Detección de exoplanetas con redes neuronales usando la base de datos de la misión Kepler

|  |  |  |
| --- | --- | --- |
| TITULACIÓN:  Máster en Inteligencia Artificial  Curso académico:  2019-2020  Lugar de residencia, mes y año: | Alumno/a:  Martín Solís, Héctor  D.N.I:  52880566R  Director:  Guzmán Álvarez, César Augusto | Convocatoria:  Primera  Orientación:  Créditos: |

Índice

[Resumen 3](#_Toc40006078)

[Abstract 4](#_Toc40006079)

[Terminología 5](#_Toc40006080)

[Introducción 6](#_Toc40006081)

[Planteamiento 6](#_Toc40006082)

[Justificación 6](#_Toc40006083)

[Objetivos 7](#_Toc40006084)

[Objetivo principal 7](#_Toc40006085)

[Objetivos secundarios 7](#_Toc40006086)

[Estado del arte 8](#_Toc40006087)

[Misión Kepler 8](#_Toc40006088)

[Técnicas para detectar exoplanetas 8](#_Toc40006089)

[Método de velocidad radial 9](#_Toc40006090)

[Método de astrometría 10](#_Toc40006091)

[Método de tránsito 10](#_Toc40006092)

[Herramientas 11](#_Toc40006093)

[Desarrollo del proyecto 12](#_Toc40006094)

[Búsqueda elementos observados en la misión Kepler (KOI) 12](#_Toc40006095)

[Procesamiento de los KOI para obtener los datos de sus mediciones 12](#_Toc40006096)

[Procesamiento de ficheros FITS para obtener datos de flujos de luz (FLUX) 12](#_Toc40006097)

[Procesamiento de datos FLUX de todos los KOI para crear un dataset único con la información necesaria 13](#_Toc40006098)

[Montar la red neuronal, entrenar y probar 13](#_Toc40006099)

[Resultados 14](#_Toc40006100)

[Conclusiones 14](#_Toc40006101)

[Bibliografía 15](#_Toc40006102)

Resumen

Lo que se va a hacer

Abstract

Lo mismo que el resumen, pero en inglés

Terminología

* TPF (Target Pixel Files):
* FFI (Full Frame Images):
* SAP Flux (Simple Aperture Photometry Flux): intensidad de luz tras sumar los pixels calibrados dentro de la apertura óptima.
* BJD (Barycentric Julian Day): Fecha Juliana corregida por las diferencias en la posición de la Tierra con respecto al Baricentro del Sistema Solar (centro de masa del Sistema Solar).
* MAST (Mikulski Archive for Space Telescopes) [1]: portal de obtención de datos de misiones y telescopios.

Introducción

Planteamiento

La exploración del universo cercano a nuestro planeta ha ido creciendo de manera exponencial a lo largo de los últimos años. Con el lanzamiento de más sondas, satélites, telescopios espaciales, etc, se han ido descubriendo gran cantidad de planetas que orbitan estrellas con características similares a la del Sol. Estos planetas, en el ámbito de la astronomía son los denominados exoplanetas.

Con la información que se ha obtenido durante estas exploraciones, se han descubierto las características de estos planetas: gigantes gaseosas, *tierras-super-calientes* (hot-super-Earths), gigantes de hielo [2]. Todas ellas, con características extremas que no hacen posible la existencia de vida tal como la conocemos.

Hoy en día, el objetivo de la búsqueda de exoplanetas en el universo es encontrar planetas con características similares a la tierra; es decir, que tengan un tamaño similar, que orbiten una estrella similar al Sol, y que dispongan de agua en estado líquido.

En este punto es en el que nace la misión Kepler, específicamente diseñada para estudiar nuestra región de la Vía Láctea y descubrir cientos de planetas similares a la Tierra cerca de la zona habitable, determinando la fracción de los cientos de miles de millones de estrellas de nuestra galaxia que podrían tener tales planetas.

Esta, así como las demás misiones espaciales, obtienen una gran cantidad de imágenes que son inviables de tratar a simple vista por el ojo humano, por ello, se utilizan técnicas computacionales que ayudan al procesamiento de esas imágenes para obtener gran cantidad de datos. Así mismo, tal cantidad de datos es manejable de manera manual por lo que se analizan con programas que permiten detectar patrones o anomalías en ellos.

Justificación

Las redes neuronales han demostrado su capacidad para resolver problemas de detección de patrones en conjuntos de datos utilizando un modelo jerárquico, ajustando millones de parámetros para detectar dichos patrones que serían imposibles de detectar a simple vista.

