Classification Analysis of Mushrooms For California Garden

Outline

Introduction

Data Set

Modeling

1

3

5

2

4

6

Problem Statement

Workflow

Comparison
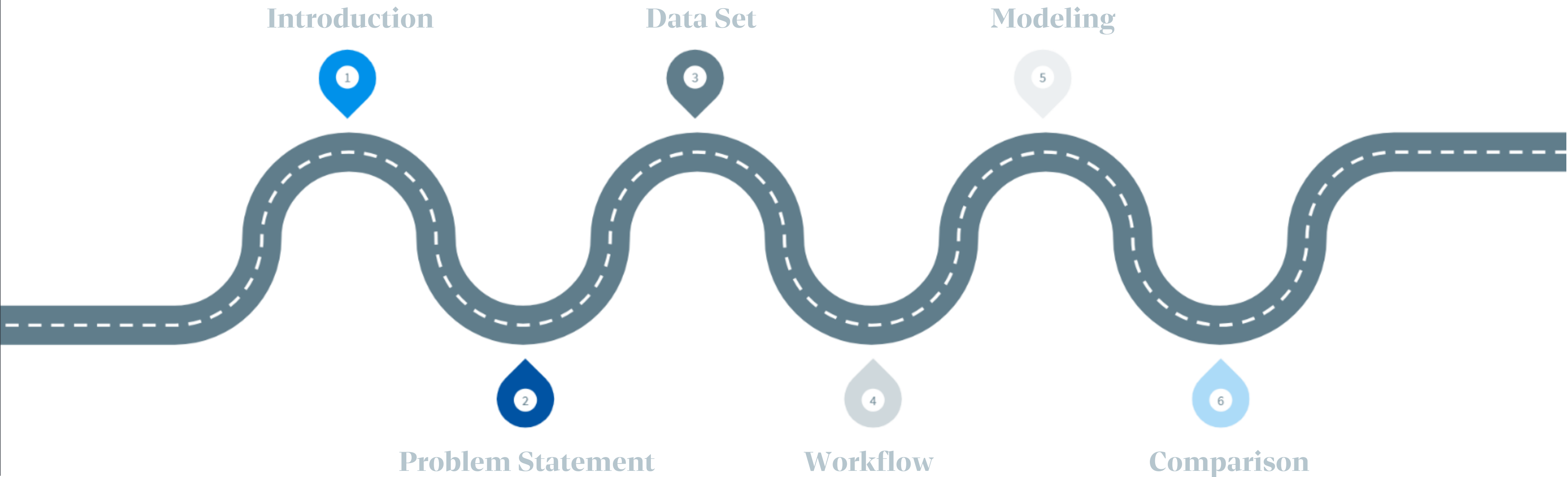
Introduction

A person filed a case against California Garden Company because he ate canned poisonous mushrooms. The company lost the case and compensated the client with a huge amount of money and incurred many losses. From here, I decided to seek help from a Data scientist so as not to make the same mistake and lose.

