Exo-AI Explorer – Project Documentation

AI-Powered Exoplanet Classification Using NASA Mission Data

Exo-AI Explorer automates the classification of exoplanet candidates using data from NASA's Kepler, K2, and TESS missions. It turns weeks of manual analysis into real-time AI-powered predictions.

Makes It Unique

I built a custom machine learning system that combines multiple classifiers and delivers results through an easy-to-use web interface.

Who It's For

- NASA researchers and mission scientists

- Astronomy students and universities

- Citizen scientists and volunteers

How It Works

Data Pipeline, we used data from:

- Kepler Confirmed Planets

- K2 Candidate Catalogs

- TESS Objects of Interest

In total, we processed 206,065 entries.

Data Cleaning & Preparation

- Replaced missing or invalid values

- Scaled features for better model performance

- Ensured only reliable, validated data was used for training

Key Features (Inputs for the Model)

1. Planetary Radius

2. Orbital Period

3. Transit Duration

4. Transit Depth

5. Stellar Temperature

6. Equilibrium Temperature

7. Stellar Gravity

8. Detection Confidence Score

Machine Learning System

We combined four different models for better accuracy and reliability:

1. Random Forest

   - 200 trees, balanced class weighting

2. Extra Trees Classifier

   - 150 trees, tuned for diversity

3. Gradient Boosting

   - 100 estimators, slower but precise

4. Conservative Random Forest

   - A simpler, more cautious version of Random Forest

Weighted Voting

We have more influence to the stronger models:

- Optimized Random Forest → 1.4x
- Extra Trees → 1.3x
- Gradient Boosting → 1.2x
- Conservative RF → 1.0x

Handling Data Imbalance

Model Validation

Testing Approach

- 3-fold cross-validation
- Manual iteration for better control
- No black-box automation

Results

- Balanced Accuracy: Over 80%
- Per-Class Recall:
  - False Positives: 85%
  - Candidate Planets: 82%
  - Confirmed Exoplanets: 89%

A confusion matrix and classification report are generated during training for transparency.

Web Interface

Built with Gradio

- 7 interactive sliders for adjusting planetary parameters
- Built-in confidence scoring
- Instant results, even on standard CPUs

User Experience

- Preloaded examples: Earth-like, Hot Jupiter, False Positive
- Color-coded results for quick interpretation
- No installation required—runs in any browser

Deployment

- Hosted on Hugging Face Spaces
- Publicly accessible via a single URL

Saving & Exporting the Model

- Model saved as a `.pkl` file using joblib
- Metadata exported in JSON format, including:
  - Accuracy scores
  - Cross-validation results
  - Feature names and class distribution
  - Model weights

Quality Tiers

- ≥ 85% accuracy → Production-ready
- ≥ 75% accuracy → Research-grade
- < 75% accuracy → Needs retraining

Research Foundation

Our approach is based on published studies showing that:

- Ensemble models outperform single classifiers
- Feature-based models are more interpretable than deep learning for this task
- Transfer learning principles allow adaptation to new missions

Key Contributions

1. Accessibility– Brings AI-powered analysis to the astronomy community
2. Speed– Cuts analysis time from weeks to seconds
3. Education– Useful for students and researchers
4. Scalability – Ready for future missions like PLATO or Roman

What’s Next

- Real-time data ingestion from TESS alerts

- 3D visualizations of planetary systems

- Mobile app development

- Adding atmospheric composition features

- Integration with JWST data

Thank you for taking the time to learn about Exo-AI Explorer. We’re excited to contribute to the future of exoplanet discovery.