Financial Management

In propose work we have used various AI based algorithm to forecast future expenses and can recommend new expenses to user based on current expenses. Using propose model user can know about future expenses and based on income and expenses he will plan his budget.

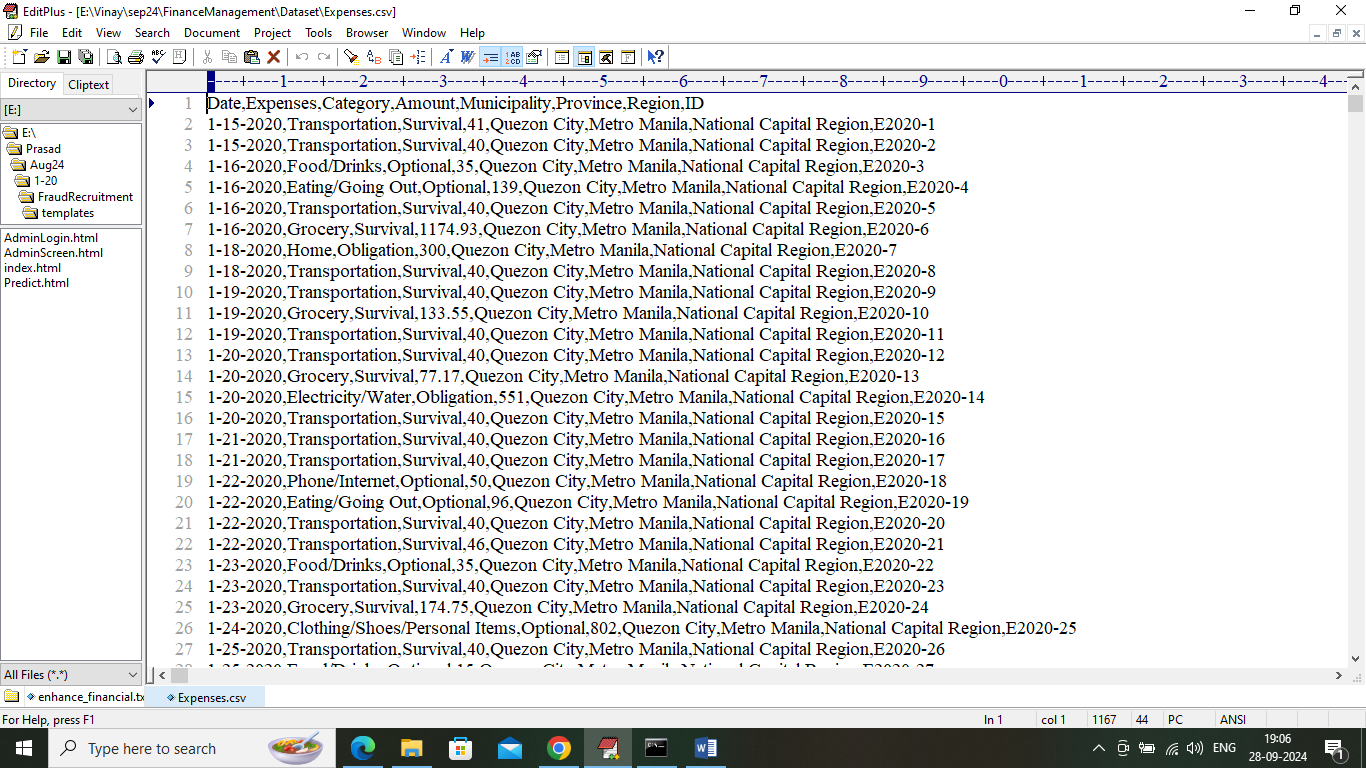
For accurate expenses forecasting we have employed multiple algorithms such as Random Forest, Decision Tree and CNN2D. Each algorithm performance is measured using different metrics such as MSE (mean square error), RMSE (root mean square error) and R2Score (accuracy of regression models).

MSE and RMSE refers to difference between original and predicted values so the lower the difference the better is the algorithm. R2score refers to accuracy between 0 and 1 and the values close to 1 will be consider as best R2score for the algorithm.

To train and test above algorithm performance we have used Expenses dataset from KAGGLE repository which can be download from below URL

<https://www.kaggle.com/code/daverowelvalois/visualizing-my-expenses/input>

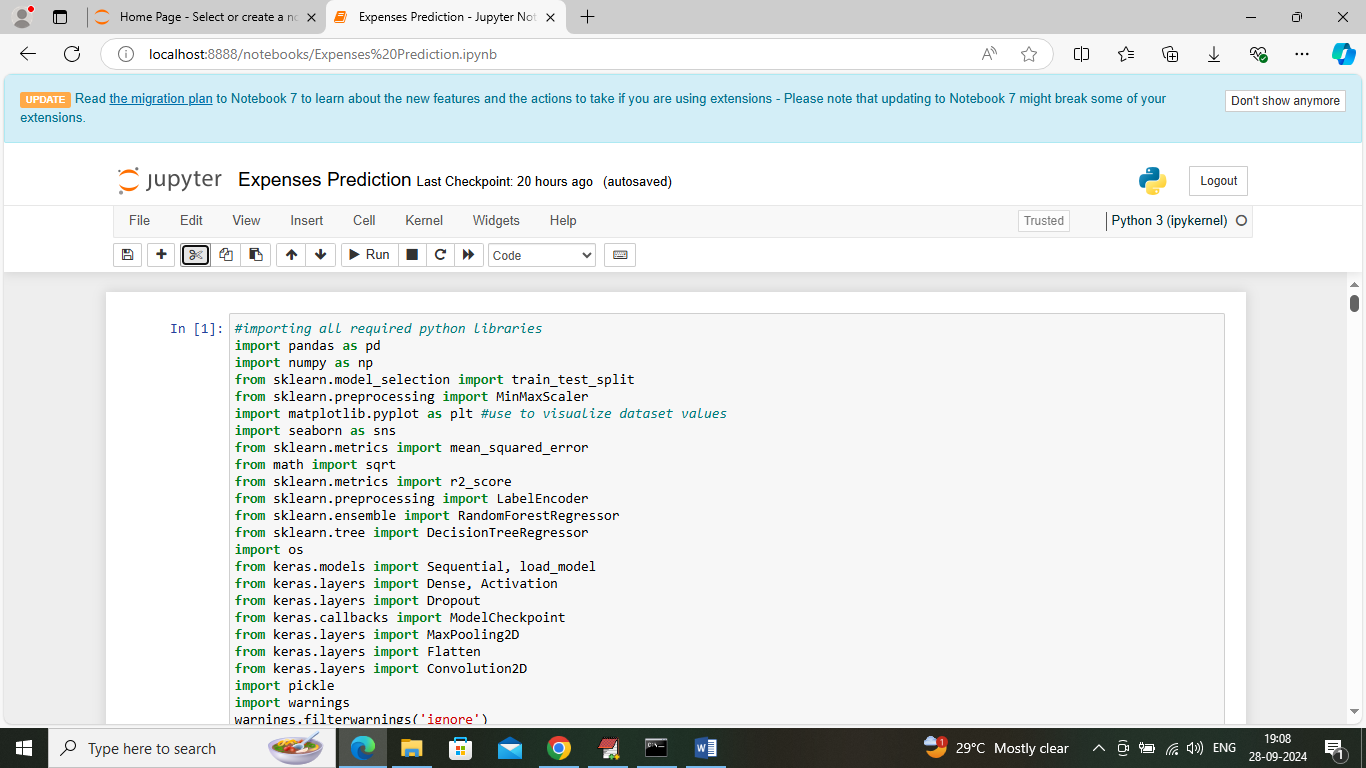
in below screen showing dataset details



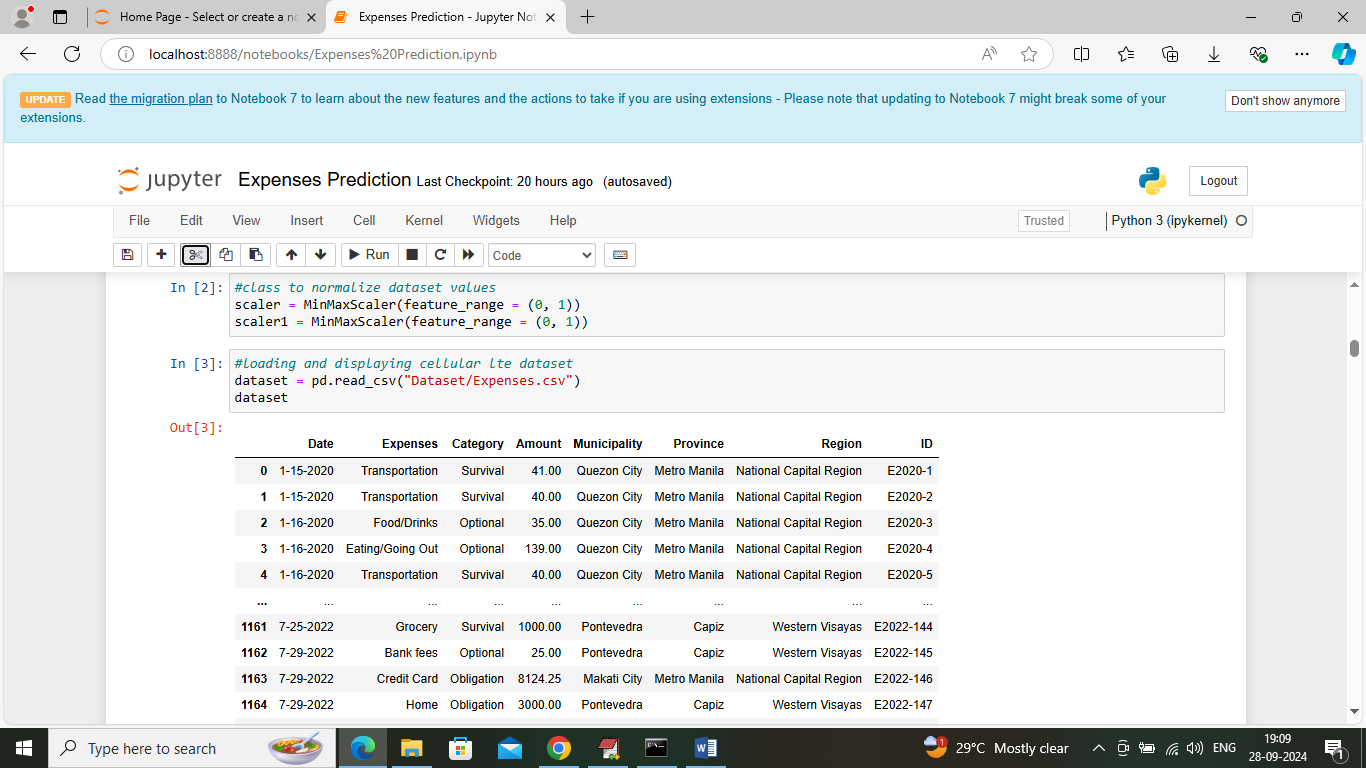
In above dataset screen first row contains dataset column names and remaining rows contains dataset values and in values we have columns like Expenses category name, amount and other details. So by using above dataset will train and test all algorithm performance.

