Hypothyroidism Analysis and Prediction

Kairavi Dave

Introduction

Hypothyroidism is a condition where the thyroid gland does not produce enough thyroid hormones, leading to various health issues.

Understanding the patterns and factors associated with hypothyroidism is crucial for early diagnosis and effective treatment.

The objective of this project is to analyze the dataset on thyroid disease and provide insights that can help healthcare professionals better manage and predict hypothyroidism cases.

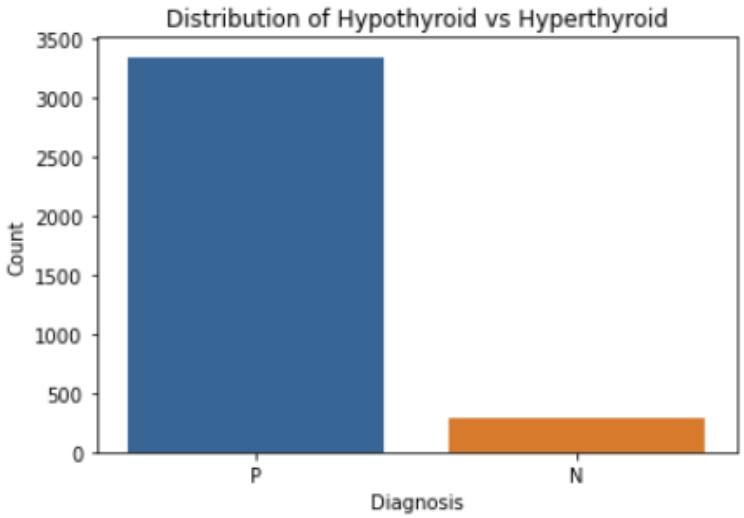
Data Preparation

DATA SOURCED FROM KAGGLE

DATA CLEANING AND PREPROCESSING USING SQL

THE CLEANED DATA WAS THEN EXPORTED FOR FURTHER ANALYSIS IN PYTHON.

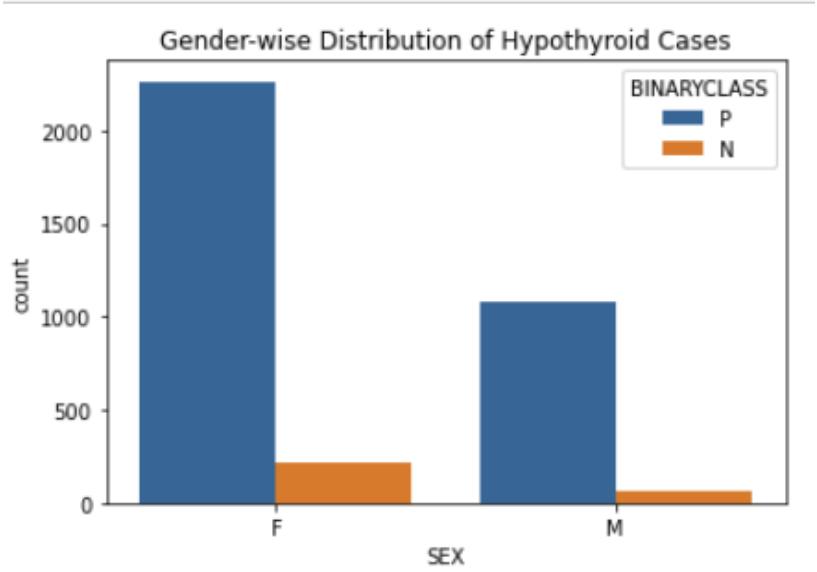
Exploratory Data Analysis (EDA)



Distribution of Hypothyroid vs. Non-Hypothyroid :

The data shows a significant number of hypothyroid cases (3480) compared to non-hypothyroid cases (526).

