**Habitable or Non- Habitable Planet Prediction**

**1. Use Case: The Problem to Solve**

SpaceX has the intention to explore new planets that can potentially be settled by humans. One of the main problems in interplanetary exploration is identifying which exoplanets are habitable or suitable for colonization. The aim of this project is to design a machine learning-based prediction model to classify exoplanets as habitable or non-habitable using various planetary and stellar attributes.

**Objective:**  
The main aim is to develop a tool for the classification of exoplanet data in real-time, thereby helping SpaceX in making informed decisions about target planets for future missions.

**Monday:** <https://view.monday.com/7622629609-8797aff4010080e9dc8444bed05e3f3b?r=use1>

**AWS EC2 URL**: <http://ec2-44-203-48-40.compute-1.amazonaws.com:8080>

**2. Architecture Diagram:**

**A diagram of a process

Description automatically generated**

**Step 1: Data Acquisition**

* **Dataset Source:** NASA Exoplanet Archive, which provides detailed planetary and stellar attributes for known exoplanets.

A screenshot of a computer

Description automatically generated

**Name:** The name or designation of the exoplanet; example: 16 Cygni Bb, 23 Librae b.

**Mass:** Mass of the exoplanet-most often given in terms of times more massive than Jupiter or Earth.

**Period:** Orbital period of the exoplanet or time taken by it for one orbit around the host star.

**Discovery Method:** Techniques used for the discovery of an exoplanet. Common techniques include radial velocity, imaging, and timing.

**Distance:** How far away the exoplanet is from its host star, in light years.

**Host Star Temperature:** Surface temperature of the host star measured in Kelvin.

**Habitability:** A classification based on a set of attributes whether an exoplanet is Habitable or Non-Habitable

**Resources:** <https://exoplanetarchive.ipac.caltech.edu/docs/data.html>

**Step 2: Data Preprocessing**

* **Cleaning:** Handled missing data, outliers, and inconsistencies.
* **Feature Engineering:** Created new features that may help improve the model, such as normalized values or ratios of different attributes.
* **Data Splitting:** Split the dataset into training and testing sets.

**Step 3: Model Development**

* **Model Selection:** Used machine learning algorithm **Random Forest** from Scikit-learn to train the model.

**Step 4: Model Deployment**

* **Flask Web Application:** Deployed the trained model as a web service using Flask.
* **Real-Time Prediction:** Created simple interface for real-time exoplanet classification based on user input of planetary attributes.

**Step 5: Integration and Hosting**

* **AWS EC2:** Hosted the Flask application on AWS EC2 for reliable, scalable deployment.
* **API for Real-Time Prediction:** Implement an API that SpaceX can use to get predictions on exoplanet habitability.

**3. Technologies/Tools Used**

* **Python (Scikit-learn, Flask):**
  + **Scikit-learn** for training, evaluating, and tuning machine learning models.
  + **Flask** for creating and deploying the model as a web service.
* **AWS EC2 Instance:**
  + Host the model and web application, ensuring scalability and availability.
* **NASA Exoplanet Archive:**
  + Data source that provides detailed information on exoplanets and stars for model training.

**Folder Structure:**

A screenshot of a computer

Description automatically generated

**Main Takeaways and Values for SpaceX**

* **Improved Decision Making:**  
  By automating the classification of exoplanets as habitable or non-habitable, SpaceX can quickly identify potential planets for further exploration and colonization.
* **Scalability:**  
  The model can be used in real-time, providing a scalable solution to handle a large volume of exoplanet data as it becomes available.
* **Efficiency in Planning:**  
  SpaceX can reduce the time and resources spent on manual analysis and focus on the most promising planets for future missions.

**Future Scope & Conclusion:**

This machine learning-based exoplanet habitability classification tool provides SpaceX with a powerful resource for making more informed decisions about future space missions and colonization. By automating the analysis of planetary data and classifying exoplanets based on their habitability, SpaceX can streamline the exploration process and focus on the most viable planets for future settlement.