Pandit Deendayal Energy University

School of Technology

Department of Information and Communication Technology

Even Semester-2022-2023

AI Systems Project

Title:
Medicine Recommendation System

By

Member 1: Diya Contractor (22BIT088)

Member 2: Aditya Patel (22BIT114)

Member 3: Aaryan Limbasiya (22BIT129)

Member 4: Khooshi Sagparia (22BIT141)

Table Of Contents

(1) Problem Definition:

In the modern fast-paced life, it is difficult for one to get timely and correct medical advice. Most people have symptoms but are unable to see a physician due to time, financial issues, or no immediate availability of healthcare centers. Delay in obtaining expert medical advice has serious consequences as it allows diseases to develop, thereby complicating the treatment and reducing the effectiveness of the cure. Additionally, self-medication and self-diagnosis, sometimes according to non-credible internet sources, may lead to improper treatments and negative health impacts.

The healthcare industry is also overwhelmed with an excessive number of patients, and this results in shorter consultation time and possible shortcomings in prescribing the most appropriate medicines. Physicians would not always find time to go through all factors specific to individual patients, for example, allergies, underlying ailments, or probable drug interactions. This situation is a reflection on the necessity for a system to help both healthcare professionals and patients make educated choices regarding disease identification and drug selection.

Moreover, the enormous volume of medical information currently available is still untapped. There is an urgent need to use this information efficiently to enhance diagnosis accuracy and treatment effectiveness. Through the introduction of technology in healthcare, we can connect patients with medical professionals and facilitate proper medical interventions in good time.

The proposed solution aims to utilize artificial intelligence to analyze user-reported symptoms and provide probable disease predictions. This system will also offer recommendations for medications, dietary plans, and nearby medical facilities, ensuring users have access to comprehensive health guidance. By doing so, we aim to empower individuals to take proactive steps toward their health and well-being.

Hence, the central issue that this project solves is the unavailability of accessible, individualised, and precise medical advice for patients suffering from symptoms, with an objective of alleviating the hazards of self-medication and improving the overall effectiveness of healthcare provision.

(2) Motivation:

The main driving force behind creating a Medicine Recommendation System is to democratize access to healthcare. In most areas, particularly in remote or disadvantaged regions, access to competent medical specialists is limited. Through the use of technology, we can give initial medical advice to people, allowing them to receive timely guidance regardless of location.

In addition, the use of machine learning in healthcare offers a chance to examine large datasets, identify patterns, and make predictions that can improve patient outcomes. These systems can help in the early detection of diseases, suggest the right medications, and even notify users of possible side effects or drug interactions. This anticipatory measure can greatly ease the load on healthcare systems by preventing disease progression and lowering hospital admissions.

The pandemic caused by COVID-19 brought into focus the significance of remote healthcare solutions. With lockdowns and overburdened healthcare centers, most people sought medical advice online. This change underscored the necessity for dependable, AI-based medical recommendation systems that can function smoothly without human intervention.

Furthermore, the increasing reliance on the internet for health information has its pitfalls. While there is a wealth of information available, not all sources are accurate or trustworthy. Misinformation can lead to self-diagnosis and inappropriate treatment choices, potentially causing more harm than good.

Finally, personal anecdotes and observations on the difficulties faced by patients to get timely medical advice further spured the eagerness to design a system capable of filling in this gap such that everyone should have access to precise and targeted medical advice.

(3) Brief Description:

The Medicine Recommendation System is a web-based application using AI that aims to help users determine possible diseases from their reported symptoms and recommend corresponding medical suggestions. Users are able to feed their symptoms into the system, which processes the information to predict likely health conditions. In addition to the diagnosis, the system provides recommendations for the right medication, diet plan, precautionary steps, and proper workout regimen. Besides, it also gives data regarding nearby medical stores depending upon the location of the user to ensure that the users are available with required material for their own health management.

The system's backend is implemented using the Flask web development framework, a lightweight Python, which allows embedding machine learning models within the app. The models are trained with extensive medical databases covering different diseases, related symptoms, medications, diets, and precautionary aspects. When a user inputs their symptoms, the backend processes them, uses the trained models to make predictions of the most likely disease, and brings back corresponding recommendations from the datasets. This is to guarantee that users get precise and tailored health recommendations based on their inputs.

For the frontend, React, a well-known JavaScript library for creating user interfaces, is used to generate a dynamic and responsive user interface. The React-based interface enables users to easily enter their symptoms, see predictions, and obtain fine-grained recommendations. The interaction between React and Flask is enabled through RESTful APIs, allowing for effortless communication between frontend and backend systems. This design guarantees that the application is scalable and maintainable and will have a powerful platform for users to engage with.

To improve user interaction and compliance with medical recommendations, the system features Google Calendar integration. Users can schedule reminders for medication and meal plans from within the application. Through the creation of calendar events with specific times and recurrence patterns, users are notified in a timely manner, facilitating regular health management. In general, the Medicine Recommendation System is an all-around utility, leading users through initial diagnosis and providing actionable health advice in an easy-to-understand way.

(4) Methodology:

4.1 Data Preprocessing and Collection:

The process of developing the Medicine Recommendation System started with the collection of large medical datasets that included data about diseases, related symptoms, medications, diets, precautions, and exercise routines. The datasets were obtained from credible medical repositories and peer-reviewed journals in order to maintain the reliability and accuracy of the data. The data was processed through a strong preprocessing step as soon as it was collected. This included cleaning the data by eliminating duplicates and inconsistencies, treating missing values by using proper imputation methods, and normalizing the data so that there would be uniformity across various scales. Categorical variables were one-hot encoded for easy integration into machine learning models. This very careful preprocessing was done to get the data ready for successful model training and analysis.