Con esta premisa, y con la base de datos de imágenes de las misiones Kepler, se decide utilizar el método de tránsito en la detección de exoplanetas [3].

Objetivos

Objetivo principal

Detectar exoplanetas usando redes neuronales convolucionales a partir de la información en imágenes de las misiones Kepler de la NASA. [2]

Objetivos secundarios

* Revisar el Estado del Arte
* Obtener datos de las misiones (curvas de luz)
* Entender cómo obtener los datos de los ficheros pre-procesados (FITS)
* Extraer datos de todas las estrellas observadas por Kepler
* Tratar los datos para hacerlos compatibles a las entradas de la CNN
* Diseñar la red neuronal para procesar los datos
* Entrenar el modelo diseñado
* Revisar y analizar los resultados obtenidos
* Extraer conclusiones

Estado del arte

Misión Kepler

Kepler fue lanzada en 2009 en una órbita de seguimiento de la Tierra y observó un trozo de cielo de 100 grados cuadrados cerca de Cygnus [4] para medir las variaciones de brillo de unas 200.000 estrellas.

Los datos obtenidos de la sonda se dividen en trimestres de 90 días debido a que su ratio de giro era de 90 grados cada 90 días para mantener los paneles solares apuntando al sol.  
La obtención de imágenes se redujo solamente a las estrellas de interés con una cadencia de obtención de imágenes cada 30 minutos (cadencia larga) o 1 minuto (cadencia corta).  
La misión produjo **serie temporal de fluctuaciones de luz** para cada estrella y con estos datos, se buscaron variaciones en estas **curvas de luz** que marcaban la presencia de un exoplaneta en tránsito por delante de la estrella observada.

Además de descubrir exoplanetas, los datos de Kepler se han utilizado para estudiar la variabilidad de las estrellas y los binarios eclipsados.

Técnicas para detectar exoplanetas

Dentro de la detección de exoplanetas existen dos grandes grupos donde se clasifican estas técnicas: la detección directa y detección indirecta [5] [6].

La **detección directa** que se basa en observar directamente los planetas con telescopios ultra ponente como el Telescopio Hubble. Esta técnica permite obtener datos inequívocos del cuerpo que se está observando. No obstante, esta técnica suele ser problemática ya que la iluminación que se puede detectar de la estrella sobre la que orbite el exoplaneta observado, suele ser mucho más brillante que el exoplaneta en sí, que no emite luz propia. Esto causa que se produzcan desenfoques en el objetivo del telescopio debido a la difracción de la luz.

Para evitar estos problemas, son mucho más utilizados los métodos de **detección indirecta**, que no observan el exoplaneta en sí, sino que se infieren sus datos y características en función de la observación de la estrella sobre la que orbitan.

Método de velocidad radial

Consiste en la observación de la luz emitida por una estrella de manera directa mediante un telescopio. Este recoge su luz que se hace pasar por un prisma que amplia y proyecta todo su espectro de luz como si de un arco iris se tratara. De esta manera se pueden detectar las características propias de la estrella. Cuando se despliega el espectro de luz se pueden detectar unas líneas horizontales negras que representa que la longitud de onda de ese color ha sido absorbida por los elementos químicos que hay en la estrella emisora de la luz. Con esta espectrometría se puede determinar los elementos de la estrella, así como su temperatura y presión.

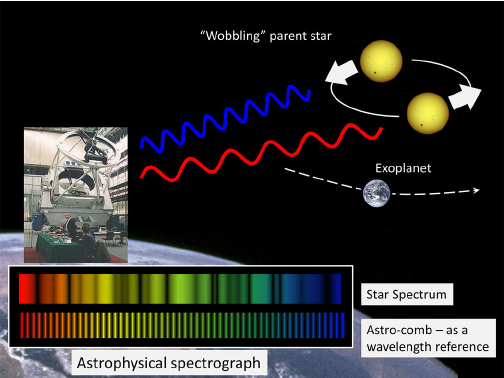


Ilustración 1: detección de exoplanetas mediante el método de velocidad radial

También con este espectro se puede estudiar qué estrellas tienen planetas grandes a su alrededor. Ya que si un planeta orbita la estrella, tira de ella con su campo gravitacional, forzando a la estrella a una pequeña órbita o bamboleo. Hace que parezca que la estrella está girando alrededor de un punto en el espacio. La estrella a veces girará hacia la Tierra y otras veces se alejará. Esto hace que las longitudes de onda se muevan hacia el extremo azul del espectro si la estrella gira hacia el planeta, o hacia la parte roja en caso contrario. Por lo tanto, si se mueven hacia adelante y hacia atrás en el tiempo, significará que tienen planetas en órbita, que, al medir la cantidad de movimiento con el tiempo, se puede determinar la masa del planeta y su órbita.

