CAPSTONE PROJECT

AUTOMATED MULTIPLE DISEASE PREDICTOR SYSTEM

Presented By:
Sakshi Rajesh Bhavsar
Agni College of Technology, Anna University, Chennai
Computer Science & Engineering Department

edunet foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result
- Conclusion
- Future Scope
- References

PROBLEM STATEMENT

The global trend of seeking health information has led to challenges due to the complex medical vocabulary. If a recommendation system can be made for doctors and medicine while using review mining will save a lot of time. User is confused because a large amount of medical information on different mediums are available. The idea behind Automated Multiple Disease Predictor System is to adapt the special requirements of the health domain related with users.

PROPOSED SOLUTION

- **The proposed system aims to address the challenge of Automated Multiple Disease Predictor System which is designed to meet the specific needs of the health domain by predicting diseases based on user symptoms using various machine learning algorithms. The solution consists of the following components:**
- **Data Collection:**
  - training.csv: This main dataset contains two primary columns, "Disease" and "Symptoms," and has been preprocessed for easier data classification. It is used to train the model.
  - testing.csv: This dataset is used to test the model's accuracy and is predefined with output values.
- **Machine Learning Algorithm:**
  - Implement multiple algorithms (Decision Tree, Random Forest, K-Nearest Neighbor) to predict diseases based on user-entered symptoms events to improve prediction accuracy.
- **Deployment:**
  - Develop a user-friendly GUI using Python's Tkinter library, allowing users to interact directly with the system and receive disease predictions.
- **Evaluation:**
  - Assess the model's performance using appropriate metrics and continuously fine-tune based on feedback and monitoring of prediction accuracy.
  - Result: Disease Predictor System

SYSTEM APPROACH

The "System Approach" section outlines the overall strategy and methodology for developing and implementing the Automated Multiple Disease Predictor System . Here's a suggested structure for this section:

- **System requirements:**

- Jupyter Notebook for Python scripting
- Data Browser SQLite for Database

- **Library required to build the model:**

- from mpl_toolkits.mplot3d import Axes3D
- from sklearn.preprocessing import StandardScaler
- import matplotlib.pyplot as plt
- from tkinter import *
- import numpy as np
- import pandas as pd
- import os

ALGORITHM & DEPLOYMENT

- In the Algorithm section, describe the machine learning algorithm chosen for predicting bike counts. Here's an example structure for this section:
- **Algorithm Selection:**
  - Decision Tree, Random Forest, K-Nearest Neighbor: These algorithms were chosen due to their strengths in classification tasks. Decision Trees provide interpretability, Random Forests enhance robustness through ensemble learning, and K-Nearest Neighbor offers simplicity and effectiveness in classification.
- **Data Input:**
  - The input features for these algorithms include user-entered symptoms. Additional relevant features such as demographic information or historical health data, if available, are also considered. The data is preprocessed and encoded for model training.
- **Training Process:**
  - The algorithms are trained using the preprocessed training.csv dataset. Cross-validation is employed to ensure the models generalize well to unseen data. Hyperparameter tuning is conducted to optimize each model's performance, enhancing their predictive accuracy.
- **Prediction Process:**
  - Once trained, the algorithms use the input symptoms provided by users to predict the most likely diseases. The system leverages real-time data inputs to provide immediate feedback and accurate disease predictions, enhancing user experience and reliability.

