

**Sehrish Mubarik**

**BSAIM-F23-001(4A)**

**Assignment no 04**

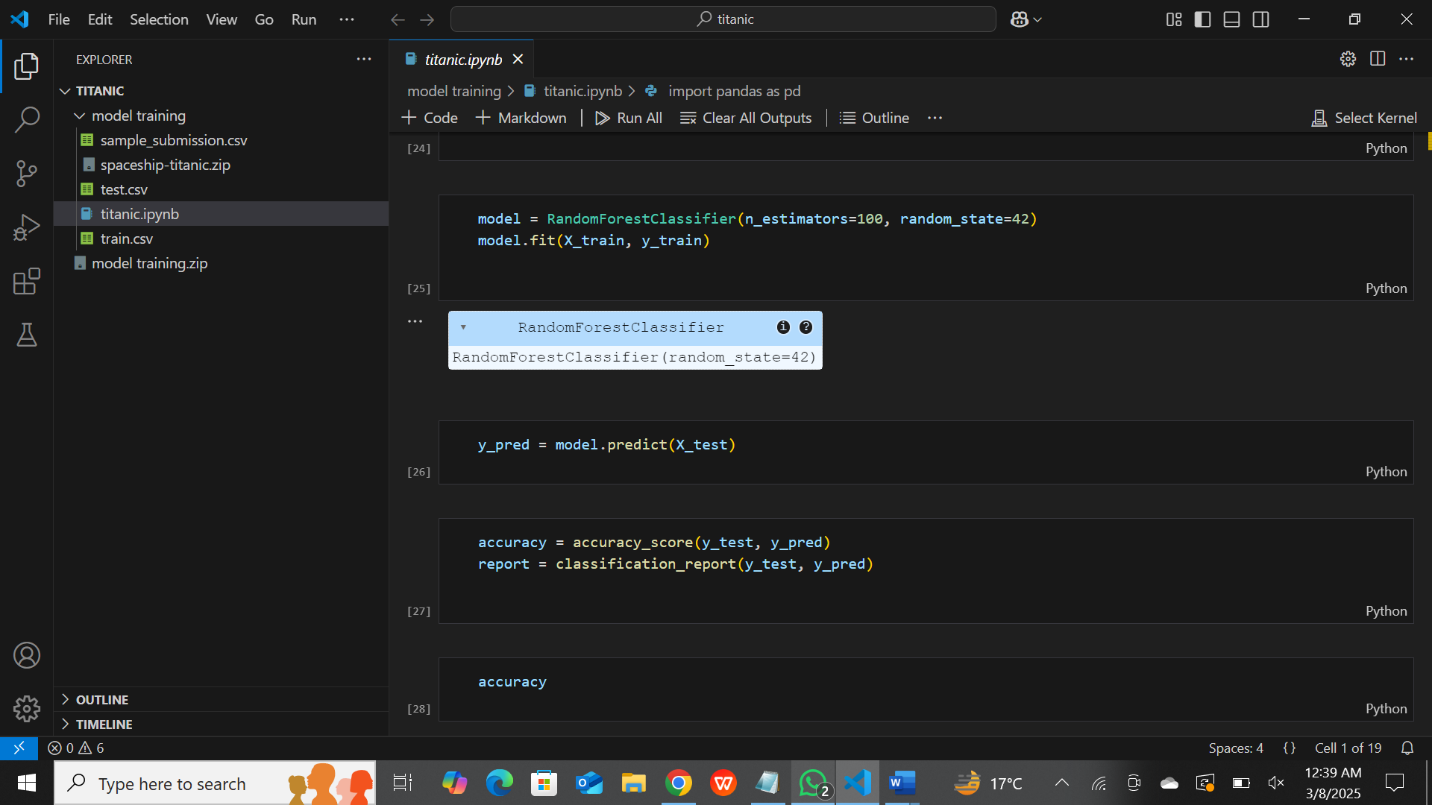
**Submitted to:**

**Sir Rasikh ali**

**Lab task 01**

**Question no 01:**

**Titanic**

****

The description of that model is:

1. **Data Loading & Exploration**

* Load the dataset (train.csv and test.csv from Kaggle's Titanic dataset).
* Display basic information
* Check for missing values and data distributions.

1. **Data Preprocessing**

* Fill missing values (e.g., using median age for missing Age values).
* Convert categorical features into numerical values using Label Encoding or One-Hot Encoding.
* Drop irrelevant columns such as PassengerId, Name, Ticket, and Cabin if they don’t contribute to prediction.

