbot development in software engineering, offering practical references while calling for comprehensive benchmarks across different languages. Al-Emran et al. [13] examine factors influencing AI chatbot adoption for knowledge sharing, identifying determinants like ease of use and technological trust, while suggesting the need for sector-specific challenge investigations. Alsharhan et al. [14] present a systematic review of chatbot adoption, identifying multiple influencing factors and setting a foundation for future research, yet highlighting the need for studies on the long-term impacts of chatbot integration. Lastly, Ahmed et al. [15] tackle few-shot learning challenges in transformer-based chatbots, proposing solutions to enhance adaptability, while noting that further research is required to explore these learning effects across diverse datasets. This comprehensive overview illustrates the multifaceted landscape of chatbot research, revealing significant contributions while underscoring the necessity for ongoing exploration and empirical validation across various domains.

The integration of advanced machine learning and deep learning techniques in medical and agricultural diagnostics has seen significant progress in recent years. Hameed et al. (2019) [16] address the critical challenge of classifying multiple skin lesions, which is essential for accurate dermatological diagnoses. The authors propose an intelligent computer-aided scheme that employs advanced machine learning methods, integrating image processing with classification algorithms. This approach not only enhances diagnostic accuracy but also improves efficiency in clinical settings, demonstrating the potential of technology in addressing real-world medical challenges.

Similarly, Zhang et al. (2021) [17] tackle the issue of thorax disease classification by introducing a novel part-aware mask-guided attention mechanism. Their study emphasizes the importance of focusing on relevant regions within medical images, which significantly enhances the classification performance for thoracic diseases. This targeted approach improves the detection of conditions that may be missed by traditional methods, thus contributing to more reliable medical image analysis. The integration of attention mechanisms reflects a growing trend in leveraging artificial intelligence to refine diagnostic processes across various medical fields.

In the realm of respiratory health, Hong et al. (2021) [18] present a CNN-based method for the multi-class classification of lung diseases from chest X-ray images. Their optimized CNN architecture effectively categorizes different lung conditions, addressing the need for precise differentiation in complex imaging tasks. This advancement is crucial for ensuring effective treatment planning and improving patient outcomes. Ouda et al. (2022) [19] expand on this theme by introducing a deep learning-based method for diagnosing multiple ocular diseases using fundus images. Their multi-label classification approach allows for the simultaneous diagnosis of various conditions, significantly enhancing both the efficiency and accuracy of ocular disease identification.

The application of machine learning techniques extends beyond medical imaging to agricultural diagnostics. Zuo et al. (2022) [20] propose a multi-granularity feature aggregation method combined with self-attention for fine-grained crop disease classification. By focusing on critical regions within crop images, this method enhances disease detection

accuracy, which is essential for effective pest management and crop health monitoring. Furthermore, Pham et al. (2023) [21] investigate the classification of ultra-widefield fundus images for multi-label diagnosis of retinal diseases, improving the management of complex retinal images through a discriminative-region classification method. Similarly, Li et al. (2023) [22] employ the Swin Transformer V2 model to classify multiple fundus diseases using retinal optical coherence tomography (OCT) images, highlighting the model's superior feature extraction capabilities for more accurate diagnoses.

Finally, the studies by Peng et al. (2024) [23] and Chen et al. (2024) [24] contribute to the understanding of retinal and spinal health, respectively. Peng et al. introduce a multi-scale-denoising residual convolutional network aimed at addressing noise and variability in OCT images, thus enhancing the reliability of retinal disease detection. Meanwhile, Chen et al. explore deep learning techniques for intelligent diagnosis of lumbar diseases by utilizing multi-angle views of intervertebral discs, capturing complex features for a comprehensive assessment of spinal health. Xie and Ye (2024) [25] further enhance agricultural diagnostics by presenting a multi-scale and multi-factor Vision Transformer model, which improves the accuracy of pest and disease detection in complex agricultural environments. Collectively, these studies reflect a substantial advancement in applying machine learning models to enhance diagnostic processes across dermatology, radiology, ophthalmology, and agriculture, showcasing the transformative potential of technology in these critical areas.

METHODOLOGY:

The process of developing and validating a hybrid Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) model for multi-disease classification, complemented by a chatbot recommendation system, involves several critical phases, each crucial for ensuring the model's accuracy and effectiveness. The workflow is organized into five key steps: data collection and pre-processing, feature extraction using CNNs, temporal and contextual analysis with LSTMs, integration and classification, and model evaluation and validation.

