STAR CLASSIFICATION

Abigail Campbell

Capstone Project

THE EXPANSE IN THE NIGHT SKY

There are between 2600 and 4500 objects visible in the night sky, with millions more detectable by telescope.

Humans have been looking at, finding patterns in, and identifying these objects for thousands of years.

Can our lab write a model to identify the stars in the night sky and help humanity farther down the path of understanding our universe?

CAN WE TRAIN A MODEL CLASSIFY A STELLAR OBJECT?

Goal:

Create an model that can correctly classify a stellar object based on telescopic measurements

Benefits:

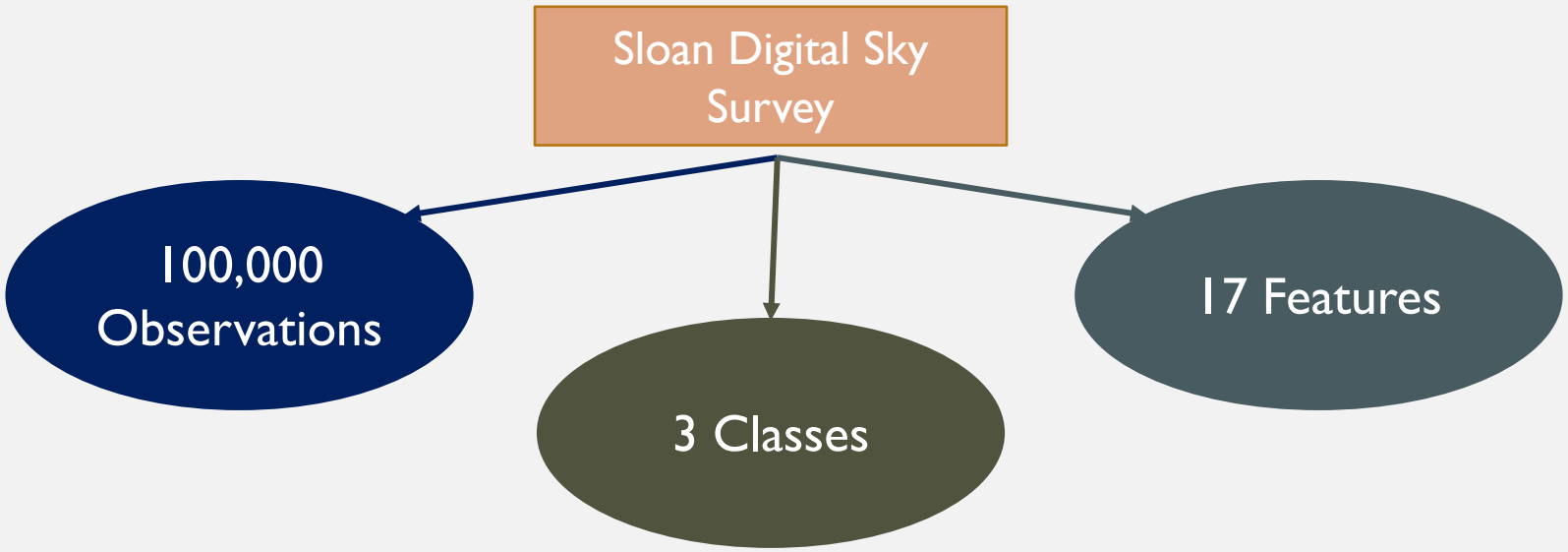
Faster object classification

Create an accurate sky map

Free up tech time

DATA SOURCE: SLOAN DIGITAL SKY SURVEY

kaggle

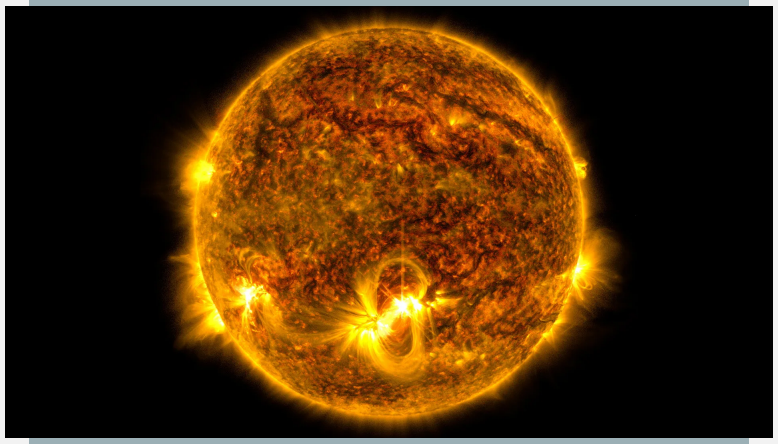


Redshift	EM Meas.	Coordinates	Class

CLASSES OF STELLAR OBJECTS

Star

luminous ball of gas held together by self-gravity



Galaxy

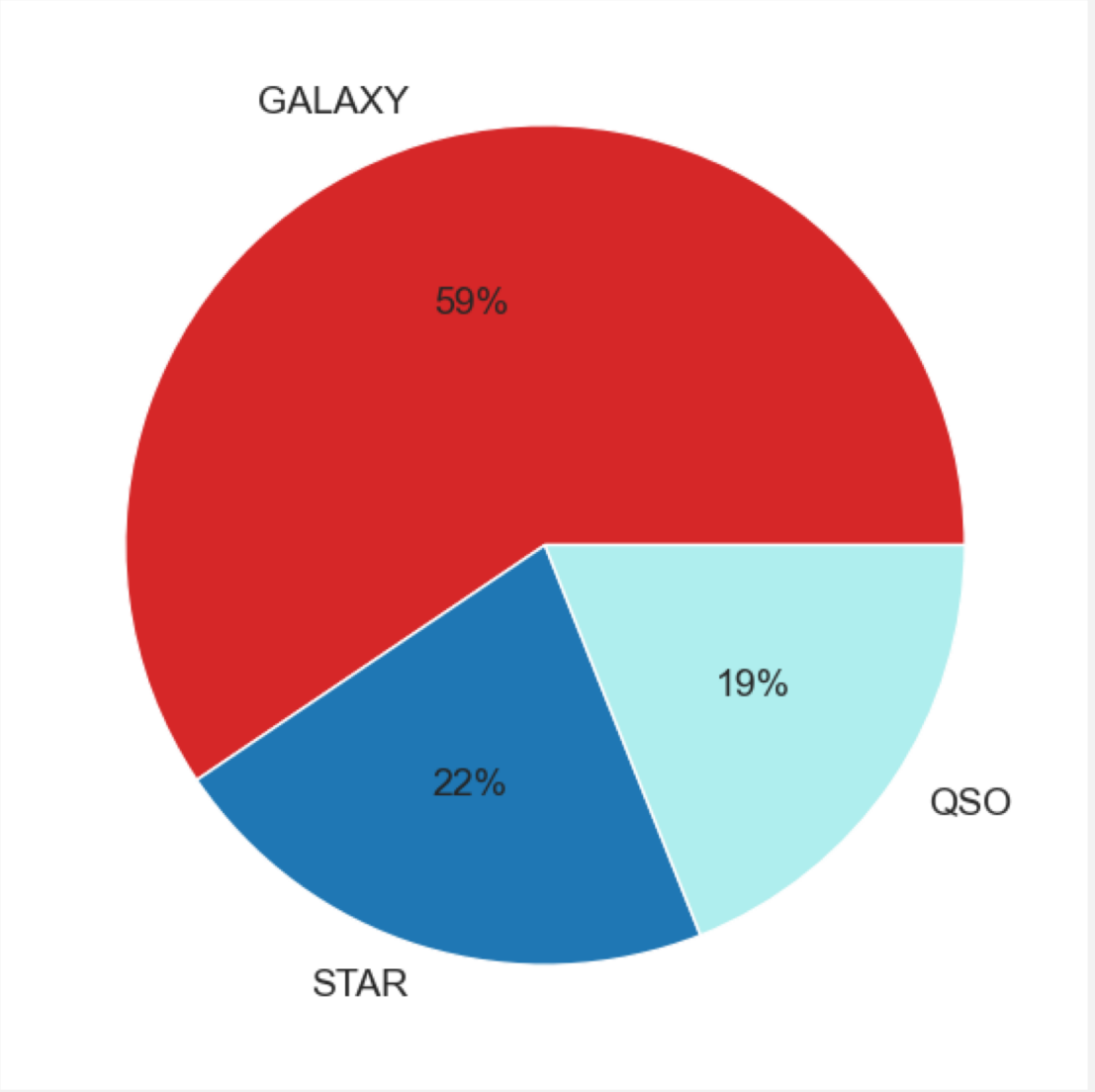
Large collection of dust, gas, and billions of stars