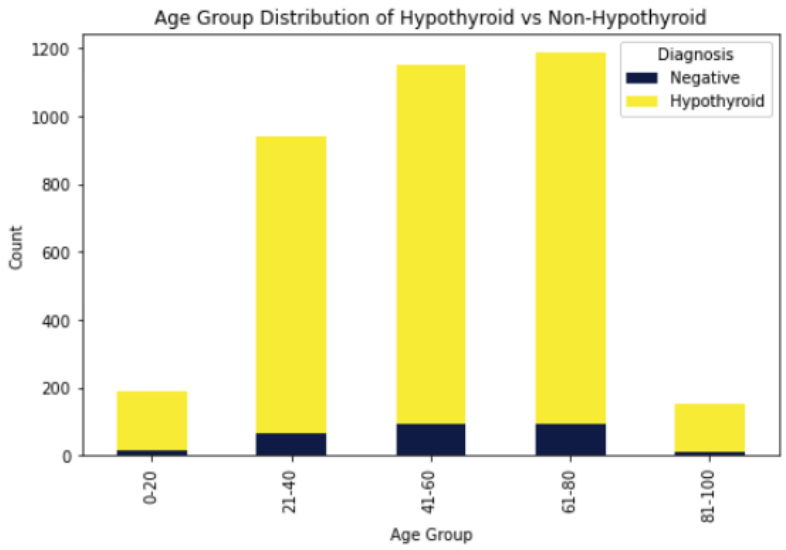
Exploratory Data Analysis (EDA)



Gender-wise Distribution of Hypothyroid Cases:

The majority of hypothyroid cases are in females (2,008) compared to males (990).

Exploratory Data Analysis (EDA)

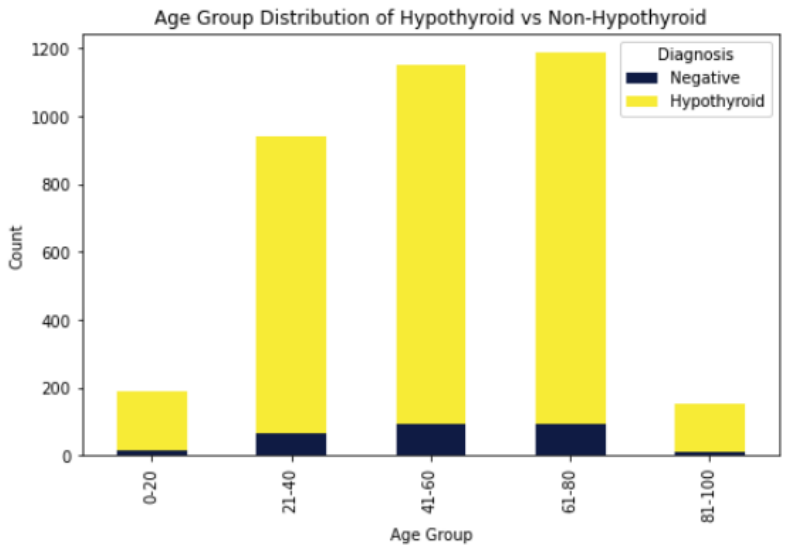


Age-Group Distribution of Hypothyroid and Non-Hypothyroid Cases:

The age group with the highest number of hypothyroid cases is 61-80 years old.

The non-hypothyroid cases are more evenly distributed across age groups.

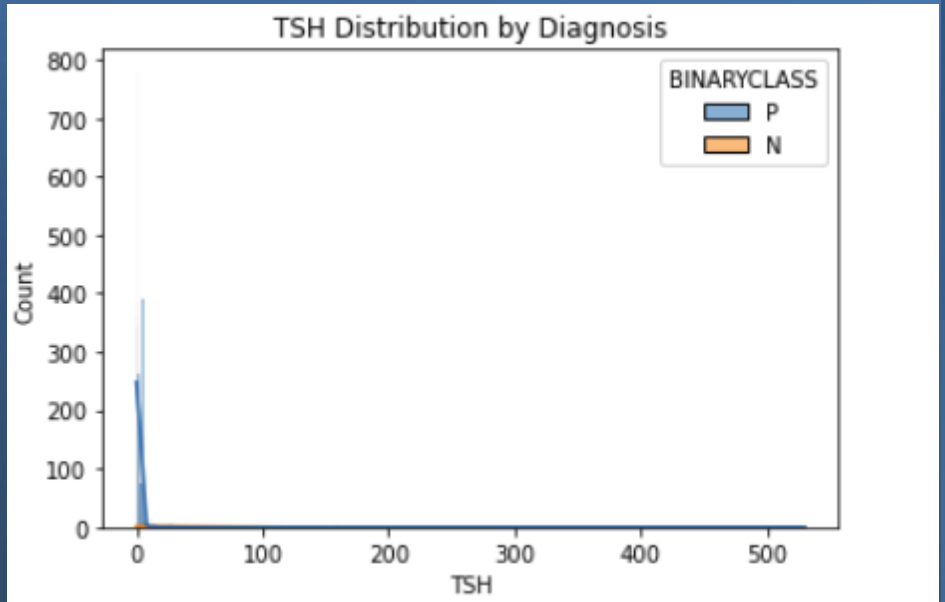
Exploratory Data Analysis (EDA)