Problem Statement

classification of mushrooms, whether edible or poisonous based on other features .
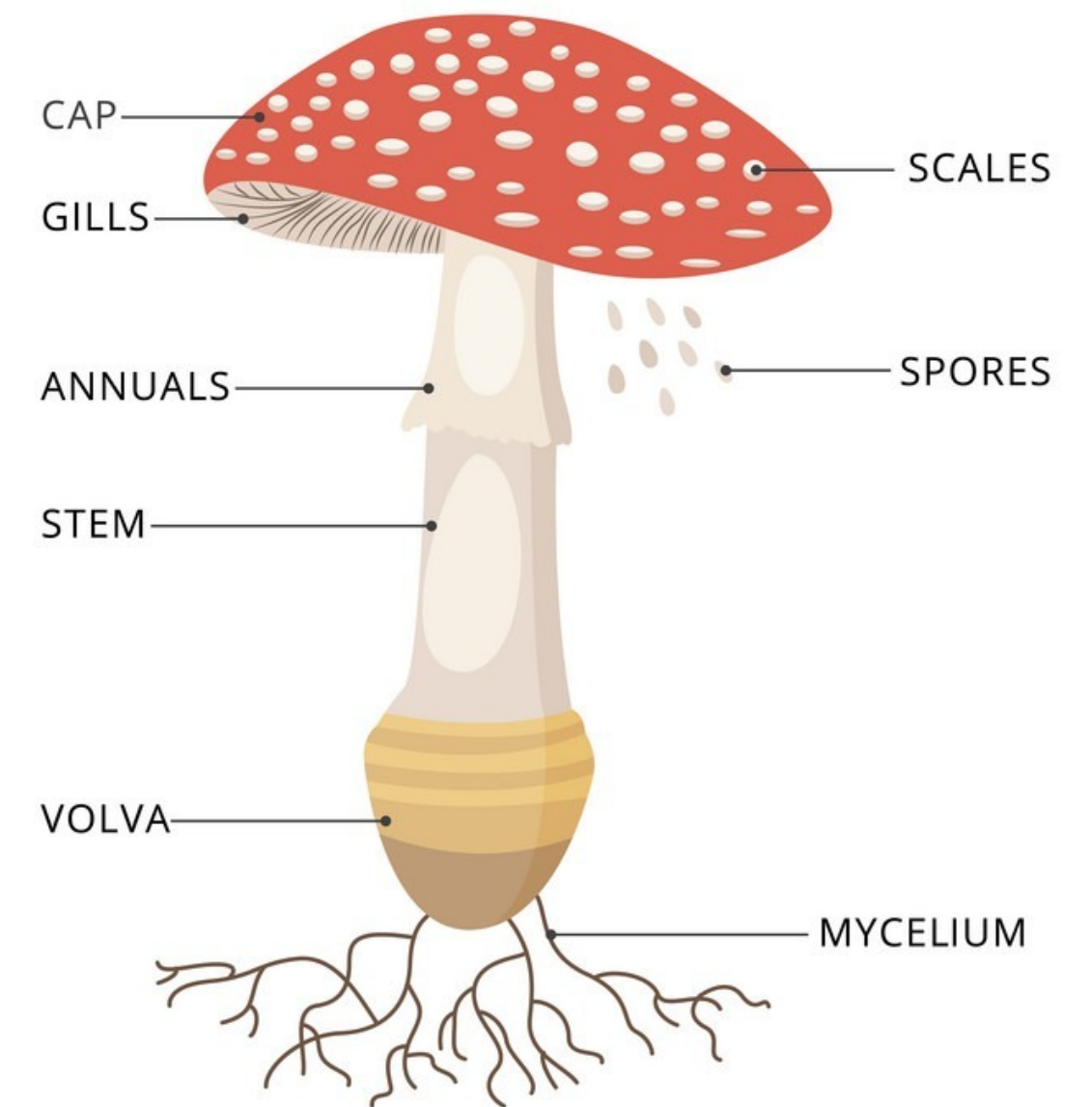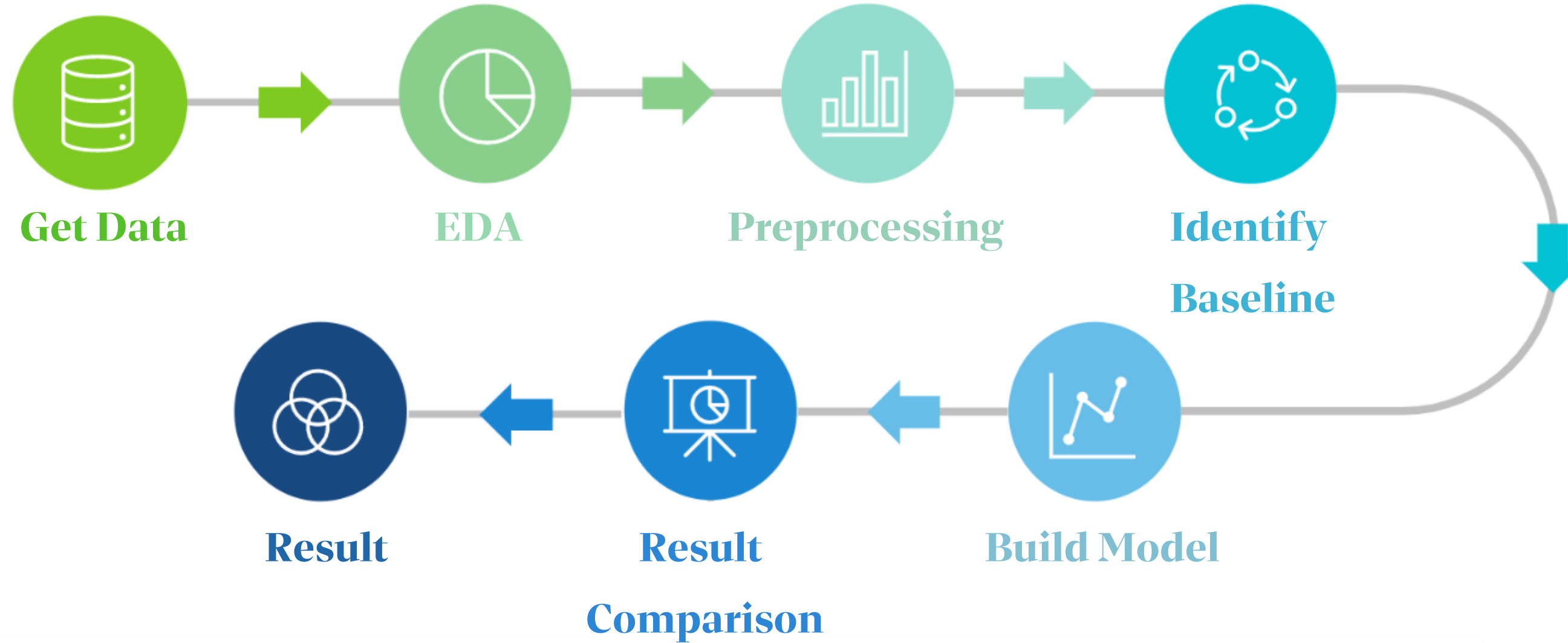
Data Set

