## Explanation for Spaceship Titanic

This code predicts whether passengers were transported in the Spaceship Titanic competition. It follows a structured pipeline for data preprocessing, feature engineering, and model training using **Support Vector** **Classifier (SVC).**

**1.Import Libraries**

* **Pandas** 🡪 Handles data manipulation.
* **Numpy** 🡪 Used for numerical operations.
* **Train\_test\_split** 🡪 Splits data into training and validation sets.
* **LabelEncoder** 🡪Encodes categorical features into numbers.
* **SVC** 🡪Machine learning model used for classification.

**2. Kaggle Dataset Link**

<https://www.kaggle.com/competitions/spaceship-titanic>

**3.Load Dataset**

Reads **train.csv** and **test.csv** into pandas DataFrames

**4.Data Exploration**

* Examines the dataset structure and missing values
* This helps understand data types, missing values and dataset size.

**5.Handle Missing Values**

**Why?**

* Missing values can cause errors in machine learning models.
* **Categorical columns:** Filled with the most frequent value(model).
* **Numerical columns:** Filled with the median value to prevent distortion.

**6.Encode Categorical Data**

**Why?**

* Machine learning models work with numerical values.
* **LabelEncoder()** converts categorical values into unique numbers.
* Handles unseen categories in test data by assigning them **-1.**

**7.Define Features and Target Variable**

* **Features:** Selected columns for training.
* **Y:** Target variable (Transported converted to 0 and 1).

**8.Train Test Split**

**Why?**

* Splitting the dataset prevents overfitting and helps evaluate performance.
* **80%** training**, 20%** validation.

**9.Train the Support Vector Classifier (SVC)**

**Why?**

* **SVC** is a powerful classification model that finds the best decision boundary between classes.
* Uses the default kernel to separate data.

**10.Model Evaluation**

**Why?**

* Measures how well the model performs on unseen data.

**11.Save submission File**

* Create a CSV file with **PassengerId** and **Transported** for kaggle submission
* Saves it as **submission.csv.**