Quasar

Super luminous galactic nucleus, emitting intense radiation

A LOOK INSIDE THE DATA – CLASSIFICATIONS

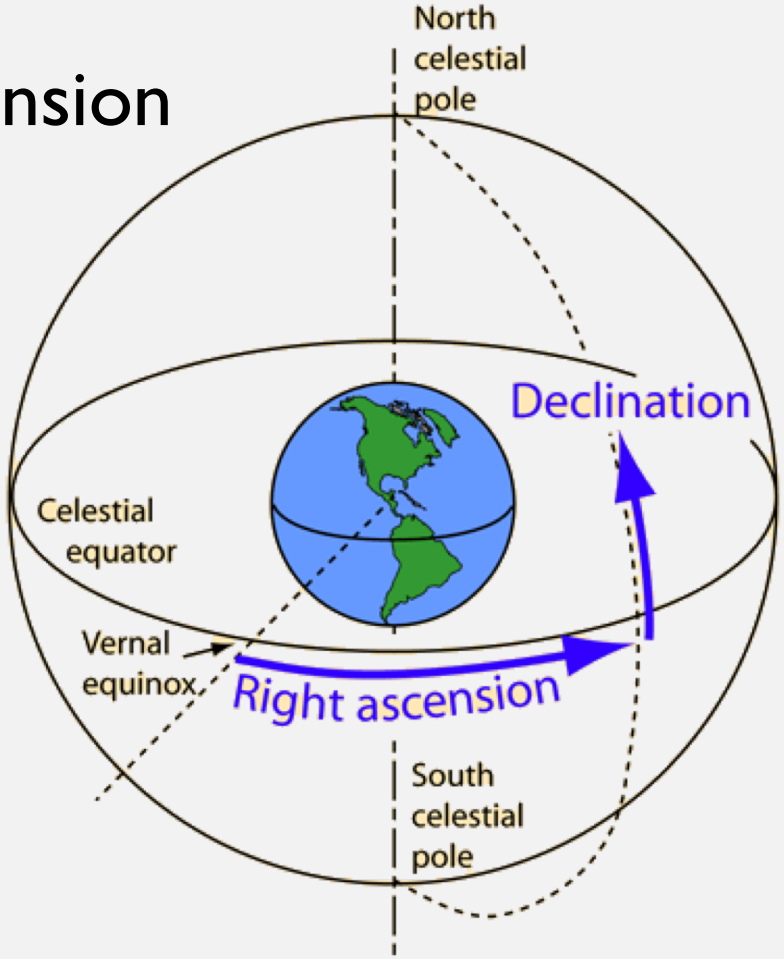
- The majority of observations are galaxies
- Star and Quasar classifications are roughly equal



