Disease Prediction and Drug Recommendation Prototype Using Machine Learning

ILAPAVULOORI RAGHUTEJA[1], B. MURALI[2]
[1]*PG Student, CN, Quba College of Engineering & Technology*
[2]*Associate professor, CN, Quba College of Engineering & Technology*

Abstract— Online recommender systems are being used increasingly often for hospitals, medical professionals, and drugs. Today, the great majority of consumers look online before asking their doctors for prescription suggestions for a range of health conditions. The medicine recommendation system gives the patient reliable information about the medication, the dosage, and any possible adverse effects. In this project, the power of machine learning techniques were used to recommend the drug based on symptoms. We have used Random Forest Classifier and Support Vector Machine (SVC) algorithm for disease prediction and drug recommendation prototype among those we got best precision score of 94% with Support Vector Classifier.

Index Terms— Online Recommender Systems, Healthcare Recommendation Systems, Medical Recommendation Systems, Drug Recommendation Systems, Machine Learning in Healthcare

I. INTRODUCTION

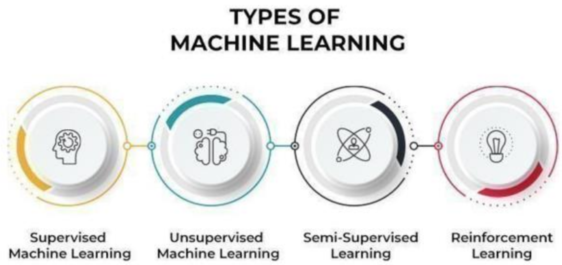
1.1 MACHINE LEARNING AND ITS IMPORTANCE

Machine learning (ML) is the study of computer algorithms that improve its input automatically through experience intelligence. Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or infeasible to develop conventional algorithms to perform the needed tasks.

Machine learning (ML) involves computers discovering how they can perform tasks without being explicitly programmed to do so. It involves computers learning from data provided so that they carry out certain tasks. For simple tasks assigned to computers, it is possible to program algorithms telling the machine how to execute all steps required to solve the problem at hand; on the computer's part, no learning is needed. For more advanced tasks, it can be challenging for a human to manually create the needed algorithms. In practice, it can turn out to be more effective to help the machine develop its own algorithm, rather than have human programmers specify every need step. The discipline of machine learning employs various approaches to help computers learn to accomplish tasks where no fully satisfactory algorithm is available. In cases where vast numbers of potential answers exist, one approach is to label some of the correct answers as valid. This can then be used as training data for the computer to improve the algorithm(s) it uses to determine correct answers.

Machine learning revolutionizes industries by enabling data-driven decision-making. In fields like finance, healthcare, and marketing, it offers valuable insights that drive strategic planning and optimize operations. For instance, financial institutions leverage machine learning algorithms to detect fraudulent activities, while healthcare providers utilize predictive models to diagnose diseases earlier and personalize treatment plans. By analyzing vast amounts of data, machine learning uncovers patterns and trends that humans might overlook, leading to more informed and efficient decision-making processes.

1.2 TYPES OF MACHINE LEARNING Early classification for machine learning approaches sometimes divided them into three broad categories depending on the nature of “feedback” available to the learning system.



II. LITERATURE SURVEY

1. Drug Recommendation System for Diabetes Using Collaborative Filtering and Clustering Approach, L. F. G. Morales, P. Valdiviezo-Diaz, R. Reategui, and L. Barba- Guaman

In this research, the authors presented a novel approach to recommending drugs for diabetes management. Utilizing collaborative filtering and clustering techniques, their system aims to provide personalized recommendations tailored toindividual patients. The study also includes an evaluation of the system's performance, assessing its efficacy in assisting healthcare providers and patients in making informed treatment decisions for diabetes.