Esta técnica está limitada por la precisión de los espectroscopios que, actualmente, los mejores no pueden detectar pequeños desplazamientos; es decir, no podrán detectar planetas pequeños (del tamaño de la tierra).

Método de astrometría

Consiste en medir con precisión la posición de una estrella, para que cualquier oscilación se pueda detectar directamente. Se realizan observaciones desde observatorios fijos desde la superficie de la Tierra, o usando telescopios espaciales como el Telescopio Hubble.

Es un método poco preciso por las mismas razones que la detección radial: es difícil tratar la posición de los planetas de menor tamaño por la precisión de los elementos. Además, la observación desde la superficie de la tierra también supone un hándicap debido a las alteraciones de las mediciones que produce la atmósfera.

Método de tránsito

Este método consiste en tomar imágenes del entorno de una estrella, y a partir de dichas imágenes, hacer mediciones de la luminosidad que muestra la estrella en el tiempo, controlando la alteración de los valores de su brillo. Un descenso en los valores de iluminación significa que un cuerpo celeste está haciendo un tránsito entre la estrella y el observador. Tales tránsitos bloquean una pequeña fracción de la luz y pueden inferirse datos de masa y órbita del planeta con el cambio de flujo de luz.

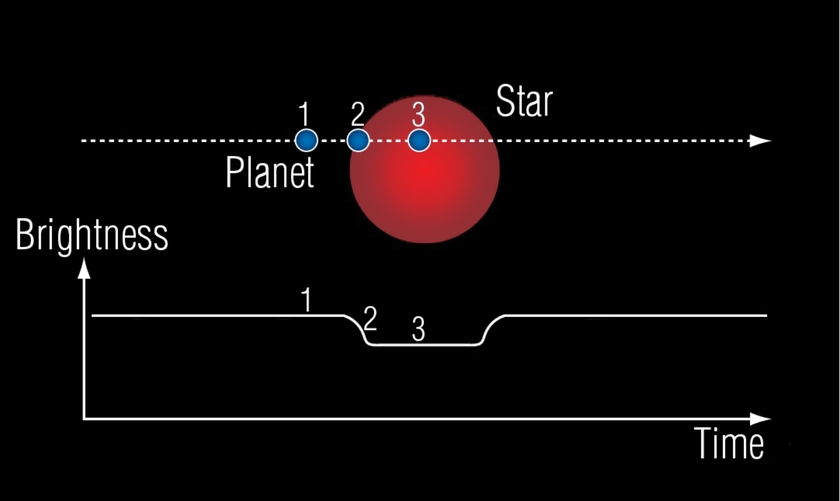


Ilustración 2: detección de exoplanetas mediante el método de tránsito

Este método es el más prometedor para detectar exoplanetas. Al tener tanto el tránsito como la información de velocidad radial, la órbita de los planetas se puede determinar exactamente, y se puede encontrar la verdadera masa y tamaño del planeta.

Herramientas

TO-DO

[4] Ekskljssjljskjsklskls

Klsjsjlksjksks

sssss

lkjsskls

Desarrollo del proyecto

Búsqueda elementos observados en la misión Kepler (KOI)

TO-DO: describir proceso para encontrar los datos de KOI en MAST

Procesamiento de los KOI para obtener los datos de sus mediciones

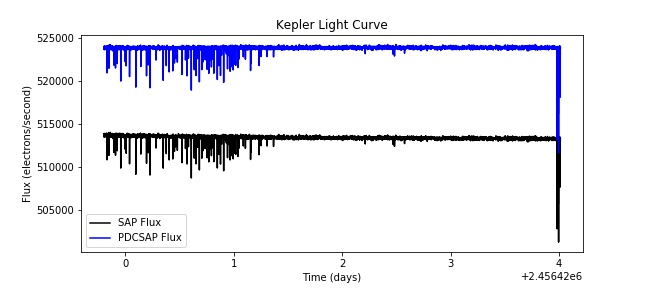
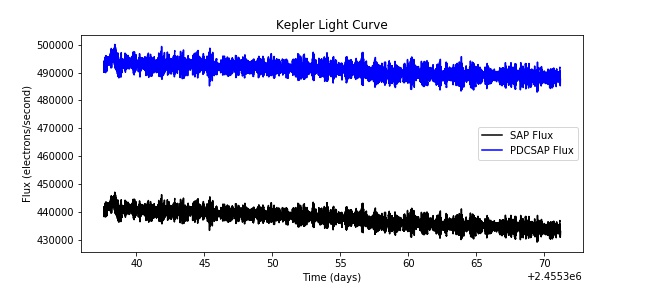
TO-DO: explicar que se hizo para tratar los KOI y obtener los ficheros FITS que tienen toda la informacion de las observaciones.

