HEART STROKE PREDICTION

P. Kriti Meher: 21B01A05D6
S. Manojna: 21B01A05H1
T. Tejasri : 21B01A05I0
V. Yaswini : 21B01A05I9

Shri Vishnu Engineering College For Women

February 9, 2024

Introduction

- Recognizing individuals at risk of heart stroke early on is crucial for taking proactive steps to prevent it.
- Using machine learning techniques provides a strong way to precisely gauge the likelihood of heart strokes.
- In this presentation, we discuss how we built and tested a simple-to-use predictive model designed specifically to predict heart strokes.

Problem Statement

- The problem at hand is to create a reliable and precise predictive model to detect individuals at a heightened risk of encountering a heart stroke. This model should effectively utilize various demographic, clinical, and personal data to deliver personalized risk assessments.

Dataset Description

- The dataset encompasses various attributes like age, gender, hypertension, heart disease, marital status, work type, residence type, average glucose level, BMI, and smoking status.
- The dataset also contains information on stroke occurrence (binary target variable).

Model Selection

- We have employed a variety of machine learning models including Logistic Regression, Decision Trees, Support Vector Machines (SVM), and Random Forests.
- Both Random Forests and SVM demonstrated notable performance in terms of accuracy.
- Decision Tree and Logistic Regression showed relatively lower accuracy compared to Random Forests and SVM.

Optimization

- For optimization we used hyperparameter tuning(grid search)

Optimization

```
[ ] from sklearn.ensemble import RandomForestClassifier
    from sklearn.metrics import accuracy_score

[ ] rf = RandomForestClassifier()

[ ] rf.fit(x_train, y_train)
```

▾ RandomForestClassifier
RandomForestClassifier()

```
[ ] y_pred_rf = rf.predict(x_test)

    accuracy_rf = accuracy_score(y_test, y_pred_rf)
    print("Model accuracy using Random Forest: {:.2f}%".format(accuracy_rf * 100))
```

Model accuracy using Random Forest: 95.43%

Optimization

+ Code + Text | Copy to Drive | Reconnect | Colab

```
[ ] grid_search = GridSearchCV(estimator=rf, param_grid=param_grid, cv=5, scoring='accuracy')
```

```
[ ] grid_search.fit(x_train, y_train)
```

GridSearchCV
estimator: RandomForestClassifier
RandomForestClassifier

```
[ ] best_params = grid_search.best_params_
    print("Best hyperparameters:", best_params)
```

```
Best hyperparameters: {'max_depth': None, 'min_samples_leaf': 4, 'min_samples_split': 2, 'n_estimators': 100}
```

```
[ ] best_rf = grid_search.best_estimator_
```

```
[ ] y_pred_rf = best_rf.predict(x_test)
```