SCREEN SHOTS

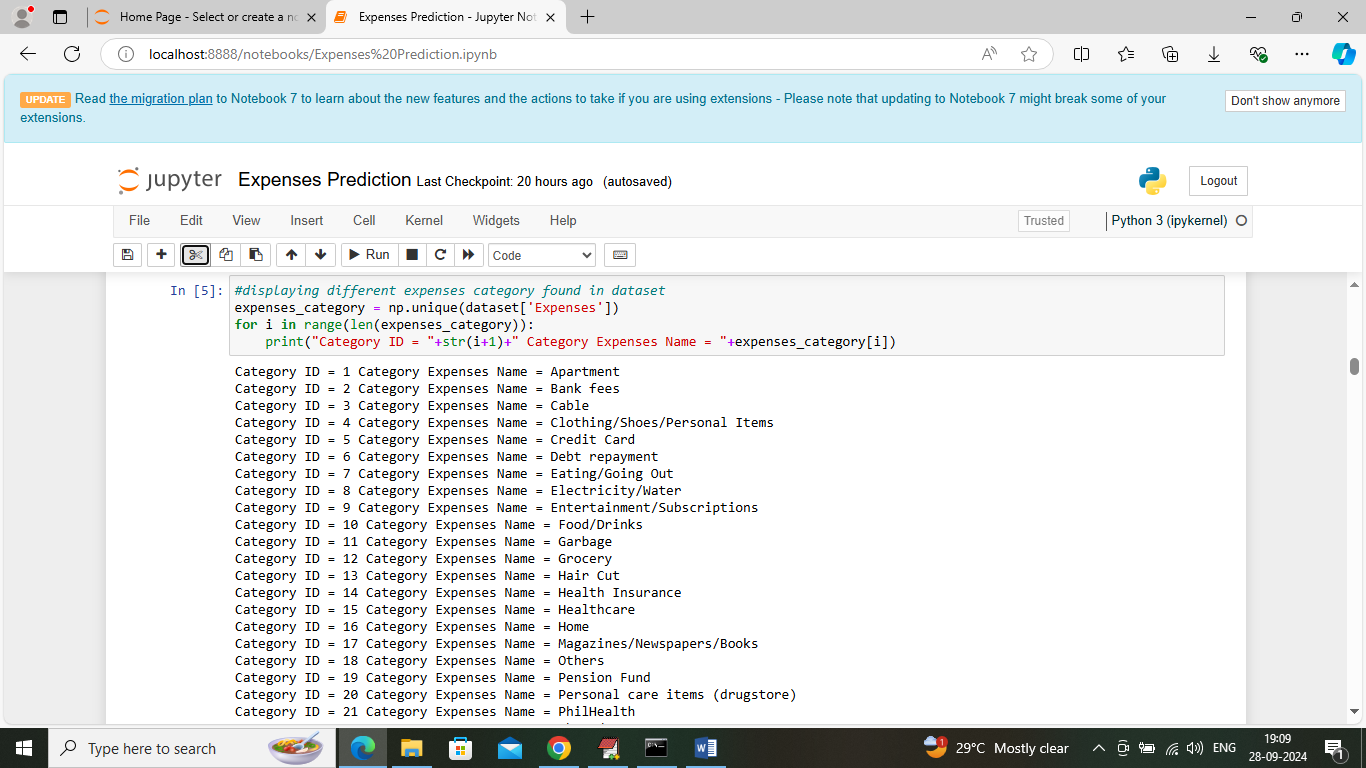
We have coded this project using JUPYTER notebook and below are the code and output screens with blue colour comments



In above screen importing required python classes and packages



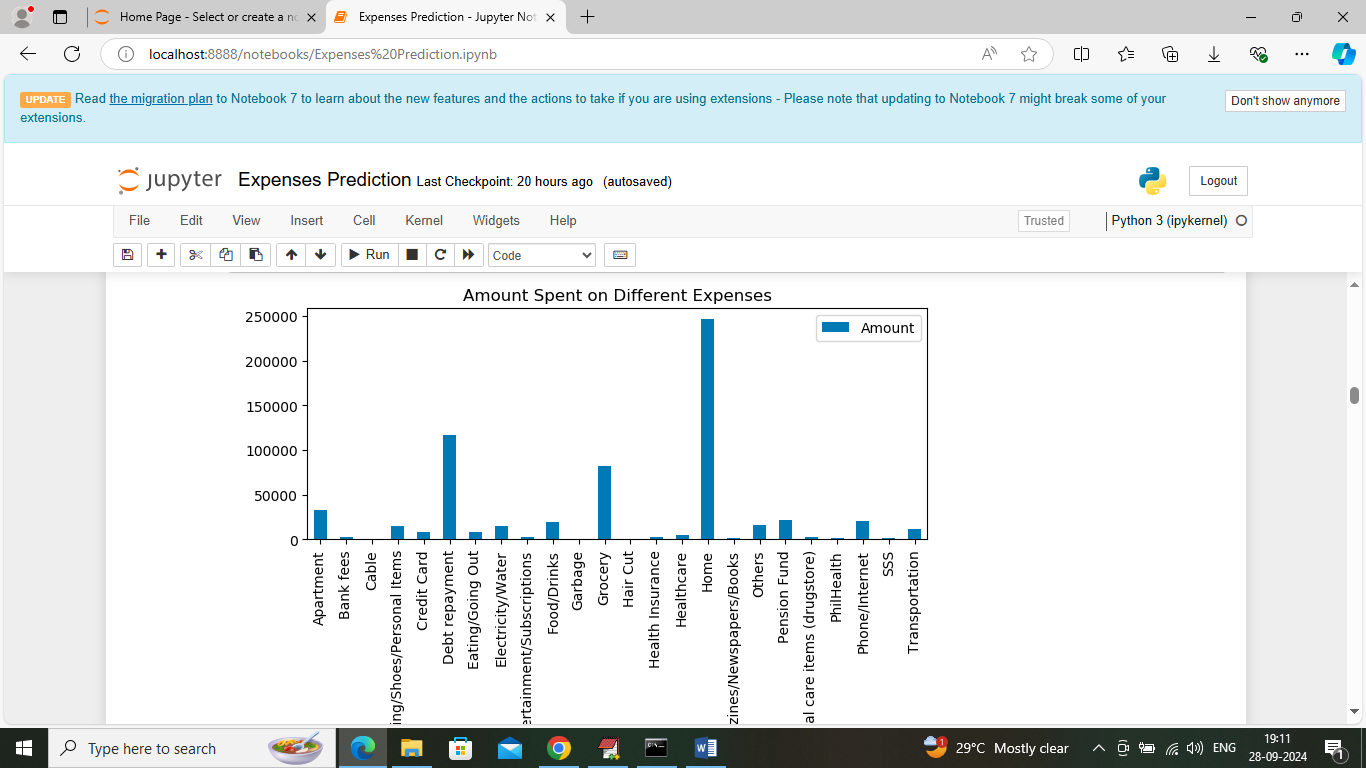
In above screen loading and displaying dataset values



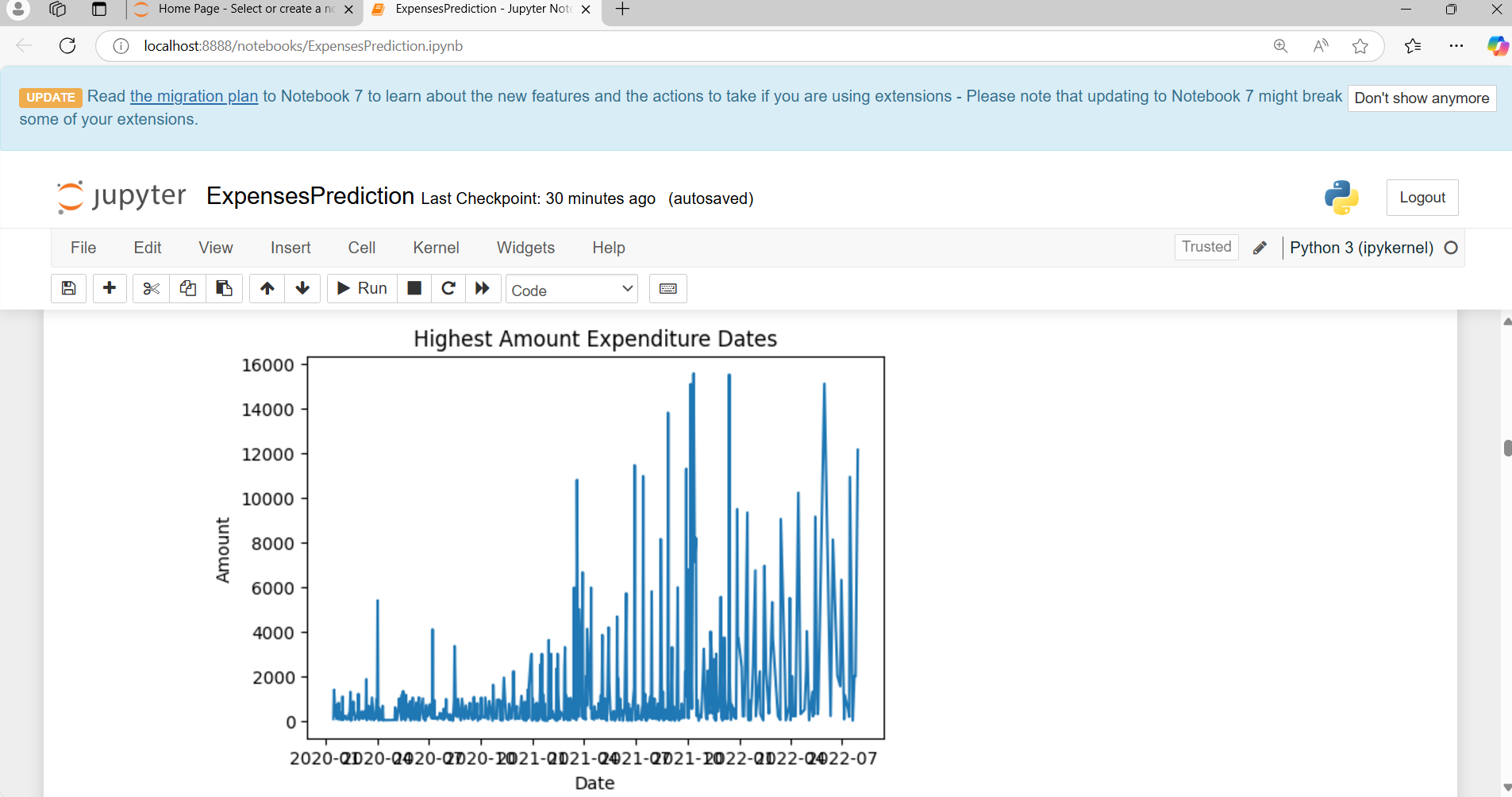
In above screen displaying category Id along with expenses category names and while recommendation you can enter any category ID and then application will recommend other category expenses names.