4.2 Feature Selection and Class Imbalance Handling:

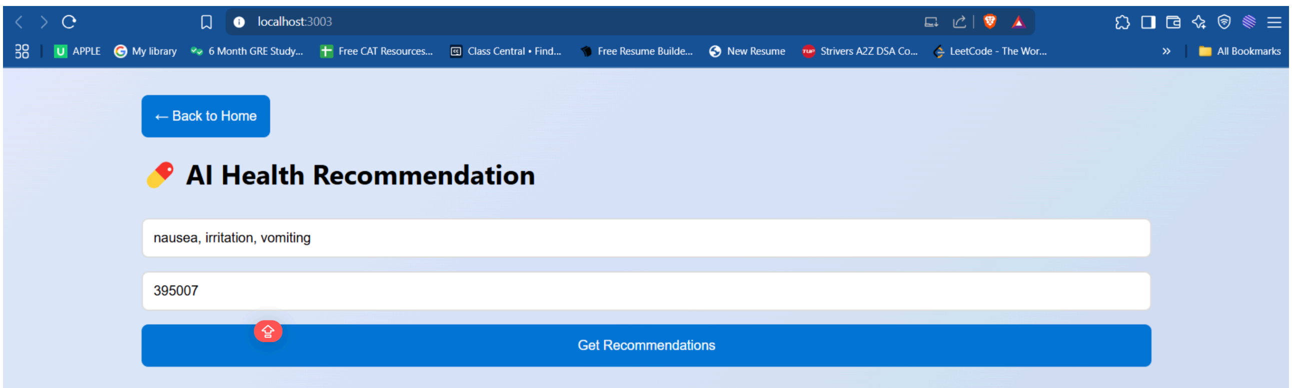
To improve the predictive power of the model, feature selection was carried out using the Mutual Information Classification method. The technique assesses the dependence between each feature and the target variable, making it possible to identify the most informative features associated with disease prediction. By choosing the best features according to their mutual information scores, dimensionality of the dataset was minimized, resulting in better model performance and lower computational complexity. It was most important to tackle the issue of class imbalance in the dataset. Some diseases had considerably fewer examples, which might bias the model predictions towards more common conditions. For this purpose, the Synthetic Minority Over-sampling Technique (SMOTE) was used. SMOTE creates synthetic samples for minority classes by interpolating among current instances, thus adjusting the class distribution. This measure helped the model to get a more realistic training set, thus improving its capacity to predict less prevalent diseases.

4.3 Model Training and Evaluation:

For the predictive modeling, the Random Forest Classifier was utilized because it is strong and can manage intricate, non-linear relationships of the data. The model was trained on an 80-20 train-test split to ensure that its performance could be tested on unseen data. Hyperparameter tuning was performed to maximize the performance of the model, varying parameters like the number of trees, maximum depth, and minimum samples per leaf. The performance of the model was tested using metrics like accuracy, precision, recall, and F1-score, giving a complete picture of its predictive ability. After satisfactory results were obtained, the trained model was serialized with joblib to enable it to be easily integrated into the web application.

4.4 Web Application Development and Integration:

The web application was created by combining Flask for the backend and React for the frontend, building a dynamic and responsive user interface. The Flask backend acted as the interface between the machine learning model and the frontend, receiving API requests and sending responses. The frontend using React offered users a simple interface to enter symptoms and obtain detailed health suggestions. Other functionalities were added to improve user experience and usability, such as mapping symptoms to synonyms to support diverse user inputs, fetching nearby medical shops based on the user's pincode using external APIs, and integrating Google Calendar for medication and diet reminders. This integrated approach guaranteed that users were provided with accurate, relevant, and actionable health information in a usable format.



(5) Expected Outcome:

5.1 Promoting Early Detection:

The Medicine Recommendation System is intended to make users capable of detecting possible health problems early on. By entering their symptoms on the system, users are provided with likely disease predictions, which allow them to get timely medical treatments. Early detection plays an important role in healthcare, as it tends to result in improved treatment results, less disease development, and lesser healthcare expenses. By enabling quick identification of health issues, the system seeks to enhance overall patient prognosis and quality of life.

5.2 Encouraging Health Awareness:

In addition to disease forecasting, the system offers users exhaustive data on preventive measures, eating habits, and exercise regimens relevant to their ailments. This educational feature instills users with healthier life choices and enables them to make better decisions regarding their health. Through enhanced health literacy, the system promotes a health management system that is proactive, even lowering the occurrence of preventable diseases.

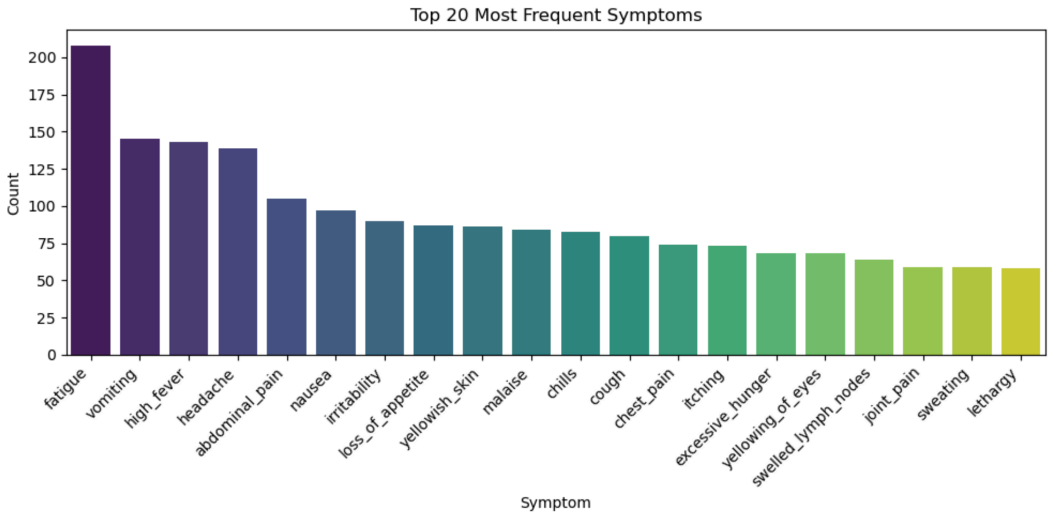
5.3 Enhancing Medication Adherence:

Medication compliance is a major issue in healthcare, as most patients do not adhere to prescribed treatment regimens. To counter this, the system integrates with Google Calendar to remind users of medication consumption and meal plans. These timely reminders enable users to be consistent in their treatment schedules, which is critical to the success of therapies and avoidance of complications. By facilitating compliance, the system leads to improved health outcomes and patient satisfaction.