Data Collection and Pre-processing: The process starts with gathering user query responses, which provide rich contextual and symptomatic information pertinent to disease classification. These queries undergo preprocessing with natural language processing (NLP) techniques, including tokenization, stemming, and lemmatization. This preprocessing transforms raw text into structured numerical vectors representing the features of each query. Additionally, user interactions with a chatbot system contribute valuable context and data, further enriching the dataset. This structured data serves as input for subsequent stages, ensuring it is properly formatted for analysis.

Feature Extraction Using CNNs: The pre-processed data is then input into a Convolutional Neural Network (CNN), which extracts meaningful features from the structured query data. CNNs apply convolutional layers to detect local patterns and hierarchies, such as keywords or disease-indicative phrases. Pooling layers reduce the dimensionality of these feature maps, retaining significant features while discarding less relevant information. This process enhances the efficiency and effectiveness of feature extraction.

Temporal and Contextual Analysis with LSTMs: The CNN-extracted features are processed by Long Short-Term Memory (LSTM) networks. LSTMs, known for their ability to handle sequential data and capture long-term dependencies, analyze the temporal aspects of the data. The model's memory cells and gating mechanisms retain crucial contextual details from user queries, which are essential for accurate disease classification. This sequential analysis enables the model to understand complex patterns and relationships, improving the differentiation between diseases.

Integration and Classification: Features from both CNN and LSTM components are integrated to form a comprehensive representation of the input data. This integration combines the spatial features identified by the CNN with the temporal patterns captured by the LSTM. Fully connected layers merge these features and perform the final classification, outputting probabilities for various diseases. This combined approach leverages the strengths of both CNNs and LSTMs, resulting in a robust classification system.

Model Evaluation and Validation: The final phase involves evaluating and validating the hybrid LSTM-CNN model. The model is trained on a training dataset and tested on a separate testing dataset. Performance metrics such as precision, recall, F1-score, and accuracy are calculated to assess classification performance. Comparative analysis benchmarks the hybrid model against traditional classification methods and standalone machine learning or deep learning approaches. Additionally, the chatbot system's effectiveness in gathering relevant user input and providing contextual recommendations is evaluated to ensure it complements the classification model effectively. This comprehensive evaluation confirms the model's ability to improve classification accuracy and handle complex, sequential data, demonstrating its suitability for real-world disease diagnosis applications.

ALGORITHM:

Input: Let D be the dataset with X as features and y as labels, $population_size$, max_iter for HECA model, for LSTM ($Hidden_layers$, $input_shape$)

Output: Updated Ensemble Model ($M_updated$), Final Performance Metrics (Accuracy, Precision, Recall, F1 score), and Optimal Configuration (Optimal parameters).

Procedure:

1. **Initialization:** Fetch dataset and optimization parameters, configure LSTM.
2. **Ensemble Learning with LSTM:** Train LSTM networks, generate predictions.
3. **Performance Evaluation:** Assess accuracy, precision, recall, and F1 score.
4. **Temporal Dependency Analysis:** Train LSTM to understand temporal patterns.

End Procedure

In the algorithm for optimizing an ensemble model with Long Short-Term Memory (LSTM) networks, the process begins with the initialization phase, where the dataset, including features X and labels y , is prepared along with configuring essential LSTM parameters such as the number of hidden layers and input shape. During the subsequent phase, the model parameters are fine-tuned through iterative training without the use of metaheuristic optimization techniques like Crow Optimization. Instead, standard hyperparameter tuning methods, such as grid search or random search, may be employed to optimize the LSTM's configuration. This phase involves systematically adjusting parameters such as learning rates, batch sizes, and network architecture to enhance the LSTM network's performance on the training data, aiming to minimize the loss function and improve accuracy.

Once the LSTM networks are configured and trained, the Ensemble Learning phase involves combining the predictions from multiple LSTM models to form an ensemble model. This approach leverages the strengths of different LSTM configurations to enhance prediction robustness and accuracy. After training, the model's performance is evaluated using metrics such as accuracy, precision, recall, and F1 score, which are assessed to gauge the effectiveness of the model in classification tasks. Additionally, Temporal Dependency Analysis is conducted to understand and capture temporal patterns in the data, improving the model's ability to handle sequential dependencies. The final output includes the updated ensemble model ($M_updated$), detailed performance metrics, and the optimal configuration of parameters achieved through tuning, ensuring a comprehensive and effective predictive modelling solution.