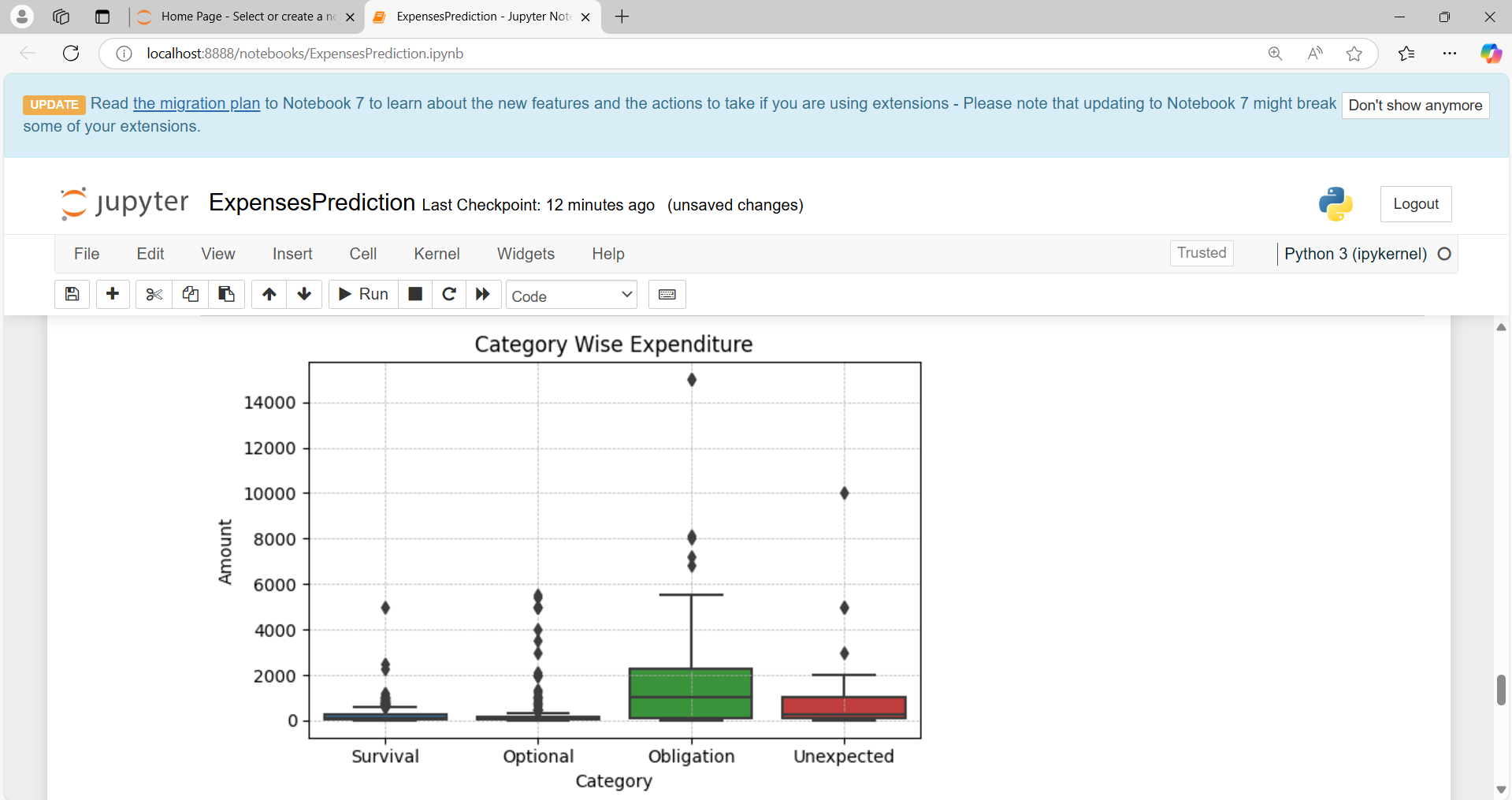
In above screen finding and displaying count of missing values in dataset but above dataset contains 0 missing values count



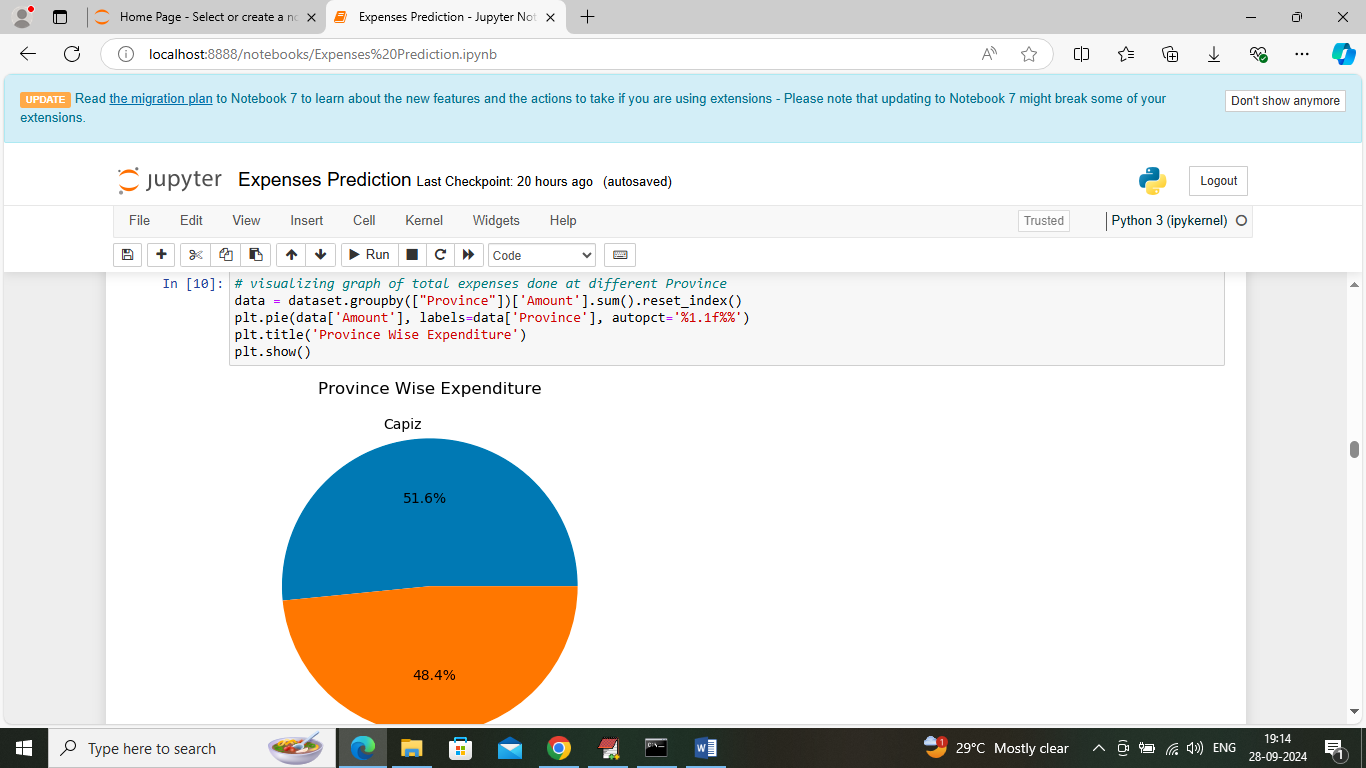
In above screen visualizing total amount spent on different expenses category where x-axis represents category names and y-axis represents expenses amount



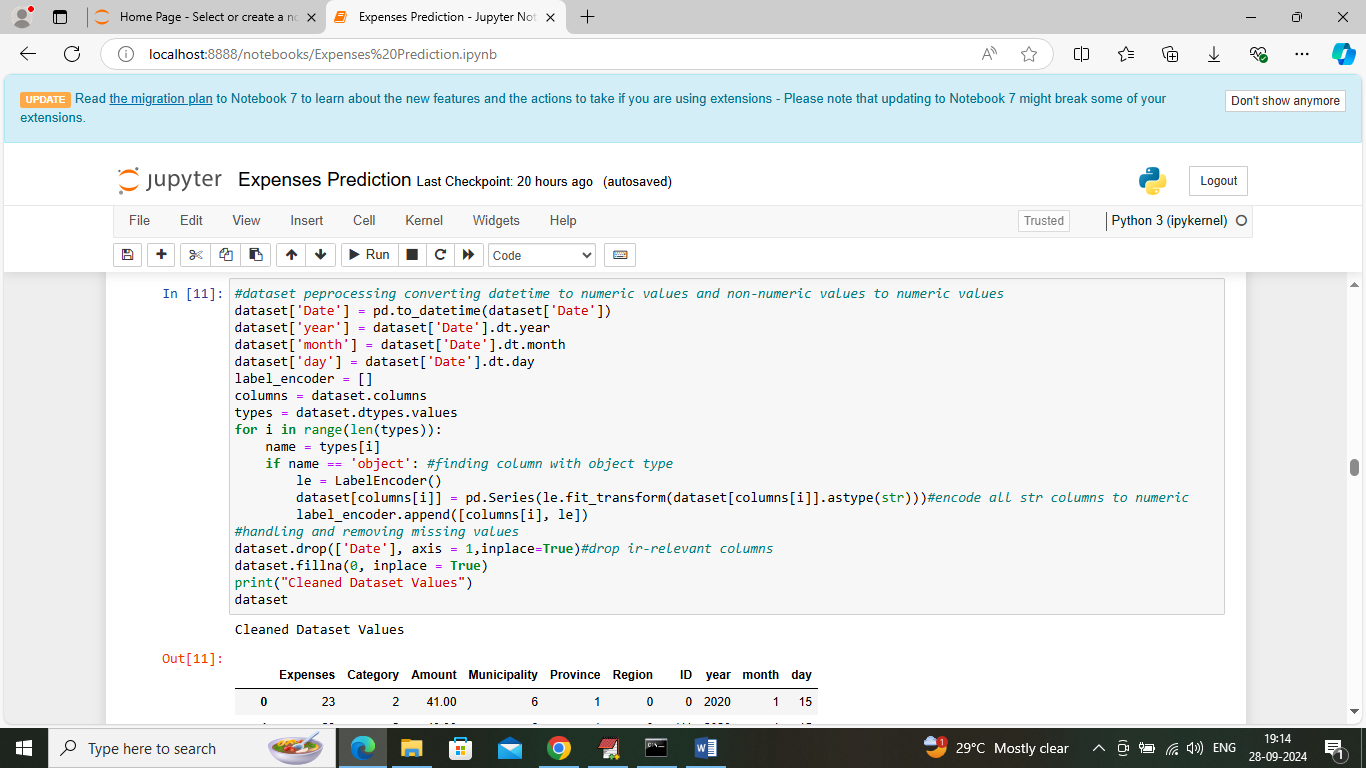
In above screen visualizing expenditure graph based on date where x-axis represents Date and y-axis represents amount