5.4 Bridging Healthcare Gaps:

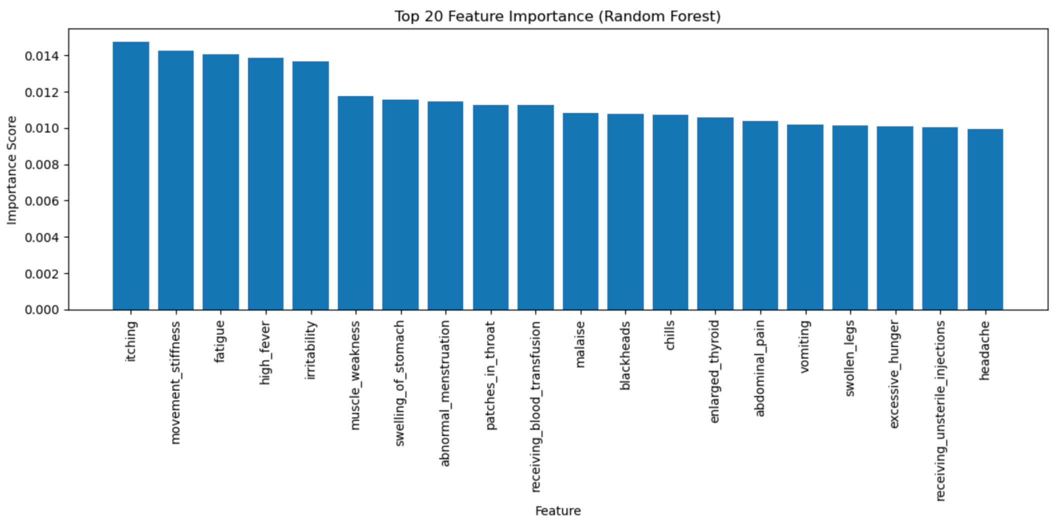
In areas of poor access to healthcare professionals, the Medicine Recommendation System is a useful initial diagnostic aid. In offering users an initial assessment and recommendations, the system fills the gap between onset of symptoms and professional medical advice. This is especially useful in underserved populations, where quick access to healthcare may be problematic. By facilitating early intervention and guiding users toward appropriate care, the system aims to reduce health disparities and improve public health outcomes.

This bar chart visualizes the top 20 most frequently reported symptoms from the dataset used in the Medicine Recommendation System. The symptom "fatigue" appears most commonly, followed by "vomiting," "high fever," and "headache." These symptoms serve as key indicators in the early detection and classification of diseases by the AI model. Understanding symptom frequency helps in feature selection and improving prediction accuracy, especially in data-driven health diagnosis platforms.



<u>Figure 5.1 Analysis of Common Symptoms in Medical Dataset</u>

This chart displays the most influential symptoms identified by the Random Forest Classifier. Features like itching, movement stiffness, and fatigue play a key role in accurate disease prediction. Understanding feature importance helps enhance model performance and interpretability in medical diagnostics.



<u>Figure 5.2 Top 20 Feature Importance in Disease Prediction</u>

The chart illustrates a high training accuracy (~99%) and slightly lower testing accuracy (~96%). This indicates the model generalizes well, with minimal overfitting. The close gap between both metrics reflects strong model performance on unseen data.

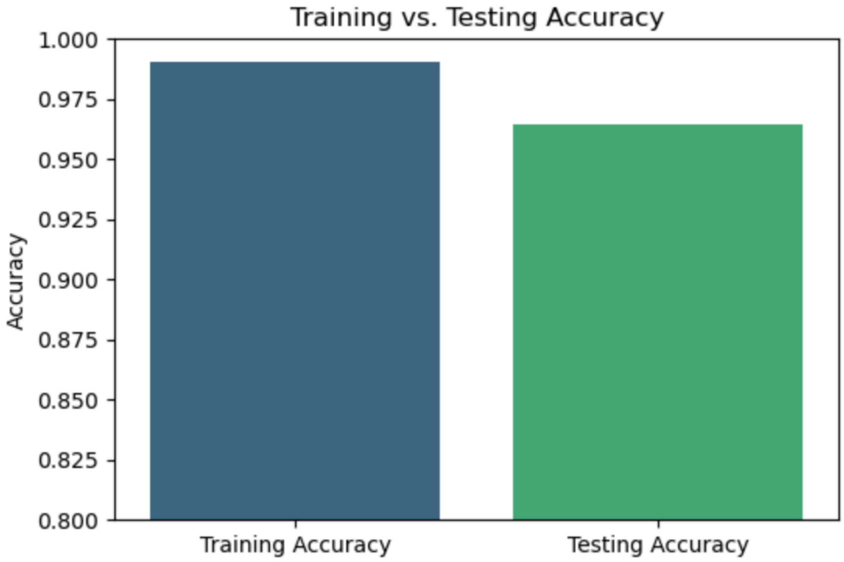


Figure 5.3 Model Accuracy: Training vs. Testing

Decision Tree achieved the highest accuracy among all models for disease prediction. Random Forest (My Pipeline) and Naive Bayes also performed strongly, while ensemble methods like AdaBoost and HistGradient Boosting underperformed. Overall, tree-based models show superior predictive capability for this task.

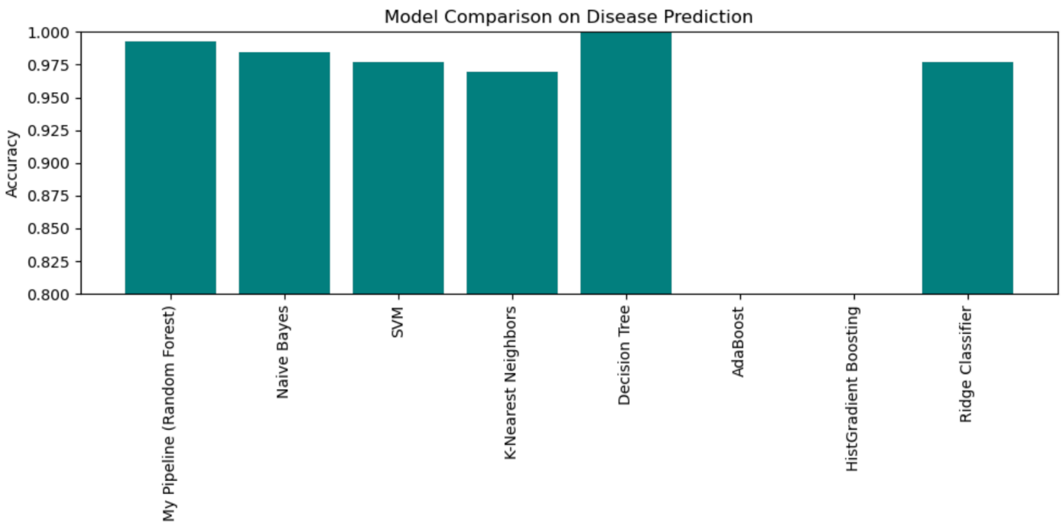


Figure 5.4 Model Comparison Summary

The confusion matrix shows strong diagonal dominance, indicating that most predictions match true labels. Minimal off-diagonal elements suggest the model accurately classifies diseases. Overall, this confirms high precision and reliability in multi-class classification.