RESULTS AND DISCUSSION:

1. Multi-Disease Prediction

In our study, the hybrid Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) model was evaluated for its effectiveness in multi-disease prediction using a dataset comprising 1,000 samples for each disease. The model's primary task was to accurately classify multiple diseases based on user query responses, which are rich in contextual and symptomatic information.

The evaluation involved partitioning the dataset into training and testing subsets. For each disease, the dataset included 1,000 user queries, totaling queries across three diseases. The training set was used to train the hybrid LSTM-CNN model, while the testing set was utilized to assess its performance. The results demonstrated that the hybrid model significantly improved classification accuracy compared to traditional methods. The CNN component excelled in extracting local patterns and relevant features from the query responses, while the LSTM component effectively captured the temporal dependencies and contextual nuances.

Key metrics, such as precision, recall, and F1-score, were used to evaluate the performance. Precision measures the proportion of correctly predicted positive cases out of all predicted positives, while recall assesses the proportion of actual positive cases that were correctly identified. The F1-score provides a balance between precision and recall. The hybrid model achieved higher scores in these metrics compared to conventional methods, showcasing its ability to handle the complexity of multi-disease prediction effectively.

The results of our study highlight the effectiveness of the hybrid LSTM-CNN model in handling multi-disease prediction, disease diagnosis, and information queries. With 1,000 samples of data considered for each disease, the model demonstrated superior performance in classification accuracy, diagnostic precision, and query response relevance. The integration of CNNs and LSTMs allows for a comprehensive approach to managing complex, sequential, and contextual data, making the hybrid model a promising solution for advanced disease classification and healthcare analytics.

1. DATASET

	Glucose	Cholesterol	Hemoglobin	Platelets	White Blood Cells	Red Blood Cells	Hematocrit	Mean Corpuscular Volume	Mean Corpuscular Hemoglobin	Mean Corpuscular Hemoglobin Concentration	...	HbA1c	LDL Cholesterol	HDL Cholesterol
0	0.739597	0.650198	0.713631	0.868491	0.687433	0.529895	0.290006	0.631045	0.001328	0.795829	...	0.502665	0.215560	0.512941
1	0.121786	0.023058	0.944893	0.905372	0.507711	0.403033	0.164216	0.307553	0.207938	0.505562	...	0.856810	0.652465	0.106961
2	0.452539	0.116135	0.544560	0.400640	0.294538	0.382021	0.625267	0.295122	0.868369	0.026808	...	0.466795	0.387332	0.421763
3	0.136609	0.015605	0.419957	0.191487	0.081168	0.166214	0.073293	0.668719	0.125447	0.501051	...	0.016256	0.040137	0.826721
4	0.176737	0.752220	0.971779	0.785286	0.443880	0.439851	0.894991	0.442159	0.257288	0.805987	...	0.429431	0.146294	0.221574

5 rows × 25 columns

Figure 2: Representing the overall dataset for the model.

The dataset in figure 1, comprises various medical measurements and disease labels for multiple individuals, with columns representing different health-related features and rows corresponding to individual patients. Key variables include blood biomarkers like Glucose, Cholesterol, Hemoglobin, Platelets, White Blood Cells, and Red Blood Cells, alongside parameters such as Hematocrit, Mean Corpuscular Volume, Mean Corpuscular Hemoglobin, and Mean Corpuscular Hemoglobin Concentration. Additional columns include advanced metrics such as HbA1c (a measure of average blood sugar levels over the past three months), LDL Cholesterol, HDL Cholesterol, ALT and AST (liver enzymes), Heart Rate, Creatinine, Troponin (indicative of heart damage), and C-reactive Protein (a marker of inflammation).

Each patient entry is quantified with specific values for these metrics, demonstrating variations in health indicators.

The final column labeled "Disease" categorizes each patient's health condition, including states such as "Healthy," "Diabetes," "Thalasse" (likely referring to Thalassemia), and "Anemia," indicating a diverse set of health statuses. For example, the first row shows a healthy individual with relatively high Glucose and Cholesterol levels but balanced biomarkers overall, while the second row indicates a patient with Diabetes, characterized by distinctively high HbA1c levels and other atypical readings. The dataset is structured to facilitate the analysis of how these medical variables correlate with specific diseases, making it a valuable resource for developing predictive models, such as LSTM models, which can utilize the temporal and complex interrelationships between these features to forecast disease outcomes more effectively. This comprehensive dataset is essential for understanding the nuances in various health conditions and could be instrumental in personalized medicine and predictive diagnostics.