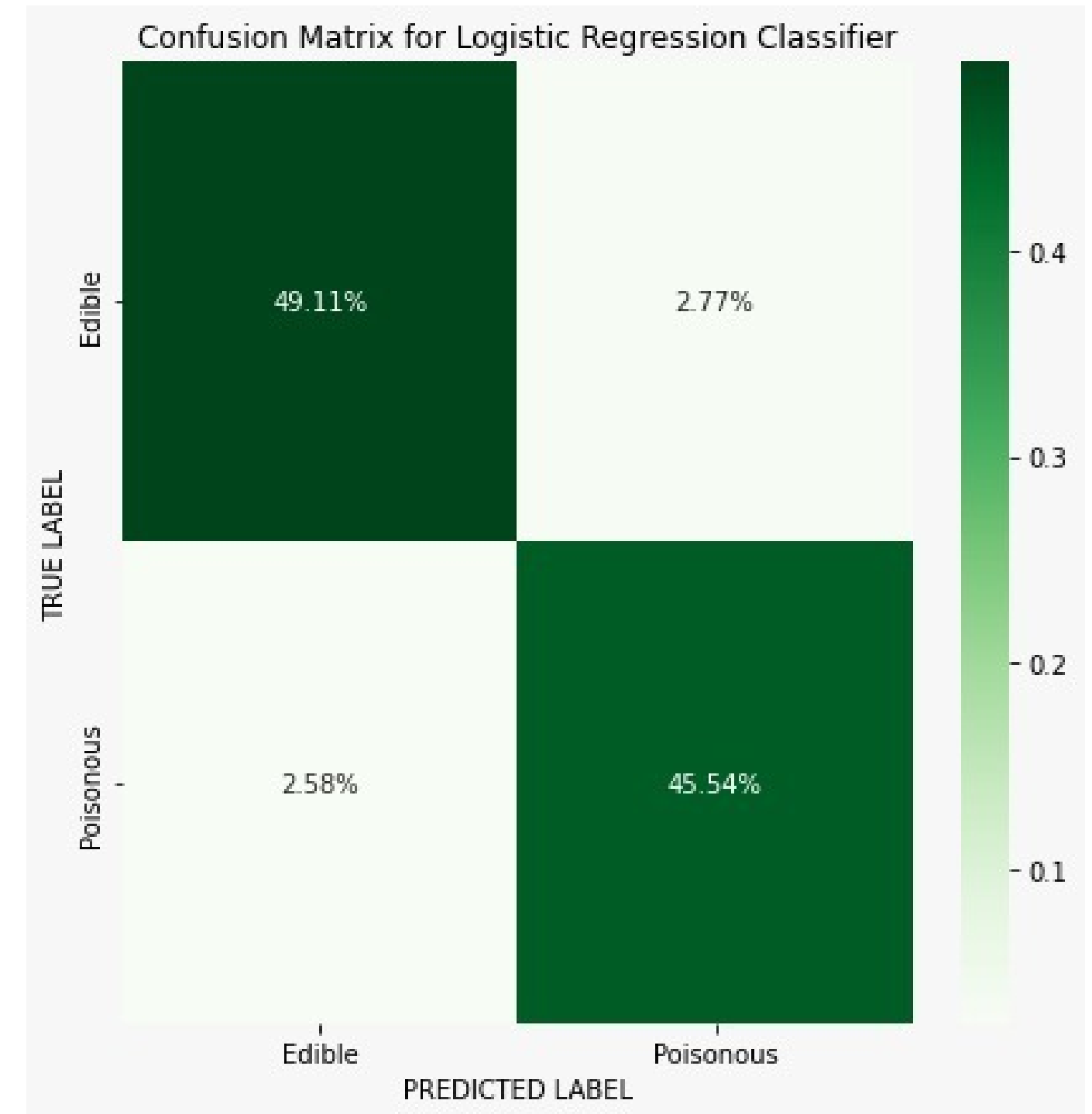Kaggle

8125

observation

23

features

FLY AGARIC ANATOMY

CAP

GILLS

ANNUALS

STEM

VOLVA

SCALES

SPORES

MYCELIUM

Workflow

Get Data → EDA → Preprocessing → Identify Baseline

Result ← Result Comparison ← Build Model

Modeling

Logistic Regression