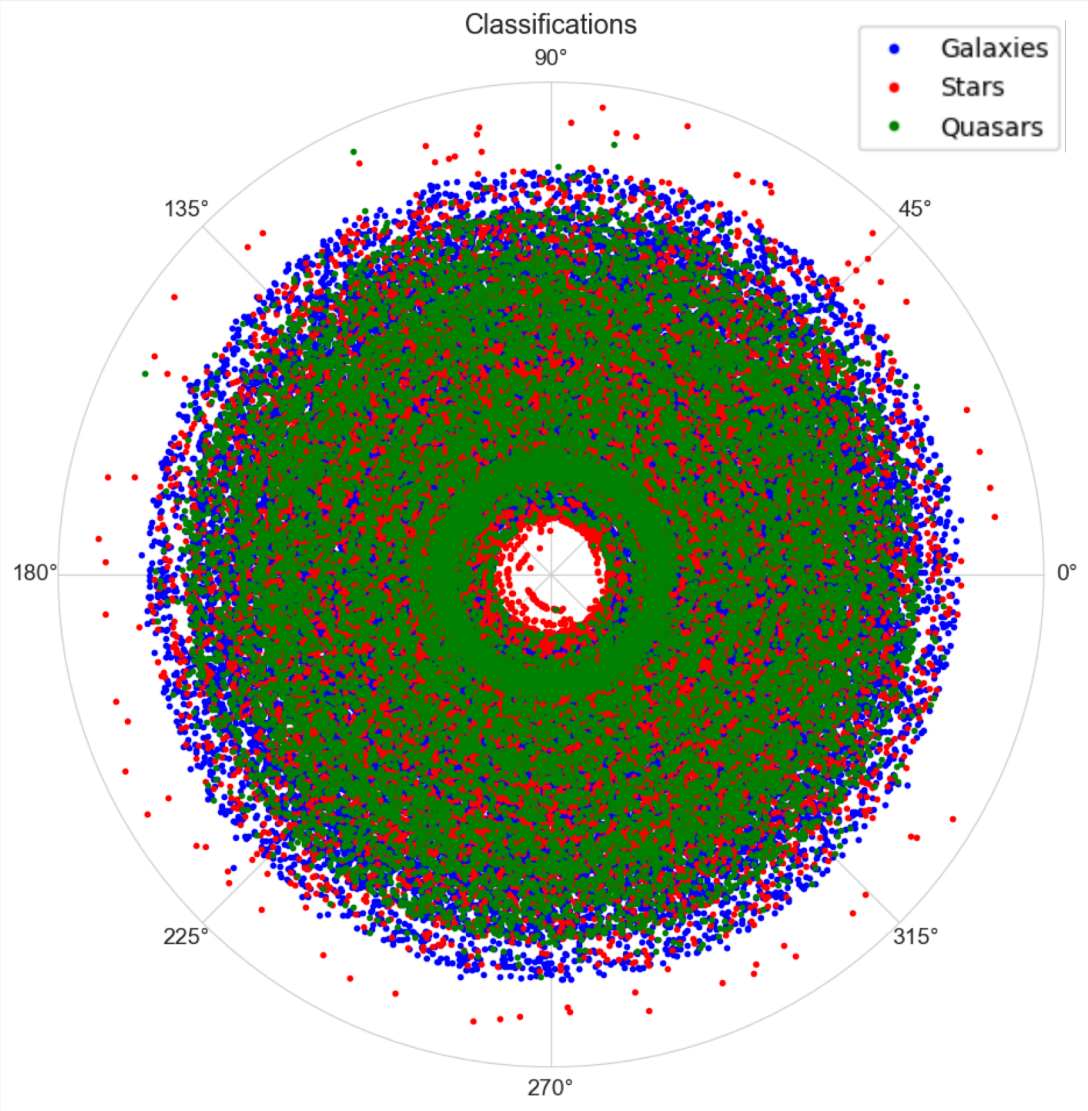
A LOOK INSIDE THE DATA – OBJECT LOCATIONS

Delta: Declination angle

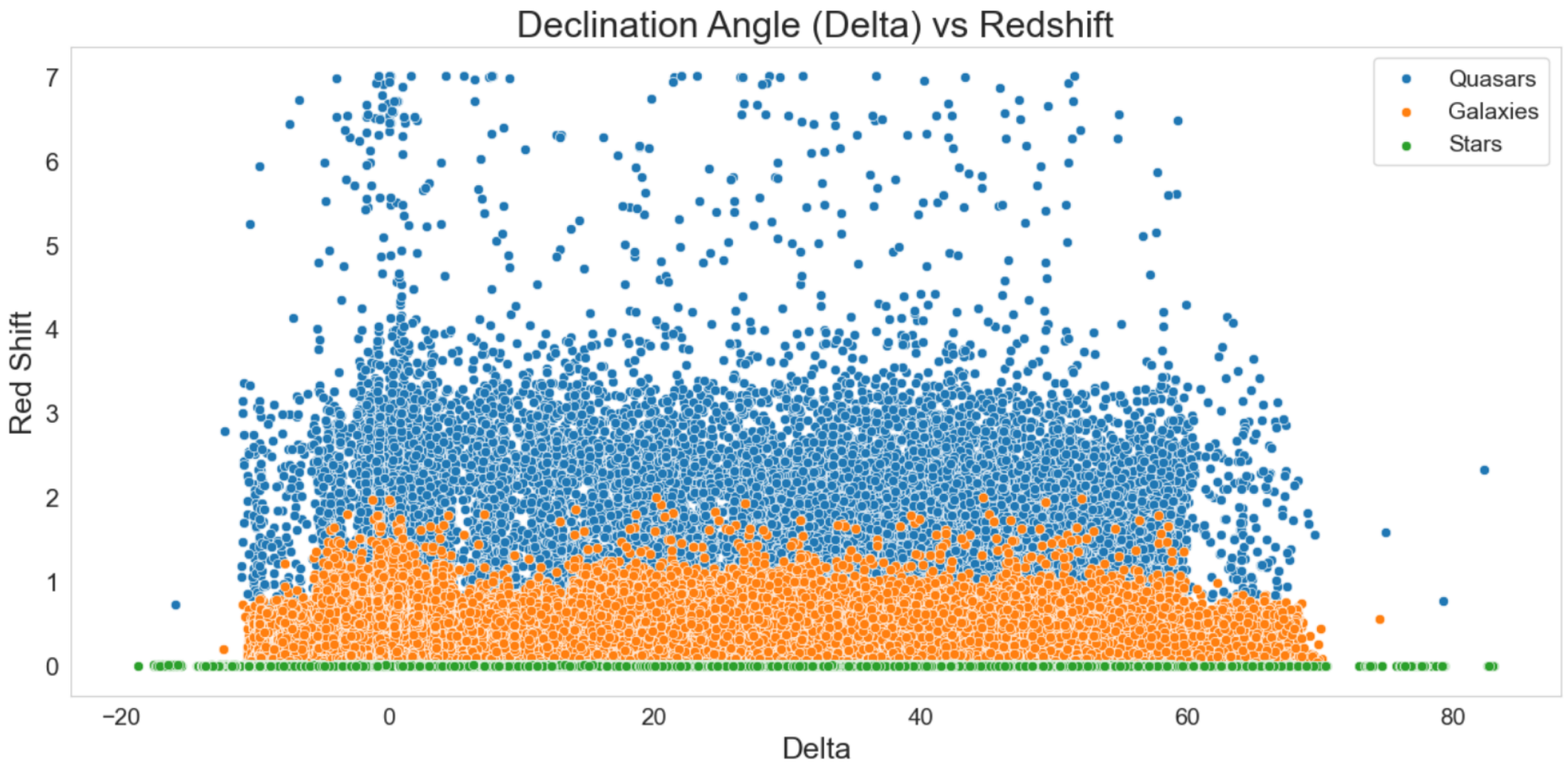