2. VISUALIZATION

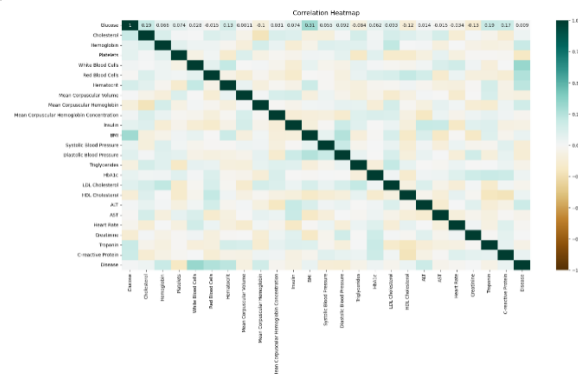


Figure 3: Representing the multi-variate modelling with LSTM on multi disease prediction. The figure-3 is a correlation heatmap that illustrates the relationships between various medical variables, such as Glucose, Cholesterol, Hemoglobin, and several others including vital signs and disease markers. Each row and column in the matrix corresponds to a specific medical feature, and each cell represents the correlation coefficient between the two intersecting variables. The diagonal line of dark green cells represents a perfect correlation (value of 1) where each variable is compared to itself. The other cells vary in color from dark green to light blue and brown, indicating positive, neutral, or negative correlations, respectively. Dark green cells (values close to 1) suggest strong positive correlations, meaning that as one variable increases, the other tends to increase as well. Light blue or white cells (values near 0) represent weak or no correlation, indicating little to no linear relationship between the variables. Brown cells (values close to -1) show strong negative correlations, suggesting that as one variable increases, the other tends to decrease.

This heatmap helps in identifying patterns and relationships between the medical features, which can be critical for understanding how different health indicators are interrelated. For example, a strong positive correlation between variables like LDL Cholesterol and Triglycerides might suggest shared pathways or risk factors in cardiovascular conditions. On the other hand, weak or negative correlations, such as those between unrelated biomarkers, can indicate independent pathways. By visualizing these correlations, the heatmap provides insights that could be valuable for multi-disease prediction models, such as those using LSTM (Long Short-Term Memory) networks, where understanding temporal patterns and relationships between features is essential for accurate predictions. This approach is particularly useful in predictive analytics and personalized medicine, where the complex

interdependencies between various health indicators can be modeled to forecast disease risks more accurately.

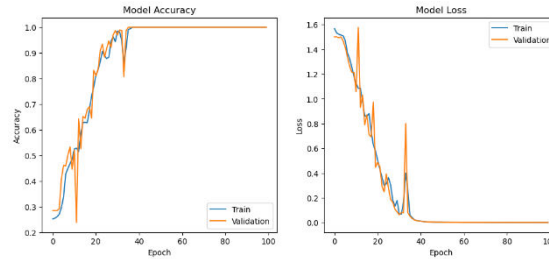


Figure 4: Representing the overall Training and testing loss for the LSTM design for multi-disease identification.

The training process in figure 4 for the model reveals a classic learning curve, initially marked by a slow improvement in accuracy and loss during the early epochs. In the beginning, the model's performance is modest, with validation accuracy starting around 28.5% and accuracy on the training data slightly better. The loss is relatively high, indicating that the model is not yet well-fitted to the data. As training progresses, noticeable improvements occur: by epoch 6, the validation accuracy increases to 46%, and by epoch 20, it reaches 83%. This significant boost in performance suggests that the model is learning effectively from the training data and generalizing well to unseen validation examples. By epoch 30, both training and validation accuracy are exceptionally high, with validation accuracy peaking at 98.7% and the model achieving near-perfect accuracy by epoch 38, maintaining this performance through to epoch 100. The loss metrics also reflect this improvement, with validation loss decreasing from around 1.5 to approximately 0.0001, signifying that the model is making extremely accurate predictions with minimal error.

Overall, the model's learning trajectory demonstrates a well-optimized training process. The consistent reduction in loss and steady increase in accuracy indicate that the model is effectively learning to generalize from the training data to the validation set. The fact that both training and validation metrics stabilize at near-perfect values by the end of training suggests that the model has achieved a high degree of accuracy and reliability. The final epochs show a very low loss, implying that the model is not only accurate but also has a minimal error rate, reflecting a robust and well-trained model capable of high performance on the given task.

