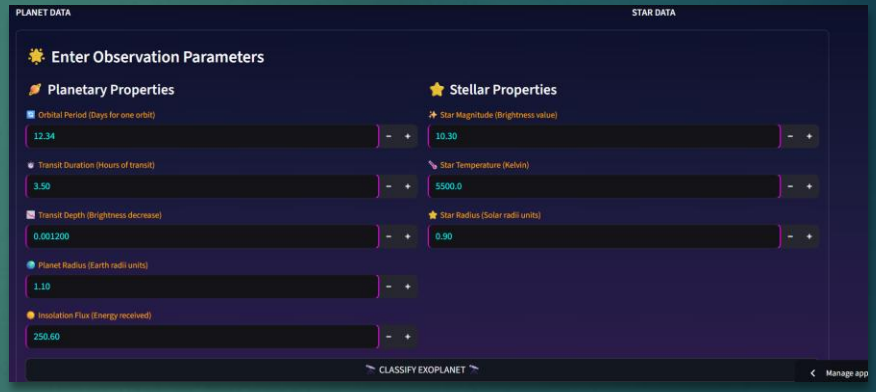
Exoplanet Classification using NASA TESS Data

- Machine Learning for Space Discovery – NASA Space Apps Challenge
- Team **Lone Orbit** – COMSATS University Islamabad
- Project: **StellarSynth**
- Exploring planets beyond our solar system using NASA's data



Why Classify Exoplanets?

- Exoplanets orbit stars beyond our solar system.
- NASA's data includes Confirmed, Candidate, and False Positive planets.
- Manual classification is complex and time-consuming.
- Goal: Build a model to classify exoplanets automatically.

NASA TESS Data Overview

- Datasets: TESS from NASA Exoplanet Archive.
- Contains ~7699 entries and 75 attributes (9 important ones).
- Checked duplicates, handled missing values, encoded categorical labels.
- Clean, structured data ready for machine learning.

Our AI Pipeline

- Models Tried: Logistic Regression, Random Forest, XGBoost, SVM.
- Model Chosen: XGBoost
- Steps: Preprocessing, Train-Test Split (80/20), Label Encoding.
- Trained and evaluated models for accuracy and F1-score.
- Pipeline ensures consistency and reproducibility.

Model Evaluation Results

- XGBoost achieved 68.1% accuracy (best performance).
- Random Forest achieved 67.5% accuracy.
- Other models performed moderately well.
- Metrics compared: Accuracy, F1-score.

AI Empowering Space Discovery

- XGBoost identified subtle patterns in exoplanet data.
- Helps NASA automate and prioritize new candidates.
- Demonstrates AI potential in astronomical research.
- Turning Data into Discovery.

Beyond This Mission

- Integrate JWST and future telescope datasets.
- Further enhance the interactive web-based dashboard.
- Extend to deep learning for higher precision.
- AI Helping Humanity Explore Beyond!