ACCURACY: 0.9464615384615385

RECALL: 0.9462915601023018

Confusion Matrix for Logistic Regression Classifier

KNN (K-nearest neighbors)

ACCURACY: 0.9963076923076923

RECALL: 1.0

Confusion Matrix for KNN Classifier
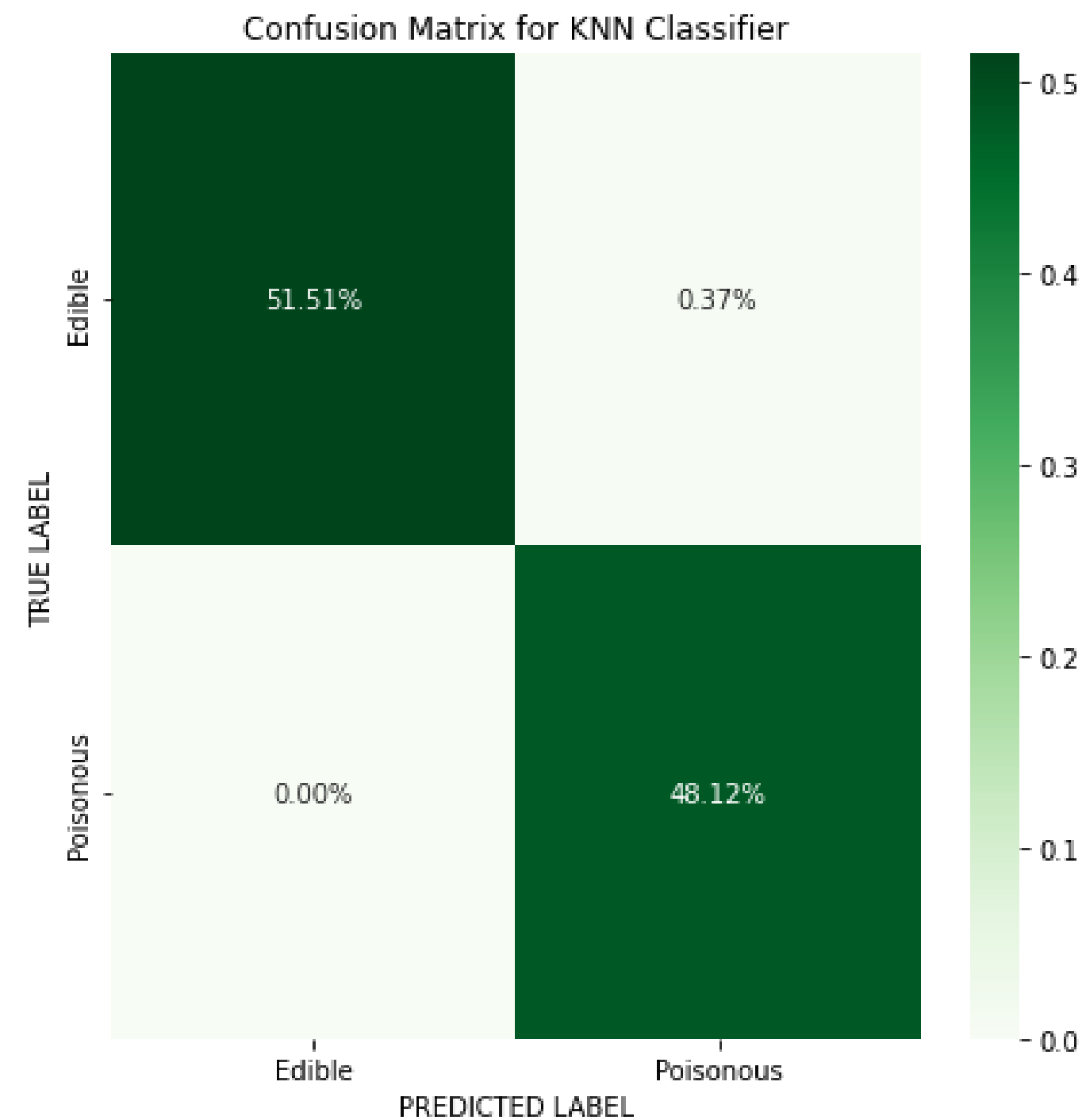
SVM (Support Vector Machine)

