

ASTRO-MATRIX

AI/ML EXOPLANET DETECTOR

“We build a system to filter the cosmic noise and find the needles: the potential exoplanets.”

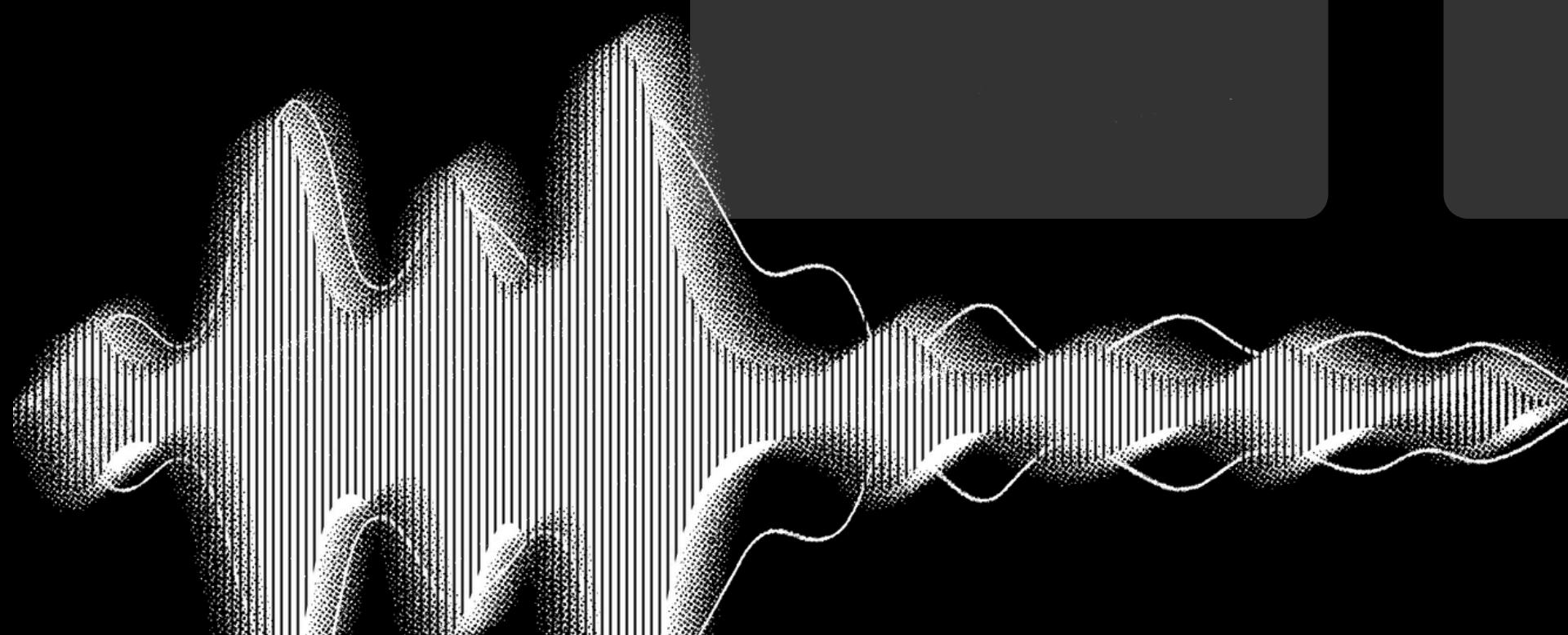
THE SIGNAL VS. THE NOISE

CHALLENGE

To Classify KOIs (Kepler Objects of Interest)

PROBLEM

High volume, subtle transit signals, and overwhelming False Positives.





FEATURE-BASED MACHINE LEARNING (RANDOM FOREST)

DATA ACQUISITION:

Used lightkurve and astroquery to download raw Kepler light curves and their official NASA KOI disposition labels.

