

Exoplanets Hunting

Project Report

Introduction:

- Exoplanets are planets that orbit stars outside our solar system.
- These planets are further studied by astrophysicists for possibilities of alien life, and to deepen our understanding of the universe.
- A KOI is a target identified by the Kepler Project that displays at least one transit-like sequence within Kepler time-series photometry that appears to be of astrophysical origin and is initially consistent with a planetary transit hypothesis.
- NASA's Kepler's Objects of Interest (KOI) table can be used to train models to recognize candidates for exoplanets using the properties of KOI objects as features

Aim:

In this project, we aim to analyze data collected by NASA's Kepler telescope on astronomical objects to find out the typical characteristics of an exoplanet. We will then build machine learning models to predict whether a given KOI is an exoplanet or not.

Data:

- Data was collected from NASA's exoplanet archive: <https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=koi>. NASA updates data on Kepler's findings regularly on this website. The dataset consisted of 8054 rows and 49 columns.
- Data wrangling methods (like assigning numerical values to categorical variables, dealing with nan and infinity values) were used to clean the data.

Analysis and findings:

The following conclusions were made after analysis and visualization of the data:

- There are almost equal candidates and non-candidates and among the candidates, almost two thirds are confirmed exoplanets (2652 out of 4029)
- Most exoplanets have an orbital period of less than 100 days, higher orbital periods usually indicate that the KOI is not an exoplanet.
- Most exoplanets have transit epoch time in between 125 and 200 BJD, a KOI with a higher time is most probably not an exoplanet.
- Most KOIs with impact parameter above 1 are not exoplanets.
- Most exoplanets have a transit duration of less than 10 hours and any KOI with a duration higher than 15 hours is most probably not an exoplanet.
- The higher the stellar flux lost at the minimum of the planetary transit in ppm, the less likely that the KOI is an exoplanet.
- The higher the planet radius, the less likely that the KOI is an exoplanet, most exoplanets have radius below 5 earth radii, KOIs with radius higher than 15 earth radii are most probably not exoplanets.
- Most exoplanets have an equilibrium temperature of below 1500K and a KOI with higher equilibrium temperature is most probably not an exoplanet.

- Most exoplanets have relative insulation flux below 1000, and any KOI with higher flux is most probably not an exoplanet.
- Most exoplanets have relative transit depth below 250 and any KOI with higher transit depth is most probably not an exoplanet.
- The photo-spheric temperature of the star of an exoplanet is most likely in between 3000-7000k and higher temperature KOIs are mostly not exoplanets
- Most exoplanets have stellar surface gravity between 3 and 5
- Most exoplanets have stellar radius less than 1.5 solar radii and a koi with higher stellar radius is most probably not an exoplanet
- The probability of a KOI being an exoplanet is inversely proportional to the KIC right ascension.

Predictive Modelling:

The following algorithms were used to build models to predict whether a KOI is a candidate exoplanet or not:

1. Logistic Regression (Accuracy: 0.796, Recall: 0.884, F1 Score: 0.818, Precision: 0.760)
2. KNN (Accuracy:0.784, Recall: 0.795, F1 Score: 0.792, Precision: 0.788)
3. Decision Tree Classifier (Accuracy: 0.99, Recall:0.988, F1 Score:0.991, Precision: 0.994)
4. Random Forest Classifier (Accuracy: 0.992, Recall: 0.985, F1 Score: 0.992, Precision:1.0)

Conclusions:

- 2652 exoplanets have been confirmed so far by NASA (as on 14-08-2022)
- There are almost equal candidates and non- candidates and among the candidates, almost two thirds are confirmed exoplanets.
- Impact parameter, transit duration, flux and radius (planetary and stellar) are some of the most important deciding factors on whether a koi has an exoplanet or not.
- We were able to predict with a 99% accuracy whether our observations is an exoplanet or not.
- While all models had over 78% accuracy, random forest classifier was the most accurate model.