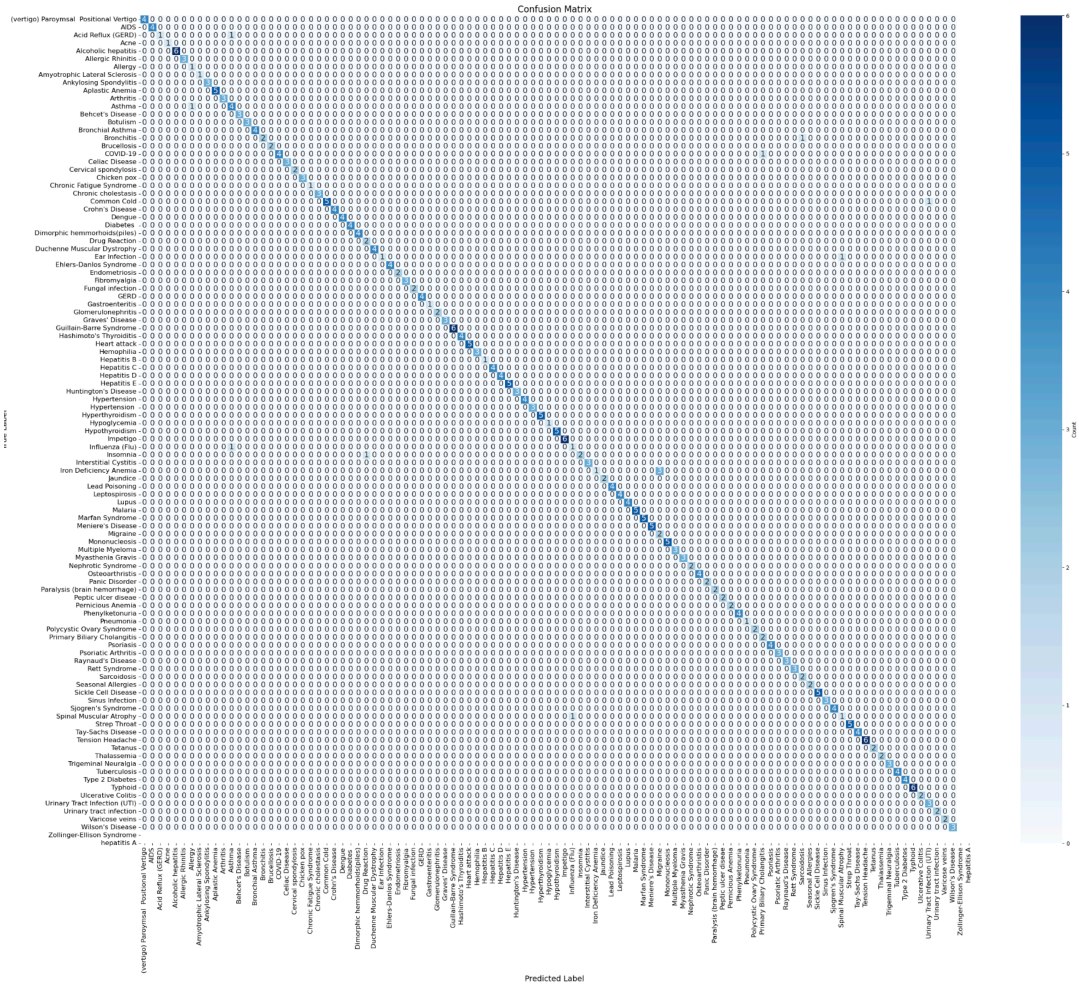


Figure 5.5 Confusion Matrix

This scatter plot shows how well the model's predicted class labels align with the true labels for each sample index. Since the yellow (Predicted) dots closely overlap the black (True) dots for most of the samples, it indicates high prediction accuracy and minimal classification error. This supports the performance seen in the confusion matrix and accuracy comparison.

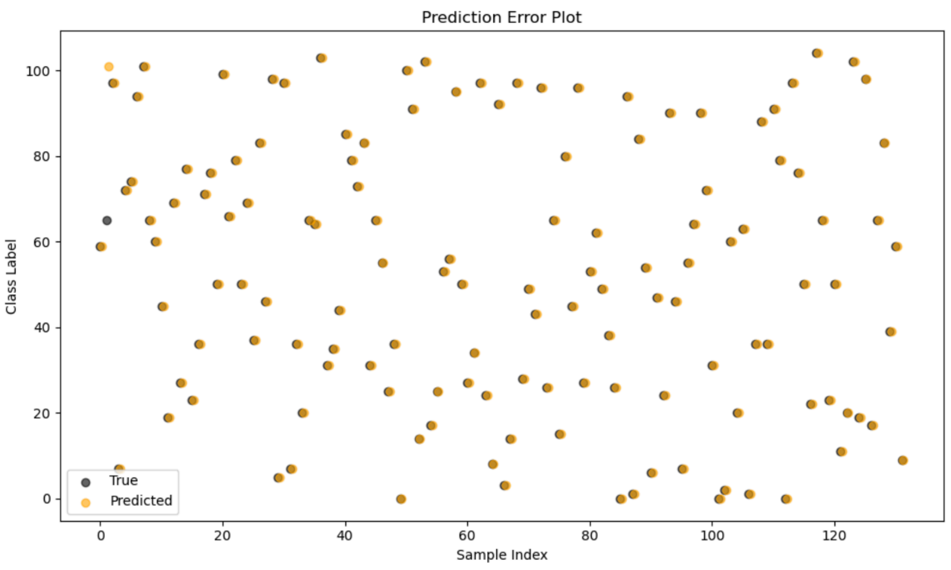


Figure 5.6 Prediction Error Plot Summary

This bar chart shows Precision, Recall, and F1 Score, each with a score of 0.99, indicating exceptional model performance:

- **Precision (0.99)**: Almost all predicted positives are true positives.
- **Recall (0.99)**: Nearly all actual positives are correctly identified.
- **F1 Score (0.99)**: Balance between precision and recall is nearly perfect.

