Exo X Hunter – AI-Powered Exoplanet Discovery

Executive Summary:

Our team, *Exo X Hunter*, is participating in the NASA Space Apps Challenge under the theme **"A World Away: Hunting for Exoplanets with AI."** The search for exoplanets is one of the most exciting frontiers in astronomy, as these worlds may hold answers to the age-old question: *Are we alone in the universe?* However, the vast amount of data collected by satellites such as NASA's Kepler and TESS makes manual classification of exoplanet candidates time-consuming and inefficient.

Our solution is a machine learning (ML) model trained to analyze satellite light curve data, detect transit events, and classify whether an exoplanet candidate is likely valid or a false signal. Beyond detection, we designed an **interactive website** to present these results in an engaging, story-driven way.

The **main page** includes an overview of our mission and vision, as well as a "book-style gallery" of discovered exoplanets. Each exoplanet appears as a card with details such as orbit, radius, and classification confidence. The sixth card is special: it acts as a gateway to a **detective-themed game** inspired by *Hunter X Hunter*. In the game, the user plays the role of an interstellar detective, receives training instructions, and is presented with random satellite data. The player decides whether the signal corresponds to a planet or not, earning points and building their own collection of discovered worlds.

By combining AI analysis, interactive visualization, and gamification, our solution makes the science of exoplanet discovery accessible, fun, and educational. It empowers both experts and the public to experience the thrill of hunting for worlds beyond our solar system.

Problem Definition:

NASA missions such as Kepler and TESS have produced terabytes of stellar observation data. Within these massive datasets lie subtle patterns — tiny dips in brightness caused by planets transiting in front of their host stars. Detecting these patterns is essential to identifying new exoplanets, but the process is extremely challenging. The signals are often faint, noisy, and can be confused with false positives from stellar variability or instrument noise.

The problem is twofold: first, **how to efficiently classify exoplanet candidates at scale** using satellite data; second, **how to present these findings** in a way that communicates the excitement of discovery and educates the public. Addressing this problem aligns directly with the challenge goal of using AI to support the search for exoplanets while creating innovative outreach tools that broaden participation in space science.

Background & Literature Review:

Exoplanet detection through the transit method has been central to modern astronomy. NASA's **Kepler Space Telescope** alone has confirmed over 2,600 planets, and the **TESS mission** continues this legacy. Traditionally, astronomers analyze light curves manually or with rule-based methods, but with datasets containing millions of stars, this approach is unsustainable.

Machine learning has emerged as a powerful solution. Recent studies (e.g., Shallue & Vanderburg, 2018) demonstrated the use of convolutional neural networks to classify Kepler light curves with high accuracy. Other research applied gradient boosting and recurrent networks to improve classification robustness. These models outperform human vetting in speed and often in reliability.

However, existing tools remain largely technical and inaccessible to the general public. Citizen science platforms such as **Zooniverse's Planet Hunters** have shown that non-experts can contribute meaningfully, but they lack integration with modern AI pipelines.
Our solution builds on this foundation by combining **state-of-the-art ML models** for classification with a **gamified, story-driven interface**. By framing discovery as an interactive detective mission, we aim to bridge the gap between expert analysis and public engagement, making exoplanet science both rigorous and inspiring.

Methodology:

1. **Data Preparation:** Collect and preprocess NASA Kepler and TESS light curve datasets. Perform detrending, normalization, and noise reduction, then merge them into a **combined dataset** for training and evaluation.
2. **Model Training:** Implement a **stacked solution** combining multiple models for robustness. The base models include:
    1. **Multilayer Perceptron (MLP):** for capturing non-linear temporal patterns in light curves.
    2. **XGBoost & LightGBM:** for gradient boosting on extracted features, ensuring fast and high-performing classification.
3. **Stacking Strategy:** Predictions from the base models are combined in a meta-classifier to improve overall accuracy and reliability.
4. **Evaluation:** Models are compared using precision, recall, F1-score, and ROC-AUC to ensure both scientific validity and scalability.
5. **Backend**: Store processed results in a database accessible via API.
6. **Frontend Visualization**: Build a React-based website to display results: mission info, gallery of discovered planets, interactive charts.
7. **Gamification**: Integrate a game designed around player-as-detective, where random datasets are shown and users classify them. Scoreboards track progress, and successful detections add new cards to the gallery.
8. **Story Integration**: Apply *Hunter X Hunter* theme for design, animations, and detective narrative.

This methodology ensures technical rigor while aligning with the challenge's requirements for AI, visualization, and creative engagement.