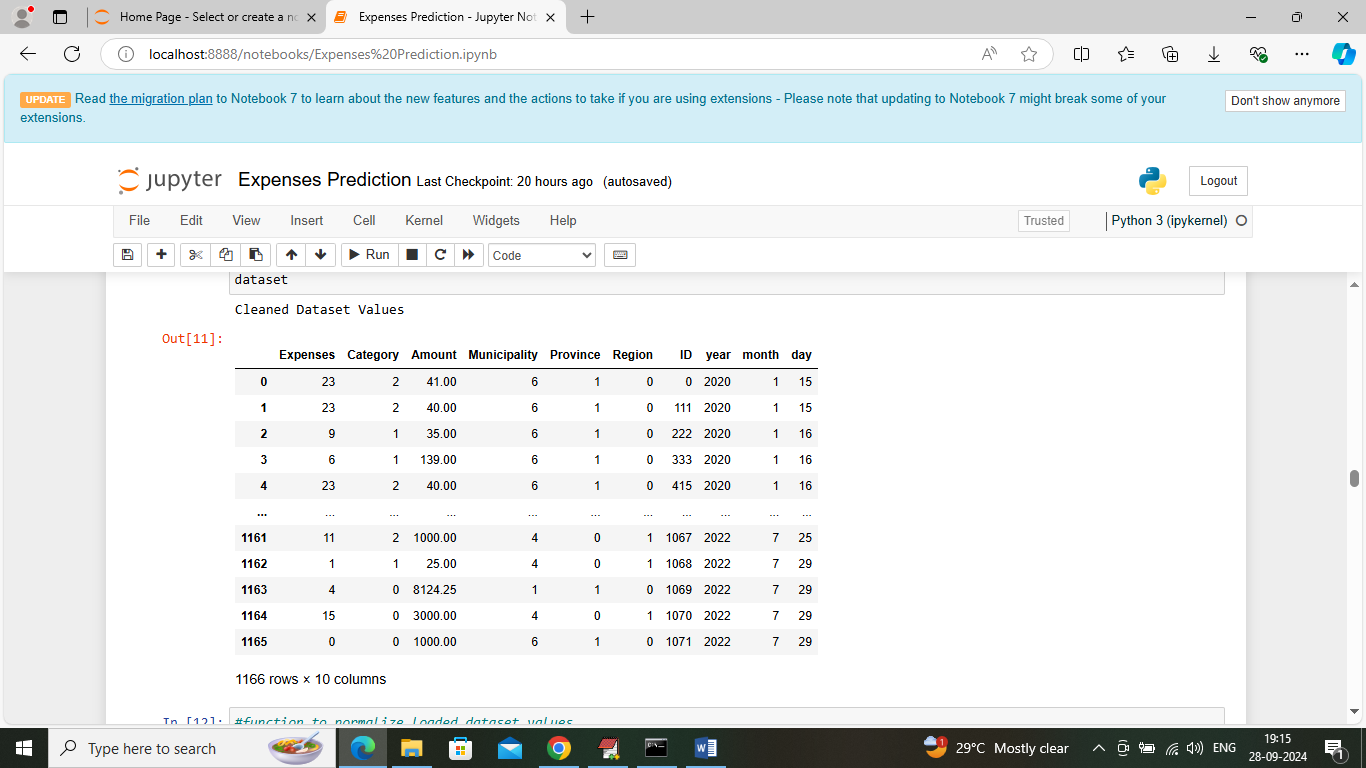
In above screen visualizing graph of amount spent on different categories where x-axis represents category name and y-axis represent amount



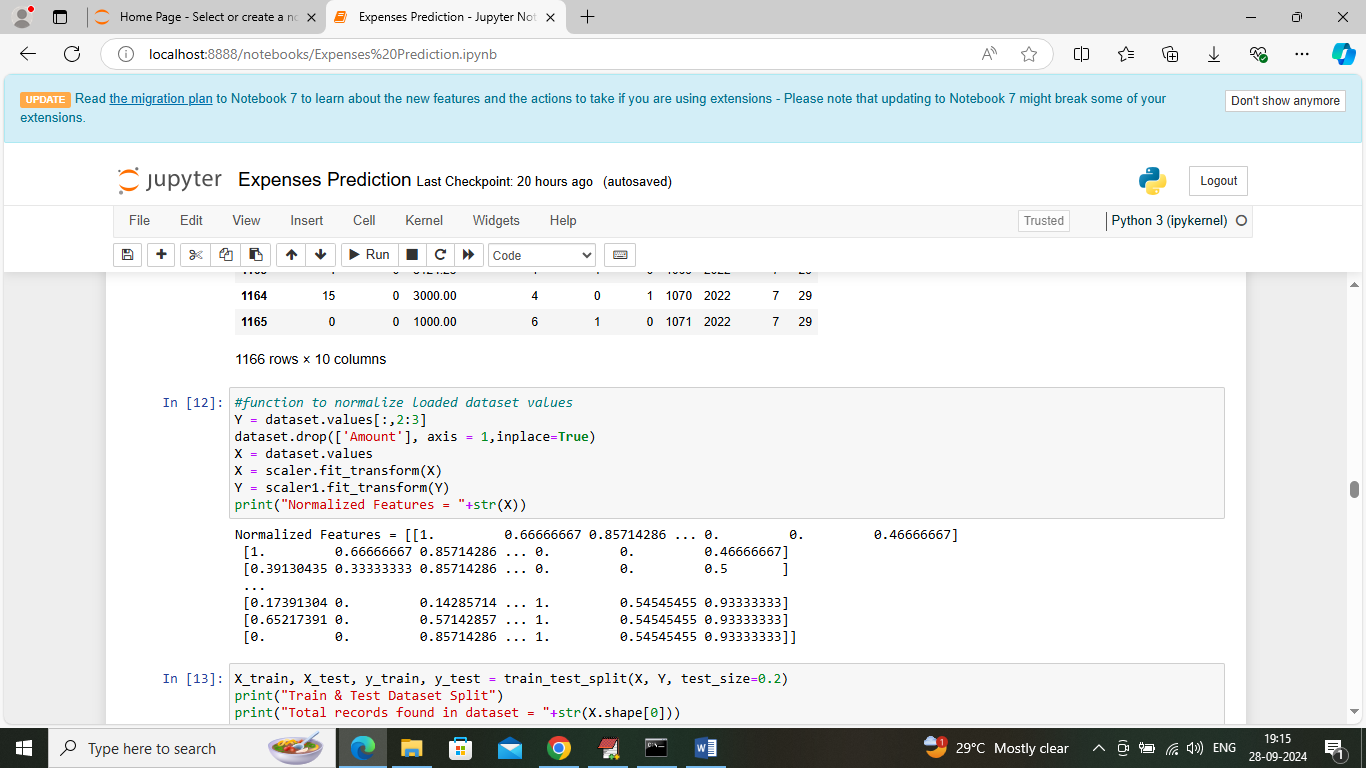
In above screen visualizing graph of total amount spent on different area or province



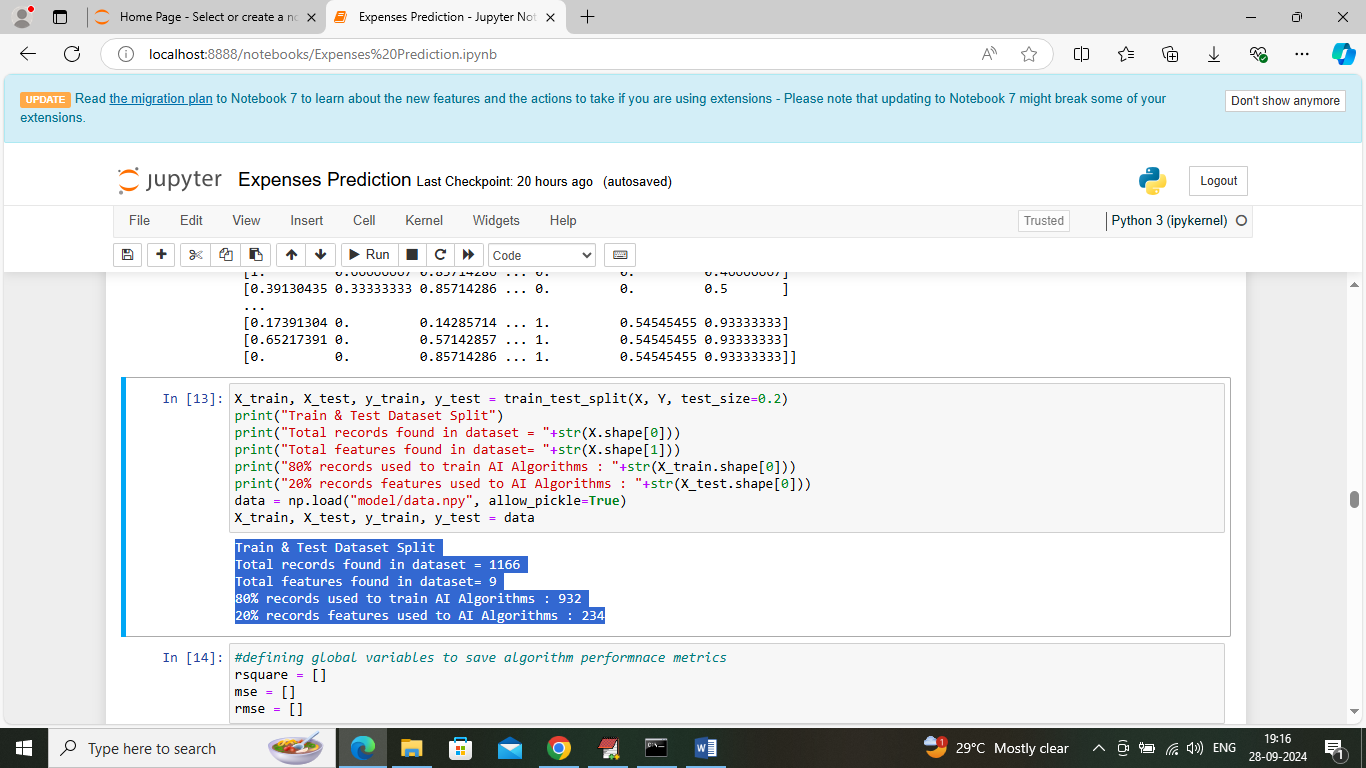
In above screen applying dataset label encoding to convert all non-numeric values to numeric values and after conversion will get below output



