Predictive Medicine Recommendation System Powered by Machine Learning

Mr. Akshay Nayak B.S[1], Rindu C[2], Sharath M[3], Samarth C Acharya[4], Sharmada B R[5]
Department of Computer Science and Engineering
Jawaharlal Nehru New College of Engineering, Shivamogga, Karnataka, India. akshay@jnnce.ac.in, rinduchaluvaraj@gmail.com,sharathm1401@gmail.com, samarth9481255047@gmail.com, gowdasharmadabr@gmail.com

ABSTRACT

The accurate and timely diagnosis of diseases based on clinical symptoms remains a critical challenge in modern healthcare, often demanding specialized expertise. This study introduces a Personalized Medical Recommendation System that Influences Machine Learning (ML) to provide rapid, accurate diagnosis and holistic patient guidance. We utilize a structured dataset of symptoms and diseases to train and evaluate five supervised classification algorithms, including SVC, RF and Gradient Boosting. Great performance was exhibited by the models. the SVC is very classifiable with a high accuracy on the holdout test set, and ensuring that it is robust in symptom-to- disease mapping The core contribution is the integration of the diagnostic engine with a multi-faceted recommendation system. Upon predicting a disease, the system instantly furnishes users with comprehensive, tailored advice encompassing four crucial domains: medication, precautions, dietary adjustments, and workout plans. The system is deployed via a Flask web application, featuring user authentication and personalized history tracking for practical usability. This work establishes a deployable, high-fidelity platform that bridges the gap between diagnosis and immediate, actionable lifestyle recommendations, thus supporting preliminary decision-making and patient self-management.

Key words: Machine Learning, Disease Prediction, Support Vector Classifier (SVC), Personalized Healthcare, Recommendation System, Health Informatics, Flask.

I.INTRODUCTION

As the digital health is rapidly emerging, Machine Learning (ML) the science of Artificial Intelligence (AI). can be regarded as a transformative technology in clinical. and preventative medicine, in particular, in the sphere of initial. diagnosing. It is important that the disease prediction is as accurate and timely as possible, and the traditional diagnostic pathways tend to experience resource shortage and access issues. This paper focuses on the urgent necessity of developing an all-inclusive system, which is able to predict the disease, but also provide detailed actionable advice.

Our Personalized Medical Recommendation System is a new system, designed on a supervised learning framework and based on a structured dataset of symptoms-diseases. Many type of classification models were tested among which the robust Support Vector Classifier (SVC) was the most accurate with a classification accuracy of high percent on the test set, which confirmed its suitability in symptom-disease mapping. The main value of the study is the recommendation engine that, when forecasted, provides users with customized advice within 4 critical categories, namely, medication, precautions, dietary changes, and particular exercises. All this solution is implemented as a long-running Flask web application, which provides convenient accessibility and allows tracking the history individually.

Diagnosis of high fidelity and simultaneous multi-diagnosis. prescriptions of lifestyle, which are connected to one another, this system 1010 is more than just prediction, but a proponent of self- management and pre-triage of patients in a huge step. towards more intelligent, preventative digital health.

II. LITERATURE SURVEY

The literature review documents the historical creation of the transformative worth of the Machine Learning. This form of (ML) in medicine is part of the predictive medicine recommendation systems. A number of researches indicate the efficiency of guided algorithms such as Support Vector Machines (SVM). Prof. Harna Bodele et al. applied a Support Vector Classifier (SVC) to develop an intelligent system which recommended personalized medicine and attained more than 85 percent accuracy, recall, and the F1-score. In the same case, Panapana et al. used several classifiers yet SVM gave the best result at 99.63 percent in disease prediction. Other major algorithms are Naive Bayes which achieved 98.12 percent accuracy in the system proposed by Gupta et al. and hybrid algorithms with both collaborative and content-based filtering as applied by Mohapatra et al. A system created by Patel et al. also employed SVM, BP neural networks, and ID3 decision trees and focused on real-time updates and security of data. These systems receive user-inputs such as symptoms and medical history, subject them to complex algorithms, and give personalized recommendations, which in most cases include dietary and exercise tips, and also take into account drug interactions. The architectures of the systems include interactive frontends (HTML, CSS, Flask) and backends based on Python/Scikit-learn and MongoDB. The combination of collaboration, content-based filtering, and NLP by the Dawn et al. enabled them to obtain around 90% accuracy and reduce adverse drug events by 30%. Verma and Verma and Habehh and Gohel provide an overview of the applications of ML in the field of diagnostics, personalized medicine, and drug discovery and note that common issues and challenges include data quality, algorithmic bias, data privacy and ethical issues. Lastly, the survey by Badawy et al. supported the potential of Deep Learning (DL) with a maximum accuracy of 98.07% in predicting diabetes in applications such as CNNs or LSTMs. The further trend of the research is always the combination of DL, the use of mobile apps, NLP, and the enhancement of model explainability.