FEATURE ENGINEERING (THE INNOVATION):

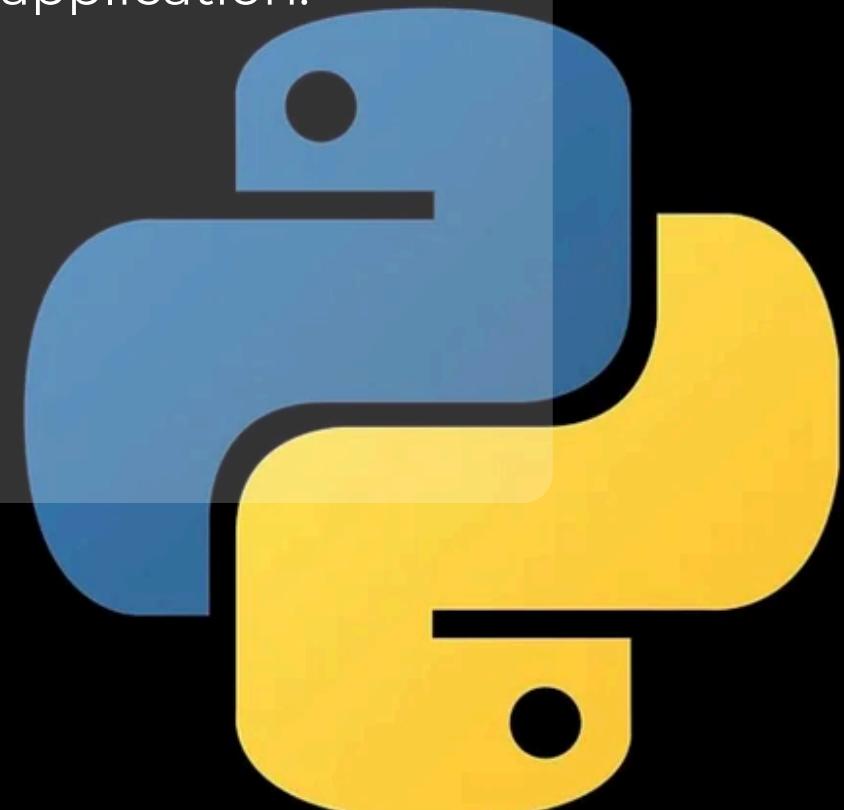
Instead of raw CNN input, we extracted 7 robust features that quantify star stability and periodicity.

CLASSIFICATION:

Trained a highly stable RandomForestClassifier (scikit-learn) on these features.

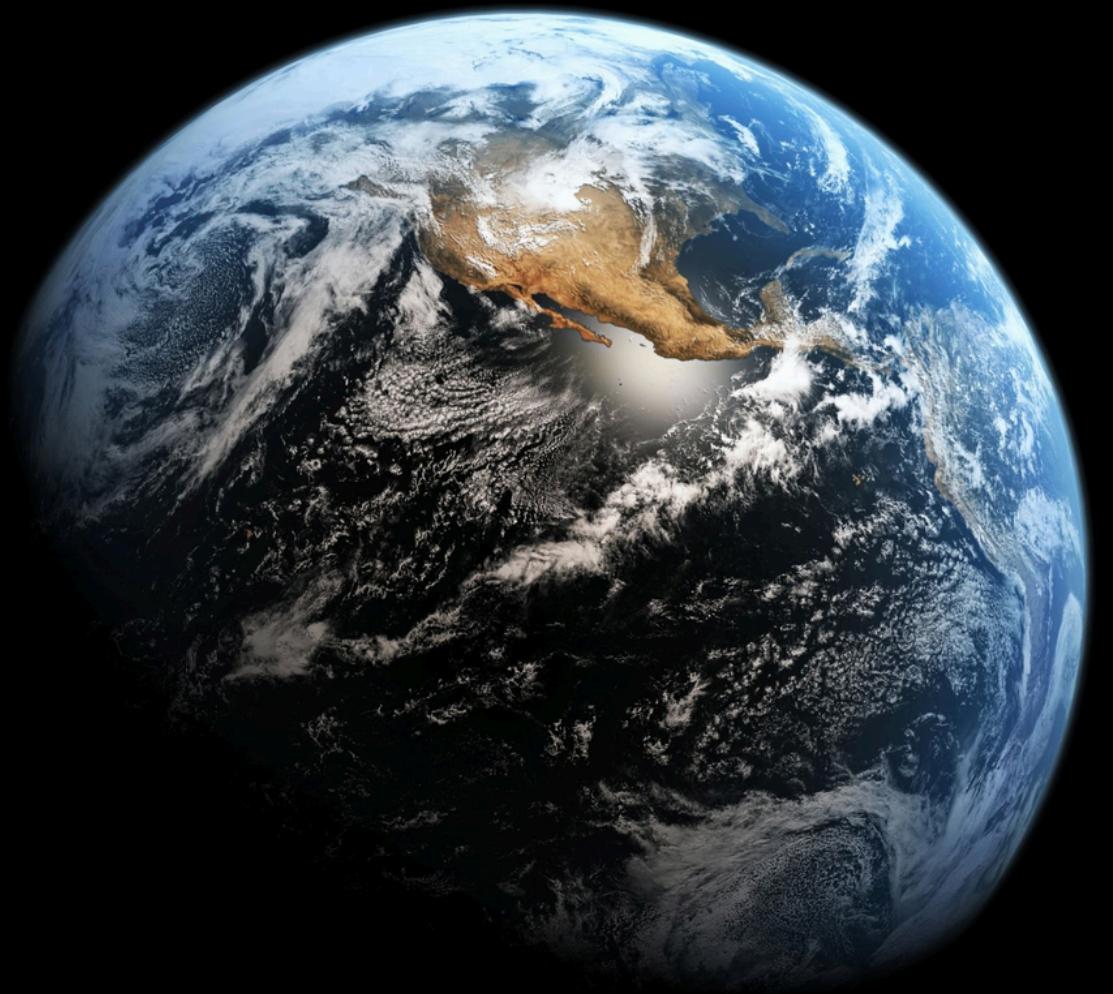
DEPLOYMENT:

Deliver results via a live, interactive Streamlit web application.





WHAT MAKES A PLANET?



STAR STABILITY

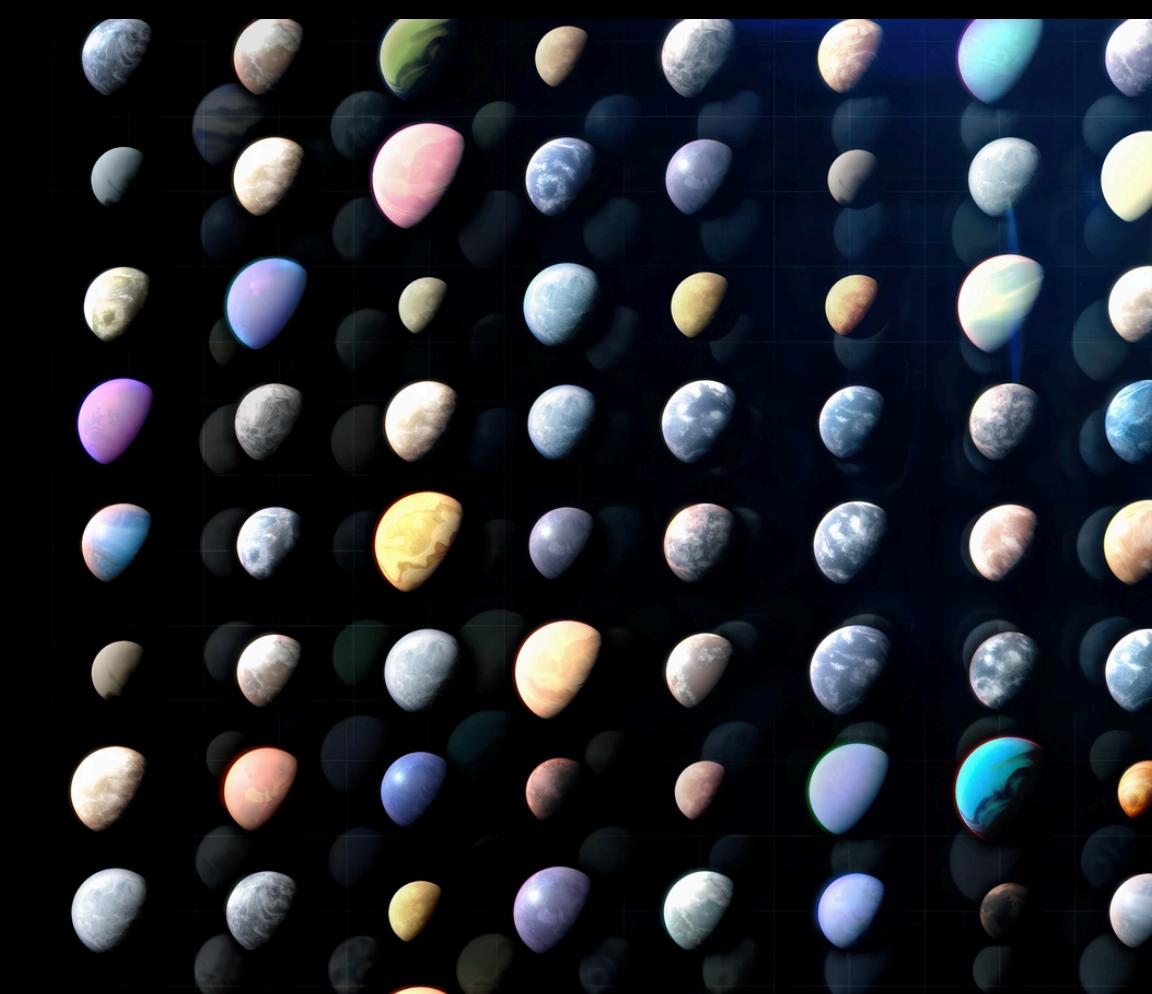
Mean, Std Dev, Variance, Min, Max Flux

PERIODICITY

Lomb–Scargle Power, Best Frequency, False Alarm Probability (FAP)

CORE MESSAGE

“We prioritize periodicity—the signal of a repeating orbit—to teach the model what a planet looks like.”



TRAINING OF THE MODEL

The screenshot shows a Jupyter Notebook interface with a dark theme. On the left is a file explorer titled 'Files' containing files like 'sample_data', 'Untitled0.ipynb', 'exoplanet_app.joblib', and 'exoplanet_hybrid_model.joblib'. The main area displays the output of a Python script:

```
1. Loading embedded Kepler feature data (INSTANT)...
✓ Data collection complete. Total labeled samples: 32

--- 2. Training Hybrid Ensemble Model (RF + XGBoost + Meta) ---
/usr/local/lib/python3.12/dist-packages/xgboost/training.py:183: UserWarning: [16:30:02] WARNING: /workspace/src/learner.cc:738: Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=i, fobj=obj)
Training complete.

✓ Hybrid Ensemble Model successfully saved to exoplanet_app.joblib
-----
👉 Now download 'exoplanet_app.joblib' from Colab's file explorer and upload it to GitHub.

--- CLASSIFICATION REPORT (FOR SLIDE 5 SCREENSHOT) ---
Use this exact output for your Metrics Slide:
      precision    recall   f1-score   support
False Positive (0)    0.00     0.00     0.00      3
