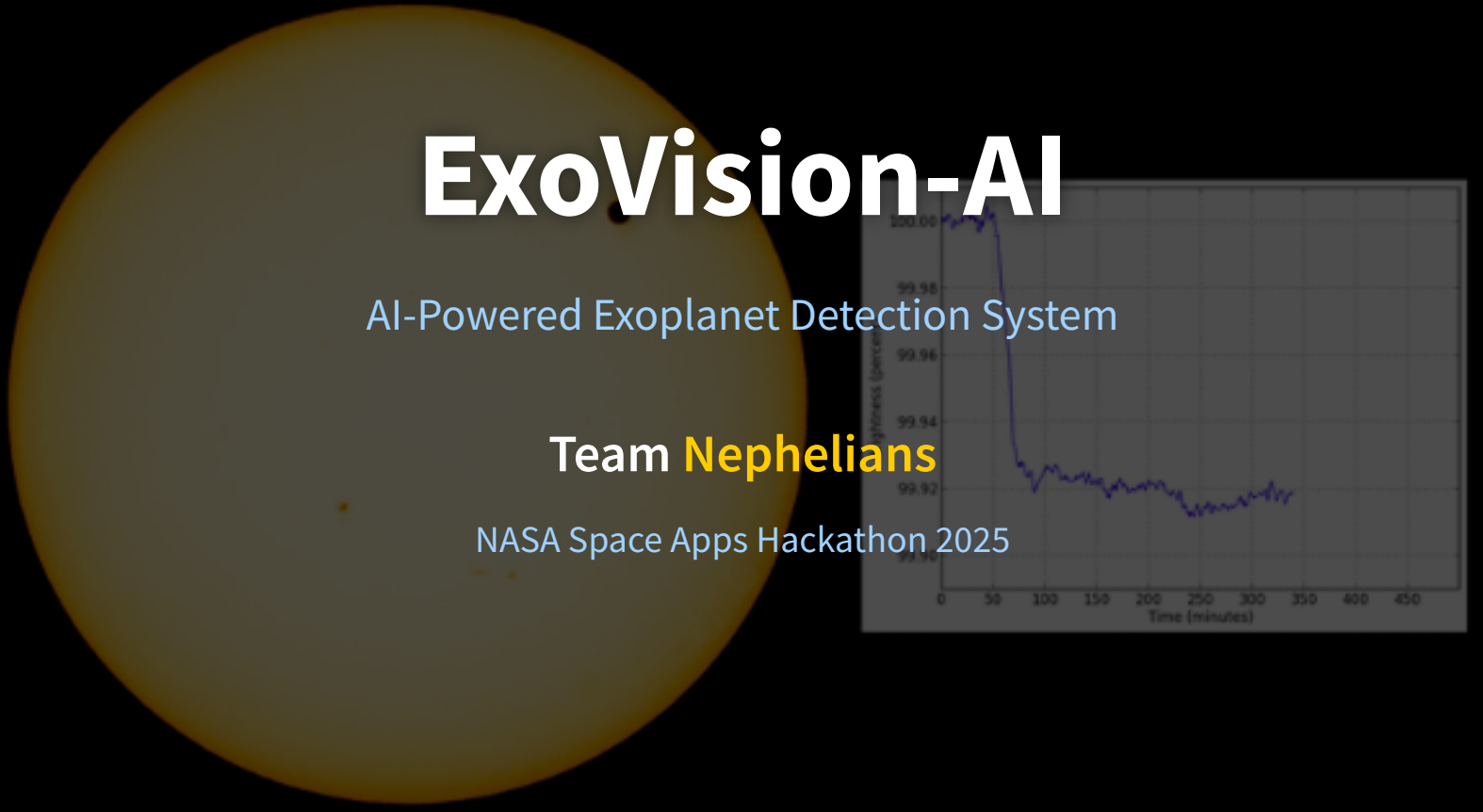
Step 1: find exoplanets



ExoVision-AI

AI-Powered Exoplanet Detection System

Team Nephelians

NASA Space Apps Hackathon 2025

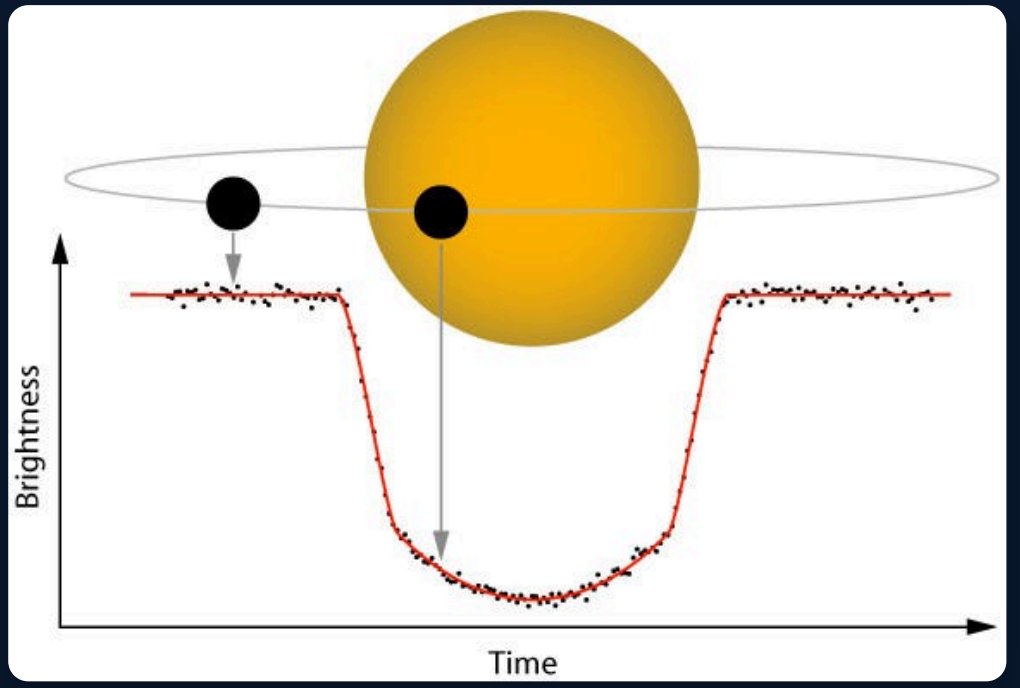
Project Overview & Problem Statement

The Challenge

- Thousands of **exoplanet candidates** identified by Kepler/K2 missions
- Many remain **unconfirmed** due to false positives
- Causes: binary stars, stellar variability, instrumental noise

Our Solution

- **AI-powered classification** of exoplanet candidates
- Accelerates discovery by reducing manual verification time
- Provides confidence scores for candidate validation



Data & Methodology

Dataset

- NASA **K2 Planets and Candidates Catalog**
- **4,004 entries** with 28 carefully selected features

