MA678 Final Project Proposal

Jordan Stout

Personal Statement

Astronomical data is abundant and available. Petabytes of data on stars, exoplanets, galaxies, asteroids, nebulae, pulsars and more are being collected and are free to download from online databases. I have always been interested in physics and space and want to be able to combine that interest with the topics we have been learning about in MA678. Working with astronomical data is beneficial also because it leads into learning about topics outside of regression models such as photometry, spectroscopy, time series analysis, cosmology, and image processing.

Question

Are we able to accurately predict the metallicity of a star given other parameters such as radius, temperature, and luminosity? Metallicity is a measure of how much metal is present in a star compared to lighter elements. It is important because it gives us an insight into the age of the star, as younger stars have a lower metallicity. Traditionally, the metallicity of a star is calculated using photometry where astronomers will analyze its color and brightness across different wavelengths. Who's to say it's not possible by other means?

Data Source

I have built the dataset myself using the ESA Gaia Archive. I wrote an ADQL query to construct a dataset that includes 20000 stars and their respective metallicity, magnitudes, distances from Earth, luminosity, surface gravity, and temperature. The Gaia Archive also offers many other parameters that I may add later as needed.

Proposed timeline

EDA: 11/22
Data Processing: 11/27
Modeling and Validation: 12/2
Write Up: 12/10