e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:12/December-2024 Impact Factor- 8.187 www.irjmets.com

DISEASE PREDICTION USING SYMPTOMS

Jyoti Kanjalkar[*1], Prasad Mandi[*2], Mayur Chaudhari[*3], Mansi Bhandari[*4], Shreya Sidnale[*5]

[*1]Head Of Department, Department Of CSE(Artificial Intelligence & Machine Learning) Vishwakarma Institute Of Technology, Pune, India.

[*2,3,4,5]Student, Department Of CSE(Artificial Intelligence & Machine Learning) Vishwakarma Institute Of Technology, Pune, India.

ABSTRACT

In our disease prediction project, we aim to create a system that can predict diseases based on symptoms reported by patients. The system uses advanced machine learning techniques to analyze symptom data and generate predictions. By accurately identifying potential diseases early on, the system can help healthcare professionals make informed decisions and provide timely treatment to patients.

To achieve this goal, we collect and preprocess a dataset containing symptom information and corresponding disease diagnoses. We then train machine learning models, such as Support Vector Machine (SVM) and Logistic Regression, using this data. These models learn from the patterns in the symptom data and can predict the likelihood of various diseases based on new symptom inputs.

The project also focuses on developing a user-friendly interface that allows healthcare professionals to input patient symptoms easily and receive prediction results quickly. This interface enables seamless interaction between users and the prediction system, facilitating efficient disease diagnosis and treatment planning.

Overall, our disease prediction project has the potential to revolutionize healthcare by enabling early disease detection and intervention. By leveraging the power of machine learning, we aim to improve patient outcomes and contribute to better healthcare delivery.

I. INTRODUCTION

This project aims to use data analysis and machine learning to predict health conditions like diabetes, heart disease, and stroke. We'll look at important information like age, gender, blood pressure, and blood sugar levels to understand who might be at risk. By building computer models, testing them, and making sure our predictions are easy to understand, we hope to help doctors identify and prevent these health problems early. We'll also prioritize privacy, fairness, and continuous improvement throughout the project to ensure responsible and effective use of our predictions.

The disease prediction system based on symptoms leverages machine learning algorithms to analyze patient data and predict the likelihood of specific diseases. This system integrates three distinct datasets for heart disease, diabetes, and stroke, each containing relevant symptoms and associated medical information. By training individual models on these datasets, the system can accurately assess symptoms presented by patients and provide predictions regarding the potential presence of these diseases.

Utilizing a computer-based infrastructure, the system processes patient input, which may include symptoms reported by individuals or clinical observations recorded by healthcare professionals. These inputs are then compared against the trained models, which have learned patterns and correlations between symptoms and disease outcomes from the respective datasets. Through this analysis, the system generates predictions for each disease, providing valuable insights for early detection and intervention.

II. RELATED WORK

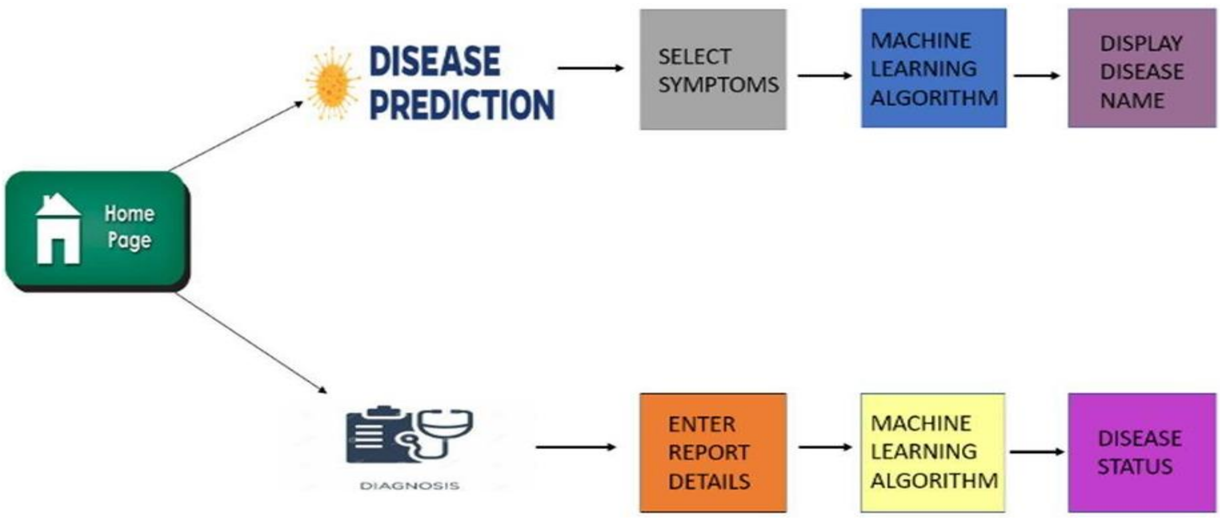
Algorithm Description:

In our disease prediction project, Support Vector Machine (SVM) and Logistic Regression are key algorithms employed for predictive modeling.

Support Vector Machine (SVM) is a supervised learning algorithm used for classification and regression tasks. In our project, SVM is utilized for binary classification tasks, such as predicting the presence or absence of a disease based on input symptoms. SVM works by finding the optimal hyperplane that best separates the data points into different classes. It aims to maximize the margin between the hyperplane and the nearest data points of each class, thereby improving the generalization performance of the model. SVM is known for its effectiveness in handling high-dimensional data and is robust against overfitting.

Logistic Regression, on the other hand, is a statistical method used for binary classification tasks. Despite its name, logistic regression is a linear model that predicts the probability of a binary outcome based on input features. In our project, logistic regression is applied to estimate the probability of a patient having a certain disease given their reported symptoms. The model calculates the log-odds of the probability using a logistic function, which maps the output to the range [0,1]. Logistic regression is interpretable, computationally efficient, and well-suited for cases where the relationship between input features and the outcome is linear or can be approximated linearly.

Architecture Design :



Classes & Methods:

1. DataLoader Class:

load_data: Method to load symptom data from a file or database.

clean_data: Method to preprocess and clean the loaded data, handling missing values and outliers.

split_data: Method to split the preprocessed data into training and testing sets.

2. FeatureEncoder Class:

encode_symptoms: Method to encode categorical symptom data into numerical format using techniques like Label Encoding or One-Hot Encoding.

scale_features: Method to scale numerical features to a standard range (e.g., using Min-Max scaling or Standardization).

3. ModelTrainer Class:

train_model: Method to train machine learning models (e.g., SVM, Logistic Regression) using the encoded and scaled data.

evaluate_model: Method to evaluate the trained models' performance using metrics like accuracy, precision, recall, and F1-score.

4. DiseasePredictor Class:

predict_disease: Method to accept symptom inputs, preprocess them, encode and scale them, and then use the trained model to predict the likelihood of various diseases.

interpret_results: Method to interpret the prediction results and provide insights or recommendations based on the predicted disease probabilities.

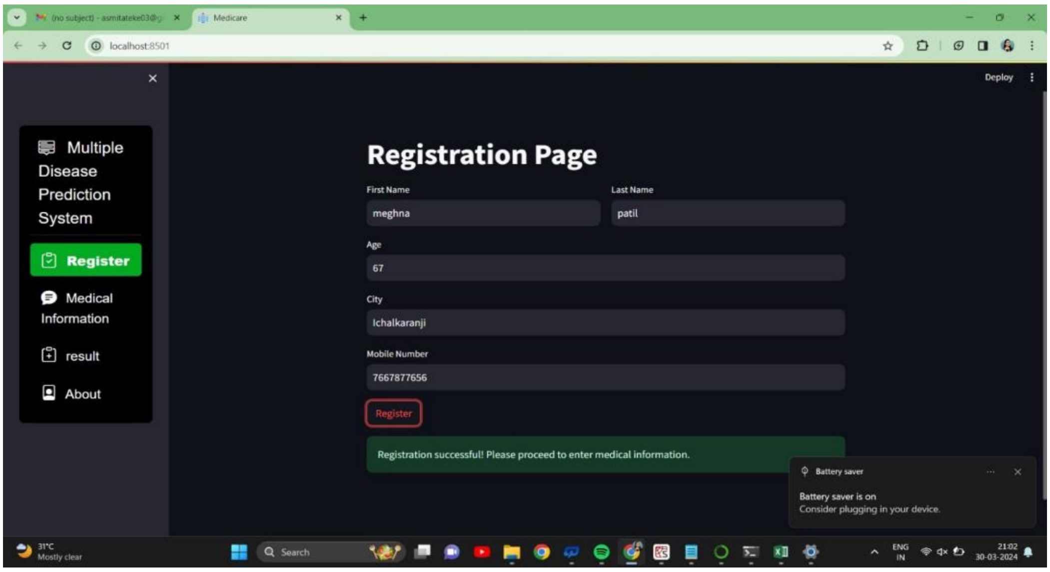
5. UserInterface Class:

get_symptoms_input: Method to interact with users and collect symptom inputs through a graphical or command-line interface.

