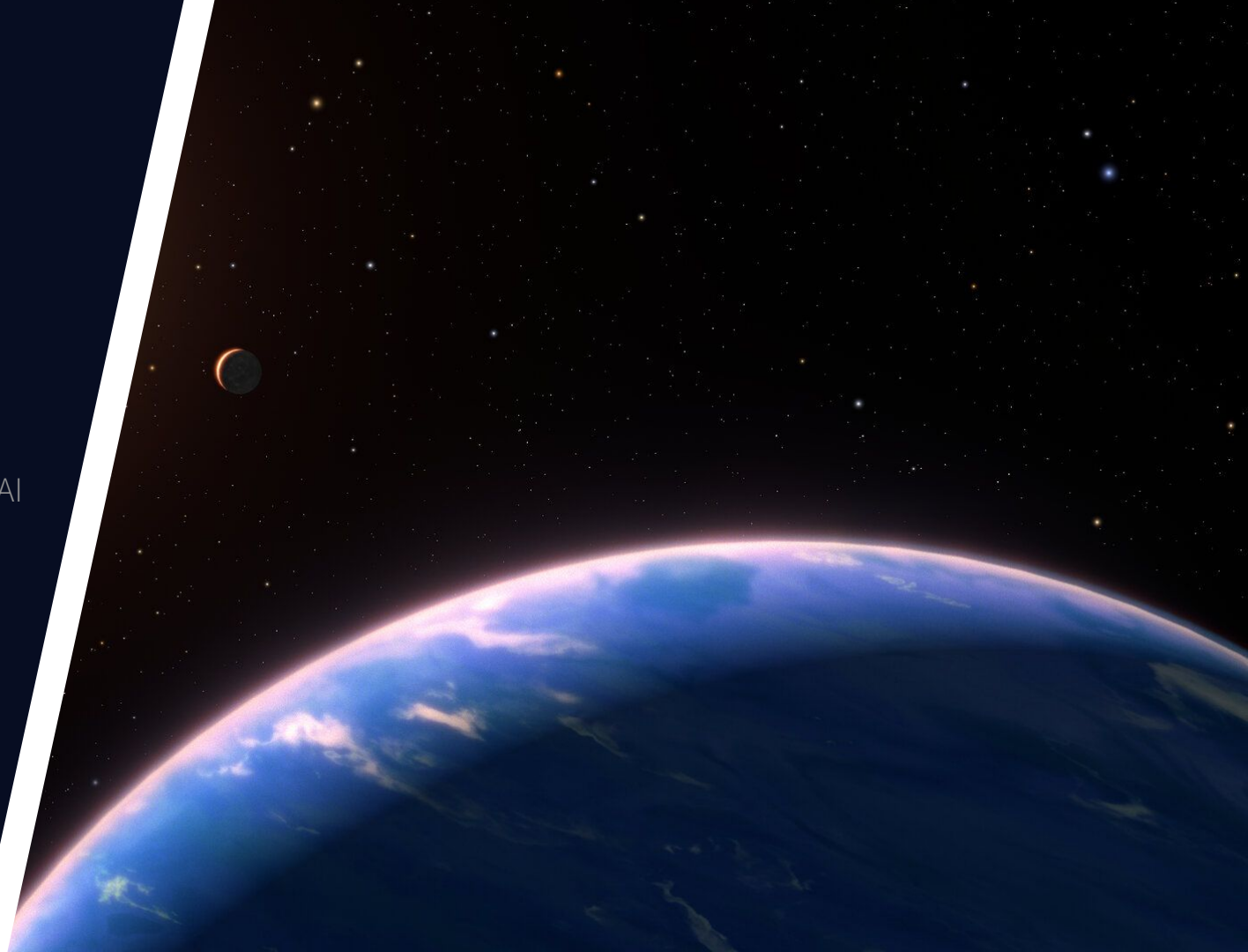
Team TerraFormers

A World Away

Hunting for Exoplanets with AI



Introduction

The discovery of exoplanets has transformed our understanding of the universe and the potential for life beyond Earth. Over the past two decades, NASA missions such as Kepler have collected enormous amounts of light curve data using the transit method. While these missions have identified thousands of exoplanets, much of the classification work was done manually by astrophysicists. As the volume of astronomical data increases, traditional manual methods become unsustainable. This project proposes an artificial intelligence and machine learning (AI/ML) pipeline to automate the identification of exoplanets from open-source NASA datasets, combined with a command line interface to make exoplanet exploration accessible to both researchers and enthusiasts.