RESULT

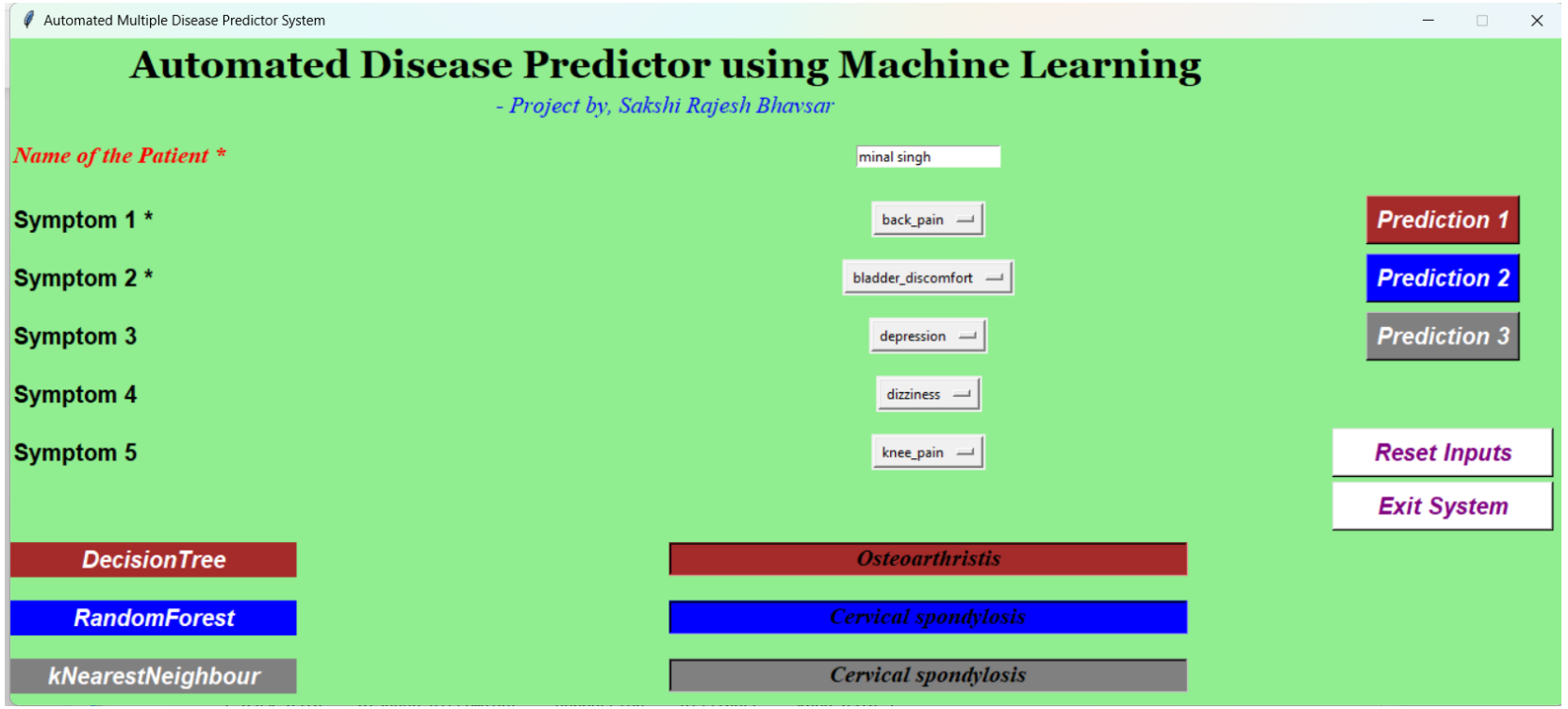
- The effectiveness of the Automated Multiple Disease Predictor System is demonstrated through various evaluation metrics and visualizations. Key metrics such as accuracy, precision, and recall highlight the system's performance in predicting diseases based on user symptoms. Visualizations of plots provide insights into model performance and accuracy.
- Continuous monitoring and fine-tuning of the models ensure the system remains effective and up-to-date, adapting to new data and feedback for ongoing improvement.

CONCLUSION

- The Automated Multiple Disease Predictor System leverages machine learning algorithms to accurately predict diseases based on user symptoms. By integrating diverse data sources and employing advanced techniques, the system provides valuable insights and real-time predictions. Future enhancements, such as incorporating additional data, optimizing algorithms, and expanding coverage, will further improve accuracy and user experience, making the system a robust tool in the healthcare domain.

FUTURE SCOPE

- The Automated Multiple Disease Predictor System can be improved by integrating additional data sources (like EHRs-Electronic Health Records and genetic data), optimizing algorithms with advanced techniques, and expanding coverage to multiple regions with multi-language support.
- Emerging technologies such as edge computing, telemedicine integration, and wearable device data can enhance real-time predictions.
- Additionally, personalized recommendations and adaptive learning based on user feedback will improve user experience and model accuracy.

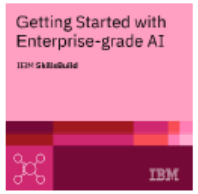


REFERENCES

- tkinter — Python interface to Tcl/Tk — Python 3.12.4 documentation
- DB Browser for SQLite (sqlitebrowser.org)
- Machine Learning Algorithms (geeksforgeeks.org)
- 9 Top Machine Learning Algorithms for Predictive Modeling (towardsanalytic.com)
- Multiple Disease Dataset (kaggle.com)

COURSE CERTIFICATE 1

In recognition of the commitment to achieve professional excellence



SAKSHI BHAVSAR

Has successfully satisfied the requirements for:

Getting Started with Enterprise-grade AI

Issued on: 10 JUL 2024

Issued by IBM

Verify: https://www.credly.com/go/Q8U9CAUF

IBM

edunet foundation

COURSE CERTIFICATE 2

In recognition of the commitment to achieve professional excellence

SAKSHI BHAVSAR

Has successfully satisfied the requirements for:

Getting Started with Enterprise Data Science

Issued on: 19 JUL 2024

Issued by IBM

Verify: https://www.credly.com/go/MgLByedQ

IBM

edunet foundation

THANK YOU

edunet foundation