IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

MULT-DISEASE DETECTION USING MACHINE LEARNING AND STREAM- LIT

Akash Sharma[1], Rohit Singh Bani[2], Sonam Mishra[3]

[1,2,3] Department of Data Science, NIET College, Greater Noida, India
[4]Head of Department of Computer Science and Engineering, NIET College, Greater Noida, India

Abstract: A comprehensive effort called Multiple Disease Prediction utilizing Machine Learning, Deep Learning, and Streamlit aims to forecast multiple diseases, such as diabetes, heart disease and Parkinson's disease. Support Vector Machine (SVM) and logistic regression are two machine learning algorithms that are used in this project. Streamlit Cloud and the Streamlit library are used to deploy the models, offering an intuitive interface for illness prediction. There are three disease options available on the application interface: diabetes, Parkinson's diseases After selecting a specific disease, the user is prompted to enter the important parameters needed for the predictive model. The application quickly produces the disease prediction result after the parameters are supplied, showing whether or not the person is impacted by the condition.

INTRODUCTION

The goal of the "Multiple Disease Prediction using Machine Learning, Deep Learning and Streamlit" project is to forecast the following three conditions: diabetes, heart disease, and Parkinson's disease. Machine learning algorithms, such as Support Vector Machine (SVM) for diabetes and Parkinson's disease and Logistic Regression for heart disease are used in the construction of the prediction models. Streamlit Cloud and the Streamlit library are used in the application's deployment. In order to prepare the data for training and testing the prediction models, the project first gathers pertinent data from Kaggle.com. A distinct machine learning algorithm that is best suited for that particular disease handles each forecast of that disease. three options are available on the program interface, each of which represents a different ailment. The program asks for the parameters needed by the matching model to estimate the disease outcome when a user chooses a certain disease. When the user enters the necessary parameters, the application uses the input to display the prediction result.The process of creating interactive and user-friendly web apps is made easier by the Streamlit library. Through the use of machine learning techniques and Streamlit to streamline the deployment process, this project seeks to deliver precise and easily navigable predictions for a variety of ailments. With the help of the application's user-friendly interface, users may input parameters relevant to an illness and receive prediction results, which makes early detection and proactive healthcare management easier

METHODOLOGY

The following succinctly describes the Multiple Disease Prediction project's methodology:

1. Data Collection: Kaggle.com, a well-known website for dataset access, is the source of the data. Particularly, information on diabetes, heart illness, lung conditions, Parkinson's disease, and breast cancer is gathered.
2. Data Preprocessing: To guarantee the quality and appropriateness of the acquired data for training the machine learning models, preprocessing is applied. This covers managing missing values, eliminating duplicates, and carrying out feature scaling or data normalization.
3. Model Selection: For every illness prediction task, a different set of machine learning algorithms is selected. The algorithms for different diseases, such as Support Vector Machine (SVM) and Logistic Regression, are chosen based on their effectiveness and fit for the particular prediction tasks.
4. Training and Testing: Sets for training and testing are created from the preprocessed data. The testing data is used to assess the models' performance after they have been trained using the training data. The evaluation metric used to assess each model's performance is accuracy.

PROBLEM STATEMENT

Create a machine learning application that predicts a variety of diseases, such as diabetes, heart disease, lung disease, Parkinson's disease, and breast cancer, by utilizing Support Vector Machine (SVM) and Logistic Regression. Based on the trained models, the application should enable users to enter pertinent parameters for a particular disease and provide a precise estimate of whether a given person is impacted by the disease. By enabling early disease identification and prediction through machine learning algorithms and optimizing the prediction process with an intuitive user interface, the initiative seeks to enhance healthcare outcomes.

INPUT AND OUTPUT

Input Design: User input in the form of disease-specific parameters is needed for the Multiple Disease Prediction system. The system asks for the necessary parameters when the user chooses a certain ailment from the choices menu. It should be ensured by the input design that the user can effortlessly supply the necessary data. three disease selections are available on the application's user interface menu: diabetes, Parkinson's disease and heart disease. The application asks for the necessary parameters for a given disease prediction when the user clicks on one of the diseases. The factors that are sought should be relevant and essential for precise illness prediction, according to the input design. An intuitive method should be available for the user to enter the parameters.

Output Design: Whether or not a person is impacted by the chosen disease is predicted by the Multiple Disease Prediction system. The output design ought to display the outcome in an easily comprehensible manner. Once the user has supplied the settings, the system ought to show the output. The output ought to be shown on the user interface so that the user may quickly understand the outcome of the prediction. Overall, the output design clearly displays the prediction result on the user interface, and the input design makes sure the user can enter the required parameters for illness prediction.

RESULTS

Heart Disease Prediction: Achieved an accuracy of 83.90%.

Diabetes Prediction: Attained an accuracy of 77.27%.

Parkinson's Disease Prediction: Obtained an accuracy of 82.05%.

CONCLUSION

To sum up, in order to create a disease prediction system, our project used machine learning algorithms including Support Vector Machine (SVM) and Logistic Regression. Diabetes, heart disease and Parkinson's disease were the three ailments that the system concentrated on. To guarantee data quality, we gathered information from Kaggle.com and preprocessed it.The application interface of the system is user-friendly, featuring a menu that provides choices for every illness. The user is prompted to provide the pertinent parameters for the prediction model when they select a certain ailment. The system shows the anticipated disease outcome as soon as the parameters are entered. The results show how well the machine learning algorithms forecast the chosen diseases based on the accuracy rates attained. It is crucial to remember that the accuracy values could change based on the particular dataset and model training procedure. All things considered, this research shows how machine learning and streamlit libraries may be used to create disease prediction models.

REFERENCES

[1] Support Vector Machine (SVM): Corinna Cortes and Vladimir Vapnik (1995). Support-vector networks. Machine Learning, 20(3), 273-297.

[2] Logistic Regression: Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). Applied Logistic Regression (3rd ed.). John Wiley & Sons

[3] Streamlit: Streamlit Documentation. https://docs.streamlit.io/

[4] Kaggle: Kaggle website. https://www.kaggle.com/

[5] Data sources: You can provide the specific datasets you used from Kaggle.com, mentioning the authors or contributors of the datasets.

[6] Zhang, Y., & Ghorbani, A. (2019). A review on machine learning algorithms for diagnosis of heart disease. IEEE Access, 7, 112751-112760.

[7] Arora, P., Chaudhary, S., & Rana, M. (2020). Prediction of diabetes using machine learning algorithms: A review. Journal of Ambient Intelligence and Humanized Computing, 11(6), 2575-2589.

[8] Kaur, H., Batra, N., & Rani, R. (2020). A systematic review of machine learning methods for breast cancer prediction. Journal of Medical Systems, 44(11), 1-15.

[9] Gupta, D., & Rathore, S. (2021). A comprehensive review on machine learning algorithms for kidney disease diagnosis. Journal of Medical Systems, 45(1), 1-17.

[10] Saeed, A., & Al-Jumaily, A. (2020). Machine learning techniques for Parkinson's disease diagnosis using handwriting: A review. Computers in Biology and Medicine, 122, 103804.