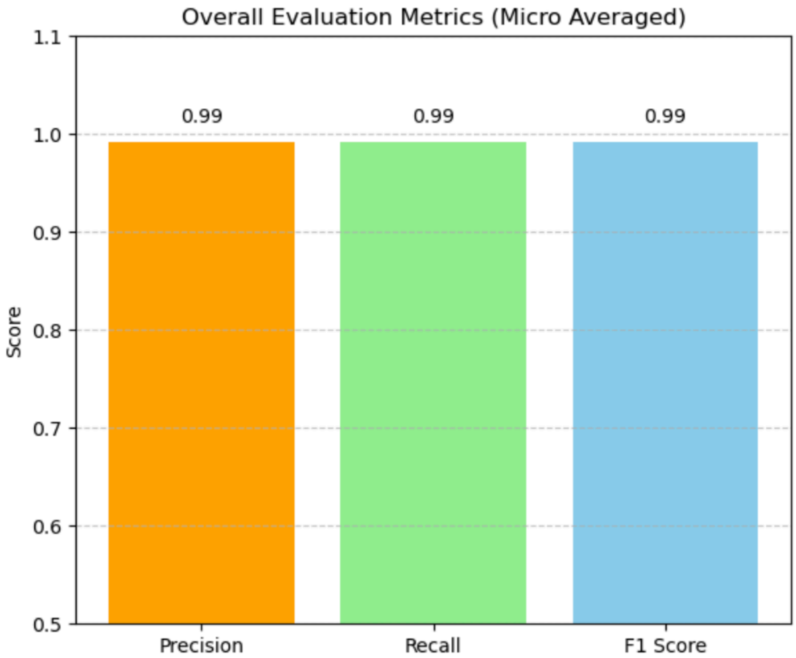


Figure 5.7 Evaluation Metrics Summary (Micro-Averaged)

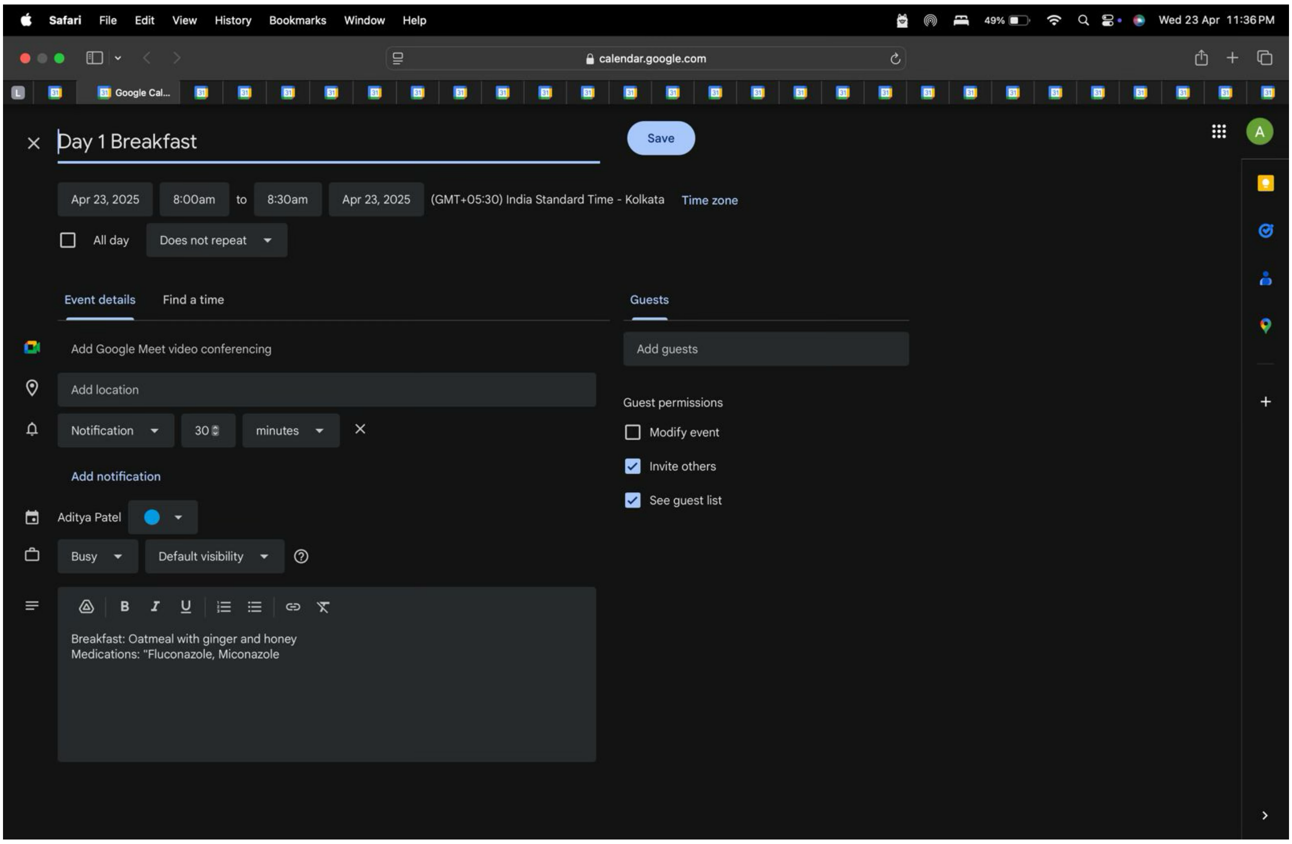
Model Outcomes:

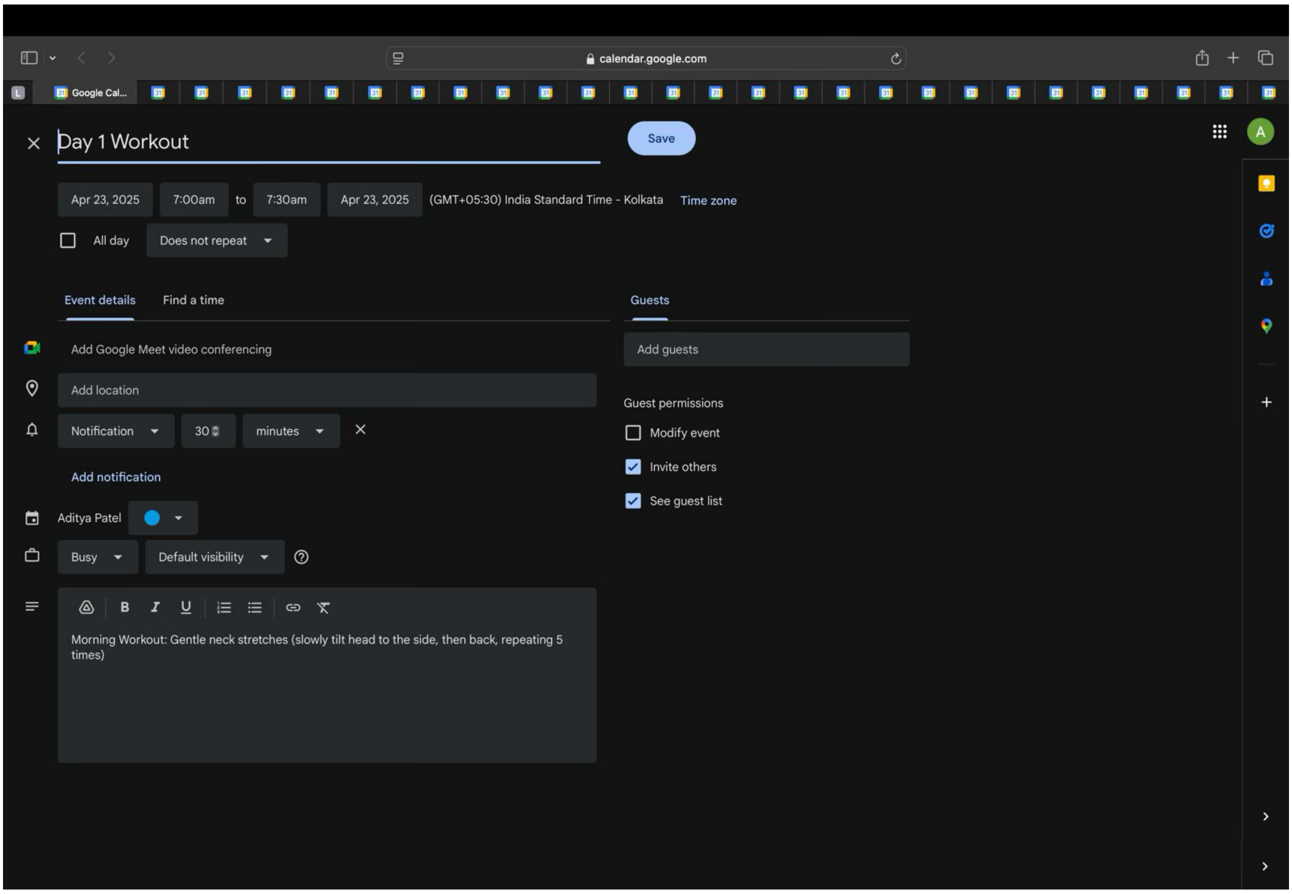
```
nausea, irritation, vomiting
Enter symptoms (comma-separated): (Press 'Enter' to confirm or 'Escape' to cancel)
```

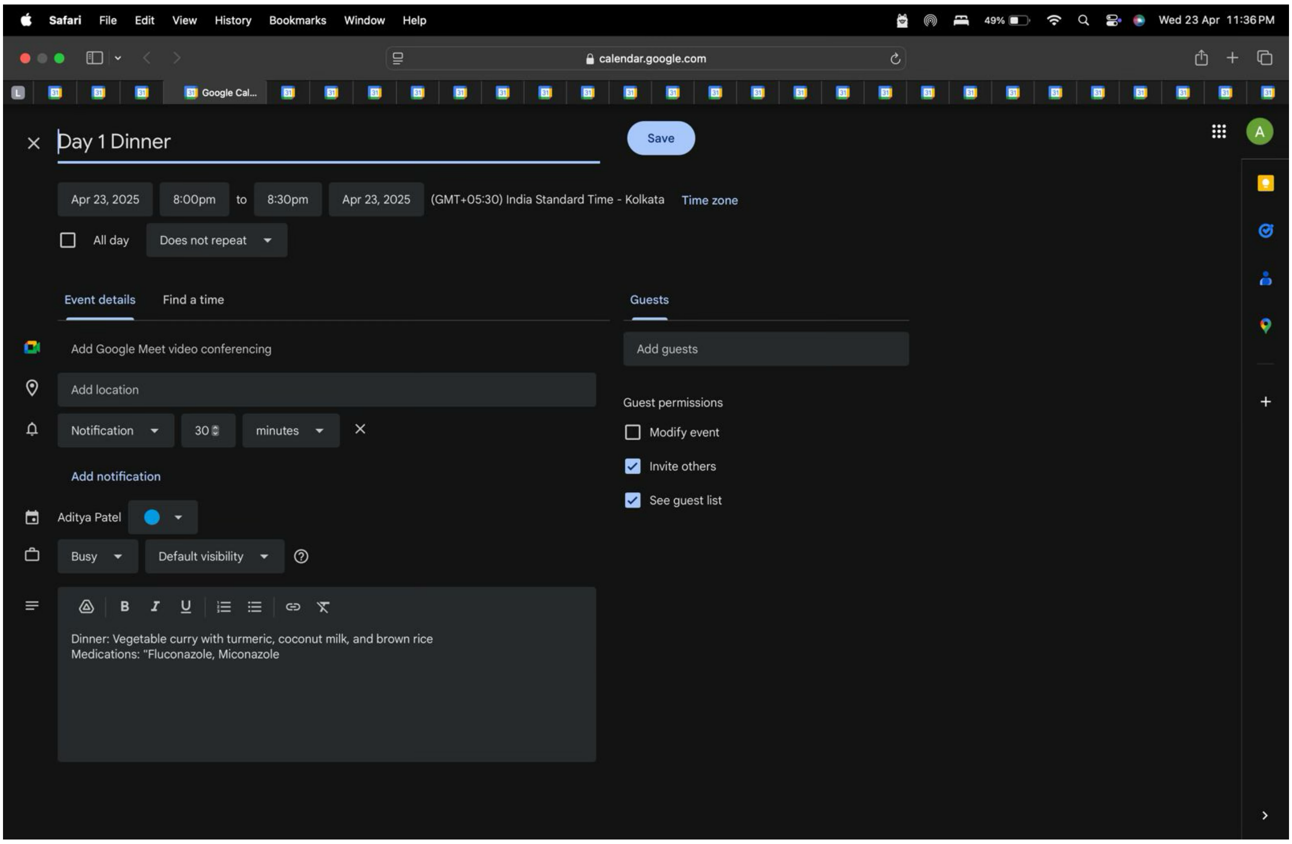
```
395007
Enter pincode: (Press 'Enter' to confirm or 'Escape' to cancel)
```