2. Recommender Systems in the Healthcare Domain: State-of-the-art and Research Issues, T. N. T. Tran, A. Felfernig, C. Trattner, A. Holzinger

The paper explored the application of recommender systems in the healthcare domain, presenting an overview of the current state-of-the-art and discussing various research issues in this area.The authors delve into the role of recommender systems in healthcare, highlighting their potential to assist healthcare professionals in making informed decisions, improving patient outcomes, and enhancing overall healthcare delivery. They discuss how recommender systems can leverage patient data, medical knowledge, and other relevant information to provide personalized recommendations for treatments, interventions, and healthcare services. Furthermore, the paper examines the challenges and opportunities associated with the development and implementation of recommender systems in the healthcare domain. It addresses issues such as data privacy, security, ethical considerations, and the integration of recommender systems into existing healthcare workflows and infrastructures.

III. SYSTEM REQUIREMENT SPECIFICATIONS

3.1 HARDWARE REQUIREMENTS SPECIFICATIONS

System	:	Pentium i3 Processor
Hard Disk	:	500 GB
Monitor	:	15'' LED
Input Devices	:	Keyboard, Mouse
Ram	:	8 GB

3.2 SOFTWARE REQUIREMENTS SPECIFICATIONS

Operating System	:	Windows 10
Pro Programming Language	:	Python 3.10.9
Web Framework	:	FLASK
Frontend	:	HTML, CSS, Java Script

IV. SYSTEM ANALYSIS

4.1 EXISTING SYSTEM

The goal of recommender systems is to give customers individualized recommendations and offer solutions to reduce the growing problem of information overload online. The health related information supplied through online comments or surveys includes veiled assumption designs that come from completely unrelated medical sources and assist the pharmaceutical sector. This system employs clustering approach likely involved grouping patients or symptoms into clusters based on similarity, which can be useful for understanding patterns in the data but might not directly recommend specific drugs for specific symptoms.

DISADVANTAGES OF EXISTING SYSTEM

Lack of Specificity: Clustering groups patients or symptoms based on similarity, which may not always lead to specific and accurate drug recommendations for individual patients with unique symptoms or conditions

Over-Specialization: Some recommender systems may become too specialized, catering to the immediate preferences of users without considering the long-term or evolving interests. This can result in users being pigeonholed into narrow categories and missing out on a broader range of recommendations.

Popularity Bias: Recommender systems may exhibit a bias towards popular items or those with a higher number of interactions. This bias can make it challenging for less popular or niche items to gain visibility, perpetuating a cycle where popular items receive more recommendations, further solidifying their popularity.

Potential for Misclassification: Clustering may misclassify patients or symptoms into incorrect clusters, leading to inaccurate drug recommendations. Limited Adaptability: Clustering models may not easily adapt to new data or changes in medical knowledge, potentially leading to outdated or ineffective recommendations over time.

4.3 PROPOSED SYSTEM

In the realm of healthcare, timely and accurate drug recommendations during medical emergencies can significantly impact patient outcomes. This project presents a robust "Disease Prediction and Drug Recommendation prototype using Machine Learning approaches," implemented in Python. The system leverages powerful classification algorithms, namely the Support Vector Classifier, attaining remarkable result on both training and test datasets.

Step 1: Load the important libraries

Step 2: Import dataset and extract the X variables and Y separately.

Step 3: Divide the dataset into train and test

Step 4: Initializing the SVM classifier model >> svm_clf = svm.SVC(kernel = 'linear')