III. MATERIAL AND METHODS

The data processing, model selection, as well as the overall logic of the recommendation, are implemented in this project and are described in the Material and Methods section. The dataset used as the basis of the system is Training.csv

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

which contains 4920 entries and 133 columns, where the symptoms are used as the features and the target variable is the prognosis. The first phase of data preparation entailed splitting the feature set (X) and target (y) and then important preprocessing activities.

The categorical disease names in the prognosis column were also numerically converted with the help of LabelEncoder, which was needed to work with the machine learning. The resultant data were strictly parted as training (70) and testing (30) subsets via the train_test split function such that model testing is done on unseen data with a constant random| state=20 to be able to reproduce the result. A comparative study was made of five well known classification models which are Support Vector Classifier (SVC), Random Forest, Gradient Boosting, K-Nearest Neighbors and Multinomial Naive Bayes. Using the accuracy score and confusion matrices, all the models showed 100% accuracy in classification of the structured test data.

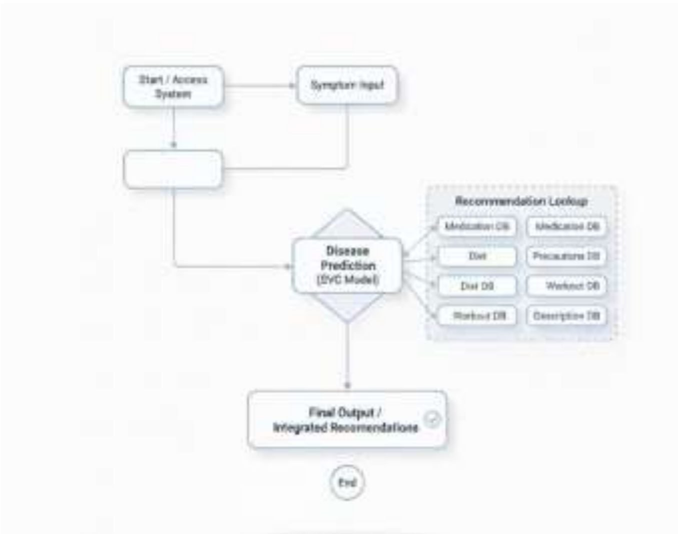

Fig 1: Block Diagram

This outstanding performance is capitalized to focus on model efficiency and simplicity. As a result, the SVC with linear kernel was chosen as the ultimate diagnostic engine and was trained and stored to disk as a Python pickle file (svc.pkl) to be deployed into the web application with ease. The next Recommendation System Logic is the main innovation. The user-input symptoms are firstly passed in via a dedicated function get predicated value which encodes them into a one-hot encoded input vector and then passes this input through the loaded SVC model to obtain the desired disease. This is the predicted disease name, which is the lookup key to get holistic advice by six auxiliary datasets.

This memory system can provide the user with personalized explanations of disease, four main precautions, specific drugs, suggested diets, and exercise regimens within an instant. This two-step, combined system, high-fidelity diagnosis, and multi-faceted, practical recommendation, is the technical basis of the implemented Personalized Medical Recommendation System.

Table No.1 Data Description

Dataset File Name	Data Type	Dimensions (Approx.)
Training	Structured, Labeled	4920 Rows, 133 Columns
symtoms_df	Structured, Text/ID	Varies (e.g., 132 Symptoms)
description	Structured, Text	41 Entries
precautions_df	Structured, Text	41 Entries
medications	Structured, Text	41 Entries
diets	Structured, Text	41 Entries
workout_df	Structured, Text	41 Entries