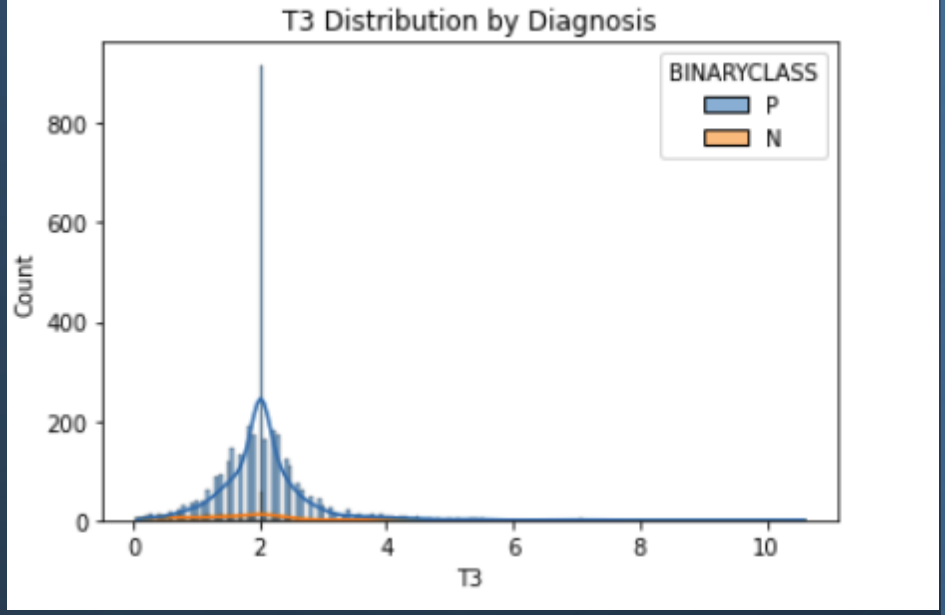


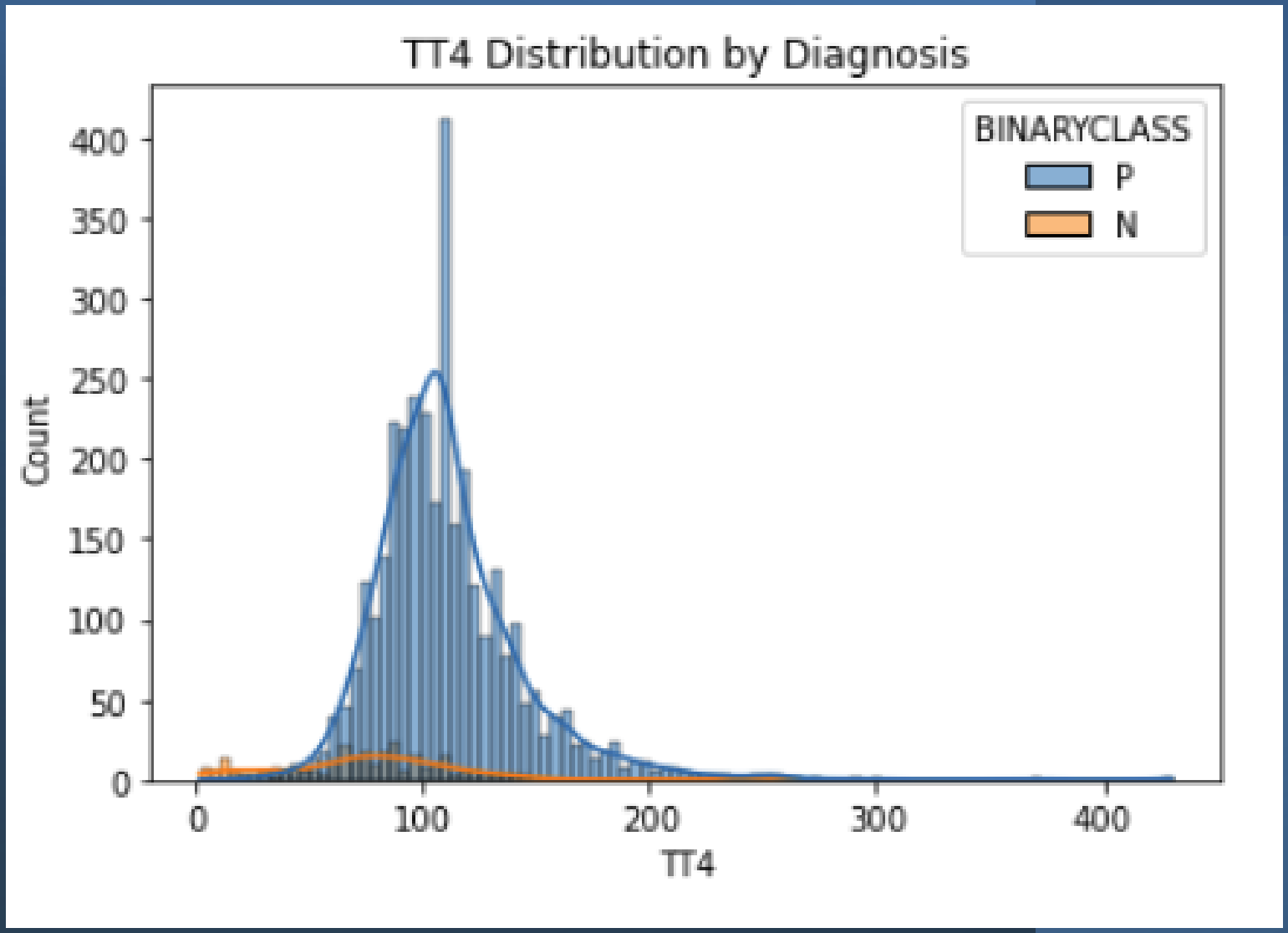
Age-Group Distribution of Hypothyroid and Non-Hypothyroid Cases:

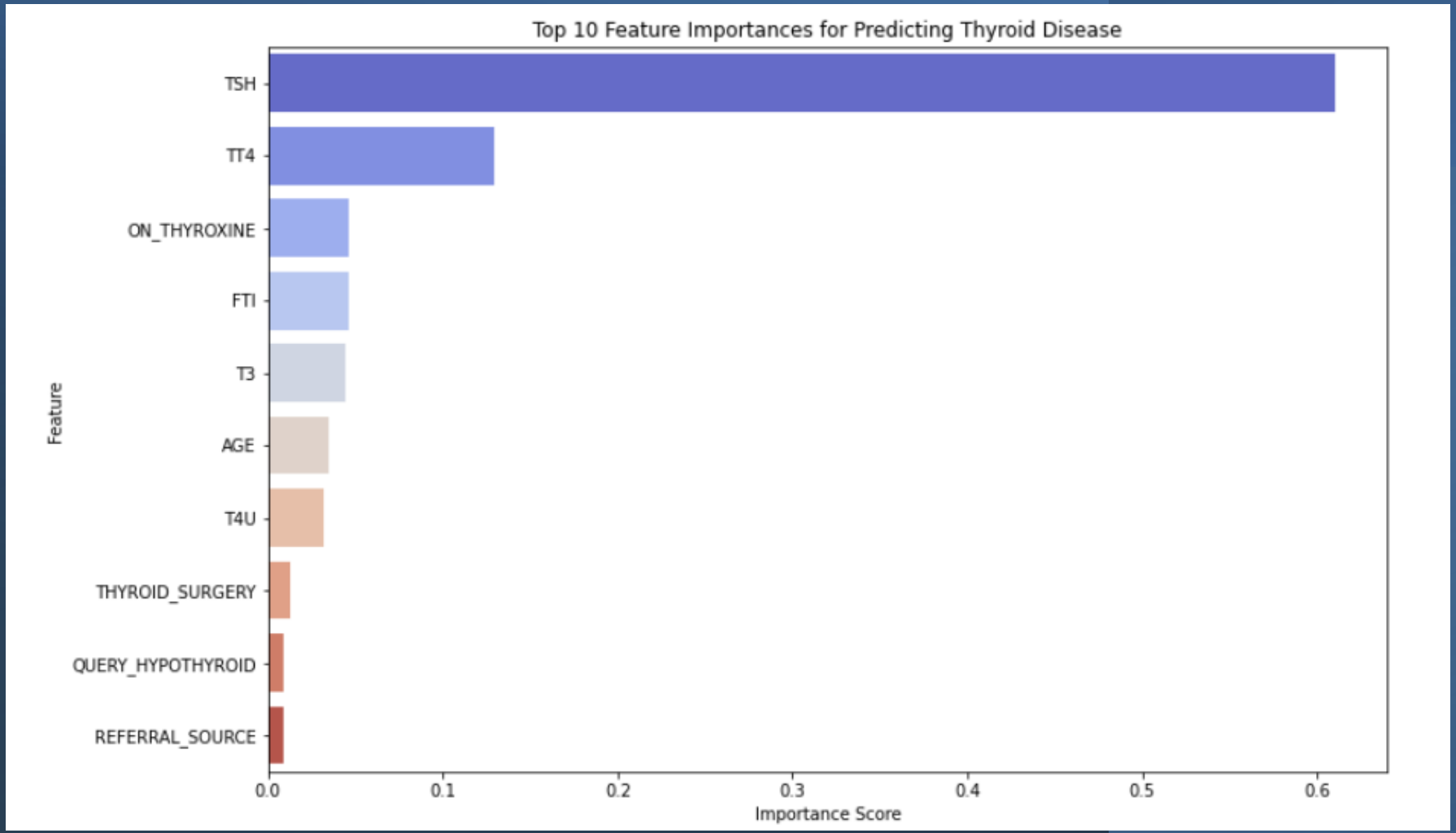
The age group with the highest number of hypothyroid cases is 61-80 years old.

The non-hypothyroid cases are more evenly distributed across age groups.


TSH Distribution by Diagnosis
BINARYCLASS
P
N
Count
TSH
0
100
200
300
400
500
800
700
600
500
400
300
200
100
0


T3 Distribution by Diagnosis
BINARYCLASS
P
N
Count
T3
0
2
4
6
8
10
800
600
400
200
0


TT4 Distribution by Diagnosis
BINARYCLASS
P
N
Count
400
350
300
250
200
150
100
50
0
0
100
200
300
400
TT4


Top 10 Feature Importances for Predicting Thyroid Disease
TSH
TT4
ON_THYROXINE
FTI
T3
AGE
T4U
THYROID_SURGERY
QUERY_HYPOTHYROID
REFERRAL_SOURCE
Feature
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Importance Score

Conclusion

Key Findings:

- 1. Thyroid diseases vary by gender, age, and medical history
- 2. Hypothyroidism is more common in females, with a significant number of cases in the 41-60 age group

Implications:

- The insights from this analysis can help healthcare professionals better understand the patterns and risk factors associated with hypothyroidism.
- Early screening and targeted interventions for high-risk groups (e.g., females, middle-aged individuals, pregnant women) can lead to earlier diagnosis and improved management of hypothyroidism.

Future Scope:

- 1. Extend analysis with machine learning models
- 2. Incorporate external datasets for broader insights

Q&A

Questions and Discussions