\*\*Summary\*\*

The Python code aims to predict investment risk in real estate based on relevant features and save the results in a CSV file. Here are the key steps:

1. Read real estate investment data from a CSV file.

2. Calculate risk scores for each investment.

3. Select features, preprocess data, and split it into training and testing sets.

4. Create three different machine learning models (Linear Regression, Gradient Boosting Regression, Random Forest Regression) to predict risk scores.

5. Train models, make predictions, and compute average risk scores.

6. Save the results in a CSV file, including risk scores and risk levels.

\*\*Detailed Descriptions of the Three Models:\*\*

1. \*\*Linear Regression Model (`linear\_model`):\*\*

- Uses linear regression algorithm to build the model.

- Attempts to predict the risk score for real estate investments by fitting a linear relationship.

- Predicted results of the model are stored in the `Predicted Score (Linear)` column.

- The predicted results are normalized and used to categorize investments into low, medium, and high risk levels based on the normalized values.

2. \*\*Gradient Boosting Regression Model (`gbt\_model`):\*\*

- Constructs the model using the gradient boosting regression algorithm.

- Improves prediction accuracy by iteratively training multiple decision trees.

- Predicted results of the model are stored in the `Predicted Score (GBT)` column.

- Like the linear regression model, the predicted results are normalized and investments are categorized into different risk levels.

3. \*\*Random Forest Regression Model (`rf\_model`):\*\*

- Builds the model using the random forest regression algorithm.

- Similar to gradient boosting, it ensembles multiple decision trees for prediction.

- Predicted results of the model are stored in the `Predicted Score (RF)` column.

- Predicted results are also normalized and used for risk level categorization.