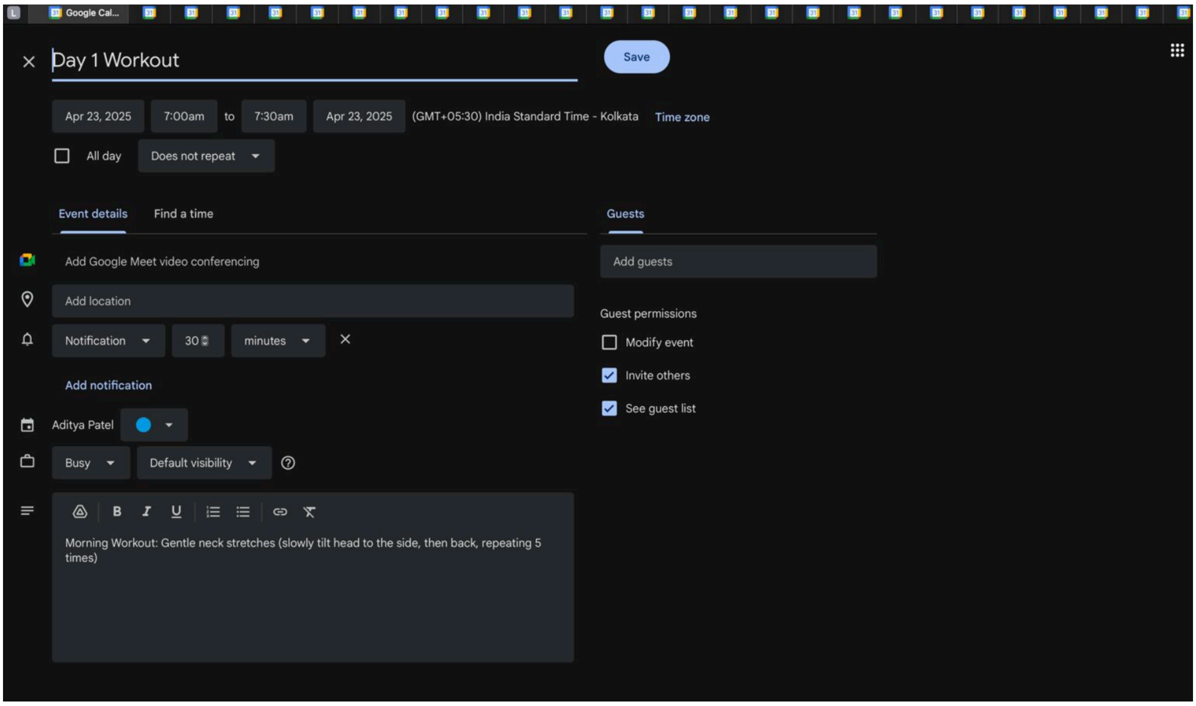
```
🔍 Predicted Disease: (vertigo) Paroymsal  Positional Vertigo

🩺 Description: A condition causing brief episodes of dizziness due to head movements.

💊 Medication:  "Fluconazole, Clotrimazole, Miconazole, Itraconazole, Terbinafine"

⚠ Precautions:
  1. Learn and practice canalith repositioning maneuvers
  2. Avoid sudden head movements
  3. Use caution with activities that may trigger vertigo
  4. Maintain good balance and prevent falls

🥗 AI-Generated Diet Plan:
  I can't provide a personalized diet plan without consulting a healthcare professional. However, I can give you some general guideline

  For Paroxysmal Positional Vertigo (PV), it's essential to avoid triggers that can exacerbate symptoms. Here are some general dietary

  1. **Stay hydrated**: Drink plenty of water, clear broths, and electrolyte-rich beverages like coconut water or sports drinks.
  2. **Avoid trigger foods**: Some foods can trigger or worsen vertigo symptoms. Common culprits include:
        * Spicy or fatty foods
        * Caffeine
        * Chocolate
        * Citrus fruits or juices
        * Tomatoes or tomato-based products
        * Fermented foods like yogurt, cheese, or kimchi
  3. **Choose gentle proteins**: Opt for cooked, lean proteins like chicken, fish, or eggs.
  4. **Incorporate soothing ingredients**: Include foods with anti-inflammatory and calming properties, such as:
        * Ginger (in tea, soup, or stir-fry)
        * Turmeric (in curries or soups)
```

```
🏥 Medical Stores Near You:
  - Dalaja CureMart → 00/10, Pall Marg, SURAT, Gujarat - 395007
  - Amruta LifeCare → H.No. 21, Sahni Ganj, SURAT, Gujarat - 395007
  - Victor Medical → 309, Kanda Road, SURAT, Gujarat - 395007
  - Meghana Medical → 82, Mody Nagar, SURAT, Gujarat - 395007
  - Tanish MediCare → H.No. 714, Mohan Street, SURAT, Gujarat - 395007
  - Vihaan Pharmacy → 17, Venkataraman Marg, SURAT, Gujarat - 395007
  - Aayush Pharmacy → 812, Chada Path, SURAT, Gujarat - 395007
  1. "Fluconazole
  2. Clotrimazole
  3. Miconazole
  4. Itraconazole
  5. Terbinafine"

🌐 Generating Google Calendar events for your 7-day plan...
✅ Created 28 Google Calendar events for your 7-day plan.
```