Alpha: Right ascension



Even distribution of objects in the sky



A LOOK INSIDE THE DATA – REDSHIFT



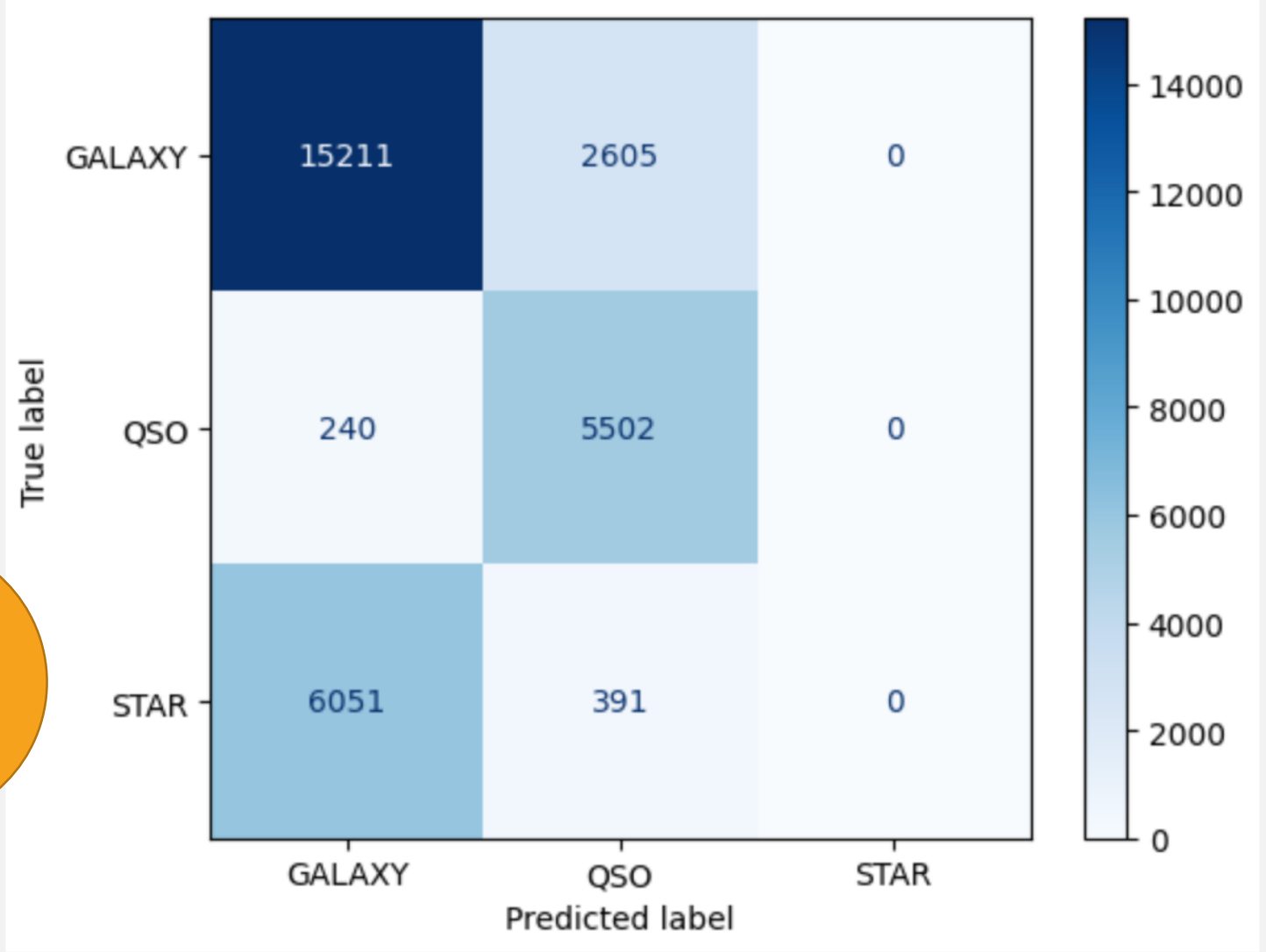
Clear redshift ranges for each class

BASE MODEL

Decision Tree

- Unable to classify stars
- Most stars mis-identified as galaxies
- More galaxies mistaken for quasars than vice versa