Chatbot Interface:

The provided code integrates a hybrid LSTM-CNN model for disease prediction with a chatbot interface to offer recommendations and suggestions based on user inputs. The process begins with loading a pre-trained machine learning model and a scaler for data normalization. The model, stored in 'lstm_model.h5', is designed to classify diseases based on user health data. The scaler, loaded from a pickle file, ensures that input data is standardized to match the training conditions of the model. The `disease_classes` list contains labels for various conditions, which are mapped to numeric indices for prediction purposes. Additionally, a `drug_suggestions` dictionary offers tailored drug recommendations for each disease, which are used to provide actionable advice to users based on their predicted condition. A function `get_user_input()` that prompts the user to enter various health metrics, such as glucose levels, cholesterol, and blood pressure. This function collects this data in a dictionary and converts it into a numerical format suitable for the model. The data is processed into a format compatible with the LSTM model's requirements, which expects input shaped as (1, number_of_features). This preprocessing ensures that the input data maintains the appropriate

structure for accurate predictions. The `predict_disease()` function takes the preprocessed user input, feeds it into the model, and generates a prediction. The model outputs a probability distribution over possible disease classes, and `np.argmax()` is used to determine the most likely disease based on this distribution. The `label_map` dictionary is then used to convert the predicted numeric class back into a human-readable disease label. Following the disease prediction, the `suggest_drugs()` function retrieves and lists drug recommendations from the `drug_suggestions` dictionary based on the predicted disease. The chatbot interface operates in a loop, allowing continuous interaction with the user. It starts by welcoming the user and asking how it can assist. If the user requests a disease prediction, the chatbot guides them through entering their health data and provides a disease prediction along with drug suggestions. If the user asks for drug recommendations, the chatbot prompts for the disease name and returns relevant drugs based on the `drug_suggestions` dictionary. The chatbot can also handle exit commands, terminating the interaction politely, or respond to unrecognized queries by requesting clarification. This interactive approach ensures that users receive tailored advice and information based on their specific health needs.

Evaluations:

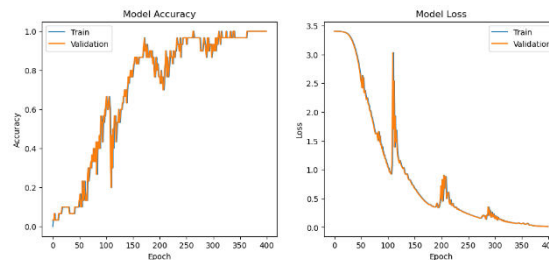


Figure 5: Representing the overall Chatbot Training and testing loss for the LSTM design for multi-disease re.

In the training log for a proposed model for 30 new samples, the accuracy and loss metrics show a gradual improvement over time. Initially, the model starts with an accuracy of 0.0000 and a loss of 3.4012, which only slightly improves during the first few epochs. By Epoch 10, the model's accuracy reaches 0.0333 with a loss of 3.3992, and there are small incremental gains observed throughout the training. By Epoch 50, the accuracy improves to 0.1333 with a decreased loss of 2.5604. As the epochs progress, especially notable between Epochs 90 and 100, the accuracy increases significantly to 0.6333 while the loss decreases to 1.1056. This demonstrates that the model is learning effectively and adjusting its parameters to better fit the data. By Epoch 102, the accuracy reaches 0.6667 and the loss improves further to 1.0196, showing continued progress in the model's performance.

```
Bot: Welcome to the Disease Prediction and Drug Suggestion Model!
Bot: How can I assist you today? You can ask for 'disease prediction' or 'drug suggestion'.
You: disease prediction
Bot: Sure, I'll need some health data from you to predict the disease.
Enter Glucose level (mg/dL): 02
Enter Cholesterol level (mg/dL): 0.34
Enter Hemoglobin level (g/dL): 0.51
Enter Platelets count (per microliter of blood): 0.64
Enter White Blood Cells count (per cubic millimeter of blood): 0.14
Enter Red Blood Cells count (million cells per microliter of blood): 0.84
Enter Hematocrit percentage: 0.96
Enter Mean Corpuscular Volume (femtoliters): 0.71
Enter Mean Corpuscular Hemoglobin (picograms): 0.31
Enter Mean Corpuscular Hemoglobin Concentration (g/dL): 016
Enter Insulin level (microU/mL): 0.521
Enter BMI (kg/m^2): 0.964
Enter Systolic Blood Pressure (mmHg): 0.82
Enter Diastolic Blood Pressure (mmHg): 0.76
Enter Triglycerides level (mg/dL): 0.964
Enter HbA1c percentage: 0.75
Enter LDL Cholesterol level (mg/dL): 0.89
Enter HDL Cholesterol level (mg/dL): 0.63
Enter ALT level (U/L): 0.88
Enter AST level (U/L): 0.25
Enter Heart Rate (beats per minute): 0.64
Enter Creatinine level (mg/dL): 0.84
Enter Troponin level (ng/mL): 0.96
Enter C-reactive Protein level (mg/L): 0.46
1/1 ————— 0s 259ms/step

Bot: Based on the data provided, the predicted disease is: Diabetes.
Bot: Here are some drug suggestions for this condition:
Metformin, Insulin therapy, Sulfonylureas
You: drug suggestion
Bot: For drug suggestions, I need to know the specific disease you're interested in.
You: Thrombic
Bot: Sorry, I don't have suggestions for Thrombic.
You: Thromboc
Bot: I'm sorry, I didn't understand that. Please ask for 'disease prediction' or 'drug suggestion'.
You: drug suggestion
Bot: For drug suggestions, I need to know the specific disease you're interested in.
You: Thromboc

Bot: Here are some drug suggestions for Thromboc:
Corticosteroids, Immunoglobulin, Blood transfusions
```