ACCURACY: 0.9926153846153846
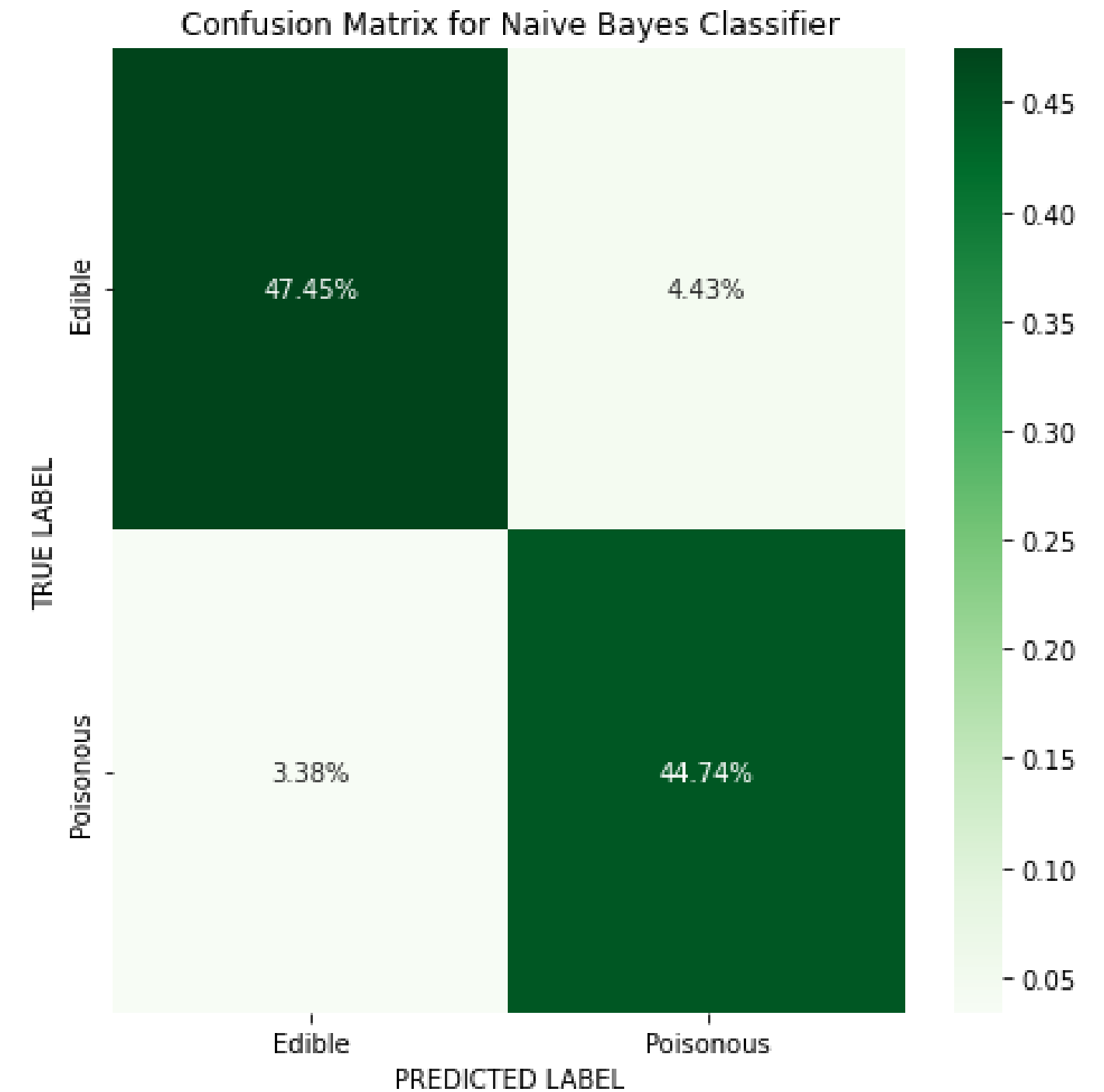
RECALL: 0.9859335038363172