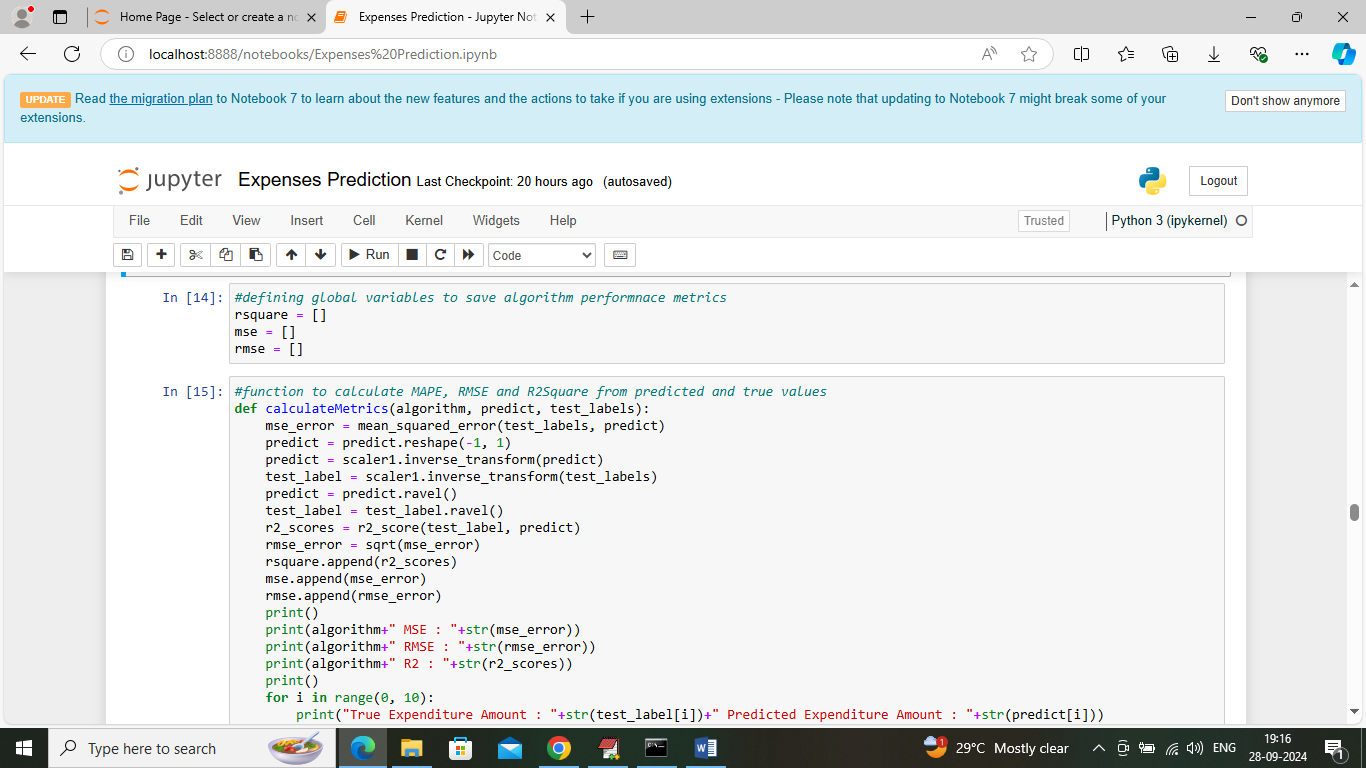
In above screen entire dataset columns converted to numeric format



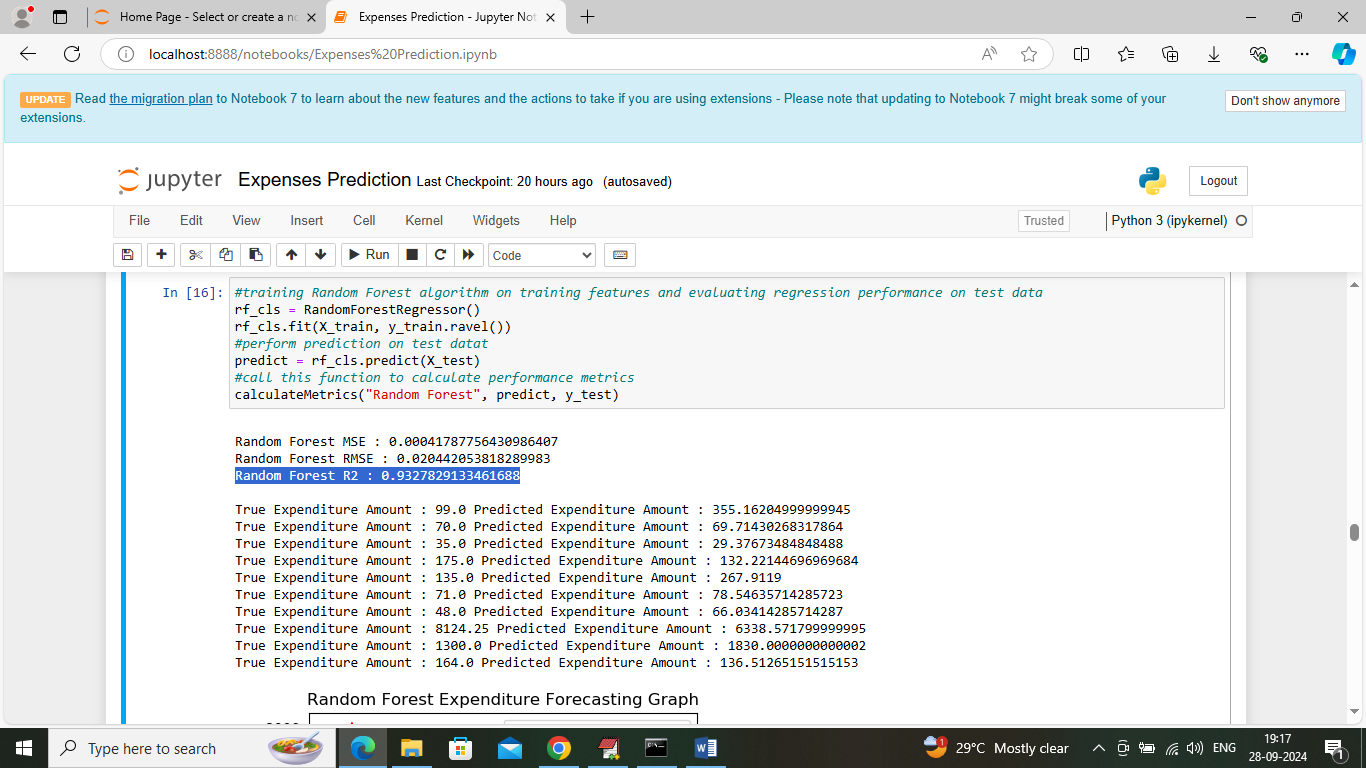
In above screen applying MINMAX algorithm to normalize dataset values



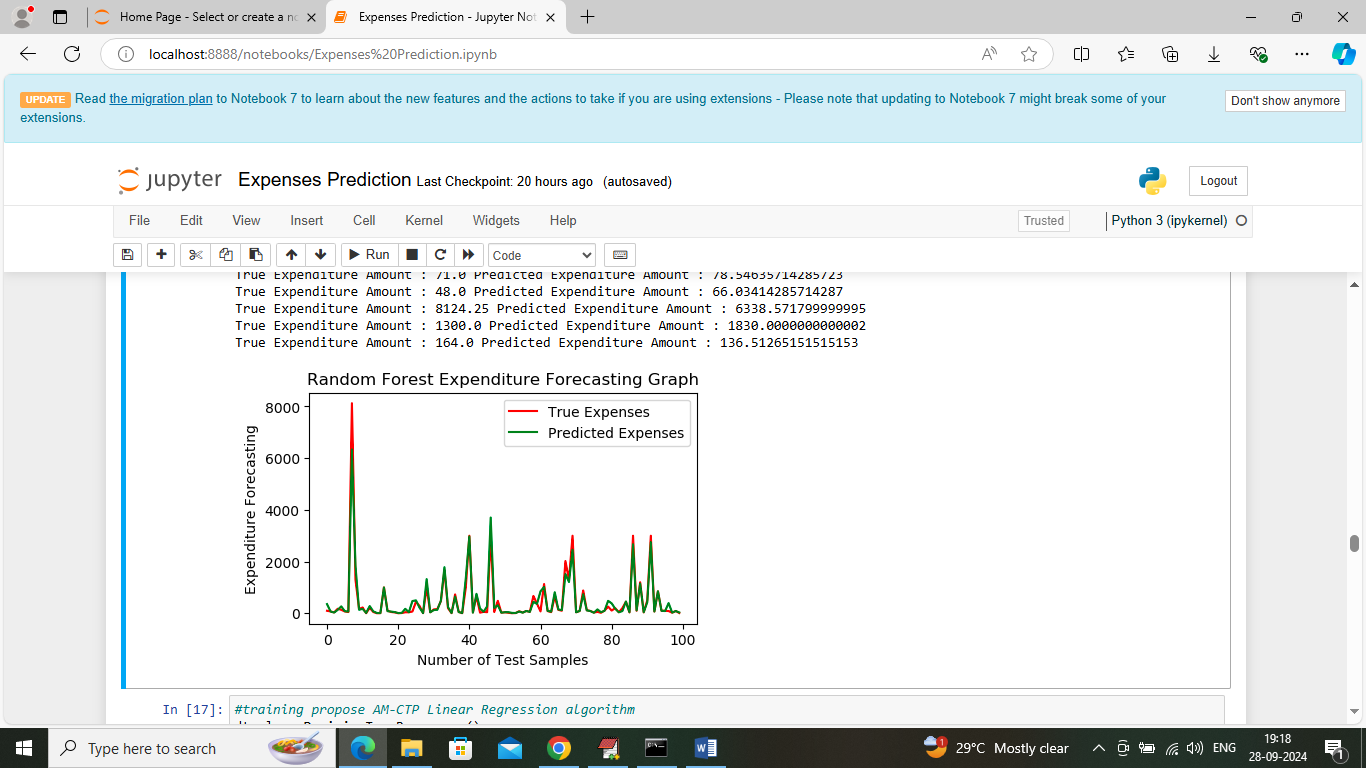
In above screen splitting entire dataset into train and test where application using 80% dataset for training and then apply trained model on 20% test data to calculate prediction R2score accuracy



