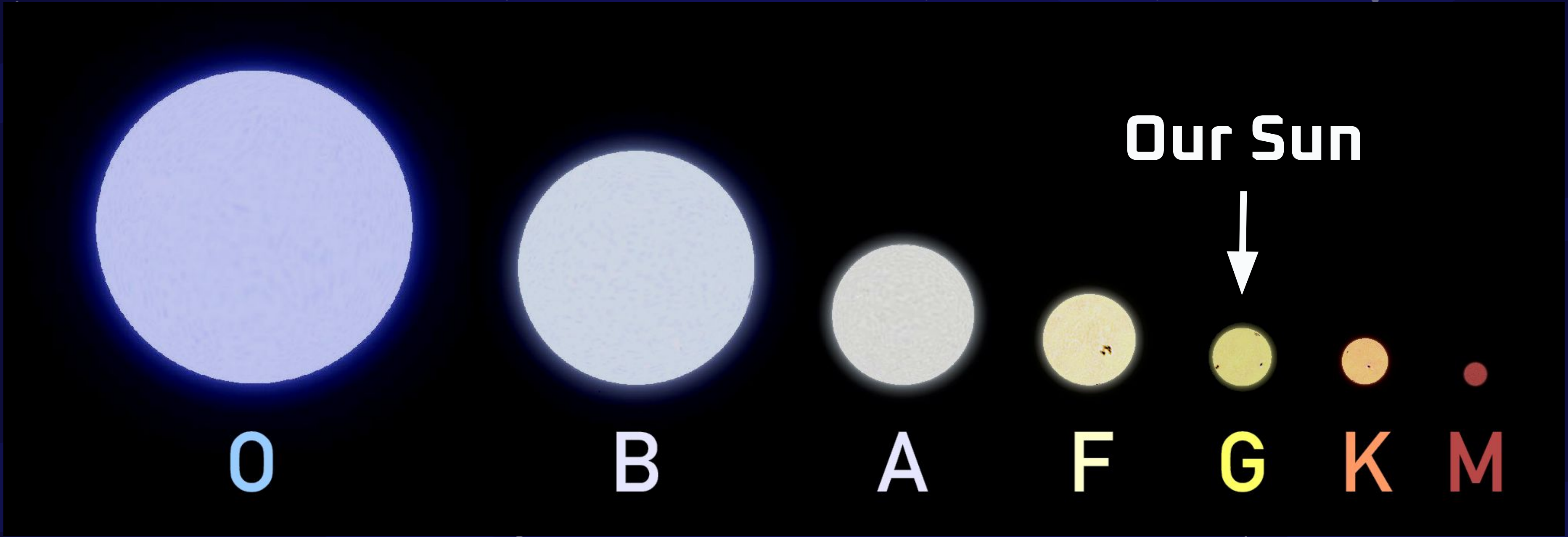
Classifying Stellar Spectra

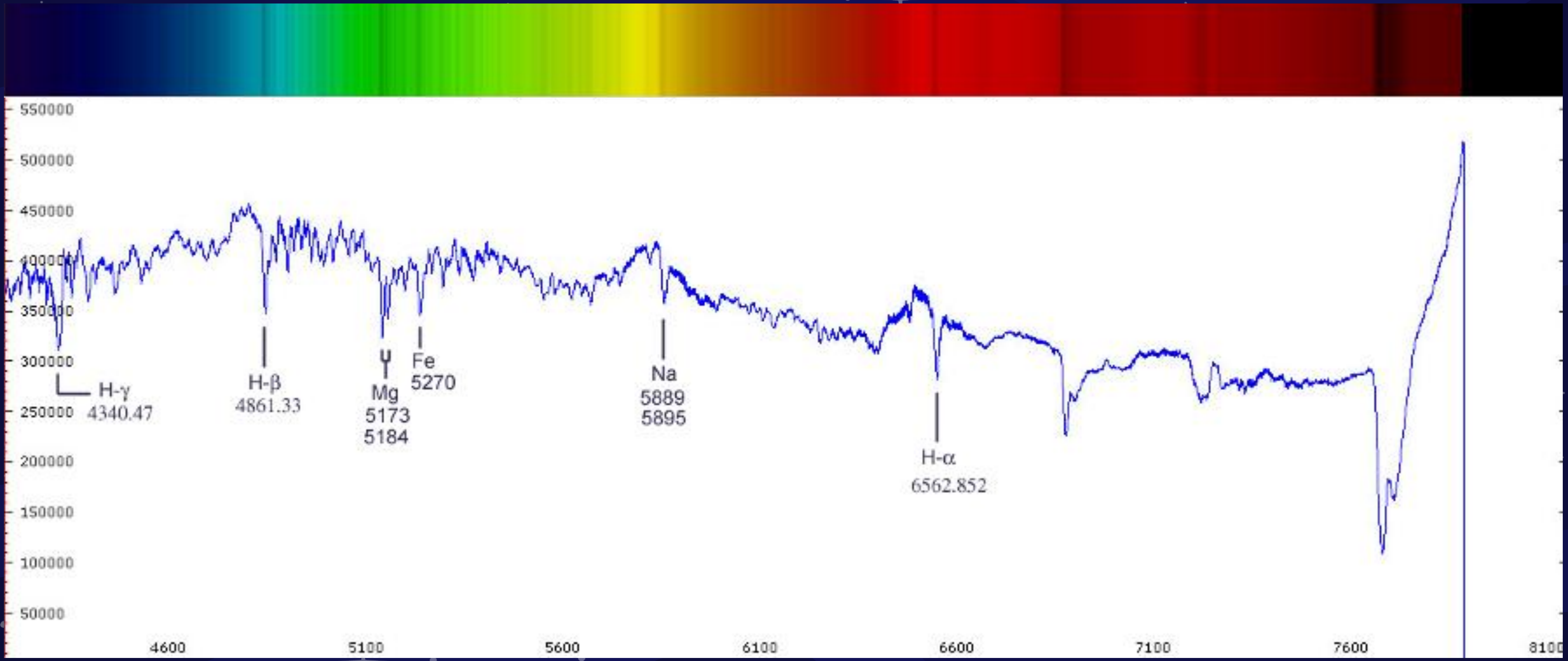
By Edith Johnston

Table of Contents

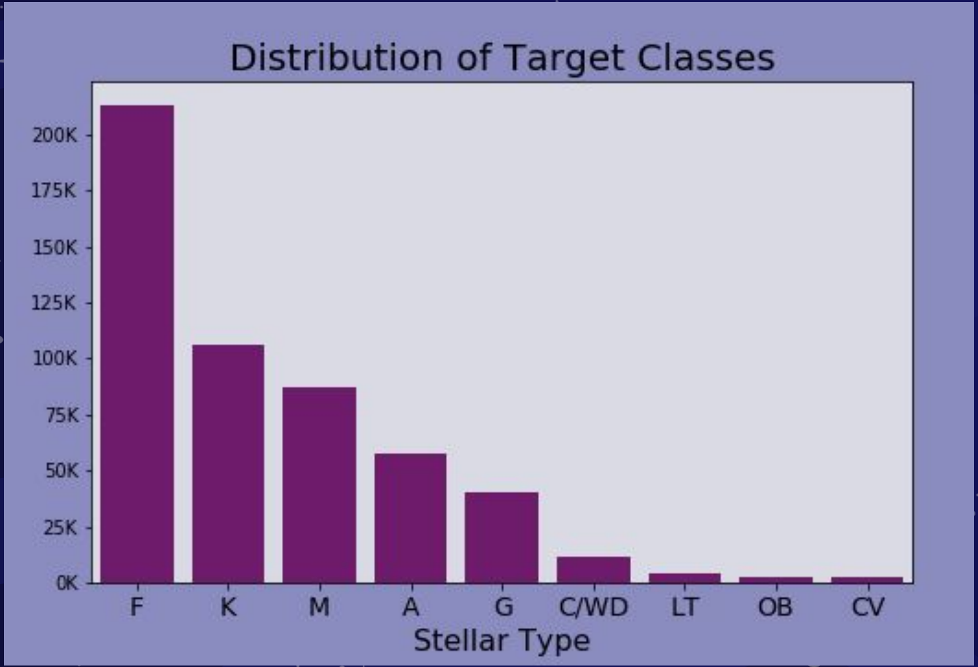
Stellar Spectral Types



Stellar Spectra



Class Distribution



- ★ Largest 5 Classes: Random Undersampling
- ★ Smallest 4 Classes: Random Oversampling

Advantages of This Model:

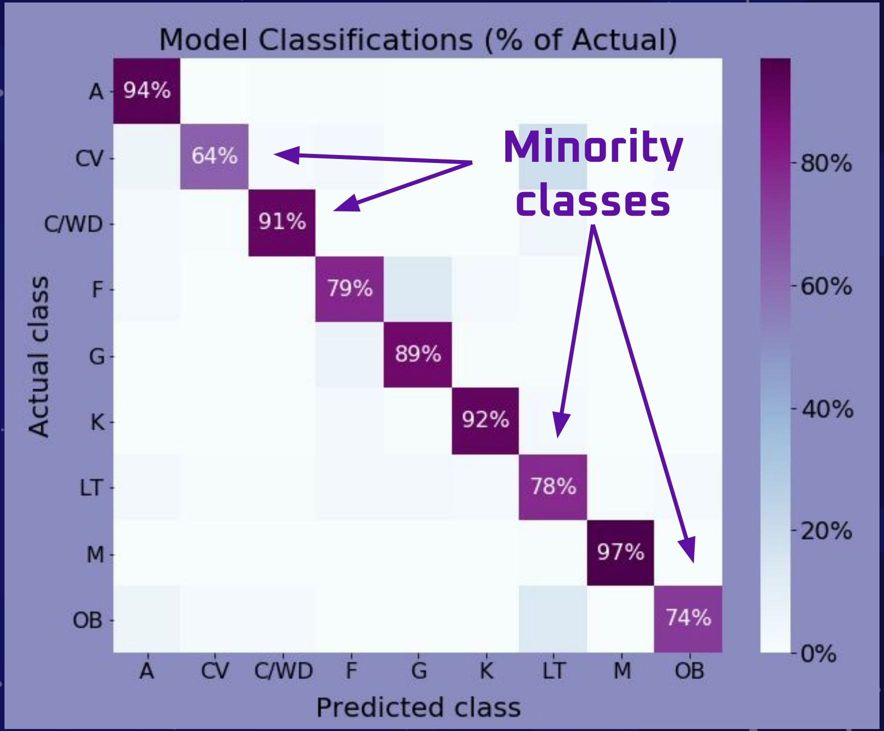
Model

Histogram-based Gradient Boosting Classification Tree

- ★ Scales well to more features
- ★ NaN handling
- ★ Faster for large sample sizes

Model Performance

Balanced accuracy score: 0.85



Correct classifications are on the diagonal

Next Steps

Adding Capability

- ★ Subclasses
- ★ Non-stellar objects
  - ○ Galaxies
  - ○ Quasars

Improving Performance

- ★ Data from other sources
- ★ Train iteratively with resampled data

Thank You
For Your Time

Project by: Edith Johnston
Github repository:
https://github.com/edithalice/stellar_classification
LinkedIn: www.linkedin.com/in/edith-johnston

The Fine Print

This research makes use of the SciServer science platform (www.sciserver.org).

SciServer is a collaborative research environment for large-scale data-driven science. It is being developed at, and administered by, the Institute for Data Intensive Engineering and Science at Johns Hopkins University. SciServer is funded by the National Science Foundation through the Data Infrastructure Building Blocks (DIBBs) program and others, as well as by the Alfred P. Sloan Foundation and the Gordon and Betty Moore Foundation.

Appendix

What's the Point of Classifying?

★ Different behavior

★ Filtering searches

  ○ Potentially habitable planets can be found orbiting F or G stars

My Project Goal

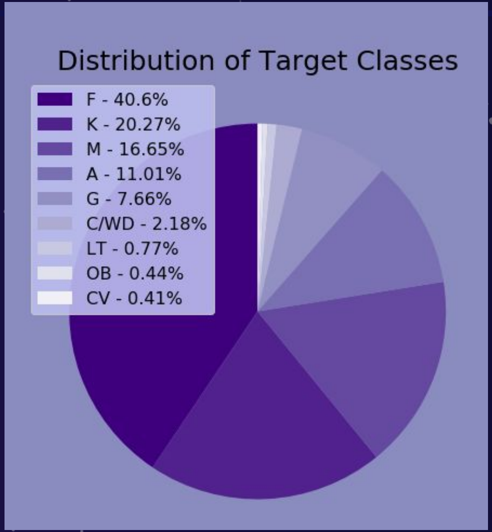
Too many stars!

⇩

Automation

Data Source:
[Sloan Digital Sky Survey]

Data Structure



Features

- 71 features with 4 sub features each
- Each feature represent elemental line index
- Sub features: global continuum fit, local value fit, error, pixel quality mask

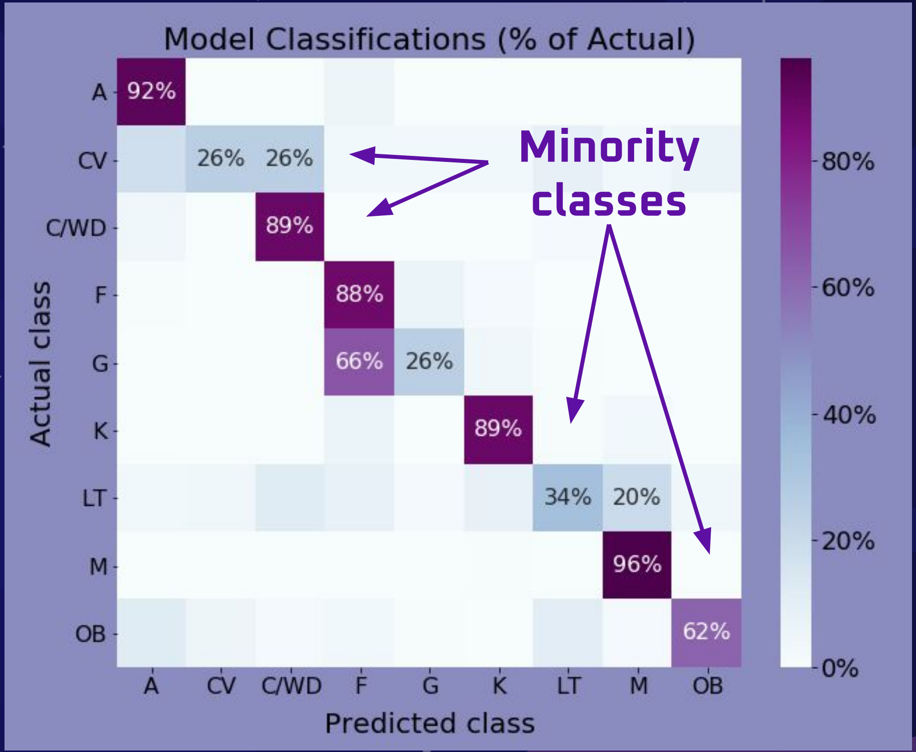
Target

- Initially, 21 classes
- After grouping by subclass, only 9 classes
- Highly imbalanced: 5 of the 9 classes contain >96% of samples