Confusion Matrix for SVM Classifier
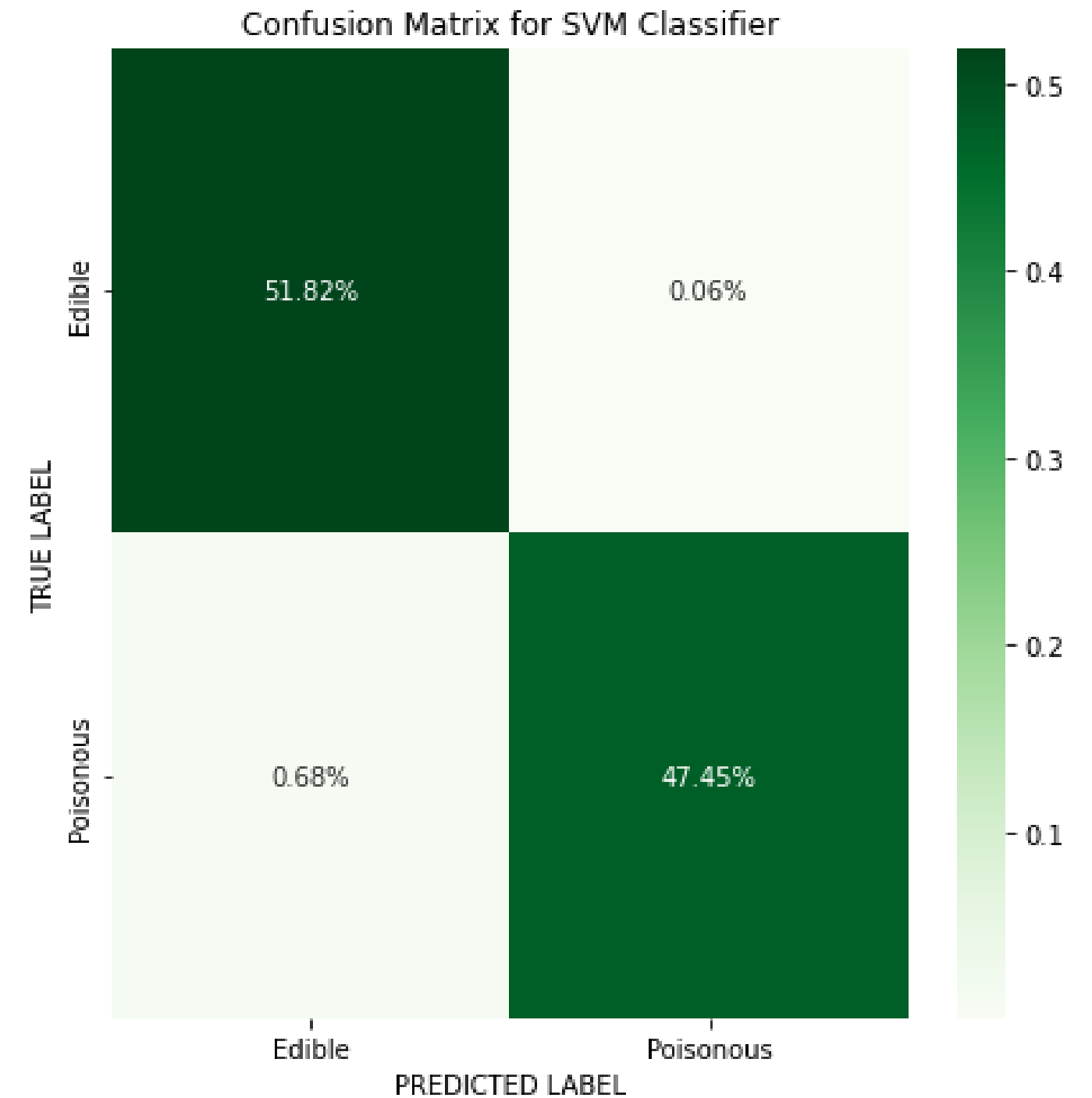
Naive Bayes
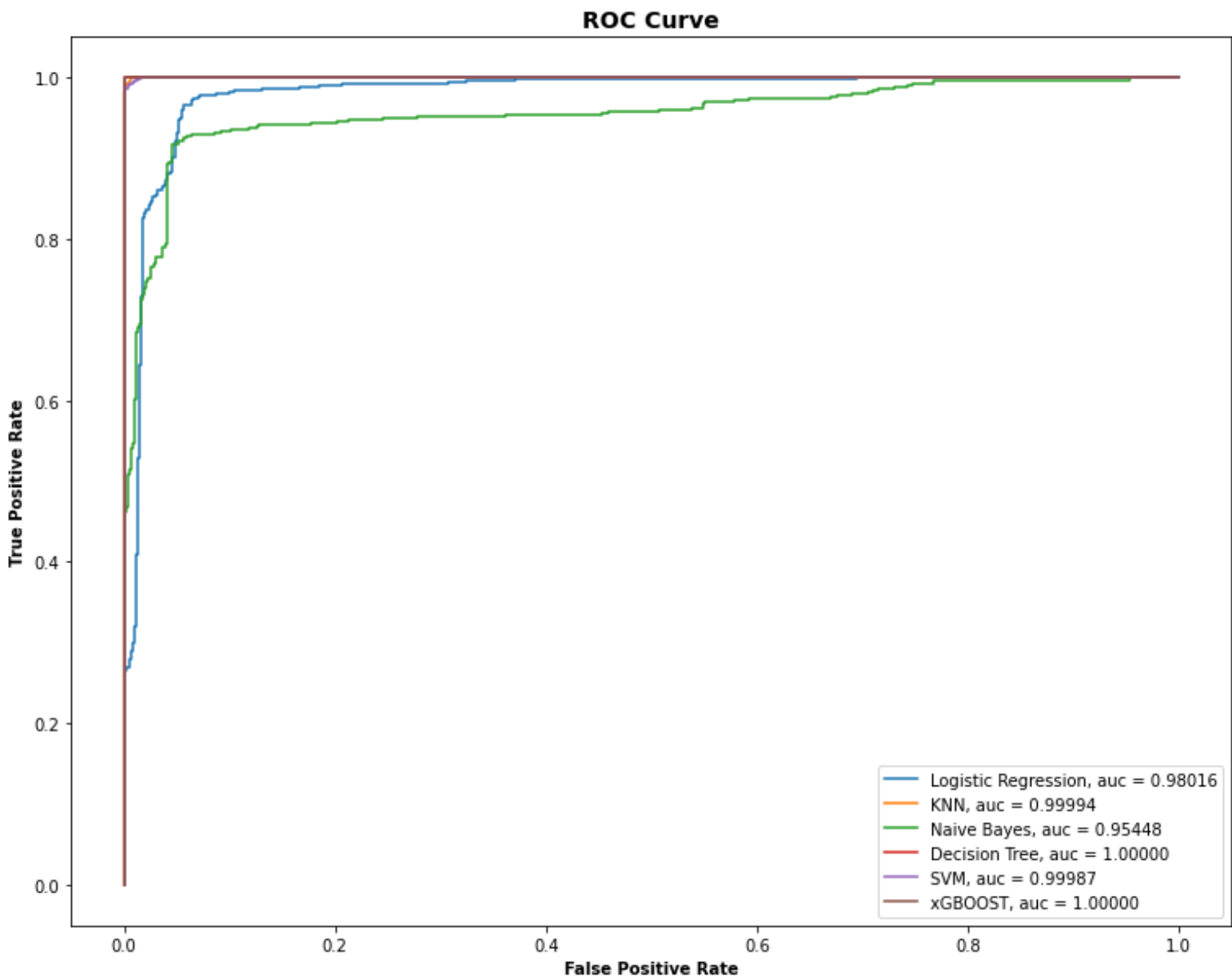