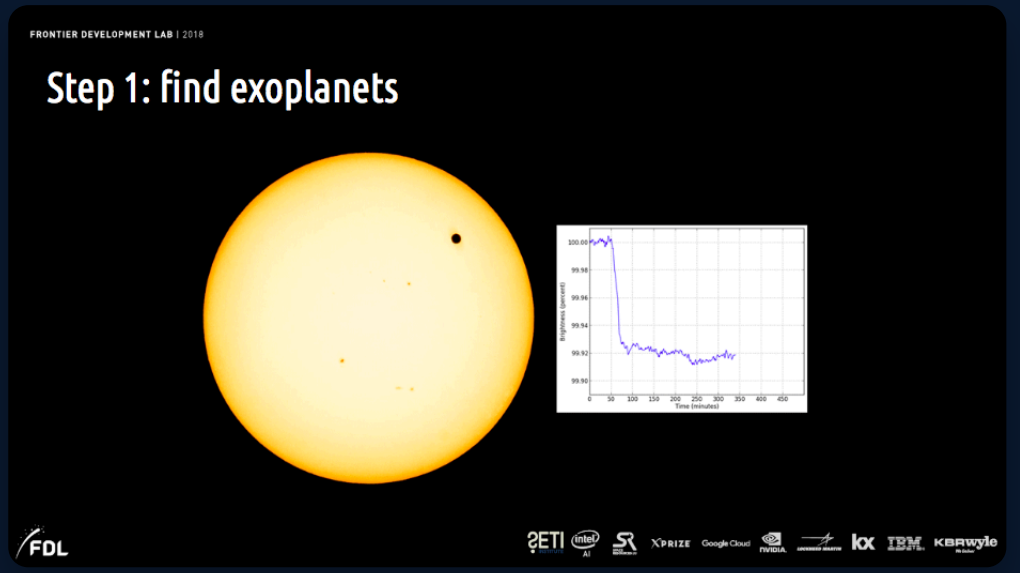
Key Features

Orbital Period | Planet Radius | Planet Mass

Stellar Temp | Stellar Radius | System Distance

Preprocessing

- Missing values imputed with **median**
- Features **standardized** to zero mean and unit variance
- Created new features: star-planet size ratio, temperature difference



Machine Learning Models

XGBoost Classifier

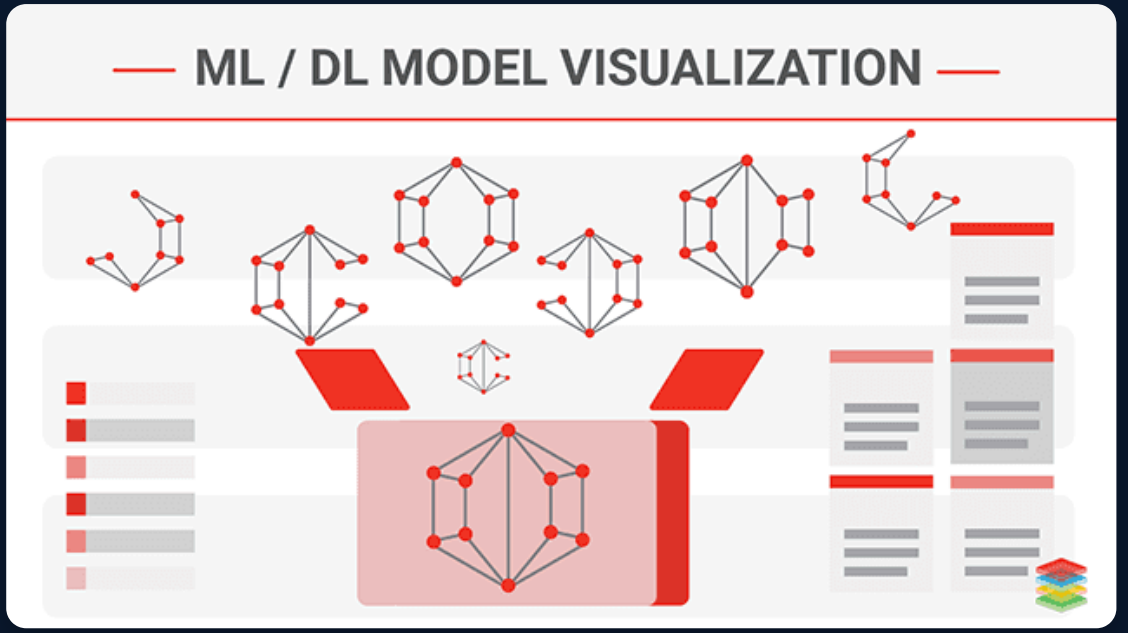
n_estimators: 300
learning_rate: 0.05
max_depth: 6
subsample: 0.8

Best performing model

Decision Tree Classifier

max_depth: 4
min_samples_split: 30
min_samples_leaf: 15
max_features: sqrt

Highly interpretable



Model	Accuracy	Precision	Recall	F1 Score
XGBoost	99.7%	99.5%	100%	99.7%
Decision Tree	99.6%	99.4%	100%	99.7%

Web Application

Flask-Based Interface

- **Two input methods** for exoplanet data
- **Real-time predictions** from both models
- Interactive **visualizations** and confidence scores