Objectives

1. Develop a high-accuracy ML model trained on the Kepler dataset to classify candidates into confirmed exoplanets, planetary candidates, or false positives.

2. Build a command line interface/web interface that allows users to explore and test new data.

3. Enable real-time interaction by providing tools for data visualization, model accuracy statistics, and adjustable hyperparameter tuning.

Data Sources

We will use publicly available NASA data repositories such as:

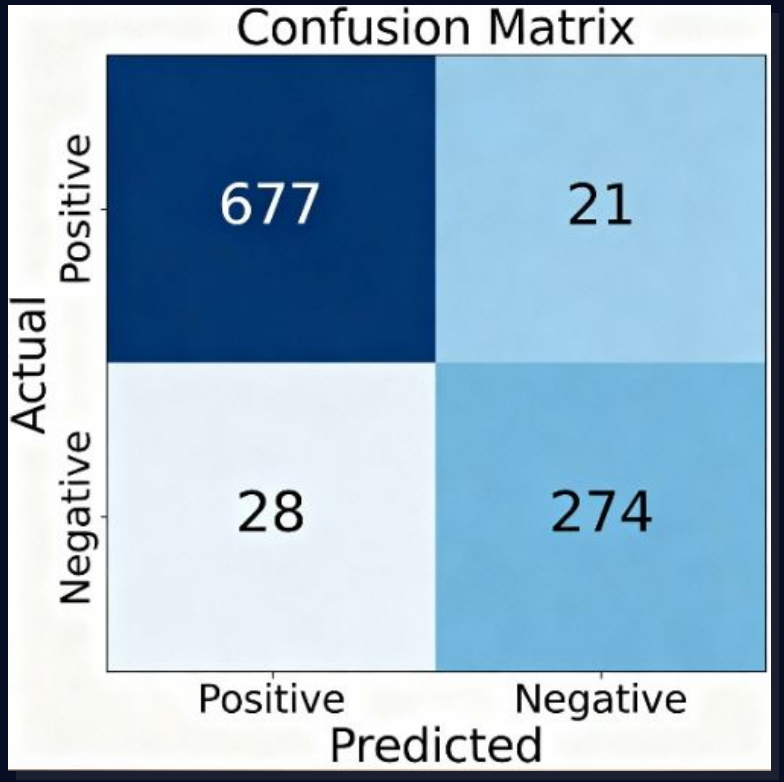
- **Kepler Object of Interest (KOI) catalogs**
  - include confirmed planets, candidates, and false positives.
- The Kepler dataset contains parameters such as orbital period, transit depth, planetary radius, stellar properties, and photometric precision.

