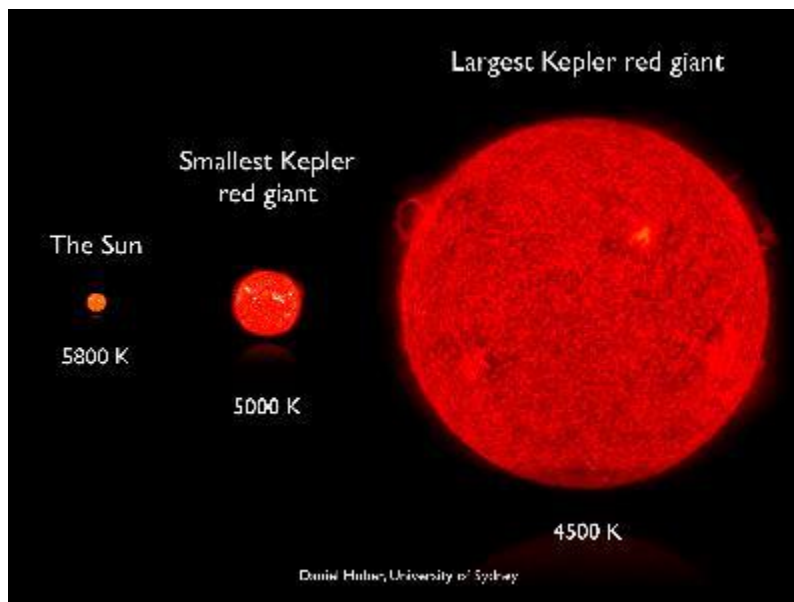


Evaluación Final

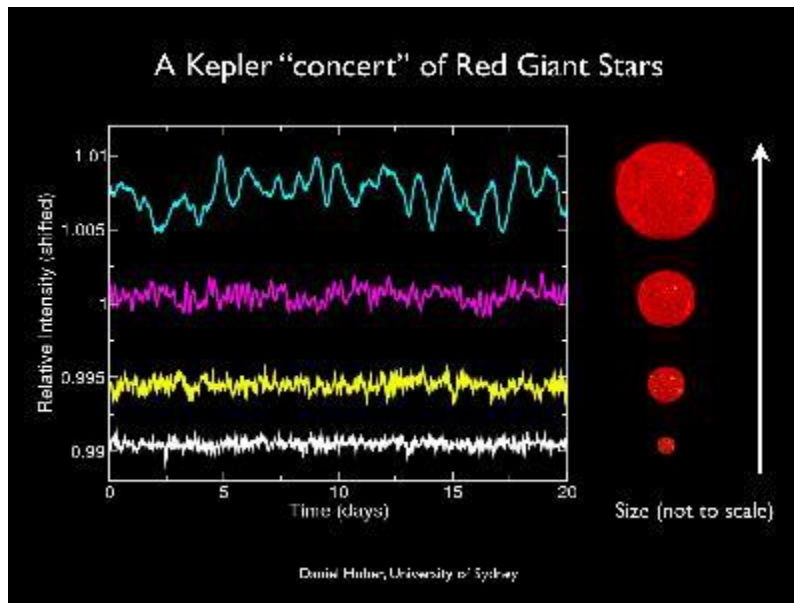
Contexto del Problema

Usted ha sido contratado como parte de un equipo de científicos que busca aplicar ciencia de datos a las investigaciones sobre astronomía. Deberá, por lo tanto poner a prueba todas las habilidades y conocimientos adquiridos a lo largo de este curso.

Su primer desafío consiste en apoyar al equipo de científicos que está estudiando a los “Gigantes Rojos” (Red Giants). Los Gigantes Rojos son estrellas gigantes de masa baja o intermedia que, tras haber consumido el hidrógeno en un núcleo, comienzan a quemar hidrógeno en una cáscara alrededor del núcleo de helio inerte, lo cual provoca un efecto de aumento de volumen de la estrella y un enfriamiento en la superficie, por lo que su color se torna rojizo.



La misión Kepler realizó mediciones sobre las frecuencias que producen las estrellas, el cual es incluso audible por el oído humano. Es más, el sonido emitido por una estrella puede dar cuenta de su tamaño. Como se observa en el siguiente gráfico, en la medida que los gigantes rojos son más grandes, las frecuencias son más bajas pero de mayor amplitud.



Lo anterior, fue clave para recolectar información de distintos gigantes rojos. Se ha confeccionado un set de datos que contiene la siguiente información:

Variable	Descripción
POP	Population class predictions for Kepler red giants: 0 = RGB (Red Giant Branch) 1 = HeB (Helium Burning)
Dnu	Delta-nu, (uHz) Mean large frequency separation of modes with the same degree and consecutive order
numax	F9.5 (uHz) Frequency of maximum oscillation power
epsilon	F7.3 Location of the l=0 mode

Productos solicitados

En su rol de ingeniero y científico de datos, se le pide que desarrolle los siguientes productos entregables:

- **Producto 1:** Análisis Exploratorio de los datos, utilizando los conceptos, técnicas y herramientas vistos en clase que sean de utilidad para llevar a cabo este producto.
- **Producto 2:** Modelo de aprendizaje de máquina que permita clasificar un gigante rojo de acuerdo a la clase RGB ó HeB.
- **Producto 3:** Una aplicación Python de consola que permita realizar la predicción de un gigante rojo a partir de sus características, utilizando el modelo entrenado anteriormente. La idea es que el aplicativo se ejecute en la consola, permita el ingreso de los parámetros de entrada (dnu, numax, etc), y el aplicativo retorne qué clase de gigante rojo se trata.

Recuerde que Usted, como ingeniero y científico de datos, debe aplicar criterio para seleccionar los conceptos, técnicas y herramientas que requiera para llevar a cabo este propósito, tomando como base los conceptos entregados a lo largo de este módulo y los adquiridos en los módulos anteriores.

Recuerde que el contexto del problema es un estudio científico, por lo cual deberá considerar, dentro de otras cosas, orden, análisis y reproducibilidad en su análisis.

Recuerde, además, que la elaboración de conclusiones es vital en un estudio de estas características.

Instrucciones

Considere las siguientes instrucciones para el desarrollo del trabajo:

- El trabajo es individual.
- Cualquier violación al código de ética podría significar la nota mínima en la evaluación.
- El plazo de entrega es de 7 días a partir de la entrega de este caso.
- Su trabajo deberá subirlo a la plataforma Moodle.
- Deberá entregar todos los códigos fuentes, notebooks y documentos que sean requeridos