1. **. Feature Engineering**

* Create new features
* Convert Fare and Age into categorical bins if needed.

1. **Model Training**

* Split data into **training and validation sets**.
* Train classification models like:
* Logistic Regression
* Random Forest
* Support Vector Machines (SVM)
* Evaluate models using accuracy, precision, recall, and F1-score.

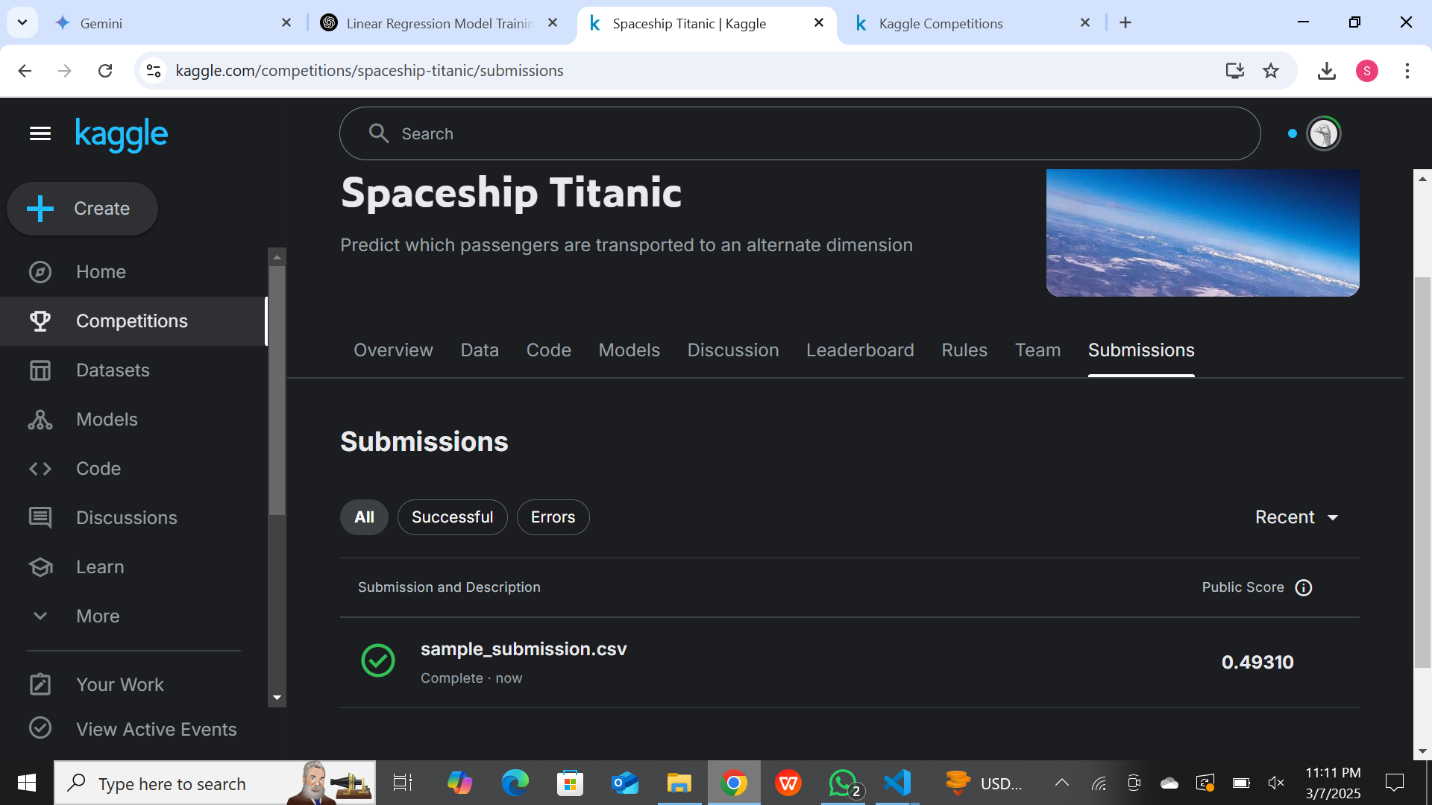
1. **Hyperparameter Tuning**

* Use GridSearchCV or RandomizedSearchCV to find the best model parameters.

1. **Prediction & Submission**

* Apply the trained model to the test dataset.
* Generate predictions and save them in submission.csv for Kaggle.

**Kaggle submission**

****