Celestial Object Classification

Presented by Parker Stratton
28 February 2022

Project Task and Purpose

Task: Build and train a Machine Learning Algorithm to classify celestial objects based on EM Spectrum emissions and Red Shift.

Purpose: Enable professional and armature scientists to automate and reduce error of classification of celestial object observations.

Dataset Summary

Dataset:

- Majority of Data:
  - Location of celestial object
- Critical Data
  - Electromagnetic emissions
  - Redshift
- What we are looking for:
  - Stars
  - Galaxies
  - Quasars

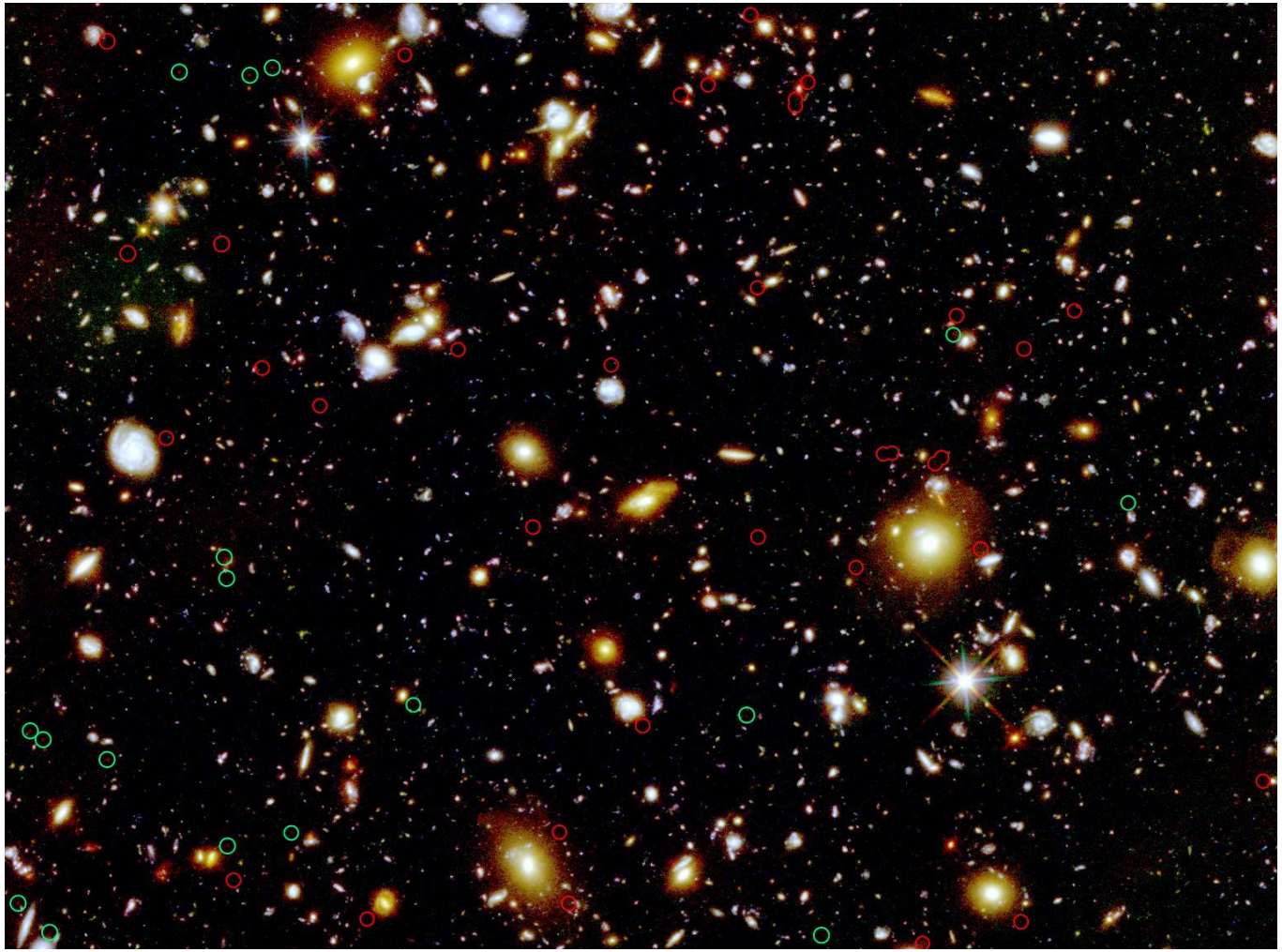


Image Credit: NASA.gov

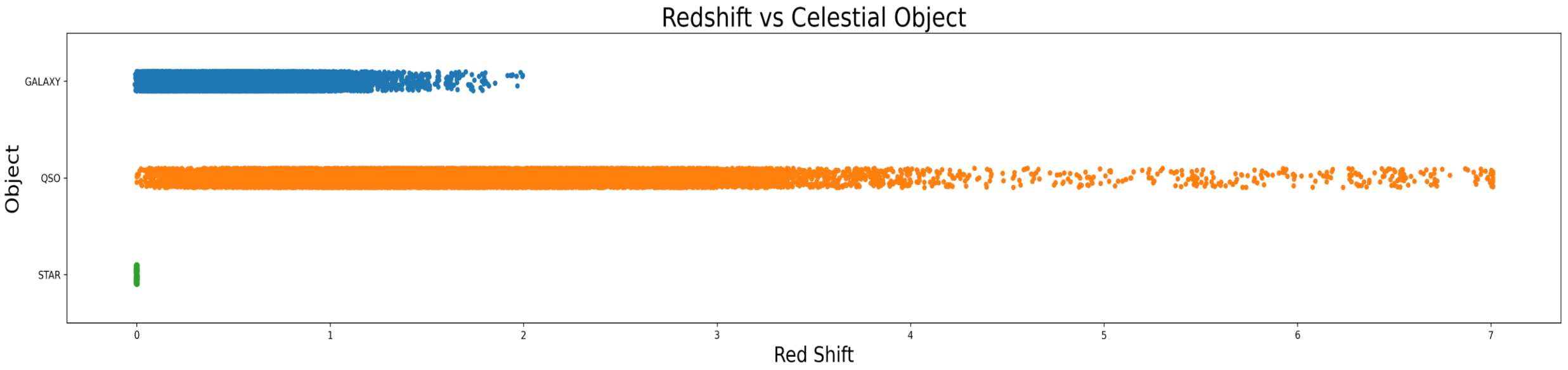
Dataset from: https://www.kaggle.com/fedesoriano/stellar-classification-dataset-sdss17

Celestial Object Classes

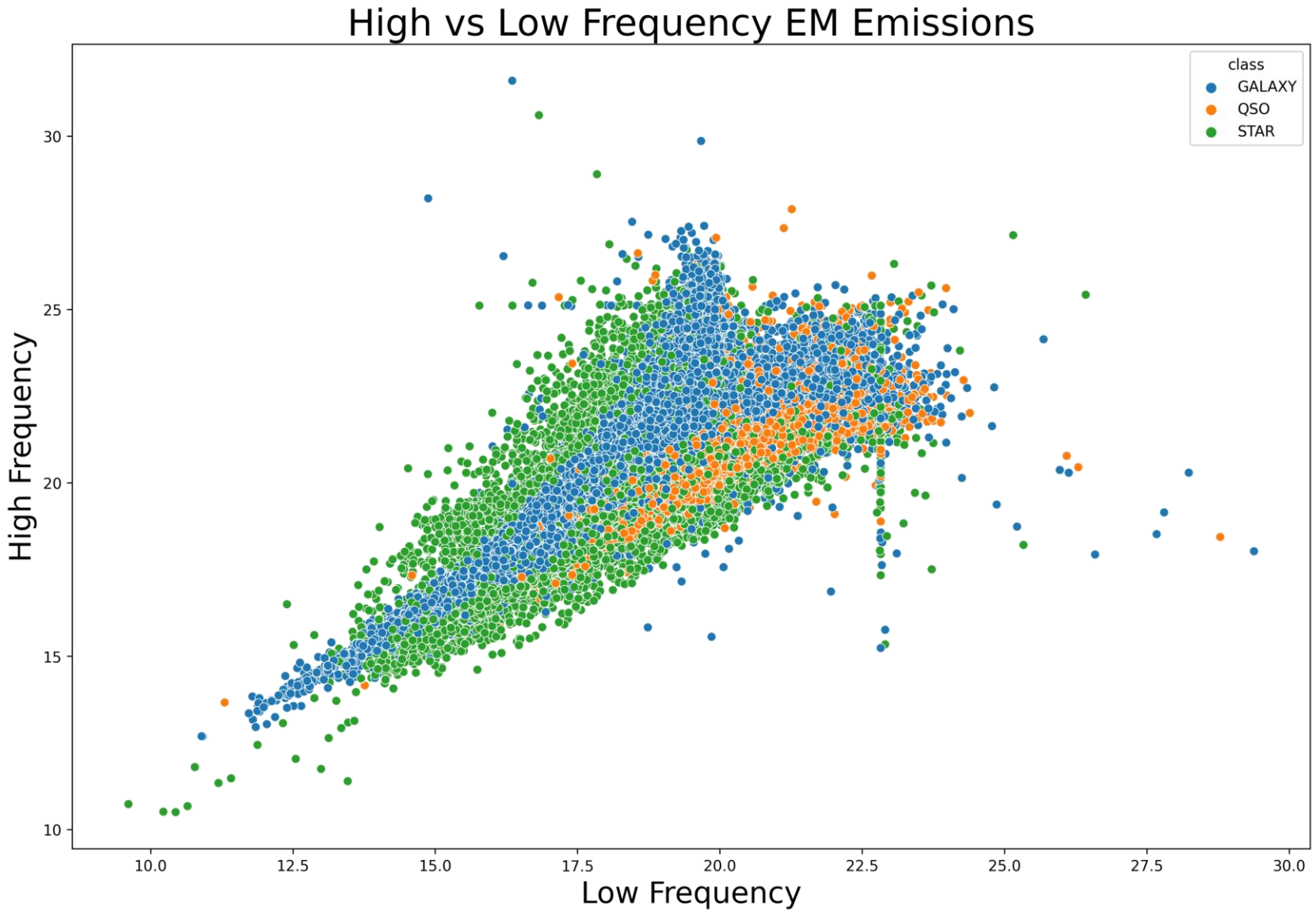
Key Trends:

Redshift

1. **Stars**
2. **Galaxies**
3. **Quasars**



Celestial Object Classes



Key Trends:

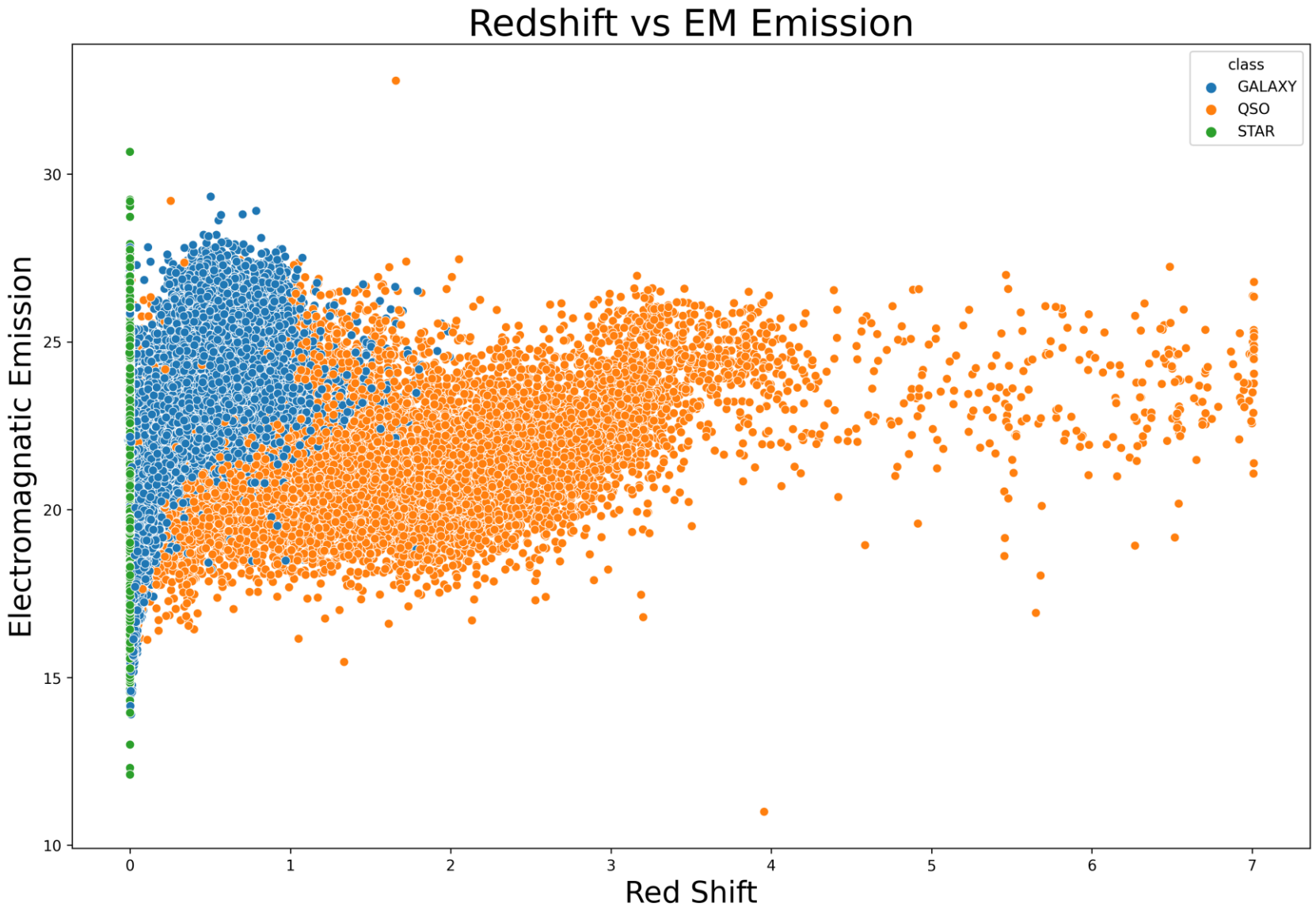
Redshift

1. Stars
2. Galaxies
3. Quasars

EM Emissions

- **Highly related**

Celestial Object Classes



Key Trends:

Redshift

1. Stars
2. Galaxies
3. Quasars

EM Emissions

- Highly related

Redshift and EM Emissions

- **Breakout of celestial object as they related to the two**

Production Model

Gradient Boosting Classifier

Reliability: 98%

Issue:

Distinguishing between Galaxies and Quasars.