Model For Comparison:

Decision Tree

Balanced accuracy score: 0.67

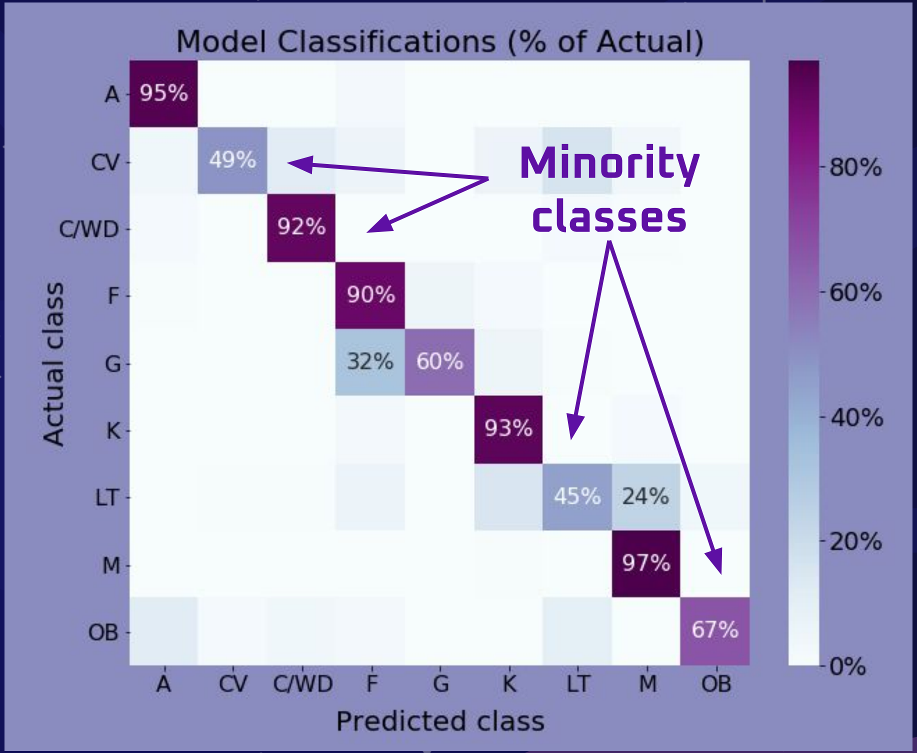


Correct classifications are on the diagonal

Model For Comparison:

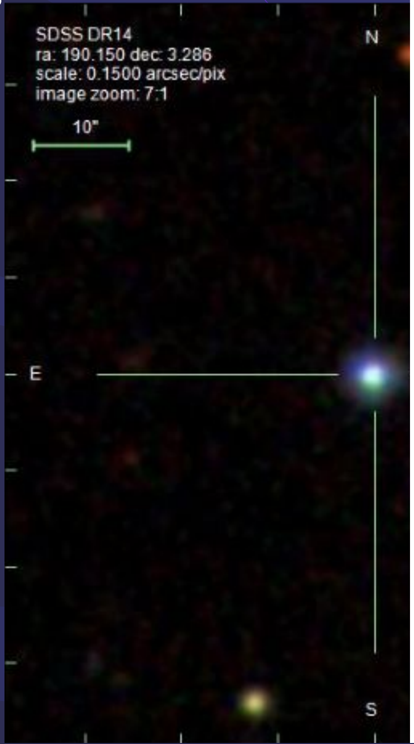
Random Forest

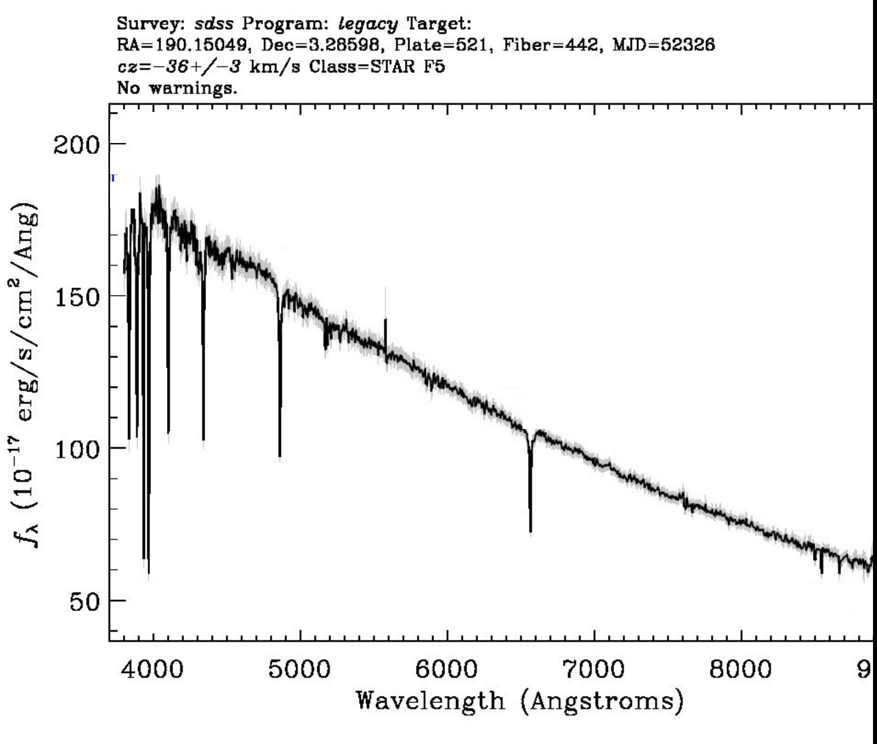
Balanced accuracy score:
0.77



Correct classifications are on the diagonal

Sample Pipeline

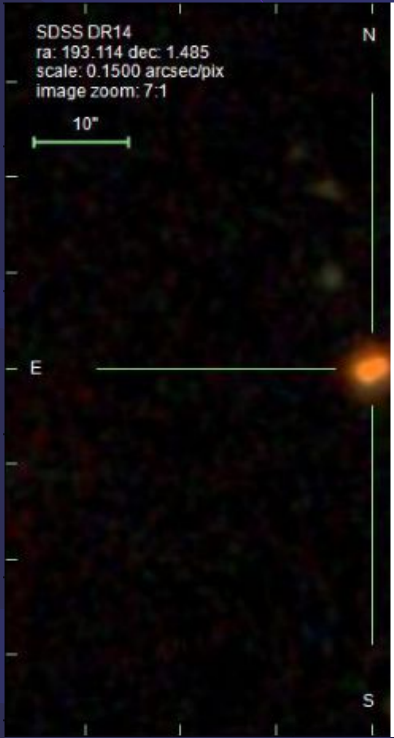


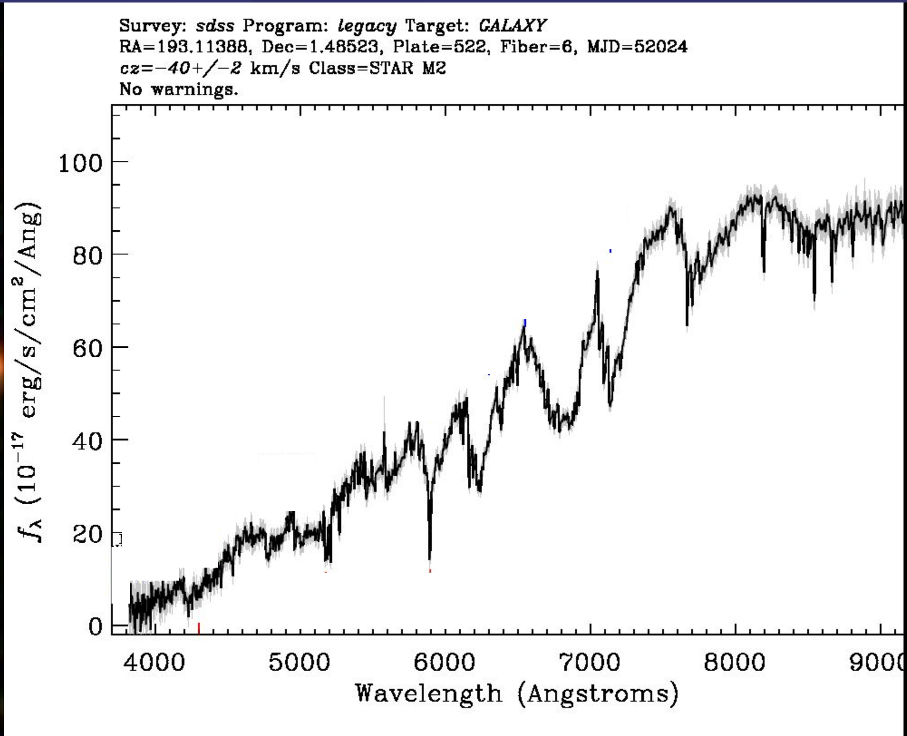


F-type star (F)

Mass: 0.8–1.4 M☉
Radius: 1.15–1.4 R☉
Luminosity: 1.5–5 L☉
Temperature: 6,000–7,600 K

