Project Documentation

Since I cannot create or host actual PDF files, I'll provide the complete documentation in a format that you can easily copy into a PDF document:

Stellar Explorer: AI-Powered Exoplanet Discovery & Education Platform

1. High-Level Project Summary

What we developed: We created an integrated educational platform that combines machine learning analysis of NASA exoplanet data with interactive learning tools. The system features specialized classification models for Kepler, K2, and TESS missions, an AI-powered educational chatbot, and a solar system exploration game for children.

How it addresses the challenge: Our project tackles the complexity of NASA's exoplanet data by making it accessible through multiple engagement layers. It transforms raw scientific data into understandable insights while providing educational content suitable for different age groups and knowledge levels.

Why it's important: This platform democratizes space science education, inspires future generations of scientists, and provides researchers with reliable tools for exoplanet classification. It bridges the gap between advanced astronomical research and public understanding of space exploration.

2. Project Details

What it does:

- Analyzes and classifies exoplanet data from three NASA missions
- Provides interactive educational content about our solar system
- Offers AI-powered explanations of complex astronomical concepts
- Engages children with gamified learning experiences

How it works:

Data Processing Pipeline: Custom preprocessing for each NASA dataset

- Machine Learning Models: Separate Random Forest classifiers for Kepler, K2, and TESS data
- Backend API: FastAPI server handling predictions and data management

- Frontend Interface: React-based web application with interactive visualizations
- Educational Components: Solar system game and AI chatbot for space education

Benefits:

- Makes NASA exoplanet data accessible to non-experts
- Provides educational content for multiple age groups
- Offers reliable exoplanet classification using peer-reviewed methodologies
- Inspires interest in astronomy and data science

Intended Impact:

- Increase public engagement with NASA's exoplanet discoveries
- Support STEM education through interactive learning
- Provide researchers with accessible analysis tools
- Inspire future generations to pursue space science

Technical Stack:

- Backend: Python, FastAPI, Scikit-learn, Pandas, NumPy
- Frontend: React.js, JavaScript, HTML5, CSS3
- AI/ML: Random Forest, Google Gemini API
- Data Processing: Custom preprocessing pipelines
- Game Development: Modern web technologies

Creativity:

- Our project creatively integrates four distinct components:
- Scientific data analysis with machine learning
- Interactive educational gaming
- AI-powered conversational learning
- Multi-mission data harmonization

3. NASA Data Used

Primary NASA Datasets:

Kepler Mission Data

- Source: NASA Exoplanet Archive
- Usage: Training classification model for confirmed vs. candidate exoplanets
- Key Features: koi_fpflag_nt, koi_fpflag_ss, koi_period, koi_depth, koi_prad, koi_teq

K2 Mission Data

- Source: NASA Exoplanet Archive
- Usage: Secondary classification model for extended mission data
- Key Features: disposition, pl_orbper, pl_rade, pl_insol, st_teff

TESS Mission Data

- Source: NASA TESS Mission archives
- Usage: Third specialized classification model
- Key Features: tfopwg_disp, pl_rade, pl_insol, pl_orbper

How NASA Data Inspired Our Project:

The vast amount of exoplanet discovery data generated by NASA missions presented both an opportunity and a challenge. We recognized the need to make this valuable scientific information more accessible to educators, students, and the general public while maintaining scientific accuracy.

4. Space Agency Partner & Other Data

NASA Resources:

NASA Exoplanet Archive (primary data source)

TESS Mission data portals

Kepler Mission data repositories

NASA open data APIs

Additional Tools & Resources:

Google Gemini API (AI chatbot functionality)

Scikit-learn (machine learning library)

React.js (frontend framework)

FastAPI (backend framework)

Pandas/NumPy (data processing)

5. Use of Artificial Intelligence (AI)

AI Tools Utilized:

Code Development & Optimization

ChatGPT/Gemini: Accelerated Python code development for data preprocessing, ML model integration, and API development

Usage: Code structure optimization, error handling, documentation

Educational Chatbot

Google Gemini API: Powers interactive space education chatbot

Function: Explains ML model decisions, answers astronomy questions, provides educational content

Voice Generation

AI Text-to-Speech Tools: Generated educational narration for solar system game

Application: Planet descriptions, interactive learning modules, accessibility features

Concept Development

AI Research Assistants: Supported initial research on exoplanet classification and educational game design

AI Content Transparency:

All AI-generated code thoroughly reviewed and modified by team members

Audio content includes metadata indicating AI generation

Educational content verified for scientific accuracy

No NASA branding used in AI-generated content

Final implementation represents original work with AI as productivity tool

Ethical Considerations:

Scientific accuracy prioritized over AI-generated content

Human oversight maintained throughout development

Accessibility and educational value as primary goals

Proper attribution for all AI-assisted components