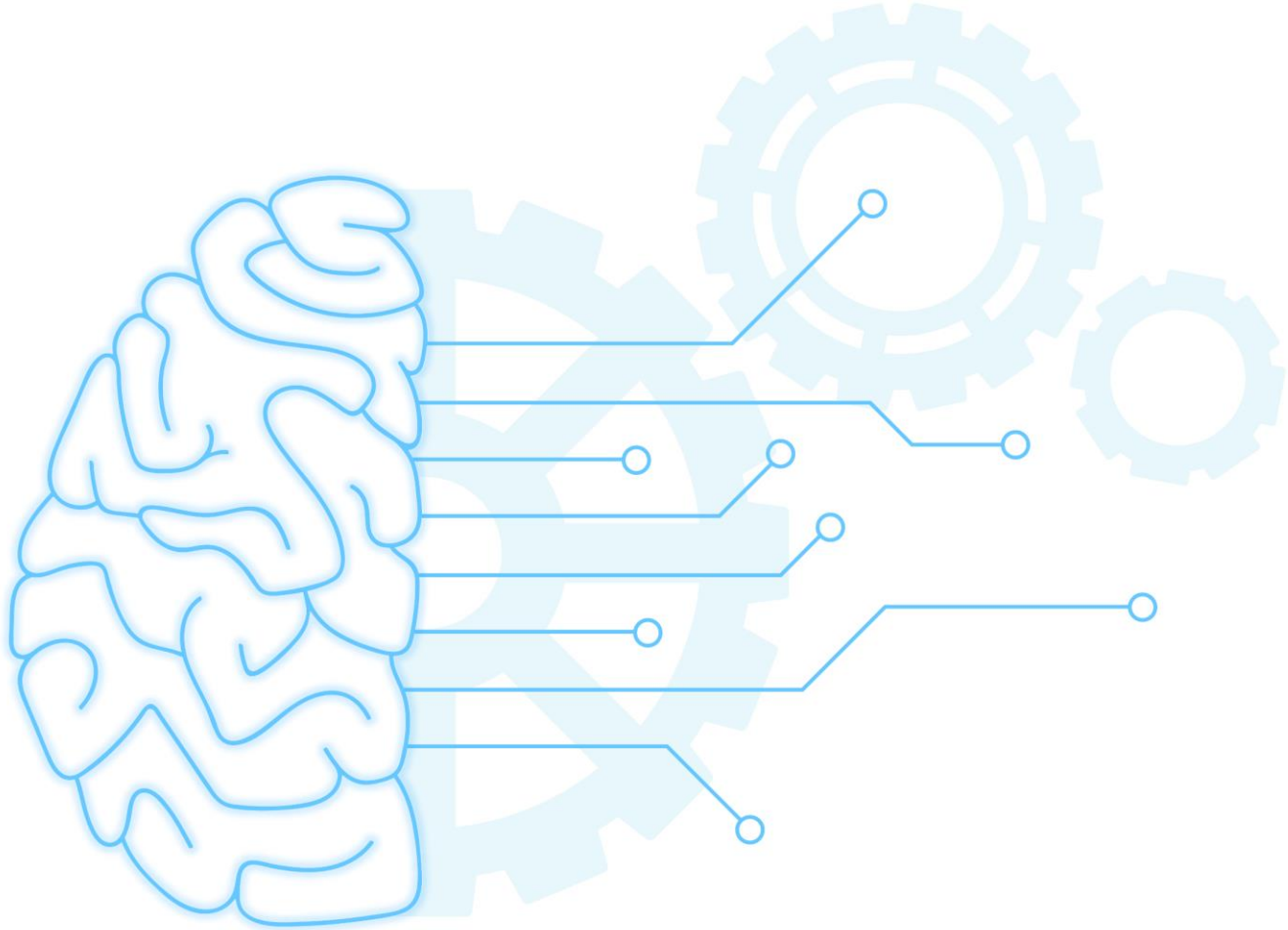


Image Credit: https://docs.microsoft.com/en-us/windows/ai/windows-ml/

Conclusion

Recommendations

1. Immediately roll out model for use by professional and amateur astronomers
   - Caveat: classification should be verified through traditional methods

2. Incorporate existing data from other observation systems to improve model

3. Continue to collect data from both professional and amateur astronomers to improve model