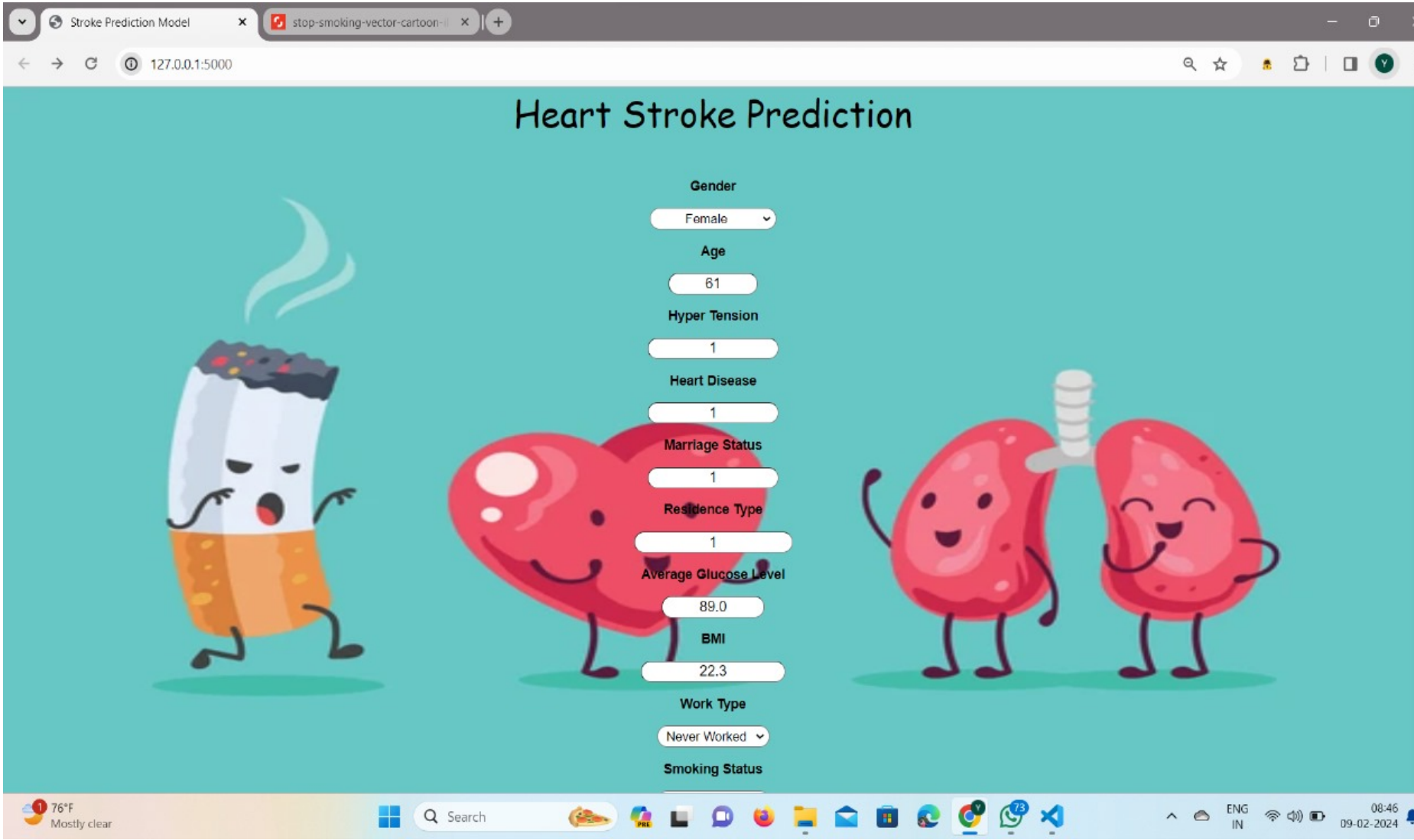
```
accuracy_rf = accuracy_score(y_test, y_pred_rf)
print("Model accuracy using Random Forest with hyperparameter tuning: {:.2f}%".format(accuracy_rf * 100))
```

```
Model accuracy using Random Forest with hyperparameter tuning: 95.50%
```