Data Preprocessing

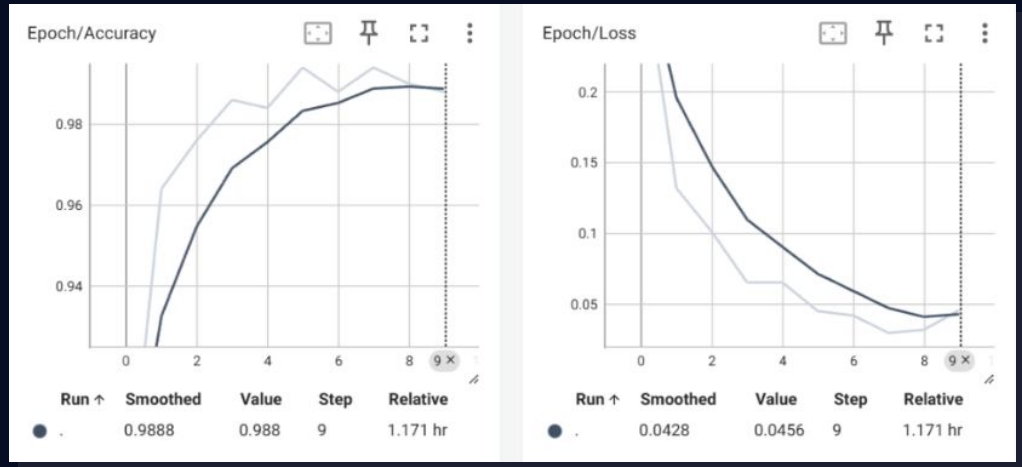
- Clean and normalize datasets (handle missing or inconsistent values).
- Balance dataset classes using oversampling or class-weighting techniques.
- Standardize input scales for ML algorithms.

Evaluation Metrics

- Confusion matrix to measure classification quality.



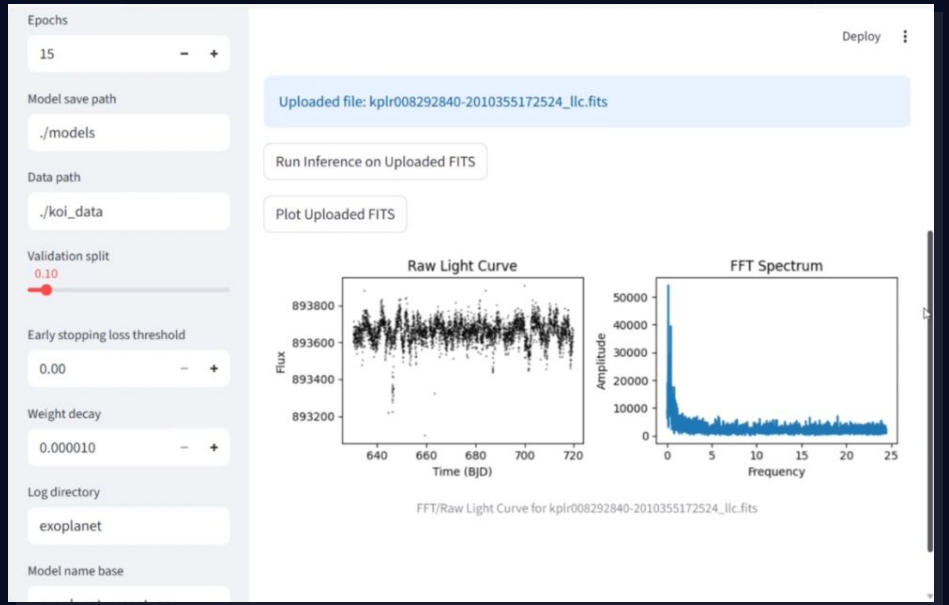
- Accuracy and loss curves for imbalanced data.



- …and cross-validation to ensure model robustness.

Expected Outcomes

- An automated classification tool achieving >90% accuracy on test datasets.
- Cli interface/web interface enabling astrophysicists and hobbyists to test and visualize new data.
- Open-source availability to encourage community contributions.
- Potential new exoplanet candidate detections from reanalyzed archival data.



Web interface showcase

Impact

This project democratizes exoplanet research by providing AI-powered tools that scale to the massive volumes of astrophysical data now available. Researchers gain a practical classification engine that accelerates discoveries, while students and amateurs gain an accessible gateway into planetary science. By uniting rigorous machine learning with cli interface/web interface, the project contributes both scientifically and educationally to the global search for worlds beyond our own.