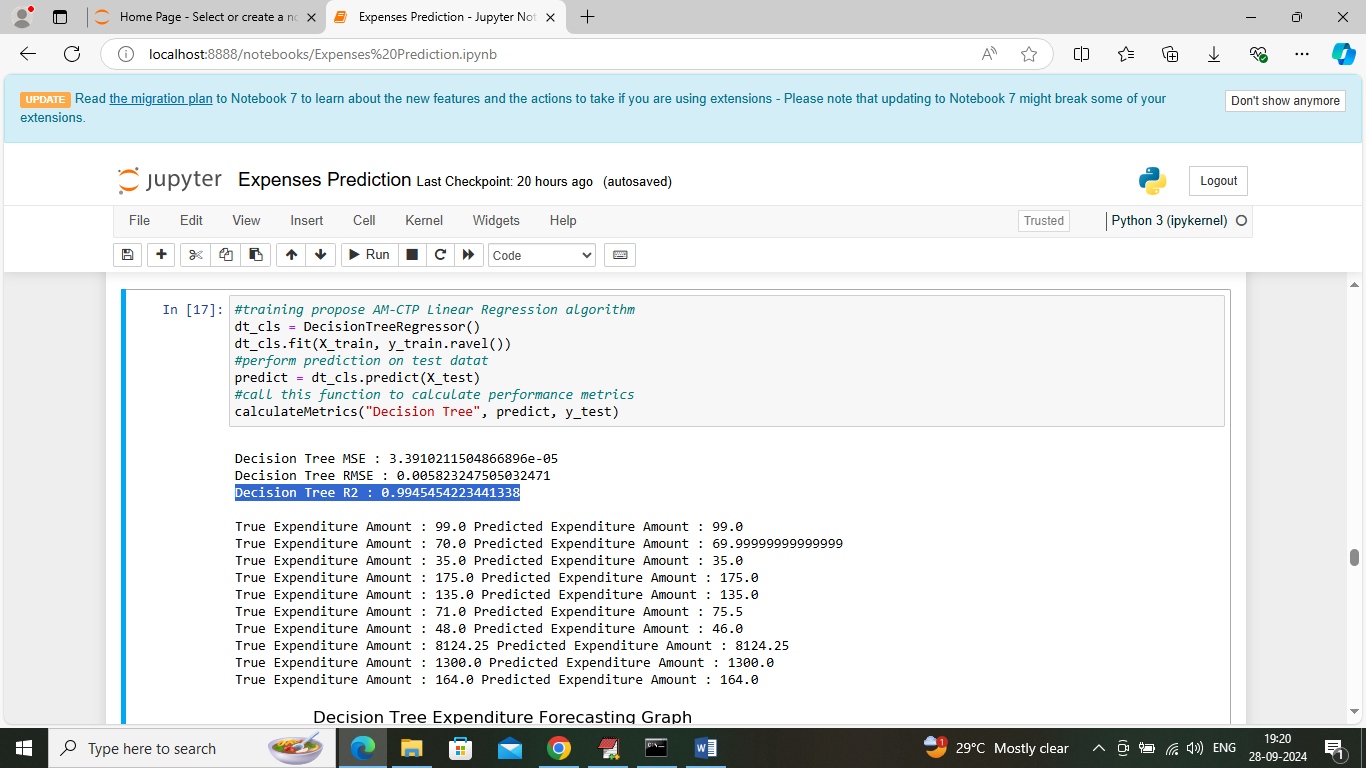
In above screen defining function to calculate RMSE, MSE and other metrics



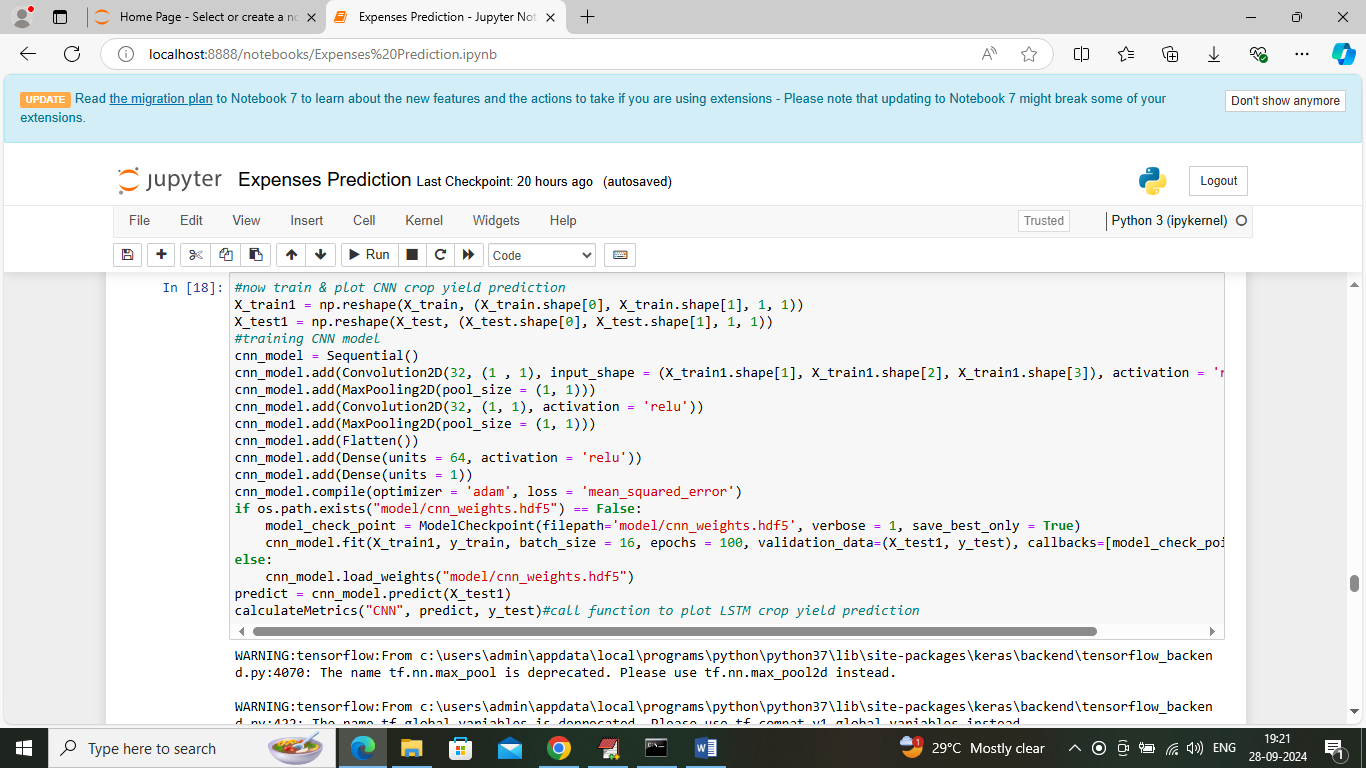
In above screen training random forest algorithm on training data and then applying 20% test data on trained model to perform prediction and then Random Forest got 93% R2score and can see RMSE and MSE error values. In next lines we can see Original and forecasted expenses from test data and below is the forecasting graph



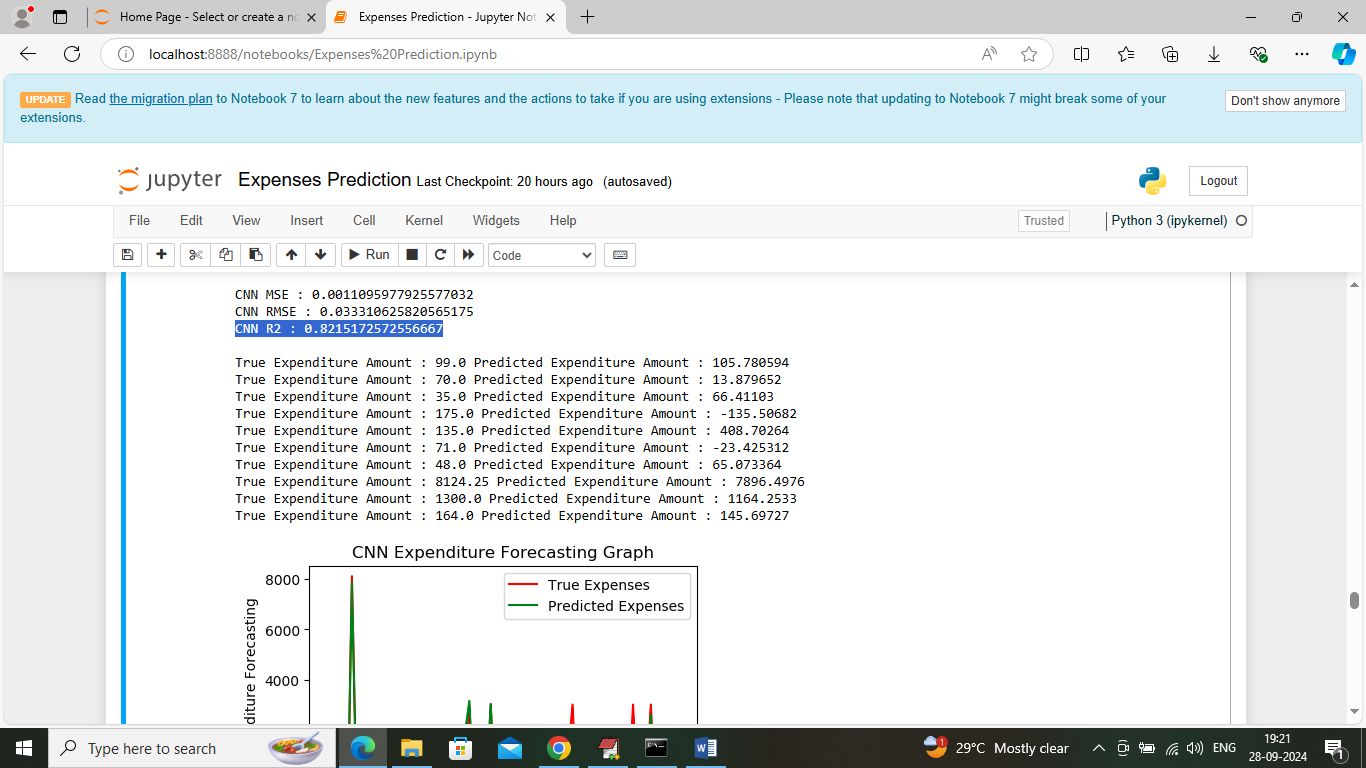
In above screen visualizing Random Forest forecasting graph where x-axis represents ‘Number of Test Data’ and y-axis represents ‘Expenses Amount’ and then red line represents true or original expenses and green line represents Forecasted expenses and can see both lines are fully overlapping with little gap si we can say Random Forest is accurate in forecasting expenses



