COSMONOVA
Hunting Exoplanets with AI
Final Report – 2025

Title of the project	**Stellar Signal: An AI-Powered Platform for Automated Exoplanet Detection**

Team Member 1	Daniel Jaber
Email	danielkhjaber@gmail.com
LinkedIn	linkedin.com/in/daniel-jaber-cs

Team Member 2	Karim Tannir
Email	tannirkarim13@gmail.com
LinkedIn	linkedin.com/in/karim-tannir-1a896429b

Team Member 3	Leen Taleb
Email	leentaleb530@gmail.com
LinkedIn	linkedin.com/in/leen-taleb-27420922a

Team Member 4	Mohammad El Saleh
Email	mohamadalsalh890@gmail.com
LinkedIn	linkedin.com/in/mohamad-el-saleh-9a8b952a7

Team Member 5	Rida Al Haj Ali
Email	ridahajali8@gmail.com
LinkedIn	linkedin.com/in/ridahajjali

Project Abstract

The exponential growth of data from space missions like Kepler and TESS has created a significant bottleneck in astronomy: the manual analysis of stellar light curves for exoplanet detection is no longer scalable. The Stellar Signal project addressed this critical challenge by developing and validating a robust, artificial intelligence-driven software solution. The primary goal was to create a hybrid ensemble machine learning model capable of automating the classification of Kepler Objects of Interest (KOI) with high accuracy and speed.

Our methodology involved constructing a rigorous data processing pipeline to clean, normalize, and engineer features from the NASA Exoplanet Archive's cumulative KOI dataset. We developed a sophisticated ensemble model, leveraging Gradient Boosting (CatBoost), which was integrated into a modular, web-based platform. This platform features an intuitive user interface built with Streamlit, allowing users to upload data and receive instant classification results with confidence scores.

The project successfully achieved its primary objectives. The final model was evaluated on a held-out test set, demonstrating exceptional performance with **93% accuracy, 93% precision, 93.5% recall, and an F1-score of 93%**. This represents a performance improvement of over 100x compared to manual analysis, reducing identification time from days to seconds. The project has delivered a functional web application that democratizes access to exoplanet research, providing a powerful tool for researchers, students, and citizen scientists. This work establishes a strong, scalable foundation for the next generation of AI-powered astronomical discovery tools, effectively bridging the gap between massive datasets and actionable scientific insights.

Project Accomplishments

This report outlines the accomplishments of the "Stellar Signal" project, aligned with its original aims.

Aim 1: Develop a High-Accuracy AI/ML Model

- **Accomplishment:** Successfully developed, trained, and validated a hybrid ensemble model exceeding the project's target of >90% accuracy.
- **Main Contributions:**
  - Processed a combined dataset from Kepler missions, involving over 4,000 records.
  - Executed comprehensive data cleaning, handling missing values, and performed feature engineering to select the most astrophysically relevant parameters (e.g., `koi_period`, `koi_depth`, `koi_model_snr`).
  - The final model achieved a benchmark **93% accuracy**, with consistent cross-validation results (variance ±1.5%) and a ROC AUC greater than 0.96, confirming high stability and reliability.
  - The model is a good fit; where there is no overfitting detected, the difference between training accuracy and validation accuracy is just 4
- **Difficulties Faced:** Initial challenges included class imbalance in the dataset, which was mitigated through strategic sampling techniques and the use of appropriate performance metrics like F1-score.

Aim 2: Create an Intuitive and Accessible Web Platform

- **Accomplishment:** Designed and deployed a fully functional, modular web application that serves as the front-end for the AI model.
- **Main Contributions:**
  - Built a responsive User Interface (UI) using Streamlit, allowing for easy user input and dynamic visualization of classification results, including confidence scores.
  - Developed a robust Python-based backend API to orchestrate the workflow between the UI, the AI engine, and the database.
  - The platform successfully lowers the barrier to entry for exoplanet research, making advanced AI tools accessible to a non-expert audience.
- **Difficulties Faced:** Ensuring the platform was both user-friendly and capable of handling computational loads required careful architectural planning and iterative user experience testing.

Aim 3: Establish a Foundation for Future Research and Discovery

- **Accomplishment:** The project has laid a strong, modular groundwork for continuous improvement and has generated significant impactful outcomes.
- **Main Contributions:**
  - The system architecture is designed for scalability, allowing for easy integration of new data from future missions (e.g., JWST, PLATO) and more complex models.
  - The project has demonstrated a tangible acceleration of the discovery process, with a 100x improvement in analysis speed.
  - It has fostered interdisciplinary collaboration between computer science and astrophysics within the team.
- **Difficulties Faced:** The "black-box" nature of the ensemble model presents a limitation for interpretability. This has been clearly documented as a key direction for future work, involving the integration of Explainable AI (XAI) techniques.

Project Deliverables

- **Title:** *Stellar Signal: An AI-Powered Platform for Automated Exoplanet Detection*
- **Online Access:**
  Technical Report (Google Drive)
  GitHub Repository