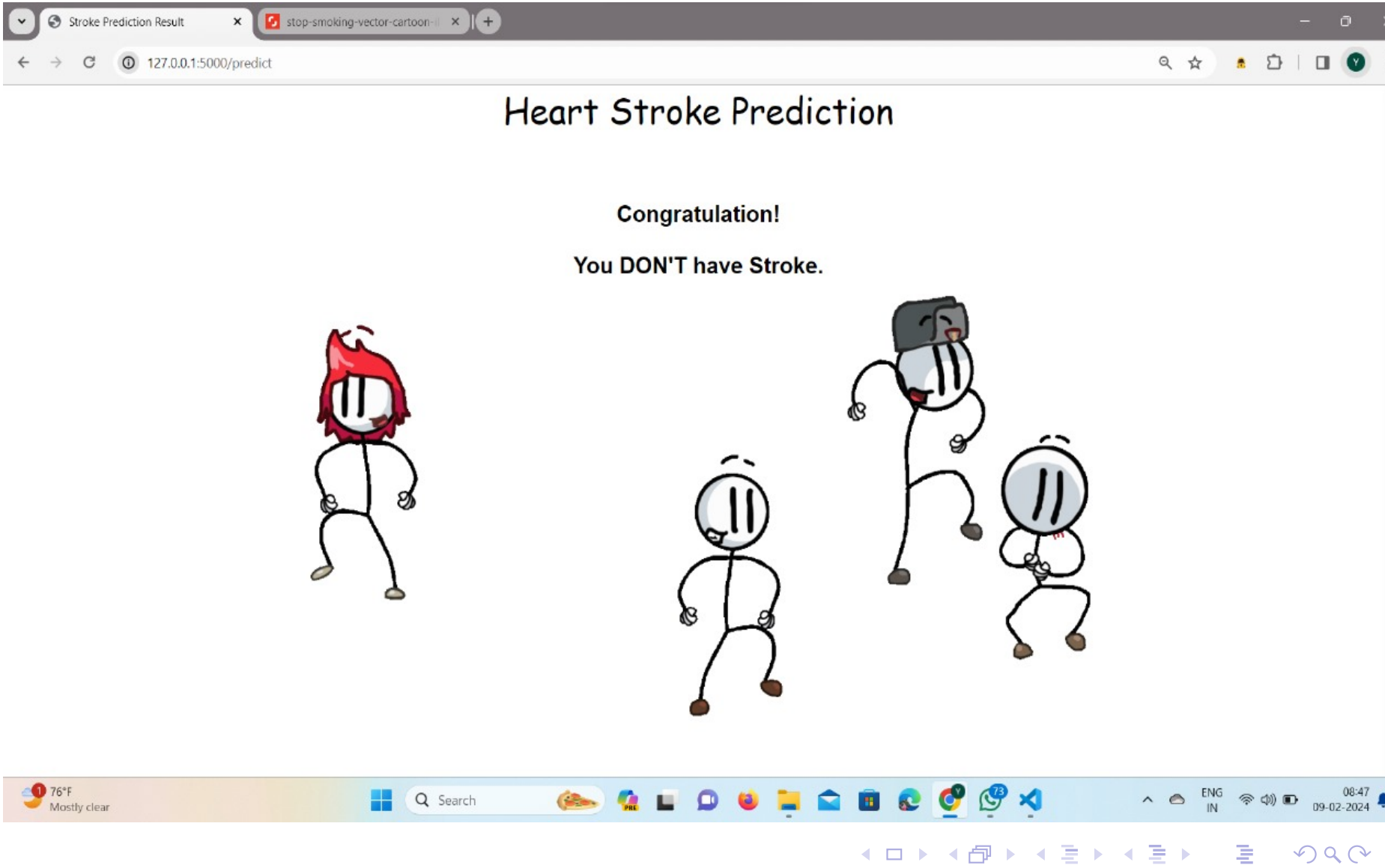
Results

- Our findings suggest that Random Forests performs exceptionally well in predicting heart strokes.
- Random Forests gives an accuracy of 95.4%.

Frontend


Stroke Prediction Model
127.0.0.1:5000
Heart Stroke Prediction
Gender
Female
Age
61
Hyper Tension
1
Heart Disease
1
Marriage Status
1
Residence Type
1
Average Glucose Level
89.0
BMI
22.3
Work Type
Never Worked
Smoking Status
76°F
Mostly clear
Search
ENG
IN
08:46
09-02-2024

Frontend


Stroke Prediction Result
stop-smoking-vector-cartoon-
127.0.0.1:5000/predict
Heart Stroke Prediction
Congratulation!
You DON'T have Stroke.
76°F
Mostly clear
Search
ENG
IN
08:47
09-02-2024

Bibliography

- www.projectpro.io
- www.kaggle.com
- Ja NaN (youtube channel)
- chatGPT
- GeeksForGeeks

Conclusion

- Detecting heart strokes early is crucial for staying ahead in healthcare.
- Easy-to-use machine learning models like Random Forests stand out as powerful tool for understanding stroke risks.