Step 5: Fitting the SVM classifier model

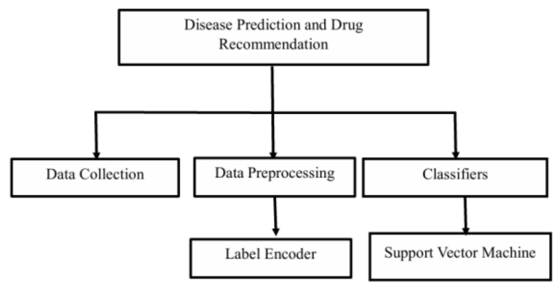
Step 6: Coming up with predictions

Step 7: Evaluating model's performance

V. SYSTEM DESIGN

5.1 ARCHITECTURAL DESIGN

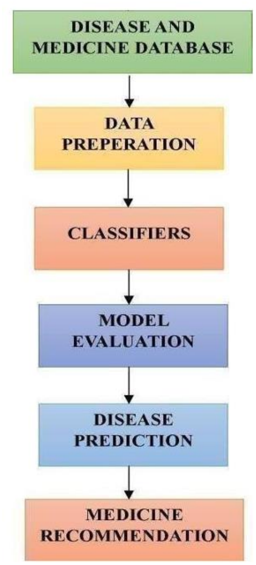
This architecture provides a framework for building a prototype for disease prediction and drug recommendation using machine learning approaches. Each component can be further refined and customized based on specific requirements and constraints. Below is the architectural design for such system.



Data Preprocessing: Data preprocessing is a critical step in preparing raw data for analysis and modeling, involving a series of operations aimed at ensuring data quality, consistency, and suitability for the intended task. Additionally, data preprocessing often involves encoding categorical variables to convert them into numerical representations, and splitting the dataset into training and testing sets for model evaluation.

WORKFLOW

The drug recommendation system based on symptoms employs a systematic workflow to provide reliable medication suggestions to users.



VI. TESTING

Testing is a crucial aspect of software development that involves verifying and validating the functionality, performance, and quality of a software product. In a paragraph, testing can be described as the systematic process of executing a program or application with the intent of finding defects or errors. This process typically involves running the software under various conditions and inputs to assess its behavior and ensure that it meets the specified requirements and standards. Testing can encompass different techniques, such as unit testing, integration testing, system testing, and acceptance testing, each targeting specific aspects of the software's functionality and performance. Ultimately, the goal of testing is to identify and resolve any issues before the software is released to end users, thereby enhancing its reliability, usability, and overall quality.

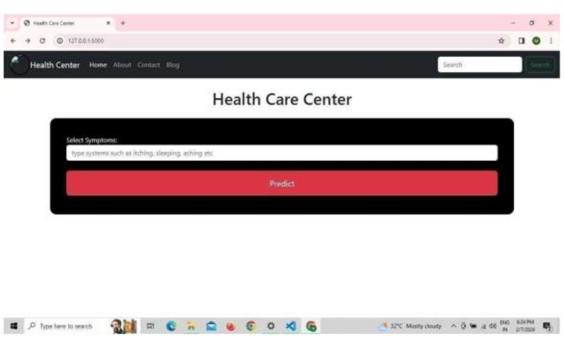
There are mainly three types of testing:

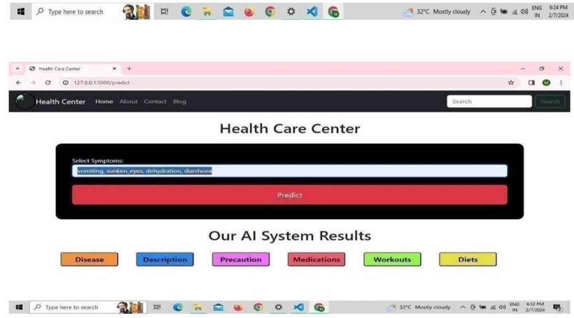
- Unit Testing
- Integration Testing
- System Testing

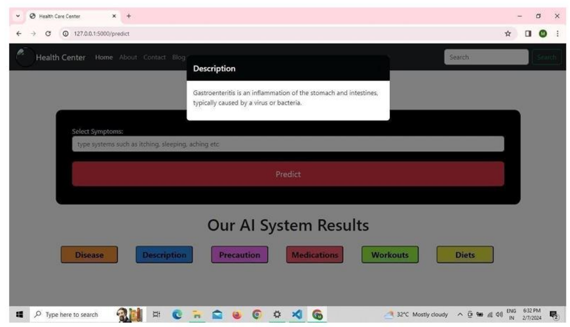
Unit Testing: Unit testing is a foundational practice in software development, focusing on validating individual units or components of code in isolation. It ensures that each unit functions correctly and produces the expected output, helping catch bugs early in the development process. By promoting modular design and providing rapid feedback to developers, unit testing enhances software quality and maintainability.

