

THYROID CLASSIFICATION

R. Bhavana : 21B01A05F9 : CSE
S. Ramadevi : 21B01A05H2 : CSE
V. Rithu Sree : 21B01A05J1 : CSE
Y. Sri Lakshmi : 21B01A05J7 : CSE

Shri Vishnu Engineering College For Women

February 9, 2024

Introduction

- Thyroid dysfunction is a prevalent condition that can have serious health implications if left untreated.
- Machine learning models provide a promising avenue for thyroid classification, aiding in the early detection and assessment of thyroid dysfunction risk.
- This presentation outlines our efforts in developing a predictive model for thyroid classification.

Dataset Description

- Our dataset includes Age, T3, TSH, TT4, T4U, FTI levels.
- The dataset also contains information on Thyroid prediction (binary target variable).
- Thyroid Classification(Hypothyroidism Hyperthyroidism) is done based on these levels.

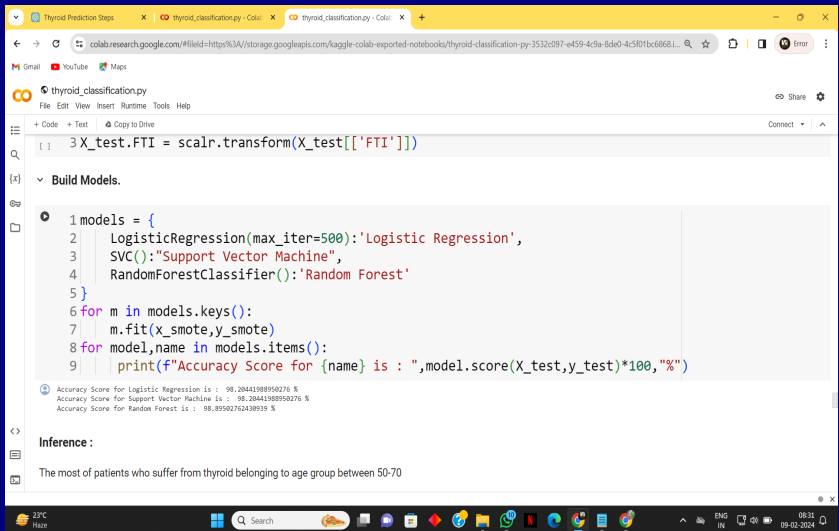
Model Selection

- We experimented with several machine learning models, including Logistic Regression, Support Vector Machine (SVM), Random Forest.
- Models showed good performance in terms of accuracy.
- We considered additional factors such as interpretability, computational efficiency, and scalability for model selection.

Evaluation Metrics

- We used various evaluation metrics to assess model performance:
 - Accuracy, Precision, Recall, F1 Score
 - Confusion Matrix
- These metrics helped us understand the strengths and weaknesses of each model.

Accuracy



```
3 X_test.FTI = scaler.transform(X_test[['FTI']])
```

Build Models.

```
1 models = {
2     LogisticRegression(max_iter=500): 'Logistic Regression',
3     SVC(): "Support Vector Machine",
4     RandomForestClassifier(): 'Random Forest'
5 }
6 for m in models.keys():
7     m.fit(x_smote, y_smote)
8 for model, name in models.items():
9     print(f"Accuracy Score for {name} is : ", model.score(X_test, y_test)*100, "%")
```

Accuracy Score for Logistic Regression is : 98.20441988950276 %
Accuracy Score for Support Vector Machine is : 98.20441988950276 %
Accuracy Score for Random Forest is : 98.89502762430939 %

Inference :

The most of patients who suffer from thyroid belonging to age group between 50-70

Results

- Our results indicate that Random Forest performs well in classifying thyroid.
- To increase efficiency we have gone through optimization technique GridSearch.

The screenshot shows a web browser window with multiple tabs. The active tab is titled 'thyroid_classification.py - Colab'. The address bar shows the file path: 'C:/Users/vidiy/OneDrive/Desktop/thyroid_classification_app/templates/results.html'. The page content is on a light blue background and displays the following text:

PREDICTION RESULT

Based on the provided information, you are classified as: hypothyroid

The browser's taskbar at the bottom shows the Windows Start button, a search bar, and various application icons. The system tray on the right indicates the temperature is 24°C, the language is set to English (IN), and the date and time are 01:23 on 09-02-2024.

Conclusion

- Early classification of thyroid is vital for preventive healthcare.
- Machine learning models, particularly Random Forest, offer effective tools for Thyroid Classification.
- Our findings contribute to the development of accurate and interpretable models for Thyroid Clasification.

Bibiliography

- Kaggle :
<https://www.kaggle.com/code/prasadchaskar/thyroid-disease>.
- Inputs from chatGPT.
- For flask usage from youtube channel OneCoders.