69%
accuracy

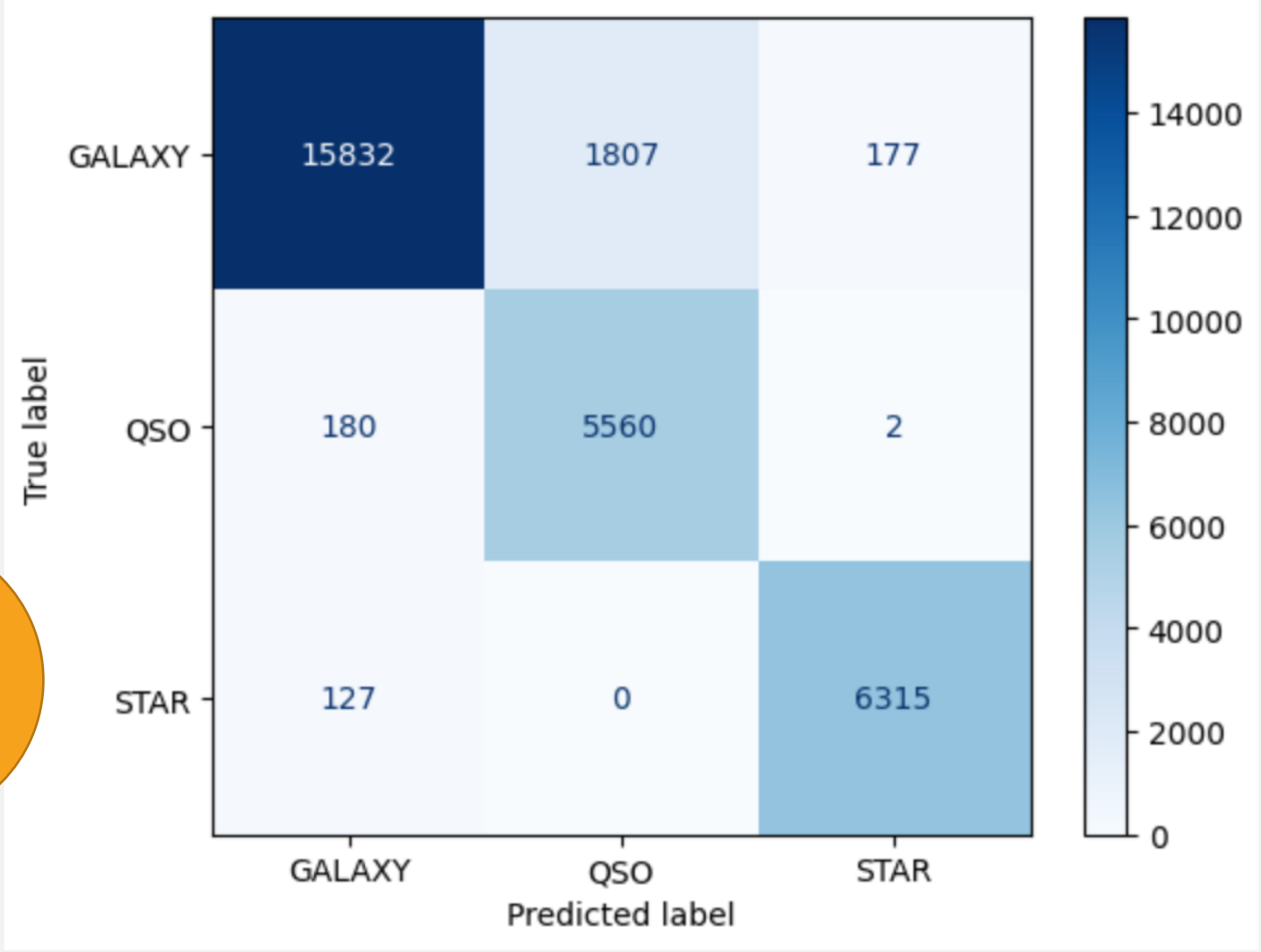


FINAL MODEL

Two Layer Neural Net

- Optimized and trained
- Cross-Validated
- Capable of classifying stars
- Mis-identifications decreased among all classes