Integration Testing: Integration testing evaluates the interaction and collaboration between different modules or units within the software system. It verifies integration points and interfaces, ensuring smooth data flow between components and detecting defects arising from interactions. By facilitating the seamless integration of software modules, integration testing contributes to the development of robust and interoperable systems.

VII. RESULTS AND SCREENSHOTS







RESULT COMPARISION

The following is the comparision table for the dataset.

ALGORITHM	PRECISION	ACCURACY	SUPPORT
VECTOR RECALL CLASSIFIER	0.94	0.93	0.81
RANDOM FOREST	0.93	0.81	0.81

CONCLUSION

Our project addresses the growing reliance on online recommender systems in healthcare, where patients increasingly seek medication recommendations prior to consulting healthcare professionals. By harnessing machine learning techniques, specifically the Support Vector Classifier (SVC) algorithm, we aimed to enhance the drug recommendation process based on symptoms. Our analysis revealed that the Support Vector Classifier algorithm achieved the highest precision of 94%, showcasing its efficacy in providing accurate medication suggestions. This underscores the potential of machine learning in improving healthcare decision- making and empowering patients with reliable information. Moving forward, our work paves the way for further advancements in personalized healthcare solutions and underscores the importance of integrating technology in modern medical practices.

FUTURE SCOPE

In the future, this online medicine recommendation system using machine learning techniques could be enhanced by incorporating additional patient data for greater personalization, such as medical history and lifestyle factors. Real-time updates from electronic health records and medical literature could ensure that recommendations are based on the most current information. Implementing a feedback loop for patients to provide input on recommended medications could improve the system's accuracy over time. Integration with telemedicine platforms would provide immediate access to recommendations from healthcare professionals. Monitoring medication adherence and integrating with wearable devices for real- time health data could further personalize recommendations. Collaboration with healthcare providers to align recommendations with clinical guidelines and expanding the system to cover a wider range of conditions and medications would benefit a larger population of patients.

REFERENCES

[1] L.F.G Morales, P. Valdiviezo-Diaz, R. Reategui, and L. Barba-Guaman, “Drug recommendation system for diabetes using a collaborative filtering and clustering approach: Development and performance evaluation,”J. Med. Internet Res., vol. 24, no. 7, Jul. 2022, Art. No. e37233.

[2] T.N.T. Tran, A. Felfernig, C. Trattner, and A. Holzinger, “Recommender systems in the healthcare domain: State-of-the-art and research issues,” J. Intell. Inf. Syst., vol. 57,no. 1,pp. 171-201, Aug.2021.

[3] C. R. Olsen, R. J. Mentz, K. J. Anstrom, D. Page, and P. A. Patel, “Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure,” Amer.Heart J., vol. 229, pp. 1-17, Nov. 2020.

[4] V. Mudaliar, P. Savaridaasan, and S. Garg, “Disease prediction and drug recommendation Android application using data mining (virtual doctor),” Int. J. Recent Technol. Eng. (IJRTE),vol. 8, no. 3, pp. 6996-7001, Sept. 2019.

[5] Q. Han, M. Ji, I. M. de Rituerto de Troya, M. Gaur, and L. Zejnilovic, “A hybrid recommender system for patient-doctor matchmaking in primary care,” in Proc. IEEE 5th Int. Conf. Data Sci.Adv. Anal. (DSAA), Oct.2018, pp. 481-490.

[6] Y. Zhang, M. Chen, D. Huang, D. Wu, and Y. Li, “iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization,” Future Gener. Comput. Syst., vol. 66, pp. 30-35, Jan. 2017.

[7] K. Feldman, D. Davis, and N. V. Chawla, “Scaling and contextualizing personalized healthcare: A case studyof disease prediction algorithm integration,” J. Biomed. Informat.,vol. 57, pp. 337-385, Oct. 2015.

[8] S. Bhat and K. Aishwarya, “Item-based hybrid recommender system for newly marketed pharmaceutical drugs,” in Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)