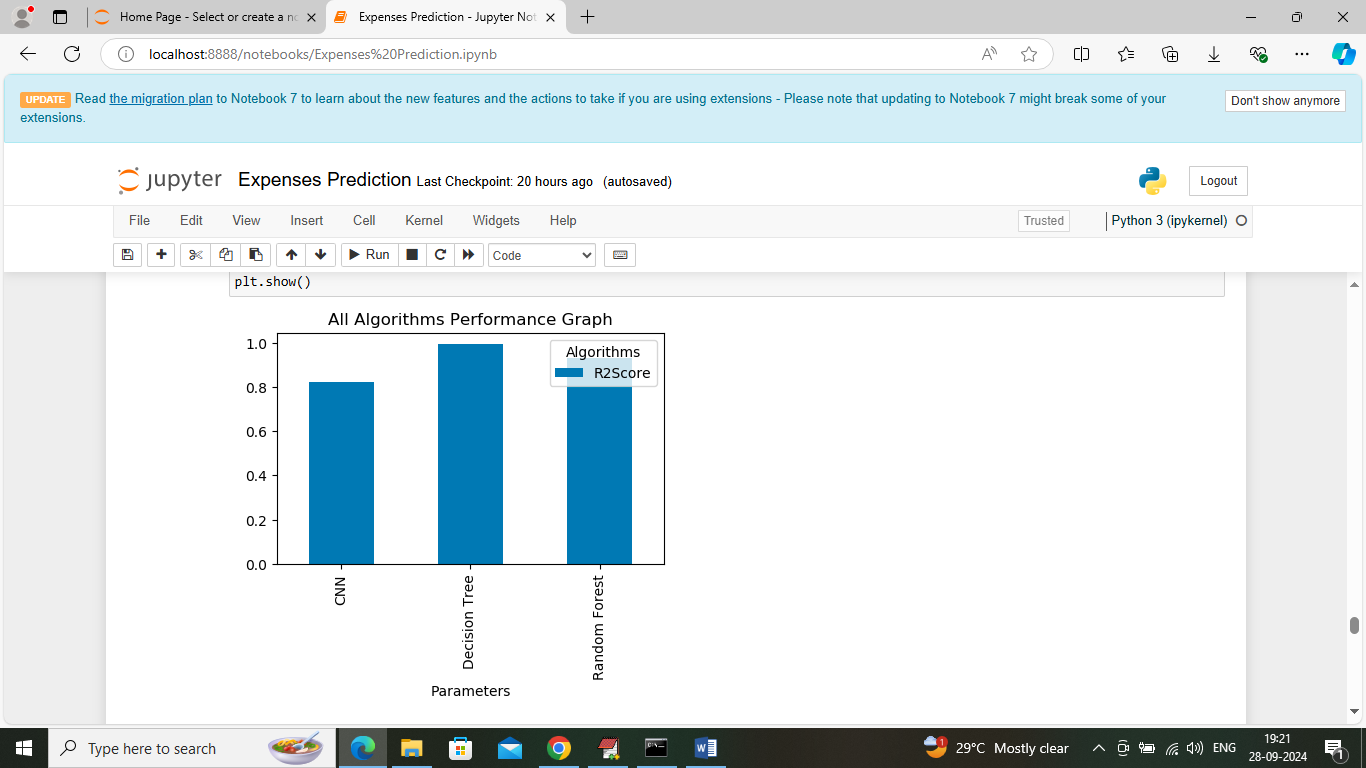
In above screen decision tree got 99% R2score and can see other metrics also along with forecasted graph



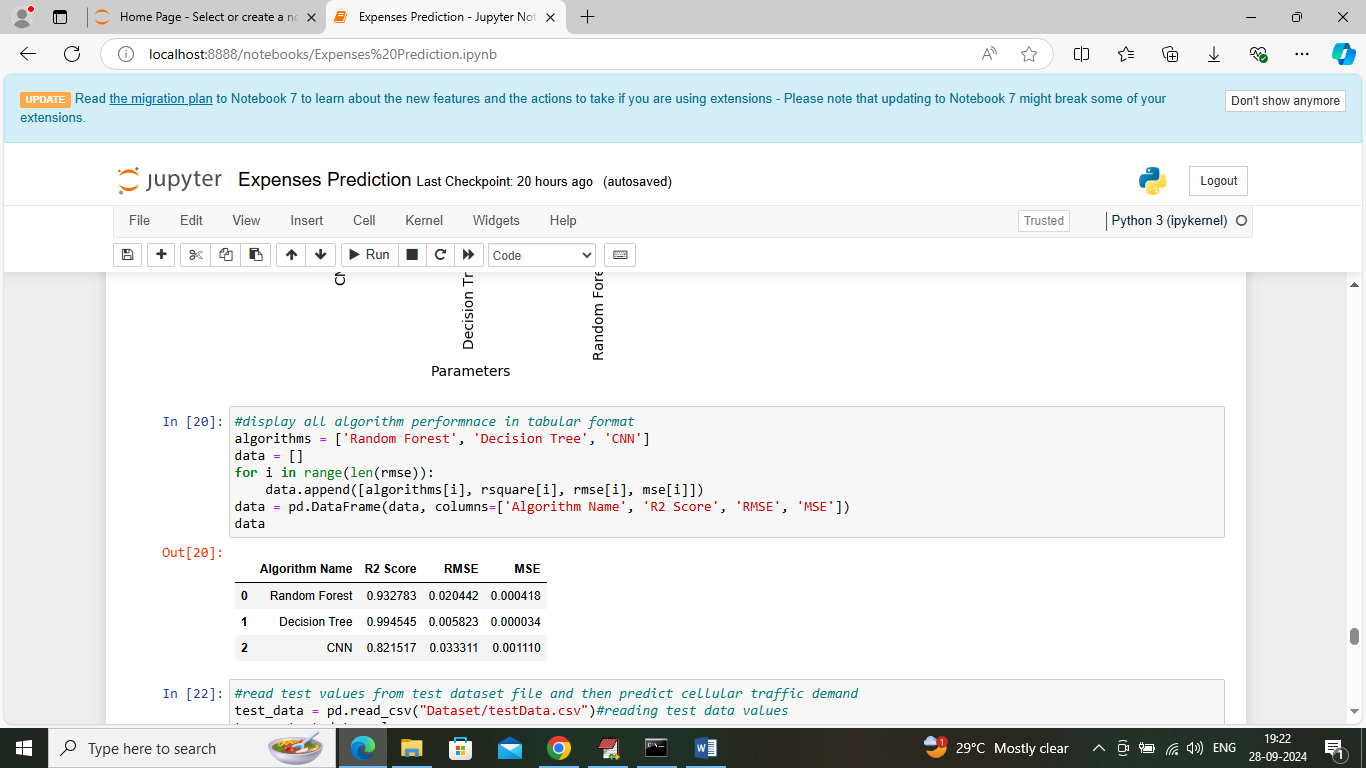
In above screen training CNN2D algorithm and after training will get below output



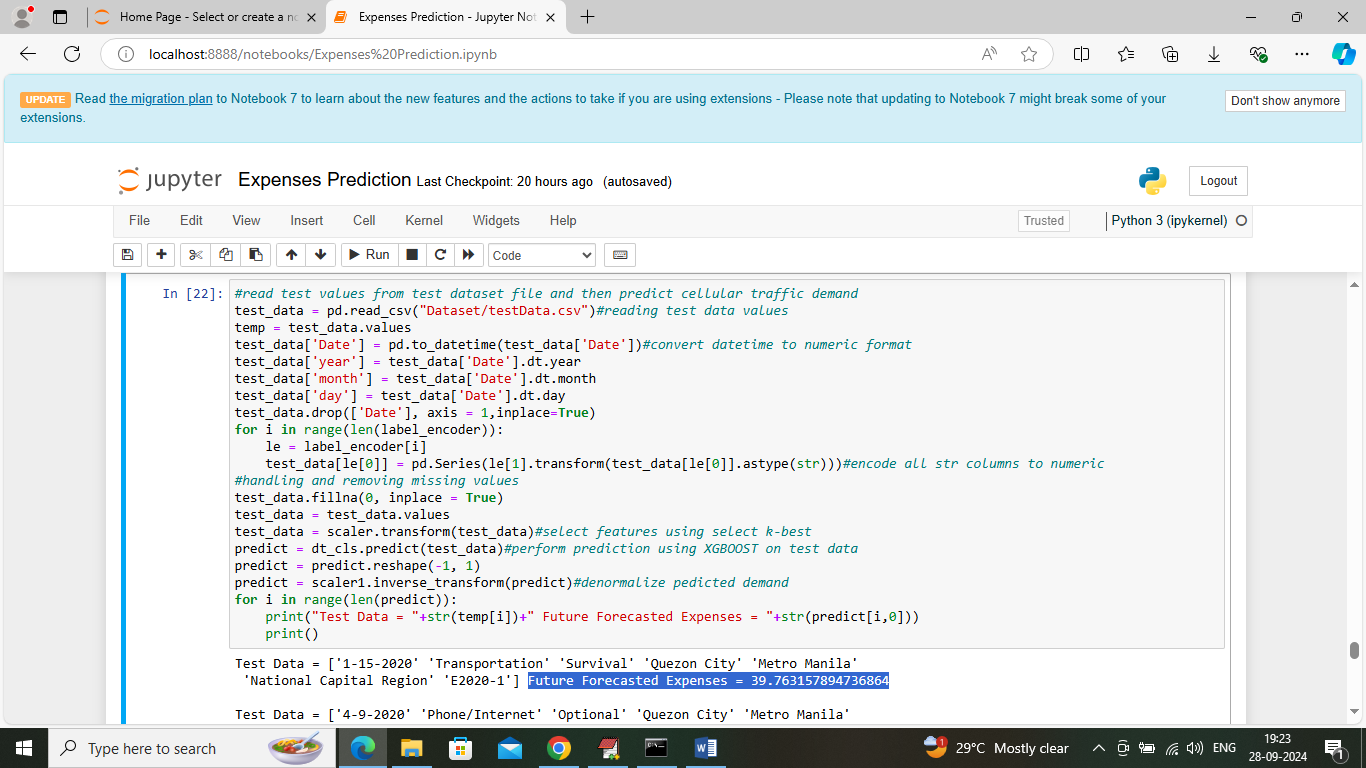
In above screen CNN2D got 82% R2score and can see other metrics also