Sample Pipeline 2



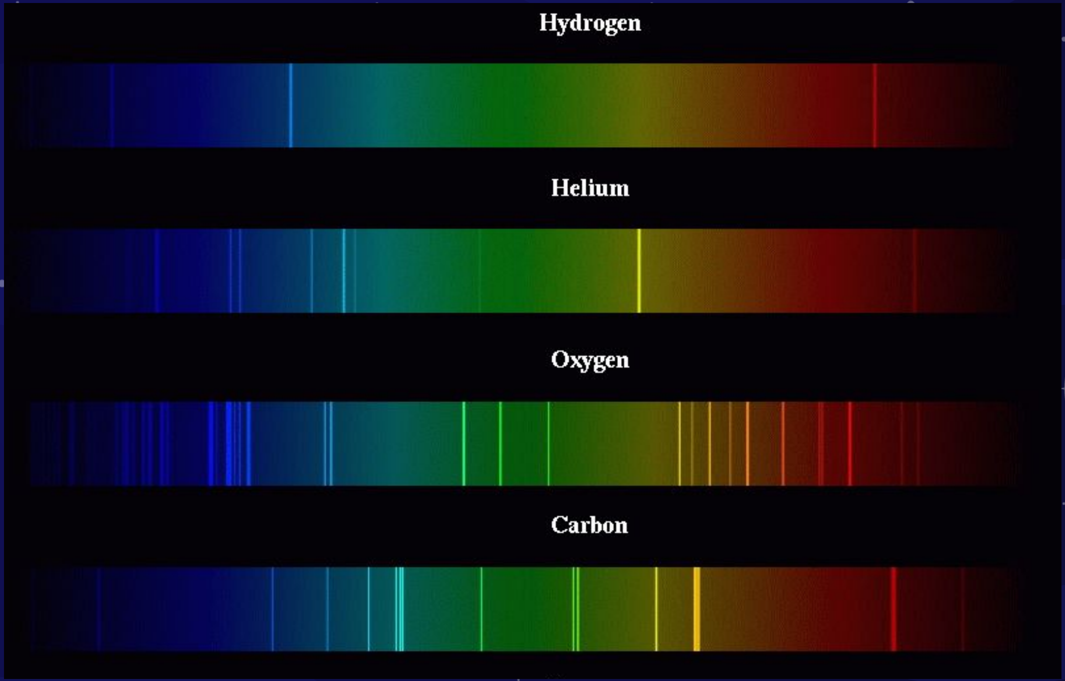


More of the Science

Spectrography

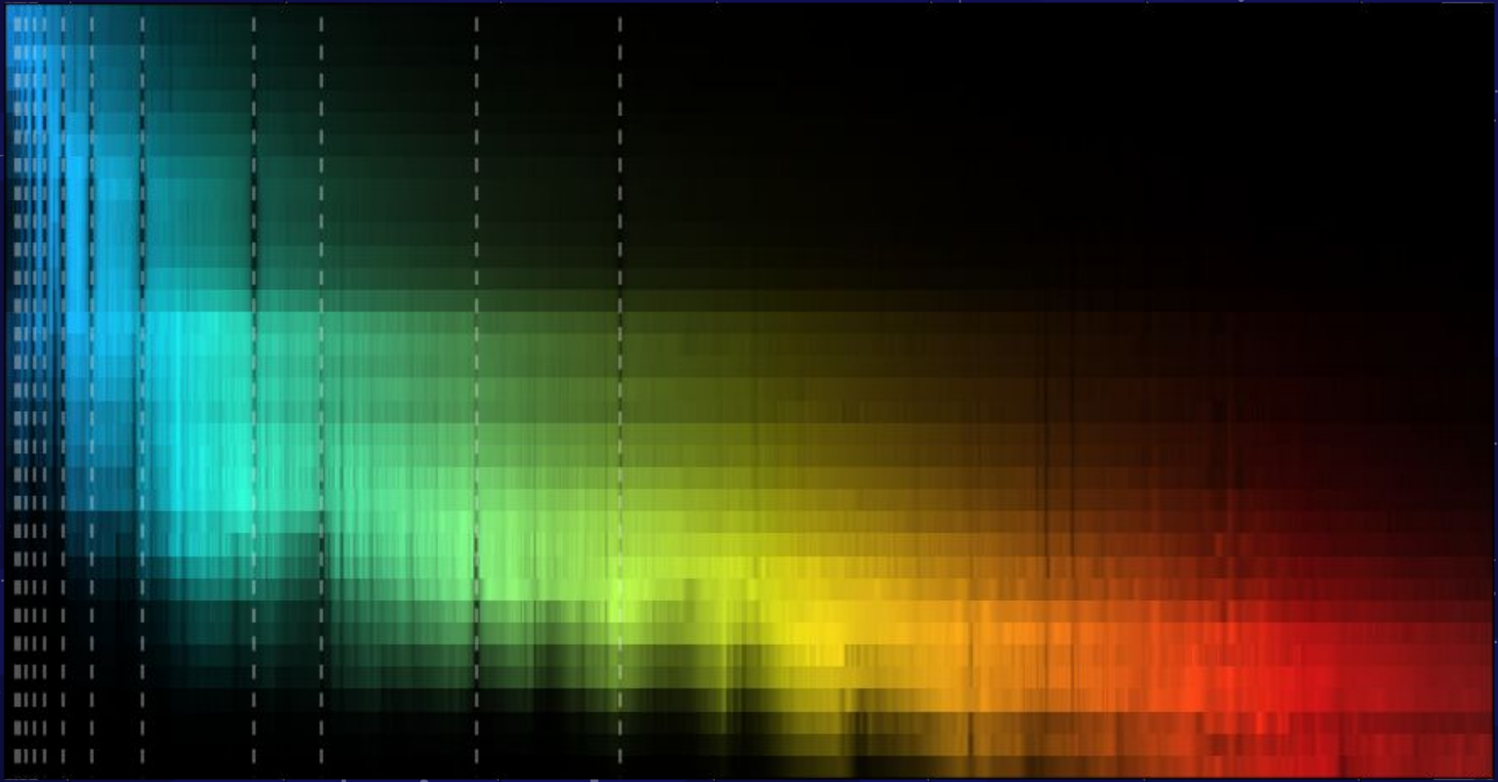
★ Different elements absorb and emit different wavelengths of light