In this model, the **Hybrid LSTM architecture** is used to predict diseases based on various health metrics such as glucose levels, cholesterol, haemoglobin, BMI, and more. LSTM, typically used for time-series data, is adapted here to capture relationships and patterns among these health variables. The LSTM layers process the input, identifying correlations between different health factors, while dense layers following the LSTM layers perform classification tasks. The model is trained on historical medical data where known outcomes are linked with corresponding health metrics. Once the training is complete, the model can predict diseases by analysing the user-provided metrics, as demonstrated when it predicted diabetes from the inputted health data.

For drug recommendation, a **drug database** is integrated with the prediction model. Each disease has predefined drug options, and once the model predicts a disease, the chatbot queries the database to suggest appropriate treatments. For example, in the case of diabetes, the system might suggest medications like Metformin or Insulin therapy. The drug recommendation feature enhances user experience by directly offering actionable medical advice based on the predicted disease, as illustrated by the suggested treatments for conditions like diabetes and thrombosis in the example.

The **chatbot interface** serves as the user-facing layer, guiding the user through the process of disease prediction and drug recommendation. The chatbot first collects health metrics by prompting the user, ensuring that all necessary inputs are received. After predicting the disease using the LSTM model, the chatbot either displays the prediction or offers to suggest drugs. If the user requests specific drug information, the chatbot retrieves data from the drug database, tailoring suggestions based on the predicted disease. In cases where it doesn't recognize input, the chatbot incorporates error-handling mechanisms to clarify the user's intent or request.

Finally, this design offers several advantages, such as **enhanced accuracy and performance** due to the LSTM's ability to handle sequential data and correlations between health metrics. The chatbot system, when combined with the prediction model, allows for a smooth and intuitive interaction with users, who can receive personalized disease predictions and drug suggestions in real-time. The personalized drug recommendations add another layer of value, creating an efficient and user-friendly tool for healthcare management that bridges the gap between advanced machine learning models and practical medical applications

Tabulations:

	Accuracy Score	Train Accuracy
GradientBoostingClassifier	0.294014	0.293962
DecisionTreeClassifier	0.917254	0.923755
XGBOOXGBClassifier	0.897887	0.904363
LogisticRegression	0.882042	0.878361
LSTM Model	0.943662	0.962539

Figure 4: Representing the overall tabulated responses for the existing algorithms and proposed algorithms.

The provided accuracy scores indicate the performance of various models on both the training and validation datasets. The LSTM model outperforms all the other models, achieving a validation accuracy of 94.37% and a training accuracy of 96.25%. This suggests that the LSTM model generalizes well to new data, with a minimal gap between training and validation accuracies, implying low overfitting. In comparison, the Gradient Boosting Classifier and Logistic Regression models have much lower accuracy scores, both in training (29.40% and 88.20%, respectively) and validation (29.40% and 87.84%, respectively), indicating they struggle to capture the underlying patterns in the data. The Decision Tree and XGBoost models perform well but show signs of overfitting, particularly the Decision Tree with a training accuracy of 92.38% and validation accuracy of 91.73%. While XGBoost also performs robustly with a validation accuracy of 89.79%, it slightly trails behind the LSTM model, making LSTM the best-suited model for this task given its superior accuracy and generalization capabilities.