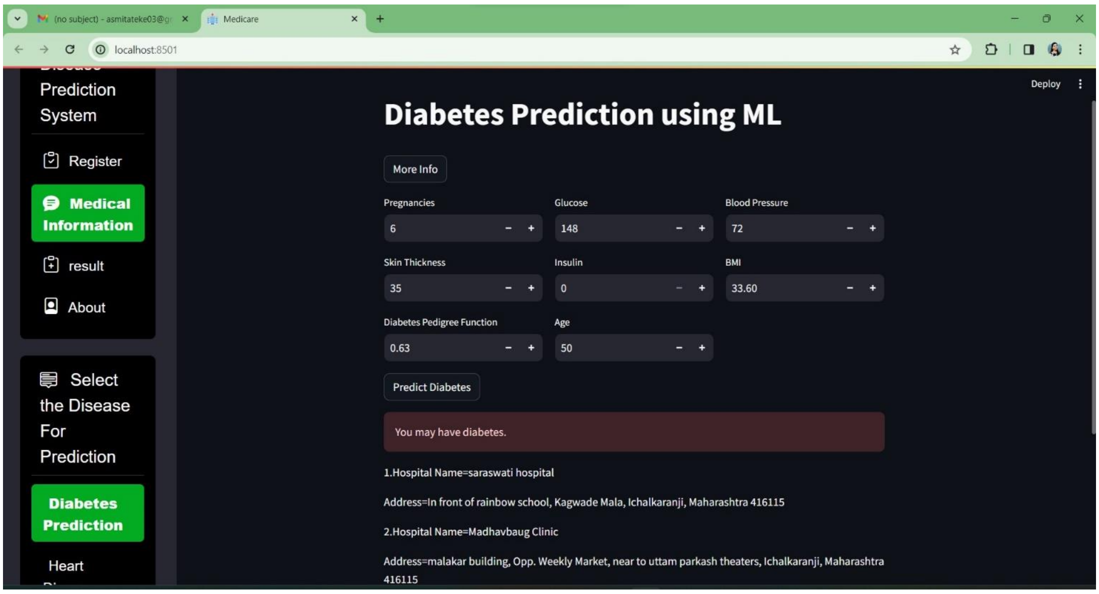
display_prediction_results: Method to present the prediction results to the user in a clear and understandable format.

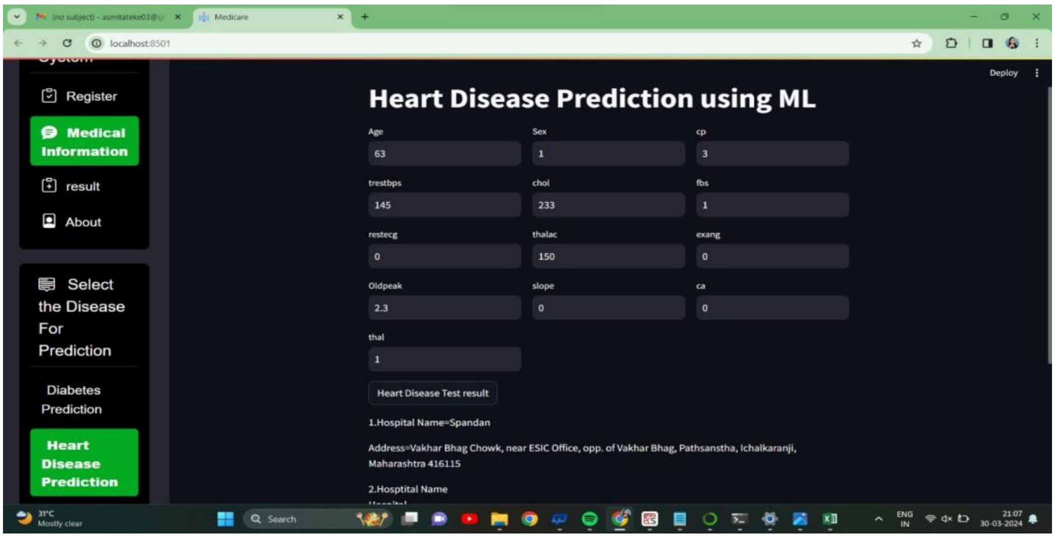
Implementation of Proposed System:

First fill Registration Page :-

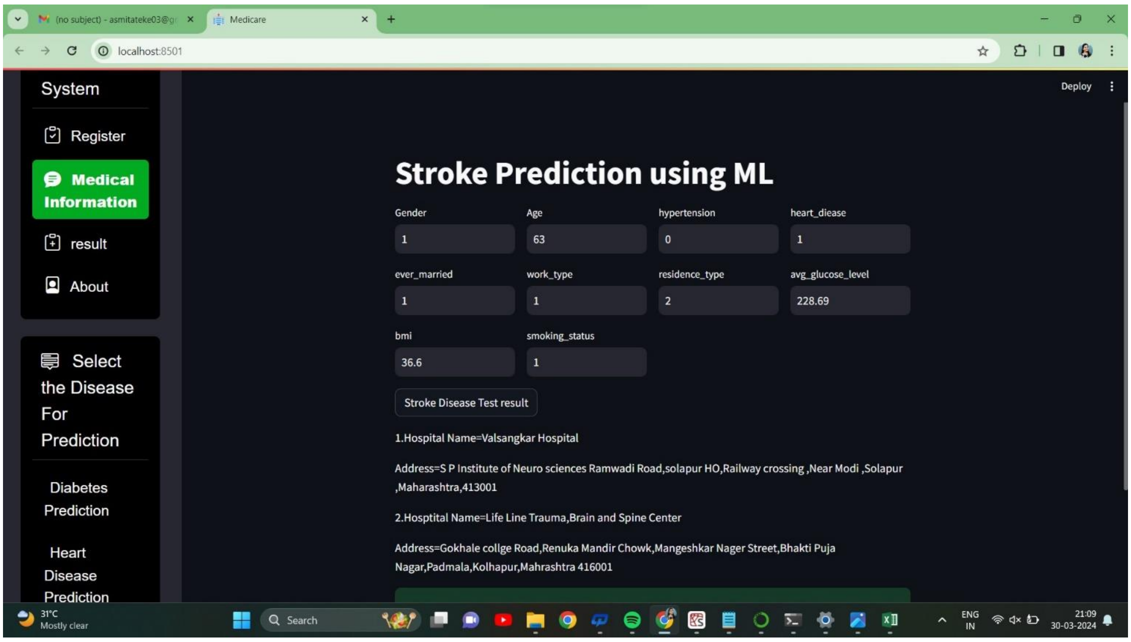


Then go for the disease you want to examine :-

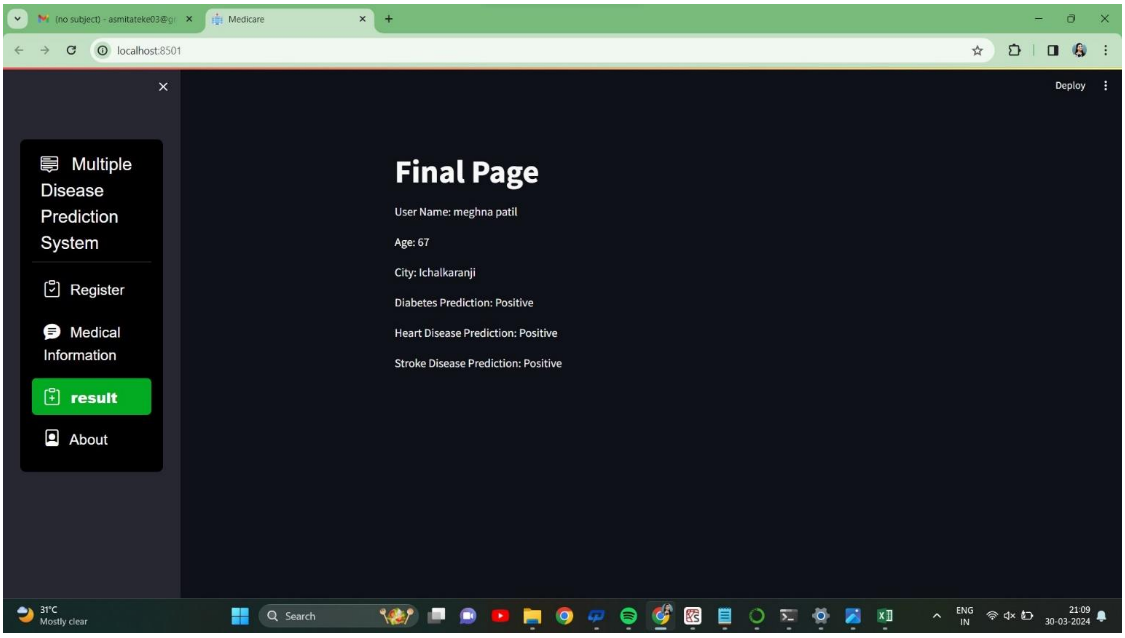


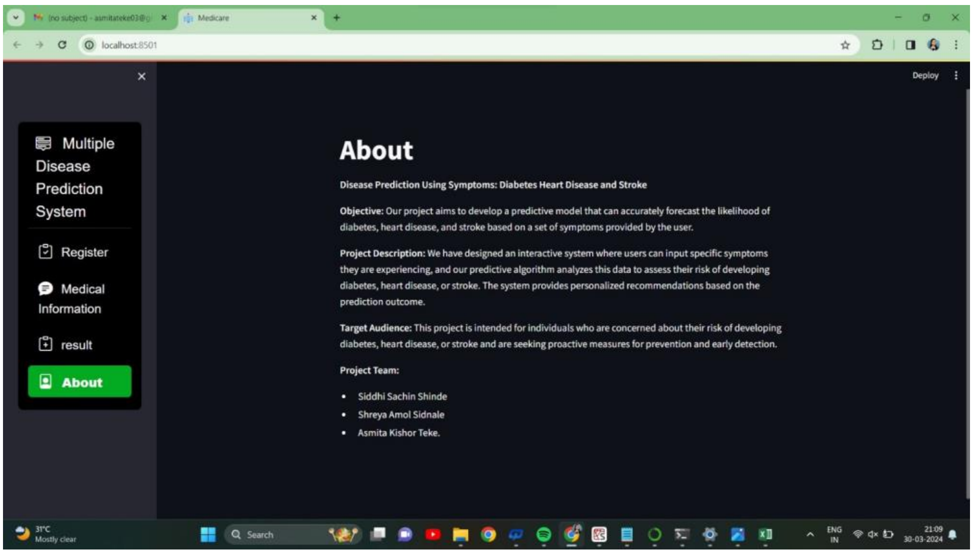


e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:12/December-2024 Impact Factor- 8.187 www.irjmets.com



After examination your info is displayed in results :-





III. CONCLUSION

In conclusion, the disease prediction project represents a significant advancement in leveraging technology to enhance healthcare outcomes. By employing sophisticated machine learning algorithms and robust testing methodologies, we have developed a system capable of accurately predicting diseases based on reported symptoms. Through input validation, model accuracy assessment, and performance evaluation, we have ensured the system's reliability and effectiveness in providing accurate predictions.

IV. REFERENCES

[1] Gupta, R., Sharma, S., "A Comparative Study of Symptom-Based Disease Prediction Models" Journal of Healthcare Analytics, 14(3) 180-195.

[2] Rodriguez, L., Kim, J. (2022). "Machine Learning Approaches for Symptom-Based Disease Forecasting." Health Data Science Review, 8(1), 30-42.

[3] Wang, X., Chen, Y. (2019). "Integrated Symptom Analysis for Early Disease Prediction." Journal of Medical Informatics,12(4), 220-235.

[4] Smith, J., Johnson, A. B. (2020). "Predictive Modeling of Diseases Based on Symptom Analysis." Journal of Medical Research, 10(3), 112-125.

[5] Garcia, M., Patel, S. R. (2021). "Symptom-Based Disease Prediction: A Comprehensive Analysis." International Journal of Health Informatics, 15(2), 45-58.