Solution:

- **How does your solution tackle the challenge?**

Our project, **Exo X Hunter**, addresses the challenge *"A World Away: Hunting for Exoplanets with AI"* by integrating **machine learning, interactive visualization, and gamification** into a unified platform.
At its scientific core, we employ a **stacked machine learning solution** trained on a **merged dataset** from NASA's Kepler and TESS missions. The ensemble combines a **Multilayer Perceptron (MLP)** for modeling temporal patterns in light curves with **XGBoost** and **LightGBM** for gradient boosting on engineered features. Their predictions are aggregated through stacking, producing reliable classifications of exoplanet candidates with confidence scores. This approach directly tackles the challenge's technical goal of applying AI to accelerate exoplanet discovery.
The results are transformed into an **interactive website**, styled around the *Hunter X Hunter* anime theme to enhance storytelling. The main page highlights our **mission and vision** while presenting discoveries

in a **book-style gallery** of planet cards. To extend engagement beyond data display, we include a **detective-themed game**, where users analyze sample stellar signals and decide whether they correspond to exoplanets. Correct identifications unlock new cards in the gallery, making the experience both educational and rewarding.

This dual approach — rigorous AI analysis combined with gamified, creative visualization — ensures that our solution not only classifies exoplanets but also **tells the story of discovery** in an accessible, inspiring way, fulfilling the challenge's objectives of **scientific innovation, creativity, education, and public engagement**.

- **What steps are necessary to implement it successfully?**

1. **Data & AI:** Preprocess Kepler and TESS datasets into a merged dataset; train and evaluate stacked models (MLP, XGBoost, LightGBM); generate classification outputs with confidence scores.
2. **Backend:** Develop APIs to serve AI predictions, merged dataset subsets, and game data.
3. **Frontend:** Build a React-based website with a book-style gallery, dashboards, and light curve visualizations.
4. **Game:** Design and implement detective gameplay mechanics, scoring system, and progression loop.
5. **Integration:** Connect AI outputs to the gallery and game, ensuring smooth interaction and feedback.
6. **Finalization:** Polish the theme, test the full workflow, and record a 5-minute demo video with narration.

- **How does your solution align with the challenge's requirements?**

- **Applies AI to exoplanet detection:** Uses a stacked ML approach (MLP, XGBoost, LightGBM) trained on a merged NASA dataset.
- **Makes discoveries interactive and understandable:** Displays AI results as an engaging book-style gallery of planet cards.
- **Engages the public through creativity:** Incorporates a detective game where players hunt for exoplanets themselves.
- **Supports education and outreach:** Teaches users how light curves reveal planetary transits through interactive storytelling.
- **Ensures reproducibility:** Provides open-source code, datasets, and methodology for transparency and future use.

Value Proposition:

Our solution, **Exo X Hunter**, delivers value on two complementary levels: **scientific impact** and **public engagement**.

For the **scientific community**, our approach accelerates exoplanet classification using a **stacked machine learning solution** trained on a **merged dataset** from NASA's Kepler and TESS missions. By combining a **Multilayer Perceptron (MLP)** with **XGBoost** and **LightGBM** in an ensemble, the system produces high-confidence predictions that reduce false positives. This allows astronomers to prioritize the most promising candidates for follow-up observations, ensuring that valuable telescope time and research resources are directed efficiently.

For the **general public and learners**, we transform complex astrophysical data into an **engaging, story-driven experience**. The *Hunter X Hunter*-themed detective game and book-style gallery invite users to step into the role of an “exoplanet hunter.” Through gameplay, non-experts learn how light curve analysis reveals planetary transits while building their own collection of discovered worlds. This interactive layer makes science accessible, educational, and memorable.

Unlike alternative approaches that focus solely on **technical analysis** (AI-only models) or rely purely on **citizen science** (manual classifications), our solution combines the **rigor of advanced machine learning** with the **excitement of storytelling and gamification**. This ensures credibility, reproducibility, and broad appeal — directly fulfilling the challenge’s objectives of **AI-driven discovery, creative visualization, and public outreach**.

In short, **Exo X Hunter** does more than detect planets — it transforms the search for new worlds into a **shared adventure**, bridging science and imagination in a way few solutions can.

Role of Team Members:

• **Mahmoud Zaki – AI & Data Analysis**
Focuses on **feature engineering and model optimization**. Fine-tunes the **XGBoost and LightGBM** components of the stacked solution, performs hyperparameter tuning, and evaluates performance with precision, recall, F1-score, and ROC-AUC. Prepares validated outputs for backend integration and game functionality.

• **Youssef El Hendawy – AI & Data Analysis**
Leads **data preprocessing and model development**. Responsible for cleaning and merging Kepler and TESS datasets, training the **Multilayer Perceptron (MLP)**, and integrating it into the stacked ensemble. Generates classification outputs with confidence scores and ensures scientific rigor in the AI pipeline.