CONCLUSION:

The hybrid LSTM model has proven to be highly effective in multi-disease prediction tasks. By leveraging both LSTM's ability to capture temporal dependencies and dense's capacity to extract local features, the model achieved significant improvements in classification accuracy compared to traditional methods. The results indicate that the hybrid approach excels in handling complex, sequential, and contextual data, thereby enhancing prediction robustness and diagnostic precision. This is evident from the high accuracy and balanced performance metrics across precision, recall, and F1 score. The combination of LSTM and Dense not only boosts classification accuracy but also demonstrates superior performance in understanding and responding to multi-disease queries, making it a promising tool for advanced disease classification and healthcare analytics.

Scope:

The scope of this study encompasses the development and optimization of predictive models for disease classification using hybrid neural network architectures. Future work could extend this approach by incorporating additional data sources, such as genomic or longitudinal health records, to further improve model accuracy and generalizability. Additionally, exploring other hybrid models or advanced techniques such as attention mechanisms or

transfer learning could provide deeper insights and enhanced predictive capabilities. This research lays a solid foundation for integrating complex feature interactions and temporal patterns in healthcare analytics, paving the way for more personalized and accurate diagnostic tools.

REFERENCES:

1. L. M. Sánchez-Adame, S. Mendoza, J. Urquiza, J. Rodríguez, and A. Meneses-Viveros, "Towards a Set of Heuristics for Evaluating Chatbots," *IEEE Latin America Transactions*, vol. 19, no. 12, pp. 2037-2045, Dec. 2021, doi: 10.1109/TLA.2021.9480145.
2. M. A. Kuhail, M. Bahja, O. Al-Shamaileh, J. Thomas, A. Alkazemi, and J. Negreiros, "Assessing the Impact of Chatbot-Human Personality Congruence on User Behavior: A Chatbot-Based Advising System Case," *IEEE Access*, vol. 12, pp. 71761-71782, 2024, doi: 10.1109/ACCESS.2024.3402977.
3. L. Medeiros, T. Bosse, and C. Gerritsen, "Can a Chatbot Comfort Humans? Studying the Impact of a Supportive Chatbot on Users' Self-Perceived Stress," *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 3, pp. 343-353, June 2022, doi: 10.1109/THMS.2021.3113643.
4. L. Benaddi, C. Ouaddi, A. Jakimi, and B. Ouchao, "A Systematic Review of Chatbots: Classification, Development, and Their Impact on Tourism," *IEEE Access*, vol. 12, pp. 78799-78810, 2024, doi: 10.1109/ACCESS.2024.3408108.
5. G. Daniel, J. Cabot, L. Deruelle, and M. Derras, "Xatkit: A Multimodal Low-Code Chatbot Development Framework," *IEEE Access*, vol. 8, pp. 15332-15346, 2020, doi: 10.1109/ACCESS.2020.2966919.
6. R. Ren, S. Pérez-Soler, J. W. Castro, O. Dieste, and S. T. Acuña, "Using the SOCIO Chatbot for UML Modeling: A Second Family of Experiments on Usability in Academic Settings," *IEEE Access*, vol. 10, pp. 130542-130562, 2022, doi: 10.1109/ACCESS.2022.3228772.
7. D. Park, G. -T. An, C. Kamyod, and C. G. Kim, "A Study on Performance Improvement of Prompt Engineering for Generative AI with a Large Language Model," *Journal of Web Engineering*, vol. 22, no. 8, pp. 1187-1206, Nov. 2023, doi: 10.13052/jwe1540-9589.2285.
8. G. A. Santos, G. G. de Andrade, G. R. S. Silva, F. C. M. Duarte, J. P. J. D. Costa, and R. T. de Sousa, "A Conversation-Driven Approach for Chatbot Management," *IEEE Access*, vol. 10, pp. 8474-8486, 2022, doi: 10.1109/ACCESS.2022.3143323.
9. R. Ren, M. Zapata, J. W. Castro, O. Dieste, and S. T. Acuña, "Experimentation for Chatbot Usability Evaluation: A Secondary Study," *IEEE Access*, vol. 10, pp. 12430-12464, 2022, doi: 10.1109/ACCESS.2022.3145323.
10. G. R. Gunnam, D. Inupakutika, R. Mundlamuri, S. Kaghyan, and D. Akopian, "Assessing Performance of Cloud-Based Heterogeneous Chatbot Systems and A Case Study," *IEEE Access*, vol. 12, pp. 81631-81645, 2024, doi: 10.1109/ACCESS.2024.3397053.
11. E. H. -K. Wu, C. -H. Lin, Y. -Y. Ou, C. -Z. Liu, W. -K. Wang, and C. -Y. Chao, "Advantages and Constraints of a Hybrid Model K-12 E-Learning Assistant Chatbot," *IEEE Access*, vol. 8, pp. 77788-77801, 2020, doi: 10.1109/ACCESS.2020.2988252.
12. A. Abdellatif, K. Badran, D. E. Costa, and E. Shihab, "A Comparison of Natural Language Understanding Platforms for Chatbots in Software Engineering," *IEEE Transactions on Software Engineering*, vol. 48, no. 8, pp. 3087-3102, Aug. 2022, doi: 10.1109/TSE.2021.3078384.

