**Objectives**

The model will be focused on the interpretation of how stars follow a pattern based on their physical features. The model will help researchers and scientists accurately classify stars based on these physical features.

**Data Description**

The dataset is called “The Stars Dataset” and consists of several features of stars. The purpose of this analysis is to prove that the stars follow a certain graph in the celestial space specifically called Hertzsprung-Russell Diagram, which is shown in **Figure 1**.

Chart, scatter chart

Description automatically generated

**Figure 1:** Hertzsprung-Russell Diagram.

The data has 6 features and 1 target:

* Absolute Temperature (K)
* Relative Luminosity (L/Lo)
* Relative Radius (R/Ro)
* Absolute Magnitude (Mv)
* Star Color (white, Red, Blue, Yellow, yellow-orange, etc.)
* Spectral Class (O,B,A,F,G,K,M)
* Star Type (Target) (Red Dwarf, Brown Dwarf, White Dwarf, Main Sequence, SuperGiants, HyperGiants)

Lo = 3.828 x 10^26 Watts (Avg Luminosity of Sun)  
Ro = 6.9551 x 10^8 m (Avg Radius of Sun)

The model will use the 6 physical features of the stars to accurately predict the Star Type.

**Data Exploration**

Brief summary of data exploration and actions taken for data cleaning and feature engineering.

**Summary of Training**

at least three different classifier models, start with a simple logistic regression as a baseline, adding other models or ensemble models. Preferably, all your models use the same training and test splits, or the same cross-validation method.

**Key Findings**

key findings related to the main objective(s) of the analysis.

**Model Recommendation**

A paragraph explaining which of your classifier models you recommend as a final model that best fits your needs in terms of accuracy and explainability.

**Suggestions**

Suggestions for next steps in analyzing this data, which may include suggesting revisiting this model after adding specific data features that may help you achieve a better explanation or a better prediction.