Input Methods

CSV Upload - Process multiple candidates at once

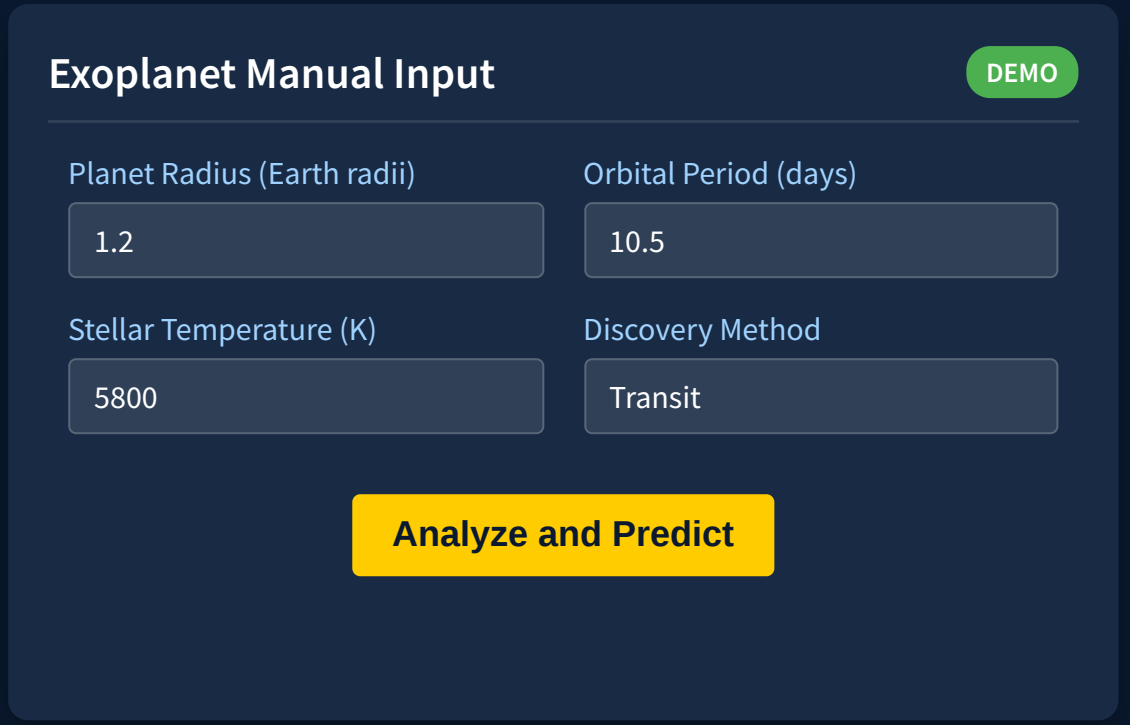
Manual Input - Enter data for single candidate

Key Features

Feature distribution charts

Correlation heatmaps

ROC curves and performance metrics



User Workflow

1. Input exoplanet data
2. AI model analysis
3. View predictions & visualizations

Results & Impact

Model Performance

99.7% Accuracy	**99.5%** Precision
100% Recall	**99.7%** F1 Score

Limitations

- Catalog-level **information leakage**
- Performance may be **inflated** by dataset biases
- Limited to **K2 catalog** characteristics

Research Impact

- **Accelerates** exoplanet discovery process
- Reduces **manual verification** time
- Guides **follow-up observations**

Future Improvements

- **Cross-validation** with different datasets
- More **rigorous scientific validation**
- Integration with **real-time telescope data**

Team & Acknowledgments

Team Nephelians

MA **Muhammad Ahmad**
Team Leader & Data Scientist

SD **Syed Darain**
ML Engineer

AA **Muhammad Ahsan**
Backend Developer

MM **Muhammad Mohsin**
Frontend Developer

AH **Ali Hassan**
Research Lead

Acknowledgments

- **NASA Exoplanet Archive** for the K2 Planets and Candidates Catalog
- **NASA Space Apps Hackathon 2025** for the opportunity

Technology Stack

- Python
- Scikit-learn
- XGBoost
- Flask
- Seaborn
- Pandas

GitHub Repository

https://github.com/c0llectorr/ExoVision-AI

Contact Email

mahmadshafi861@gmail.com