**Lab 01**

****

**Name Syed M Haider Jafir**

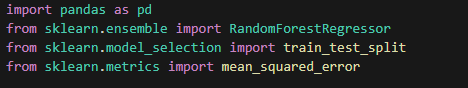
**Roll no SU92-BSAIM-S24-004**

**Section BSAI-4A**

**Subject PF (AILab)**

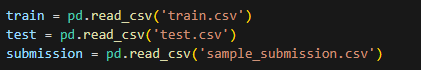
**House Price Prediction**

**Import Libraries:**

****

* pandas: for data manipulation.
* RandomForestRegressor: the machine learning model.
* train\_test\_split: to split training data for validation.
* mean\_squared\_error: to evaluate model performance.

**Load CSV Files:**



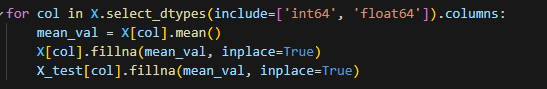
* Loads your training data, test data, and the submission template.

**Separate Features and Target:**

****

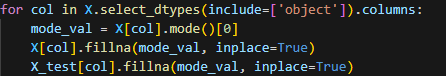
* X: all input features from training data.
* y: the target variable (SalePrice).
* X\_test: test features for prediction.

**Numeric Columns:**

****

* Fills missing numeric values with the column mean.

**Categorical Columns**

****

* Fills missing categorical values with the most frequent value (mode).

**Manual One-Hot Encoding**

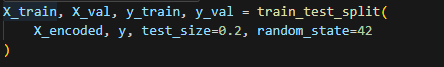
****

* Combines train and test to ensure consistent encoding.
* Converts categorical variables into binary columns.



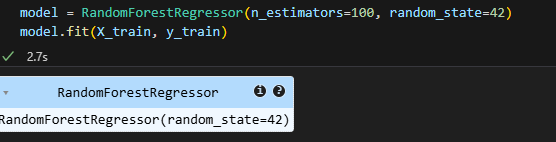
* Splits the encoded data back into training and test sets.

**Split Training Data for Validation**

****

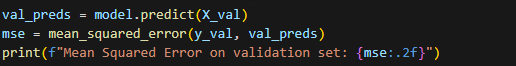
* Splits training data into 80% training and 20% validation for performance testing.

**Train the Model**



* Initializes and trains a Random Forest model with 100 trees.

**Evaluate with Mean Squared Error**

****

* Predicts on validation set and prints the MSE — a lower value means better performance.

**Final Prediction and Submission**

****

* Predicts house prices for the test set.
* Saves the results in the format required for submission.