Safari File Edit View History Bookmarks Window Help
49%
Wed 23 Apr 11:36 PM
calendar.google.com
Google Cal...
Day 1 Breakfast
Save
Apr 23, 2025
8:00am
to
8:30am
Apr 23, 2025
(GMT+05:30) India Standard Time - Kolkata
Time zone
All day
Does not repeat
Event details
Find a time
Guests
Add Google Meet video conferencing
Add guests
Add location
Guest permissions
Notification
30
minutes
Modify event
Invite others
See guest list
Add notification
Aditya Patel
Busy
Default visibility
Breakfast: Oatmeal with ginger and honey
Medications: "Fluconazole, Miconazole


calendar.google.com
Google Cal...
Day 1 Workout
Save
Apr 23, 2025
7:00am
to
7:30am
Apr 23, 2025
(GMT+05:30) India Standard Time - Kolkata
Time zone
All day
Does not repeat
Event details
Find a time
Guests
Add Google Meet video conferencing
Add guests
Add location
Guest permissions
Notification
30
minutes
Modify event
Invite others
See guest list
Add notification
Aditya Patel
Busy
Default visibility
Morning Workout: Gentle neck stretches (slowly tilt head to the side, then back, repeating 5 times)


Safari File Edit View History Bookmarks Window Help
49%
Wed 23 Apr 11:36 PM
calendar.google.com
Google Cal...
Day 1 Dinner
Save
Apr 23, 2025
8:00pm
to
8:30pm
Apr 23, 2025
(GMT+05:30) India Standard Time - Kolkata
Time zone
All day
Does not repeat
Event details
Find a time
Guests
Add Google Meet video conferencing
Add guests
Add location
Guest permissions
Notification
30
minutes
Modify event
Invite others
Add notification
See guest list
Aditya Patel
Busy
Default visibility
Dinner: Vegetable curry with turmeric, coconut milk, and brown rice
Medications: "Fluconazole, Miconazole


Google Cal...
Day 1 Workout
Save
Apr 23, 2025
7:00am
to
7:30am
Apr 23, 2025
(GMT+05:30) India Standard Time - Kolkata
Time zone
All day
Does not repeat
Event details
Find a time
Guests
Add Google Meet video conferencing
Add guests
Add location
Guest permissions
Notification
30
minutes
Modify event
Invite others
Add notification
See guest list
Aditya Patel
Busy
Default visibility
Morning Workout: Gentle neck stretches (slowly tilt head to the side, then back, repeating 5 times)

(6) Timeline/Gantt Chart:

Phase	Timeline	Duration	Key Activities
Research & Planning	Mar 01 – Mar 08, 2025	1 week	- Literature review - Requirement analysis - Dataset identification
Data Preparation	Mar 08 – Mar 12, 2025	2 weeks	- Data collection - Cleaning - Preprocessing - Feature engineering
Model Development	Mar 12 – Mar 22, 2025	2 weeks	- Model selection - Training - Validation - Performance tuning
System Integration	Mar 22 – Apr 01, 2025	2 weeks	- Developing user interface - Integrating model - Implementing extra features
Testing & Deployment	Apr 01 – Apr 07, 2025	2 weeks	- System testing - Debugging - Deployment - Feedback incorporation
Documentation	Apr 07 – Apr 17, 2025	1 week	- Project report - User manuals - Final presentations

(7) Future Scope:

1. Seamless Integration with Pharmacy APIs

To enhance user convenience, the system aims to integrate directly with online pharmacy platforms such as **PharmEasy**and **Blinkit**. This integration will enable users to:

- **Real-Time Medicine Availability:** Check the availability of prescribed medicines in nearby pharmacies.
- **Instant Ordering:** Add preferred medicines to the cart and place orders directly through the platform.
- **Automated Prescription Processing:** Utilize APIs to send prescriptions to pharmacies, streamlining the ordering process.

By leveraging these integrations, users can transition smoothly from diagnosis to treatment, reducing delays and improving adherence to medication regimens.

2. Automated Alerts to Nearby Hospitals for Critical Conditions

For cases where the system predicts a critical health condition, it plans to implement an automated alert mechanism that:

- **Identifies Nearby Hospitals:** Uses the user's pincode to locate nearby hospitals equipped to handle emergencies.
- **Sends Immediate Notifications:** Automatically alerts these hospitals with the user's critical health information, enabling prompt medical response.
- **Facilitates Emergency Services:** Coordinates with emergency services to ensure timely assistance to the user.

This proactive approach aims to bridge the gap between early diagnosis and emergency care, potentially saving lives by ensuring rapid medical intervention.

(8) Challenges Faced:

During the development of the Medicine Recommendation System, we encountered several challenges. Initially, our machine learning models exhibited overfitting, performing exceptionally well on training data but poorly on unseen data.This was primarily due to the limited and imbalanced nature of our dataset, which lacked sufficient entries for certain diseases and symptoms. To address this, we employed techniques like Synthetic Minority Over-sampling Technique (SMOTE) and generated synthetic data to balance the dataset and enhance model generalization.

Integrating the trained model into our web application also presented difficulties. Ensuring seamless communication between the frontend and backend required careful handling of data serialization and API development. We faced challenges in maintaining low latency and high responsiveness, especially when processing complex inputs. Through iterative testing and optimization, we achieved a stable and efficient integration, resulting in a robust and user-friendly application

(9) Conclusion:

the development of the Medicine Recommendation System has been a significant step toward integrating artificial intelligence into healthcare solutions. By leveraging machine learning models, we have created a platform that aids in early disease detection and provides personalized medication recommendations and diet plans. Despite challenges such as overfitting, limited data availability, and integration complexities, the project has achieved its primary objectives. Looking ahead, the planned integration with pharmacy APIs for instant medicine ordering and automated alerts to nearby hospitals for critical conditions will further enhance the system's utility. This project lays the groundwork for more accessible and proactive healthcare services, especially in underserved areas

(10) References:

- Scikit-learn: Machine Learning in Python. https://scikit-learn.org/ Wikipedia
- SMOTE: Synthetic Minority Over-sampling Technique. https://arxiv.org/abs/1106.1813
- Google Calendar API Documentation. https://developers.google.com/calendar
- Evidence-Based Medicine. https://en.wikipedia.org/wiki/Evidence-based_medicine Wikipedia
- List of Datasets for Machine Learning Research. https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research