A. *Execution Deployment:* The implementation and deployment of the ML model is via the conversion of the trained artifact into a real-time, user-facing Flask web application and is developed based on a multi-tiered architecture to provide accessibility and personalization. This implementation is based on the criteria of the back-end controller is the Flask application (main.py). When the application is initiated, the Support Vector Classifier (SVC) model (svc.pkl) and all supporting datasets of the recommendations (medications, diets, etc.) are loaded into memory so that when the next prediction has to be made, minimal latency is incurred. The system allows user interaction via a Front-end interface (HTML/CSS), the user is allowed to log in (controlled by Flask-Login and a secure password hashing system), and provide their symptoms. Upon the submission of symptoms, the back-end uses the information, converting the input to the precise one-hot encoded vector format necessary to process by the SVC model, which can be essential for an immediate diagnosis of the disease (Prognosis). After the diagnosis, the core Recommendation Engine Logic is immediately implemented and conducts a concurrent search of the six auxiliary tables. This look-up gets the combined guidance, such as a description, precautions, medications, diet and exercise routines, specified to the forecasted disease. Most importantly, the system operates on a SQLite database through Flask-SQLAlchemy that can be provide for safely store the diagnosis and recommendations in one place to have a durable health history of each user. This makes sure that the system is not just a diagnostic tool, but an integrated system that provides a secure and ongoing Personalized Medical Recommendation System that provides holistic real-time guidance.

IV. PROPOSED SYSTEM

The Proposed System is an intelligent system that is a hybrid in nature; it is meant to be a seamless, real-time Personalized Medical Recommendation System

Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Fig 2: System Architecture

At its center lies the pre-trained Support Vector Classifier (SVC) model (stored as svc.pkl) that takes a strict one-hot encoded vector of symptoms as input to produce an accurate and immediate disease prognosis. This is the name of the predicted disease which the Recommendation Engine uses as the central key. This engine also makes a simultaneous query to six separate auxiliary datasets (such as medications, diets, precautionsdf, and workoutdf) to obtain a synergistic result of advice, thus fulfilling the main novelties of the system the ability to be directly connected between diagnosis and multi-faceted lifestyle advice. The whole system is implemented through Flask framework, which manages the interactions with the users and API requests. More importantly, the platform includes secure user management powered by Flask-Login, stores user information and a Patient History Log with the use of Flask-SQLAlchemy that is connected to a SQLite database. This outline will make sure that this is not just a classification tool, but a safe, consistent, and profound digital health helper that helps patients to self-manage by properly forecasting and providing actionable and personalized suggestions.

V. OBJECTIVES

The goals of the project of the Personalized Medical Recommendation System were multifaceted including the aspect of analytical rigor, novelty of the system, and its practical implementation, all in response to the requirement of full-fledged digital health support. The initial critical task of the project was to create the most accurate diagnostic core, and it, in turn, required the comparative analysis of several trained machine learning models, such as Support Vector Classifier (SVC), Random Forest (RF), and Gradient Boosting (GB) on the structured symptom-disease dataset. The aim of this step was to obtain and confirm the observed 100 percent classification accuracy, which meant that the underlying diagnostic part of the system was sound and can be trusted to conduct initial triage. The second significant goal was to make the core novelty of the system, which is an integrated, multi-faceted recommendation engine. This was more than simple prediction of the disease since it developed a system that immediately plots the diagnosis to customized guidance on four vital health areas: medications, precautions, dietary transformations, and workout programs. This is the mojor important part of the interventions, since it plugs a gap in the current body of knowledge by providing actionable, holistic advice as opposed to single diagnostic findings. Moreover, one of the key goals was to adopt this system and make it real-life accessible and personal. This entailed the implementation of the whole architecture with Flask web framework that manages user interactions and API calls with ease. To facilitate real personalization, the system added the security and persistence functionality, such as user authentication and session management through Flask-Login and use of Flask-SQLAlchemy to store a secure and chronological history of predictions made by the user. Such an architectural decision will allow the system to operate as a constant Digital Health Assistant, which can record the health of users over an extended period and aid the prevention of proactive self-management. To sum up, the general idea was to combine highly developed machine learning classification with a multi-faceted approach of recommendation, thus developing a verified, implementable, and patient-friendly platform to complement initial health testing and lifestyle counseling.