• **Haya Mahmoud – Frontend Development**
Collaborates on **UI/UX design and styling**, ensuring smooth navigation and engaging user interactions. Implements animations, gallery transitions, and theme integration. Supports game integration on the frontend, linking visual storytelling with AI-driven outputs.

• **Zeina Dessouky – Frontend Development**
Designs and implements the **interactive React-based website**. Builds the landing page, AI results dashboard, and book-style gallery. Develops visualizations of light curves using **Plotly.js/D3.js** and ensures the *Hunter X Hunter* theme is consistently applied across the user experience.

• **Omar Dardery – Backend Development**
Develops APIs to serve AI results and datasets, manages the database for player progress and scores, and ensures smooth integration between the AI outputs and the frontend.

• **Sawsan Rany – Game Design**
Designs and develops the **detective-themed exoplanet hunting game**. Creates game rules, scoring mechanics, and progression loops where users classify stellar data. Ensures integration with the gallery (unlocking cards) and adapts the visuals, storytelling, and flow to the *Hunter X Hunter* theme for maximum engagement.

Workflow Strategy:

Our team will follow a structured and milestone-driven workflow to deliver a functional and polished solution within the 2-day hackathon timeframe. The workflow emphasizes parallel development, continuous integration, and regular checkpoints to ensure progress across AI, frontend, backend, and game design components.

◆ Steps in the Workflow

Day 1 – Problem Framing and Core Development

1. **Problem Analysis & Planning**
    - Align on challenge requirements and finalize the scope of the solution.
    - Assign responsibilities and confirm technical stack (ML framework, React, backend APIs).
    - Create wireframes for the website and flow diagrams for the game.
2. **AI Development**
    - Preprocess Kepler and TESS datasets into a merged dataset.
    - Train and validate stacked models (MLP, XGBoost, LightGBM).
    - Generate preliminary classification outputs with confidence scores.
3. **Frontend Development (Foundations)**
    - Implement the landing page, mission/vision section, and gallery skeleton.
    - Set placeholders for AI results and card-based gallery display.
4. **Backend Development (Setup)**
    - Build initial APIs for serving AI outputs and game datasets.
    - Test data flow with dummy results to connect frontend and backend.
5. **Game Design (Conceptualization)**
    - Define detective gameplay mechanics, scoring rules, and narrative elements.
    - Map integration points with gallery and backend.
6. **End-of-Day Milestone**
    - First integrated demo: AI sample outputs displayed in the gallery via backend API.
    - Game flow and UI skeleton finalized.

Day 2 – Game Development, Integration, and Finalization

1. **Game Implementation**
    - Build detective game mechanics: dataset presentation, player classification choices, scoring system.
    - Connect successful player detections to unlock new cards in the gallery.
2. **Frontend & Visualization Enhancements**
    - Add interactive light curve visualizations (Plotly.js/D3.js).
    - Apply *Hunter X Hunter*-inspired UI/UX styling, animations, and transitions.
    - Ensure smooth navigation between the landing page, gallery, and game.
3. **Backend & Integration**
    - Finalize API endpoints and connect real AI outputs to the gallery and game.

- Ensure seamless handling of player progress, scores, and unlocked planets.

4. **Testing & Refinement**
   - Validate model outputs within the website and game environment.
   - Conduct usability tests for gameplay, scoring, and gallery updates.
   - Optimize UI responsiveness and visual consistency.
5. **Finalization & Submission**
   - Polish user experience with animations, narrative elements, and theme integration.
   - Record a 5-minute demo video with transcript and voiceover narration.
   - Submit the final deliverables: website, GitHub repository, and demo video.

◆ Collaboration and Communication

- **Version Control:** GitHub for managing code contributions and integration.
- **Communication:** Slack/Discord for real-time coordination.
- **Documentation:** Google Drive for sharing assets, narration scripts, and datasets.
- **Checkpoints:** Scheduled sync meetings every 4 hours to review progress, resolve blockers, and adjust priorities.

◆ Milestones and Checkpoints

- **End of Day 1:** AI generates preliminary outputs; backend APIs functional with dummy data; frontend gallery and mission page implemented; game mechanics drafted.
- **Mid-Day 2:** Detective game functional and connected to gallery; AI outputs integrated into visualizations.
- **End of Day 2:** Fully polished solution with integrated AI, interactive website, functional game, and demo video recorded.