93%
accuracy

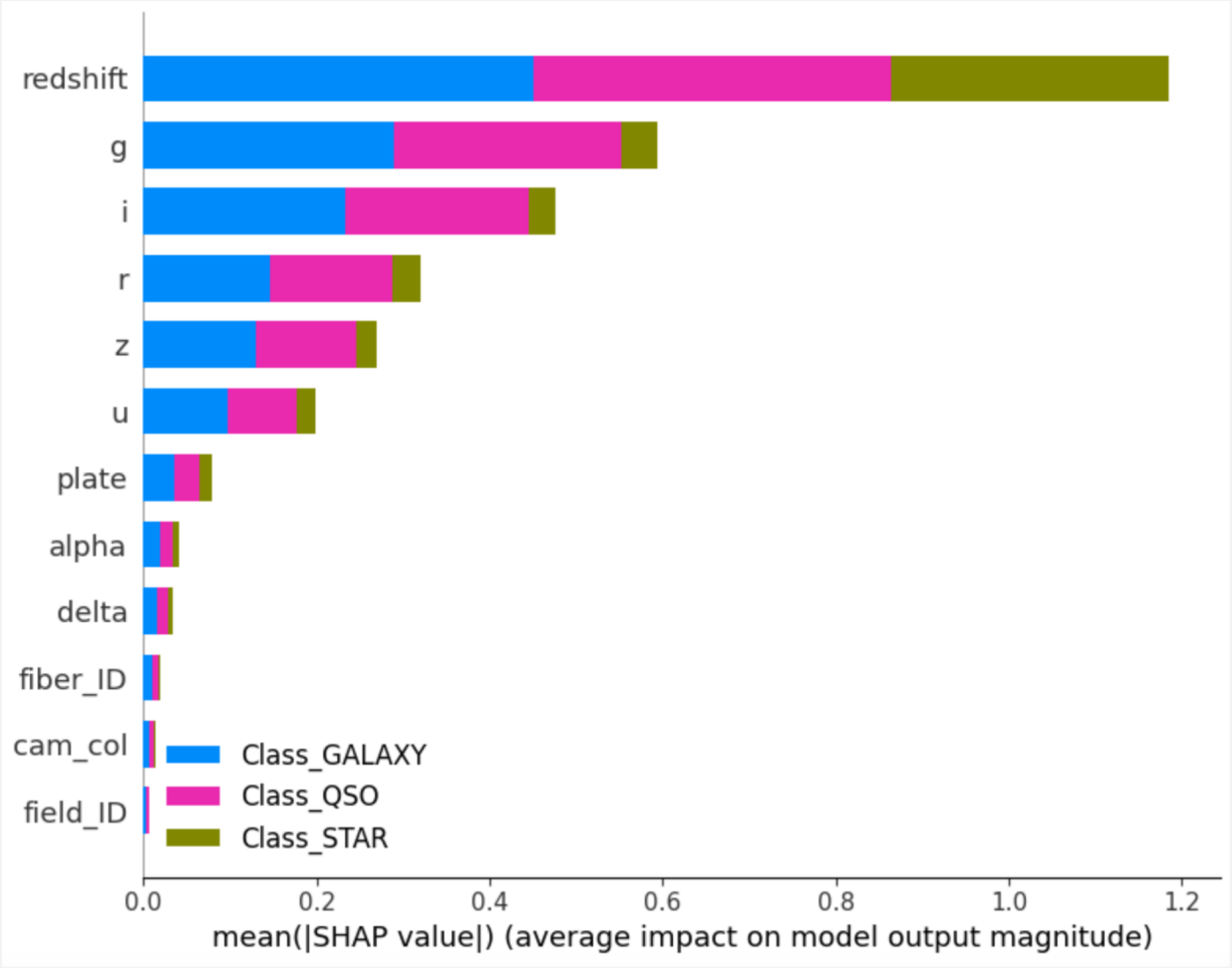


FEATURE IMPORTANCE

Most important features:
Measurements made on the objects themselves

1. Redshift
2. Green filter
3. Near infra-red filter

Information about location and equipment used to observe not as important



CONCLUSIONS

The final trained as a two layer neural network using normalized and resampled data and optimized using a grid search of parameters.

Using this final model, we are able to accurately classify objects observed in the night sky with an accuracy of 93%.

WHAT'S NEXT?

Feature Pruning

- Utilize feature importance results to simplify the data set

Integrate the neural net

- Integrate with our existing platform architecture

CONTACT INFORMATION

Abigail Campbell

abbycampbell0@gmail.com

(801) 541-2771