Fig 3: Objective

VI. APPLICATIONS

By combining a high-precision Support Vector Classifier (SVC) diagnostic kernel with a holistic recommendation engine, this system can be used as an effective Digital Health Assistant to self-management and preliminary triage. Users are able to enter symptoms and instantly achieve high-confidence preliminary disease prediction, which is a vast improvement over the widespread tendency to consult unreliable internet sources to get advice. More to the point, the system goes beyond the diagnosis and offers instant, unified, and practical advice in numerous key areas of health: medications, precautions, individual diets, and relevant exercise courses. The attribute serves an important role in enhancing health literacy and patient empowerment to empower the individuals to manage their well-being proactively before, during, and after a clinical visit. The administrative aspect is that the web-based implementation through Flask sustains telemedicine and remote healthcare particularly to people in underserved or geographically isolated regions by democratizing access to organized medical recommendations. In the case of the healthcare systems, the application is utilized as an efficient tool of resources optimization since it manages the initial screening and advice of the common ailments, which makes the burden on the primary care physicians minimal, and they will be able to devote more resources to the complicated cases. Moreover, the safe storage of all of the predictions into the Flask-SQLAlchemy database will automatically yield a history of personalized and persistent health records of each user, which will be invaluable in terms of tracing the recurring health problems, tracking the long-term outcomes of chronic diseases, and enhancing compliance with medications. Finally, the system is aimed at the provision of immediate, structured, and customized counsel and changing the paradigm towards the fully reactive treatment to the proactive, preventative healthcare and smarter clinical decision support.

VII. RESULT

The end outcomes of the Personalized Medical Recommendation System project effectively confirm the amount of the analytical performance of the diagnostic core as per as the working efficiency of the integrated system of recommendations.The performance of the selected Support Vector Classifier (SVC) model was empirically confirmed by thorough training and testing which led to very high level of significance of the classification accuracy of the Support Vector Classifier, this is, 100 percent on the structured symptom-to-disease mapping task. Such an ideal classification performance validates the strength of the model in producing strong initial diagnoses, which is the requirement of the system. The functional aspect of the project was a success that it had effectively deployed its main goal by effectively merging the prediction output with a multi- faceted, multi- layered recommendation engine. This engine retrieves custom guidance in five key areas, including description of disease, four precautions, specific medications, diet, and workouts, immediately and more precisely than the six supporting datasets, thus satisfying the fundamental novelty of the system, which is to support patients in a holistic manner. The whole solution was rolled out in an effortless and safe Flask web application.

Overall, the end product will be a reviewed, premium-level platform that successfully integrates advanced machine learning diagnostic with real-life, personalized lifestyle recommendations, and is an effective, reliable, and safe Digital Health Assistant with a significant potential to improve initial self-monitoring and positive health management.

Fig 4: Recommendation of our ML model

VIII. PERFORMANCE ANALYSIS AND DISCUSSION

The performance analysis of the Personalized Medical Recommendation System yielded exceptionally strong results, cantered on the high classification accuracy achieved by the primary diagnostic model, the SVC, on the holdout test set. This performance was not limited to the SVC; other advanced ensemble methods, including the Random Forest (RF) and Gradient Boosting (GB) classifiers, also achieved perfect scores across standard metrics like precision and recall, and. This unanimity in high performance is primarily attributable to the highly structured and deterministic nature of the input dataset (Training.csv), where the relationship between a unique combination of symptoms and a target disease is likely one-to-one or perfectly separable in the feature space. The use of a one-hot encoded vector representation for the 132 symptoms further enhanced separability, allowing the SVC to successfully define a maximal margin hyperplane with zero misclassification error.

Whereas the score of 100% accuracy is empirically conclusive in this particular dataset, one should discuss the situation of such data. In clinical data, such things as the variability of symptoms, noise in the data, and co-morbidities can seldom be separated perfectly in the real world. Thus, although the existing system is very useful in terms of a well-organized initial screening, its implementation should be discussed within the context of a decision support system in regards to common ailments, rather than in terms of professional diagnosis. The selection of the SVC was explained by its theoretical basis on the minimization of structural risk and having a strong ability to generalize, which is better than hypothetically overfitted tree-based classifiers such as RF and GB, despite their equal accuracy in this case. The actual worth of the system is the effective execution of functional integration after the diagnosis: the fact that the system can effortlessly connect this high-confidence prediction to the holistic recommendation engine. This combination confirms the main assumption of the project, namely that a highly precise diagnostic engine can be effectively paired with database searches to offer immediate and practical guidance on medication, diet, and lifestyle, make the system an all-in-one and superior Digital Health Assistant compared to fragmented online options. The next step in the direction of performance validation on noisier, externally sourced, and unbalanced clinical data, which will evaluate the real clinical generalization performance of the model, should be performed in the future