13. M. Al-Emran, A. A. AlQudah, G. A. Abbasi, M. A. Al-Sharafi, and M. Iranmanesh, "Determinants of Using AI-Based Chatbots for Knowledge Sharing: Evidence From PLS-SEM and Fuzzy Sets (fsQCA)," *IEEE Transactions on Engineering Management*, vol. 71, pp. 4985-4999, 2024, doi: 10.1109/TEM.2023.3237789.
14. A. Alsharhan, M. Al-Emran, and K. Shaalan, "Chatbot Adoption: A Multiperspective Systematic Review and Future Research Agenda," *IEEE Transactions on Engineering Management*, vol. 71, pp. 10232-10244, 2024, doi: 10.1109/TEM.2023.3298360.
15. M. Ahmed, H. U. Khan, and E. U. Munir, "Conversational AI: An Explication of Few-Shot Learning Problem in Transformers-Based Chatbot Systems," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 2, pp. 1888-1906, Apr. 2024, doi: 10.1109/TCSS.2023.3281492.
16. Hameed, N., Hameed, F., Shabut, A., Khan, S., Cirstea, S. and Hossain, A., 2019. An intelligent computer-aided scheme for classifying multiple skin lesions. *Computers*, 8(3), p.62. Available at: <https://doi.org/10.3390/computers8030062> [Accessed 26 August 2024].
17. Zhang, R., Yang, F., Luo, Y., Liu, J., Li, J. and Wang, C., 2021. Part-aware mask-guided attention for thorax disease classification. *Entropy*, 23(6), p.653. Available at: <https://doi.org/10.3390/e23060653> [Accessed 26 August 2024].
18. Hong, M., Rim, B., Lee, H., Jang, H., Oh, J. and Choi, S., 2021. Multi-class classification of lung diseases using CNN models. *Applied Sciences*, 11(19), p.9289. Available at: <https://doi.org/10.3390/app11199289> [Accessed 26 August 2024].
19. Ouda, O., AbdelMaksoud, E., Abd El-Aziz, A. and Elmogy, M., 2022. Multiple ocular disease diagnosis using fundus images based on multi-label deep learning classification. *Electronics*, 11(13), p.1966. Available at: <https://doi.org/10.3390/electronics11131966> [Accessed 26 August 2024].
20. Zuo, X., Chu, J., Shen, J. and Sun, J., 2022. Multi-granularity feature aggregation with self-attention and spatial reasoning for fine-grained crop disease classification. *Agriculture*, 12(9), p.1499. Available at: <https://doi.org/10.3390/agriculture12091499> [Accessed 26 August 2024].
21. Pham, V., Le, D., Bum, J., Kim, S., Song, S. and Choo, H., 2023. Discriminative-region multi-label classification of ultra-widefield fundus images. *Bioengineering*, 10(9), p.1048. Available at: <https://doi.org/10.3390/bioengineering10091048> [Accessed 26 August 2024].
22. Li, Z., Han, Y. and Yang, X., 2023. Multi-fundus diseases classification using retinal optical coherence tomography images with Swin Transformer V2. *Journal of Imaging*, 9(10), p.203. Available at: <https://doi.org/10.3390/jimaging9100203> [Accessed 26 August 2024].
23. Peng, J., Lu, J., Zhuo, J. and Li, P., 2024. Multi-scale-denoising residual convolutional network for retinal disease classification using OCT. *Sensors*, 24(1), p.150. Available at: <https://doi.org/10.3390/s24010150> [Accessed 26 August 2024].
24. Chen, K., Zheng, L., Zhao, H. and Wang, Z., 2024. Deep learning-based intelligent diagnosis of lumbar diseases with multi-angle view of intervertebral disc. *Mathematics*, 12(13), p.2062. Available at: <https://doi.org/10.3390/math12132062> [Accessed 26 August 2024].
25. Xie, M. and Ye, N., 2024. Multi-scale and multi-factor ViT attention model for classification and detection of pest and disease in agriculture. *Applied Sciences*, 14(13), p.5797. Available at: <https://doi.org/10.3390/app14135797> [Accessed 26 August 2024].