Exoplanet (1)        0.25     0.20     0.22      5

accuracy                  0.12      8
macro avg                 0.12     0.10     0.11      8
weighted avg              0.16     0.12     0.14      8

ROC AUC Score: 0.0000

--- FEATURE IMPORTANCE ---
mean_flux            0.154191
std_dev_flux         0.127955
max_flux             0.126490
lomb_scargle_power  0.125211
min_flux             0.120877
false_alarm_probability 0.118050
best_frequency       0.117383
variance_flux        0.109843
dtype: float64
```

At the bottom, a code cell shows:

```
[18]
import os
print(os.listdir('.'))
```

LIVE DEMONSTRATION

 Stop Deploy :

⚙ Configuration

Target Query

KIC 8311864

Stars per Batch

5 - +

Max Batches

2 - +

Flux Points per Star

1000

Label Mode

Real NASA Labels

Demo / Fake Labels

Clear Cache Before Start

Use Saved Model if Available

AI/ML Exoplanet Detector

Automatically detects Exoplanets using NASA Kepler/TESS light curves.

Features advanced periodicity analysis (Lomb-Scargle) and hybrid AI classifiers.

 Start: Fetch → Train → Predict

 Searching for 'KIC 8311864' in NASA archive...

IMPACT & FUTURE VISION



IMPACT:

1. **Speed & Trust:** Provides fast, interpretable screening (eliminates the "black box" problem).
2. **Triage Tool:** Ready for use in large data surveys like TESS or Roman.

FUTURE VISION:

Next step: Transition to a 1D CNN to learn features automatically.

COMPLIANCE:

Built using official NASA resources: lightkurve & astroquery.