Resources:

◆ Software

- **Python (TensorFlow, scikit-learn, PyTorch):** For developing and training the machine learning models, including the Multilayer Perceptron (MLP) and stacking ensemble pipeline.
- **XGBoost & LightGBM:** Gradient boosting frameworks used as part of the stacked solution to improve classification accuracy and efficiency.
- **Google Colab / Kaggle Notebooks:** Cloud-based platforms providing GPU acceleration for model training, experimentation, and dataset preprocessing.
- **React.js:** Core frontend framework for building the interactive website, book-style gallery, and seamless navigation between sections.
- **Plotly.js / D3.js:** JavaScript visualization libraries for rendering light curve data, confidence graphs, and interactive dashboards.
- **Node.js / Flask:** Backend frameworks to develop APIs that serve AI predictions, merged dataset subsets, and game data to the frontend.

- **GitHub:** Version control and collaboration platform for source code management and team synchronization.
- **Figma:** For wireframing and UI/UX prototyping of the website and game interfaces

---

Hardware

- Personal laptops for development.
- Cloud GPU environments (Google Colab, Kaggle) for ML processing.

Data

- **NASA Kepler Mission Data** – light curve data of stars for transit method analysis.
- **NASA TESS Mission Data** – time-series photometry datasets for exoplanet search.
- **NASA Exoplanet Archive** – confirmed exoplanet database used for model validation.
- **Mikulski Archive for Space Telescopes (MAST)** – archival stellar datasets.

References

1. Shallue, C. J., & Vanderburg, A. (2018). Identifying exoplanets with deep learning: A five-planet resonant chain around Kepler-80 and an eighth planet around Kepler-90. *The Astronomical Journal, 155*(2), 94. https://doi.org/10.3847/1538-3881/aa9e09
2. Jenkins, J. M., Chandrasekaran, H., McCauliff, S. D., Caldwell, D. A., Tenenbaum, P., Li, J., Klaus, T. C., & Cote, M. T. (2010). Overview of the Kepler science processing pipeline. *The Astrophysical Journal Letters, 713*(2), L87–L91. https://doi.org/10.1088/2041-8205/713/2/L87
3. Ricker, G. R., Winn, J. N., Vanderspek, R., Latham, D. W., Bakos, G. Á., Bean, J. L., Berta-Thompson, Z. K., Brown, T. M., Buchhave, L., Butler, N. R., & others. (2015). Transiting Exoplanet Survey Satellite (TESS). *Journal of Astronomical Telescopes, Instruments, and Systems, 1*(1), 014003. https://doi.org/10.1117/1.JATIS.1.1.014003
4. Twicken, J. D., Clarke, B. D., Bryson, S. T., Tenenbaum, P., Wu, H., Jenkins, J. M., Girouard, F. R., Klaus, T. C., Li, J., McCauliff, S. D., Quintana, E. V., & Smith, J. C. (2016). *Kepler data processing handbook: Photometric analysis* (Technical report). NASA Ames Research Center. https://www.osti.gov/servlets/purl/1234372
5. Thompson, S. E., Coughlin, J. L., Hoffman, K., Mullally, F., Christiansen, J. L., Burke, C. J., Bryson, S. T., Batalha, N. M., Haas, M. R., Catanzarite, J., Rowe, J. F., Barclay, T., Caldwell, D. A., Clarke, B. D., Jenkins, J. M., Li, J., Quintana, E. V., Smith, J. C., Tenenbaum, P., & Twicken, J. D. (2018). Planetary candidates observed by Kepler. VIII. A fully automated catalog with measured completeness and reliability based on data release 25. *The Astrophysical Journal Supplement Series, 235*(2), 38. https://doi.org/10.3847/1538-4365/aab4f9
6. Kovács, G., Zucker, S., & Mazeh, T. (2002). A box-fitting algorithm in the search for periodic transits. *Astronomy & Astrophysics, 391*, 369–377. https://doi.org/10.1051/0004-6361:20020722
7. NASA Exoplanet Archive. (n.d.). A service of the NASA Exoplanet Science Institute (NExScI). California Institute of Technology. https://exoplanetarchive.ipac.caltech.edu/
8. NASA. (n.d.). *Kepler mission*. NASA Science. https://www.nasa.gov/mission_pages/kepler/main/index.html
9. TESS. (n.d.). *Transiting Exoplanet Survey Satellite (mission site)*. Massachusetts Institute of Technology. https://tess.mit.edu/

10. Foreman-Mackey, D., Hogg, D. W., & Morton, T. D. (2016). Exoplanet population inference and the abundance of Earth analogs from noisy, incomplete surveys. *Astrophysical Journal*, 822(2), 1–14. https://doi.org/10.3847/0004-637X/822/2/88

Team Name: Exo X Hunter

1- **Mahmoud Zaki** – AI & Data Analysis
2- **Youssef El Hendawy** – AI & Data Analysis
3- **Haya Mahmoud** – Frontend Development
4- **Zeina Dessouky** – Frontend Development
5- **Omar Dardery** – Backend Development
6- **Sawsan Rany** – Game Design