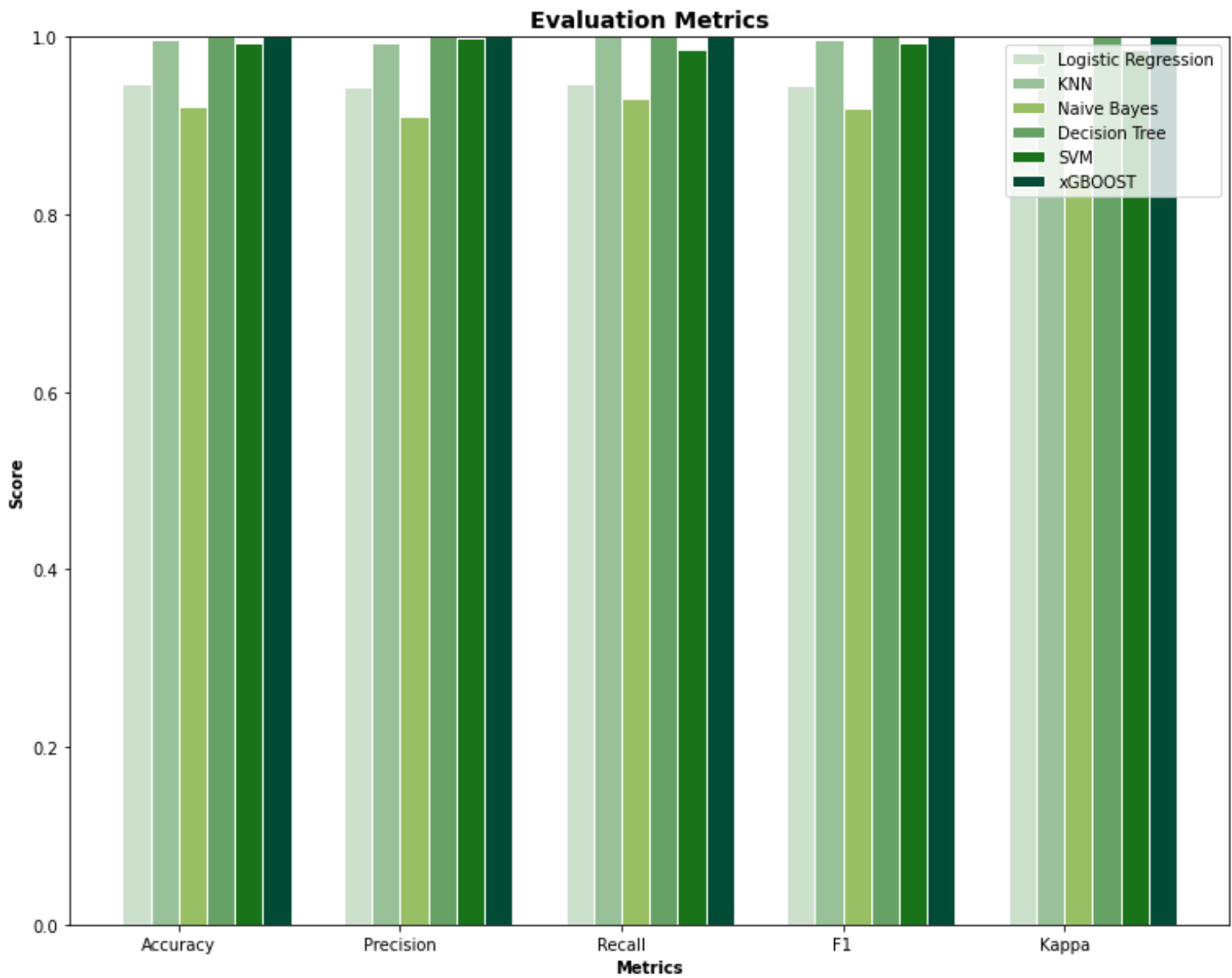ACCURACY: 0.9218461538461539

RECALL: 0.9296675191815856

Comparison

Model Comparison

Conclusion

Thank you for listening

prepared by:

Tamader AboAlhassan , Afrah Almoutiri , Moneera Alfulaij