★ Therefore we can determine presence of elements by looking for absorption lines (dark lines) and emission lines (bright lines)



Spectrography

O
B
A
F
G
K
M

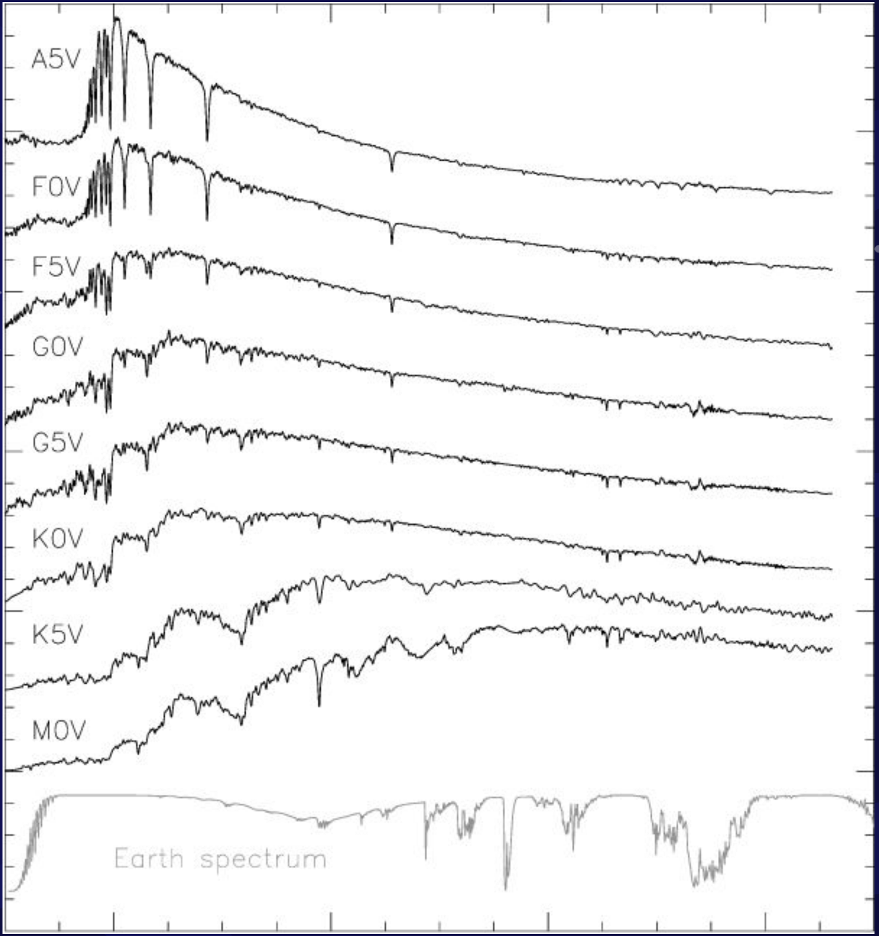


A bunch of different stellar spectra, ranging from Type O to Type M.

The different elements in these stars are causing lots of little lines.

Spectrographs

- ★ These spectrographs can also be represented as graphs like these.
- ★ Those absorptions and emission lines now show up as sharp spikes or dips on the graph.
- ★ The surface temperature of the star can be determined from these graphs by looking at highest peak.



Spectral Types

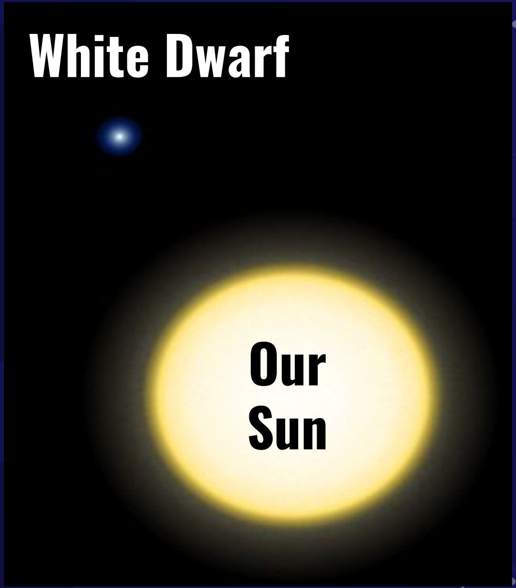
Normal spectral types range from blue hypergiants in Type O
to brown dwarfs in Y (which emit almost no light).

