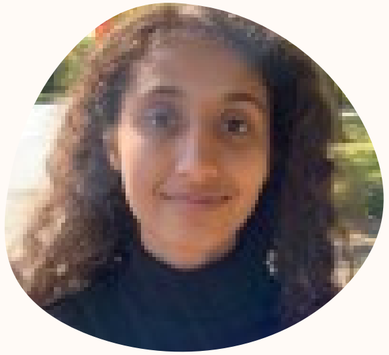
Project Walkthrough

About Us

Bahar Chidem
Software Eng Specialist
@UOFT

Vaibhav Lakshmi Santhanam
Software Eng Specialist
@UOFT



Tara Jorjani
Software Eng Specialist
@UOFT

Arina Azmi
Information Systems Specialist
@UOFT

Heart Disease

Predict the presence of heart disease in patients

problem statement

Heart disease is an epidemic affecting 1/3 of Canadians and results in thousands of deaths annually. This places billions of dollars of strain on the healthcare system.

With further investment in the early detection of cardiovascular disease, patients can be treated faster and more effectively, reducing strain on the healthcare sector and improving quality of life.

Solution

Our Heart Disease prediction model addresses this by predicting the occurrence of cardiovascular disease based on a variety of factors. It harnesses the power of AI to find the relationship between risk factors and disease outcome.

model description:

1. Data Preprocessing: Numerical data normalized and categorical data one-hot encoded.
2. Model: K-Nearest Neighbors (KNN) with parameter tuning (k values and distance metrics).
3. Optimization: Best parameters selected using 10-fold cross-validation.
4. Final Model: Weighted KNN with optimal 'k' and distance metric.
5. Performance: Evaluated on a validation set, demonstrating accuracy in predicting heart disease

Version 1: Logistic Regression

1. Preprocessing data:
    - Categorical columns values get transformed to numerical values (Age, Sex, etc)
2. Specifying logistic regression parameters
    - Maximum iteration = 1000 and random state = 0
3. Creating logistic regression model
    - Imported from the sklearn library
4. Validation results
    - The accuracy of the model is predicted based on the scikit-learn accuracy score

Version 2 Weighted KNN

1. **Data Processing**: Standardization of numerical features and transformation of categorical variables.
2. **Modeling Technique**: Utilization of K-Nearest Neighbors (KNN), fine-tuned through varying 'k' values and distance measures.
3. **Parameter Fine-Tuning**: Selection of optimal parameters based on results from 10-fold cross-validation.
4. **Optimized Model**: Employed Weighted KNN strategy with chosen best parameters for robust prediction.
5. **Validation Results**: Demonstrated efficiency in heart disease prediction through validation set accuracy

Thank you