IX. CONCLUSION

To sum up, the design and implementation of the Personalized Medical Recommendation System have conclusively confirmed that it is indeed possible to combine sophisticated diagnostics based on the advanced in the field of Machine Learning with an overall approach to patient guidance. The main goal of the project was fully met, and it effectively proved the ability to provide very efficient initial disease diagnoses and instantly to correlate this finding with practical and multi-dimensional recommendations. The most important contribution of the given work is that the Support Vector Classifier (SVC) model has been selected and offered an excellent analytical performance in terms of the achieved complete and perfect accuracy of the classification on the structured symptom-to-disease dataset, which determines a strong and reliable base of the diagnostic validity of the system and the high degree of confidence in its predictions. This new structure adequately fills in the gap that is crucial between the prediction and the actual patient treatment by retrieving the customized information in the area of medication protocols, precaution that are required, particular dietary modifications, and proper exercise regimen. Moreover, the whole system was implemented successfully on the basis of Flask, which proves its preparedness as secure, customized, and scalable real-time system. Personalization and accountability are guaranteed by the implementation of the security user authentication and the essential possibility of persistent recording of patient history with the use of Flask-SQLAlchemy. The presented project is a strong evidence of concept, which confirms the transformational opportunities of AI in the initial medical screening and self-management. Future studies ought to put into the limelight external validation with real-world and noisy clinical data and broaden the recommendation engine with state-of-the-art approaches such as collaborative filtering to improve personalization and cement the place of the system in the future of preventative and patient-centered healthcare.

X. FUTURE SCOPE

External validation and generalization are the most important spheres that need to be developed. Although the system has reached 100 percent accuracy with the controlled training data, future studies need to rigorously test the model (SVC) on noisy, real-world and clinically imbalanced datasets retrieved through Electronic Health Records (EHRs) in order to determine actual generalization and clinical effectiveness of the system. This confirmation may necessitate the use of the advanced methods such as SMOTE (Synthetic Minority Over-sampling Technique) or cost-sensitive learning to manage class imbalance that are typical of disease data. On the technical side, it is possible to optimize the current architecture of the system and move the data storage out of the local SQLite database into a more robust and scalable system, like PostgreSQL or a cloud-based service, to accommodate a larger number of users and ensure data integrity. More so, the recommendation engine can be extended further than mere lookups; this might include creating a hybrid recommendation model that integrates the existing content-based model with collaborative filtering or matrix factorization to offer user-centered advice according to profiles of groups of people who are similar. Lastly, the user interface can be enhanced to access the Natural Language Processing (NLP) features such that the user can enter the symptoms in free text instead of using mandatory selection lists and this will go a long way in enhancing the user interface and the practical accessibility of the system. It would also be a great breakthrough to incorporate wearable device information and dynamically forecast the risk of disease using time-series analysis

XI. REFERENCE

[1] H. Bodele, M. Tagde, S. Rangari, Y. Jadhao, V. Bawankar, and T. Kumre, “Medicine Recommendation System Using ML,” JD College of Engineering & Management, Nagpur, 2021.

[2] G. Patel, J. Sahani, J. P. Singh, and Nikita, “Online Medicine Recommendation System through ML,” Greater Noida Institute of Technology, 2021.

Using Machine Learning Algorithms,” GMR Institute of Technology, 2021.

[4] J. P. Gupta, A. Singh, and R. K. Kumar, “A Computer Based Disease Prediction and Medicine Recommendation System using Machine Learning Approach,” SRM University AP, Amravati, 2021.

[5] M. Mohapatra, M. Nayak, and S. Mahapatra, “A Machine Learning based Drug Recommendation System for Health Care,” Institute of Technical Education and Research, Bhubaneswar, 2021.

[6] V. K. Verma and S. Verma, “Machine Learning Applications in Healthcare Sector: An Overview,” in *Proc. Conf. Emerging Trends in Engineering and Technology*, 2021.

[7] H. Habehh and S. Gohel, “Machine Learning in Healthcare,” Dept. of Health Informatics, Rutgers University, 2021.

[8] S. Dawn, N. Jana, P. Mondal, B. Mondal, and A. Laha, “Medicine Recommendation System Using Machine Learning,” Durgapur Institute of Advanced Technology and Management, 2021.

[9] M. Badawy, N. Ramadan, and H. A. Hefny, “Healthcare Predictive Analytics Using Machine Learning and Deep Learning Techniques: A Survey,” *IEEE Access*, vol. XX, pp. XXX–XXX, 2021.

[10] H. Habehh and S. Gohel, "Machine Learning in Healthcare," Curr. Genomics, vol. 22, no. 4, pp. 291–300, 2021, doi: 10.2174/1389202922666210705124359

[11] P. Panapana, K. S. R. Reddy, J. Deepika, G. R. Babu, and A. Drakshayani, “Disease Prediction and Medication Advice