

Surveying The Sky

Sahil Gangele

Motivation

Data Science is changing Astronomy

Astronomy deals with a lot of data

Sloan Digital Sky Survey

Results

Astronomy deals with a lot of data

Sloan Digital Sky Survey

Results

Astronomy deals with a lot of data

Sloan Digital Sky Survey

Results

The Data

10,000 instances

[illegible]

[illegible]

Classes



Stars



Quasars



Galaxies

Problem / Challenges

Dealing with Celestial Objects

How can I classify celestial objects?

Feature engineering to maximize classification accuracy

Experiment of classifiers to determine optimal one

How can I classify celestial objects?

Feature engineering to maximize classification accuracy

Experiment of classifiers to determine optimal one

How can I classify celestial objects?

Feature engineering to maximize classification accuracy

Experiment of classifiers to determine optimal one

Linear SVC

Decision Tree

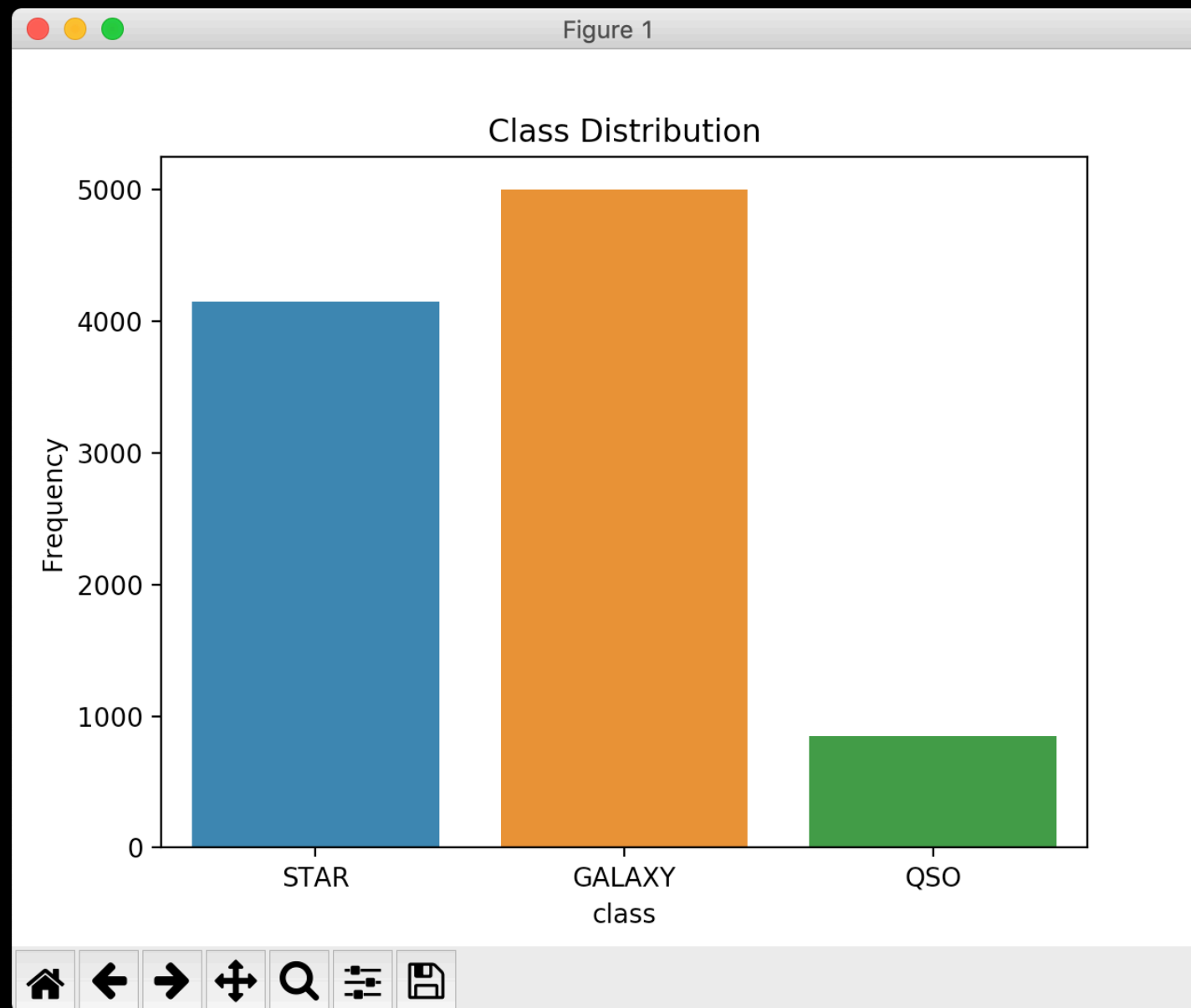
Random Forest Decision Tree

KNN

Neural Network

Naive Bayes

Imbalance



Evaluation

✗ or ✓

True Label

GALAXY
QUASAR
STAR
	GALAXY	QUASAR	STAR

Predicted Label

Business Relevance



Astronomers no longer have to analyze the data

Current Results



Figure 1

