

# Exoplanet Detection

## Preamble to the Problem

NASA's Kepler mission to search for extrasolar planets has collected data from hundreds of thousands of star-systems, and has discovered nearly 1000 confirmed exoplanets to date in addition to over 3000 unconfirmed candidates. The mission detects exoplanets using transit photometry, which detects the transit of a planet in front of a star as transient drops in stellar intensity. Raw data is collected in the form of a sequence of stellar images, which are processed into "light-curves" tracking the brightness of a star over time. An algorithm automatically searches for periodic planetary transits in these light curves, but spurious intensity dips and other noise in the data due to non-planetary stellar variability has led to high false-positive rates for detecting transits. As the initial planetary candidates found by this search method require extensive and costly subsequent validation, there is a need to reduce the error rate in exoplanet candidate identification.

## Problem Statement

Teams must come up with a machine learning algorithm to classify Kepler Objects of Interest (KOIs) as having exoplanets/not having exoplanet.

## Timeline

- Teams can start working on their algorithm immediately with the given train set.
- Teams must present their algorithms along with the documents mentioned below on the day of judging.

## Judging Criteria

The algorithm will be judged on the following points –

- 1.) True skill scores and F scores for your test data (you must make your own test data and train data from the given link)
- 2.) True skill scores and F scores of test data (provided one week before the evaluation)
- 3.) Innovative architecture
- 4.) Novelty in the algorithm

In addition to the algorithm, you also need to submit the following on the day of judging –

- 1.) A document clearly summarizing your procedure
- 2.) Code

## **Rules and Regulations**

- Only the data provided by us must be used. Use of additional data to improve accuracy can lead to disqualification.
- Plagiarism is strictly prohibited and may lead to penalties or disqualification.
- Organizers reserve right to modify rules at any point of time, with intimation to registered teams.

## **Data set:**

[Training data](#)