**CIS 508 : MACHINE LEARNING IN BUSINESS**

**INDIVIDUAL ASSIGNMENT 4**

**RESTAURANT REVENUE PREDICTION**

**OBJECTIVE:** Using the Kaggle dataset to predict the revenue of a restaurant

**STRATEGY USED:**

* Used data cleaning concepts like checking for null data and one hot encoding to handle categorical data.

***No null values are present in the data***

A screenshot of a computer

Description automatically generated

***One Hot Encoding coding snippet***

A screen shot of a computer code

Description automatically generated

* Created a new feature that calculates the number of days the restaurant was open.

A computer screen shot of a computer code

Description automatically generated

* Removed the Target column ‘Revenue’ from training data and assigned it to ‘Ytrain’ .The other columns were assigned to ‘Xtrain’

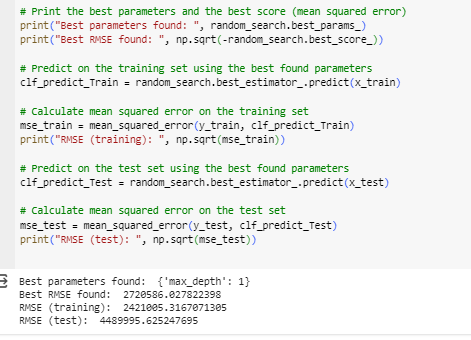
A computer code with green text

Description automatically generated

* Used various types of regressors like Decision Tree, Random Forest, Support Vector ,MLP and Gradient Boost to perform regression and calculate the RMSE for training and test data.
* Performed hyperparameter tuning on the individual regressors and obtained RMSE for test and training datasets and also obtained the best parameters for each regressor.

***For example :***

in case of Decision Tree regressor the best parameter was found and used to find the RMSE for the test and training data set



A similar process was followed for the other four regressors.

* Building a stacked model using multiple regressors like MLP, Random Forest Regressor, and Support Vector Regressor as RMSE on the testing set was comparatively lower than Gradient Boost and Decision Tree Regressor
* Used random search for hyperparameter tuning for the stacked model and found the best parameter

**Output:**

**A screenshot of a computer code

Description automatically generated**

**Kaggle Score**

**A white background with black numbers

Description automatically generated**