Spectral Types

The Itsy Bitsy Ones

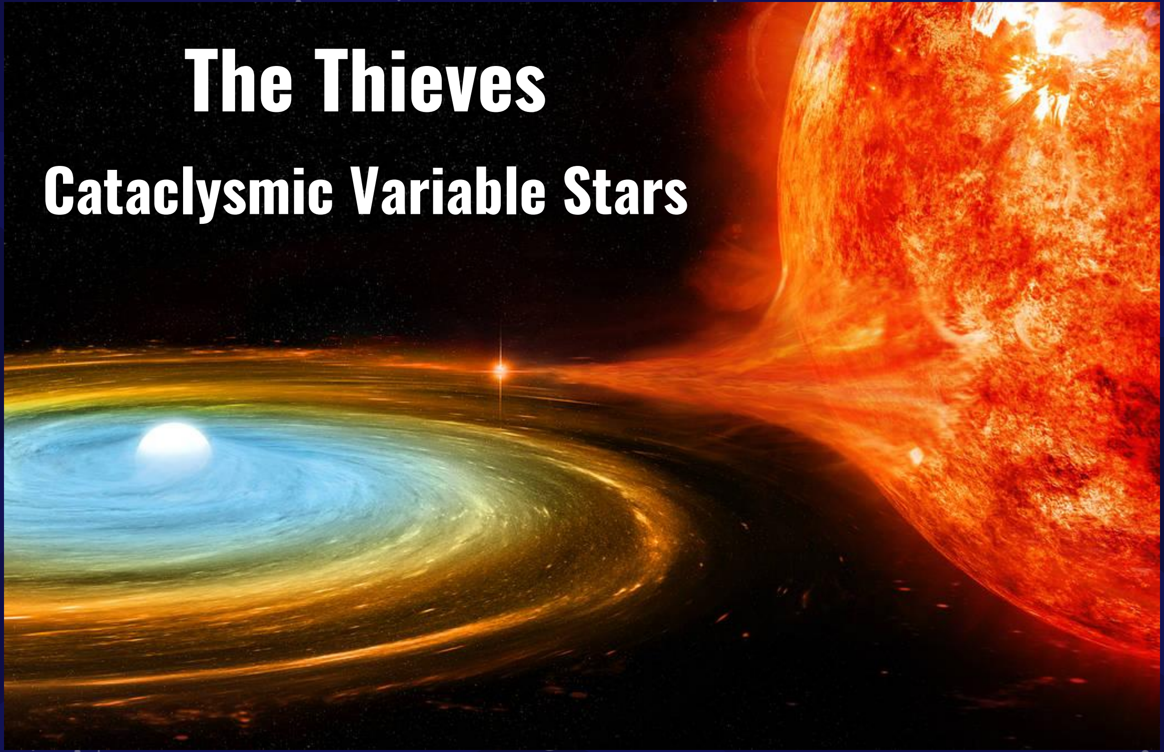
But there are also abnormal spectral types, such as white dwarfs, relatively bright but very small stars

Usually these are the core of a larger star that has blown off its outer layers in a supernova



Spectral Types

Cataclysmic variable stars are a type of star that fluctuates in brightness over time. These stars only appear in binary systems

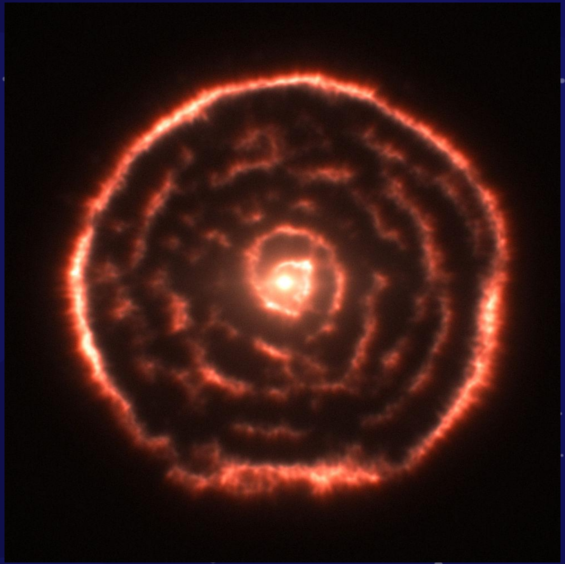


CV stars siphon material off of their binary companion, periodically building up layers and then blowing them off in a bright flash

Spectral Types

Carbon stars are the faint remnants of late stage (dying) stars that didn't have enough mass to supernova, so once fusion stops in the core, the outer layers gradually drift off

The Dying Giants



Carbon Stars