**Spaceship Titanic**

**Welcome to the year 2912, where your data science skills are needed to solve a cosmic mystery. We've received a transmission from four lightyears away and things aren't looking good.**

**The Spaceship Titanic was an interstellar passenger liner launched a month ago. With almost 13,000 passengers on board, the vessel set out on its maiden voyage transporting emigrants from our solar system to three newly habitable exoplanets orbiting nearby stars.**

**While rounding Alpha Centauri en route to its first destination—the torrid 55 Cancri E—the unwary Spaceship Titanic collided with a spacetime anomaly hidden within a dust cloud. Sadly, it met a similar fate as its namesake from 1000 years before. Though the ship stayed intact, almost half of the passengers were transported to an alternate dimension**

**To help rescue crews and retrieve the lost passengers, you are challenged to predict which passengers were transported by the anomaly using records recovered from the spaceship’s damaged computer system.**

**Help save them and change history!**

**Marking Scheme:**

**1. Data Understanding and EDA (20%)**

* **Understanding the Dataset (5%)**
  + Comprehensiveness in describing the dataset, including sources, features, and target variables.
  + Identification of missing values, data types, and overall structure.
* **Data Visualization (10%)**
  + Quality and relevance of visualizations used to explore the dataset.
  + Use of plots (e.g., histograms, bar charts, scatter plots) to identify trends, patterns, and outliers.
  + Creativity in visual representation.
* **Insights and Findings (5%)**
  + Depth of insights gained from the EDA.
  + Clear and concise summary of the key findings.
  + Identification of potential data issues or biases.

**2. Data Preprocessing (15%)**

* **Data Cleaning (5%)**
  + Handling of missing values, outliers, and incorrect data entries.
  + Justification of chosen methods for data cleaning.
* **Feature Engineering (5%)**
  + Creation of new features or transformation of existing ones to improve model performance.
  + Explanation and justification of feature engineering choices.
* **Data Scaling and Encoding (5%)**
  + Appropriate scaling/normalization of numerical features.
  + Correct encoding of categorical variables.

**3. Model Building (35%)**

* **Model Selection (10%)**
  + Justification for the choice of algorithms/models.
  + Diversity in the types of models tried (e.g., linear models, tree-based models, ensemble methods).
* **Model Training (10%)**
  + Effective use of training data and proper handling of the train-test split.
  + Implementation of cross-validation techniques.
* **Model Evaluation (15%)**
  + Comparison of different models based on evaluation metrics.
  + Interpretation of results and selection of the best model.

**4. Model Interpretability and Explainability (15%)**

* **Feature Importance (5%)**
  + Identification and visualization of important features contributing to model predictions.
* **Model Explainability (5%)**.
  + Clear explanation of how the model makes predictions.
* **Transparency and Documentation (5%)**
  + Detailed documentation of the model building process.
  + Clarity in the explanation of each step taken.

**5. Presentation and Communication (15%)**

* **Clarity and Structure (5%)**
  + Logical flow and organization of the project report or presentation.
  + Clear headings, sections, and sub-sections.
* **Visual Aids (5%)**
  + Use of visual aids (charts, graphs, tables) to support explanations.
  + Aesthetic quality and readability of visual aids.
* **Overall Communication (5%)**
  + Ability to convey complex concepts in a simple and understandable manner.
  + Engagement and persuasiveness of the presentation.

**6. Innovation and Creativity (10%)**

* **Novelty in Approach (5%)**
  + Innovative methods or techniques used in EDA, preprocessing, or model building.
  + Uniqueness of the approach compared to standard methods.
* **Creative Solutions (5%)**
  + Creative problem-solving skills demonstrated throughout the project.
  + Implementation of cutting-edge techniques or tools.