Explicar los tipos de ficheros según Cadencia, FFI, etc

Procesamiento de ficheros FITS para obtener datos de flujos de luz (FLUX)

TO-DO: explicar como se procesa los ficheros FITS para obtener toda la info de los flujos de lux

Explicar las graficas



Procesamiento de datos FLUX de todos los KOI para crear un dataset único con la información necesaria

TO-DO

Montar la red neuronal, entrenar y probar

TO-DO

Resultados

Conclusiones

Bibliografía

* Portal MAST [1] : recopila la información de las observaciones realizadas por telescopios espaciales.
* Página de la misión Kepler en el portal de la NASA [2]: contiene toda la información relevante sobre qué consiste, objetivos y metodología de la misión Kepler y K2.
* Artículo en sobre el funcionamiento del Método de Fotometría de Tránsito para detectar planetas [3]
* Artículo informativo sobre la constelación Cygnus, cerca de donde se realizó la misión Kepler [4]
* Herramientas Software para ver los datos de la misión Kepler [5]
* Astronomía Nova [6]: sienta las bases de las llamadas *leyes de Kepler* sobre el movimiento de los cuerpos celestes
* Documentación de la librería de Astroquery para Python [7]
* Información técnica sobre extracción de datos de la misión Kepler (curvas de luz, FFI, etc) [8]
* Tutorial de obtención de curvas de luz usando Astropy [9]
* Documentación técnica de Astropy [10]
* Artículo en Medium sobre la detección de exoplanetas usando un dataset de Kaggle [11]
* Dataset en Kaggle de los flujos de luz [12]
* Método de la velocidad de radio para detectar exoplanetas [15]

|  |  |
| --- | --- |
| [1] | MAST, «MAST Portal,» [En línea]. Available: https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html. |
| [2] | NASA, «KEPLER & K2 Missions Page,» [En línea]. Available: https://www.nasa.gov/mission\_pages/kepler/overview/index.html. |
| [3] | planetary.org, «Transit Photometry Method article,» [En línea]. Available: https://www.planetary.org/explore/space-topics/exoplanets/transit-photometry.html. |
| [4] | NASA, «Cygnus Skyscape,» [En línea]. Available: https://science.nasa.gov/cygnus-skyscape. |
| [5] | NASA, «NASA tools for Kepler Data Manipulation,» [En línea]. Available: https://keplerscience.arc.nasa.gov/software.html. |
| [6] | J. Keppler, Astronomia Nova, 1609. |
| [7] | A. Ginsburg, «Astroquery Docs,» [En línea]. Available: https://astroquery.readthedocs.io/en/latest/. |
| [8] | «KEPLER Data Products,» [En línea]. Available: https://archive.stsci.edu/missions-and-data/kepler. |
| [9] | S. T. S. Institute, «Tutorial for extracting light curves with astropy,» [En línea]. Available: https://github.com/spacetelescope/notebooks/tree/master/notebooks/MAST/Kepler/Kepler\_Lightcurve. |
| [10] | T. A. Developers, «Astropy Docs,» [En línea]. Available: https://docs.astropy.org/en/stable/index.html. |
| [11] | G. Garza, «Exoplanet Hunting with Machine Learning and Kepler Data,» [En línea]. Available: https://medium.com/@gabogarza/exoplanet-hunting-with-machine-learning-and-kepler-data-recall-100-155e1ddeaa95. |
| [12] | Kaggle, «Kepler labelled time series data,» [En línea]. Available: https://www.kaggle.com/keplersmachines/kepler-labelled-time-series-data. |
| [13] | O. Ginerich, The eye of the heaven. Ptolemy, Copernicus, Kepler, vol. Masters of modern physics Vol.7, 1993. |
| [14] | «LcTools,» [En línea]. Available: https://sites.google.com/a/lctools.net/lctools/lctools-product-description. |
| [15] | «KeplerPORTs,» [En línea]. Available: https://github.com/christopherburke/KeplerPORTs/. |