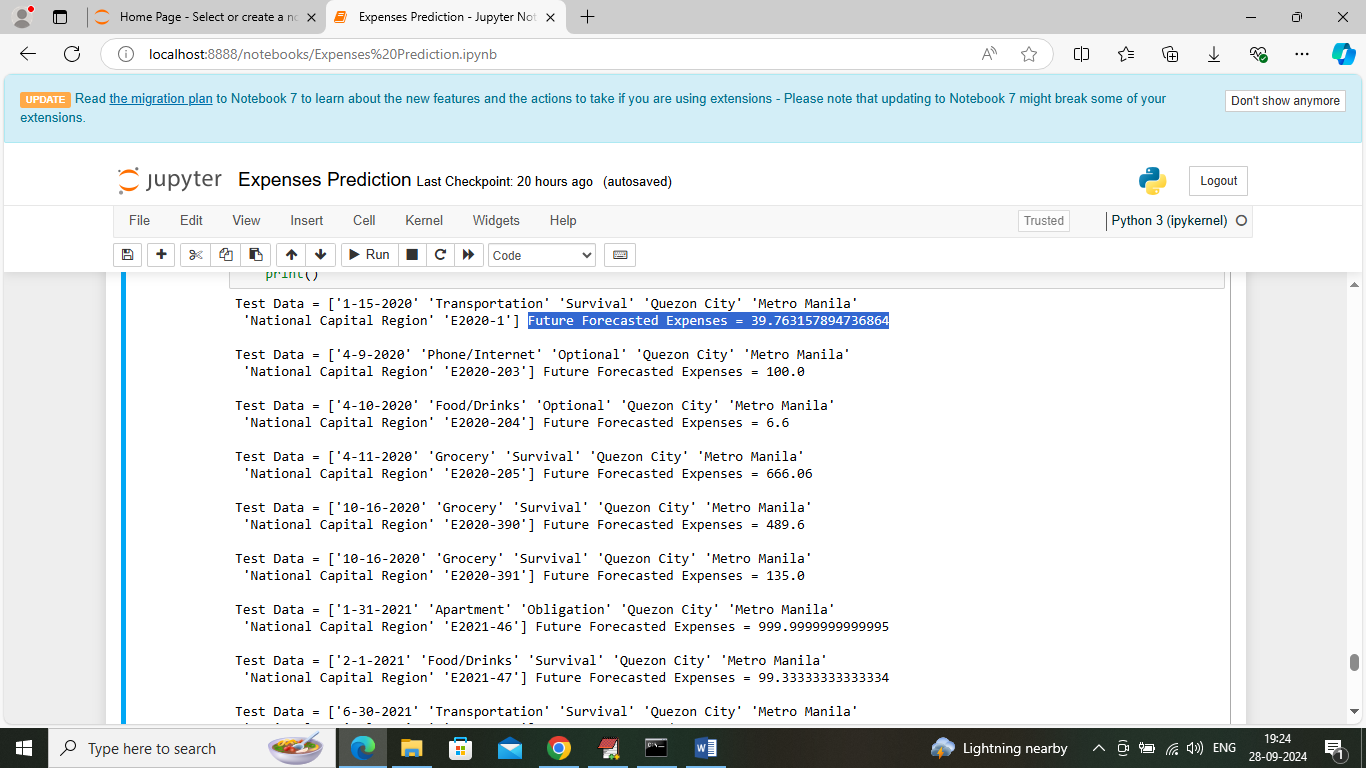
In above screen visualizing R2score graph of all algorithms where x-axis represents algorithm names and y-axis represents R2score



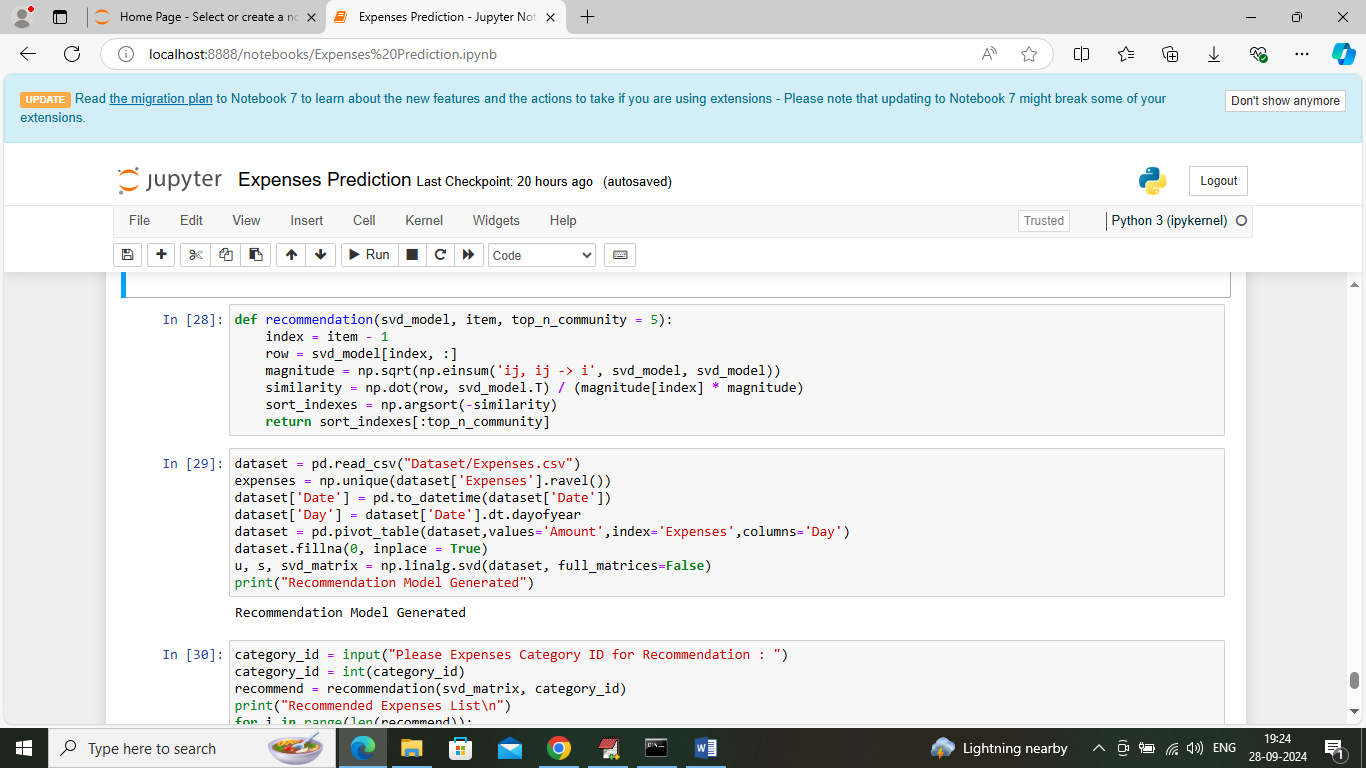
In above screen displaying all algorithm performance in tabular format



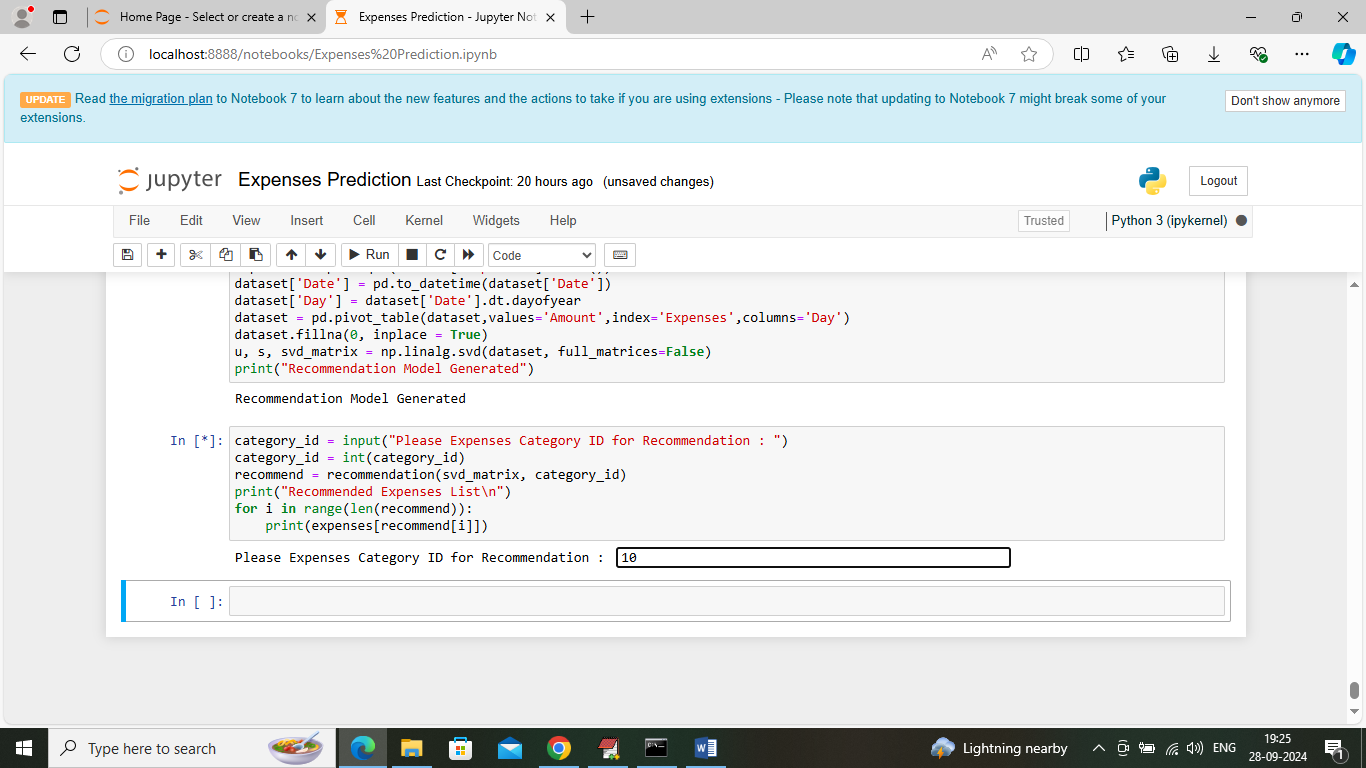
In above screen defining code which will read test data from ‘test file’ and then predict future expenses and then in output can see test data values in square bracket and blue colour text displaying predicted expenses amount



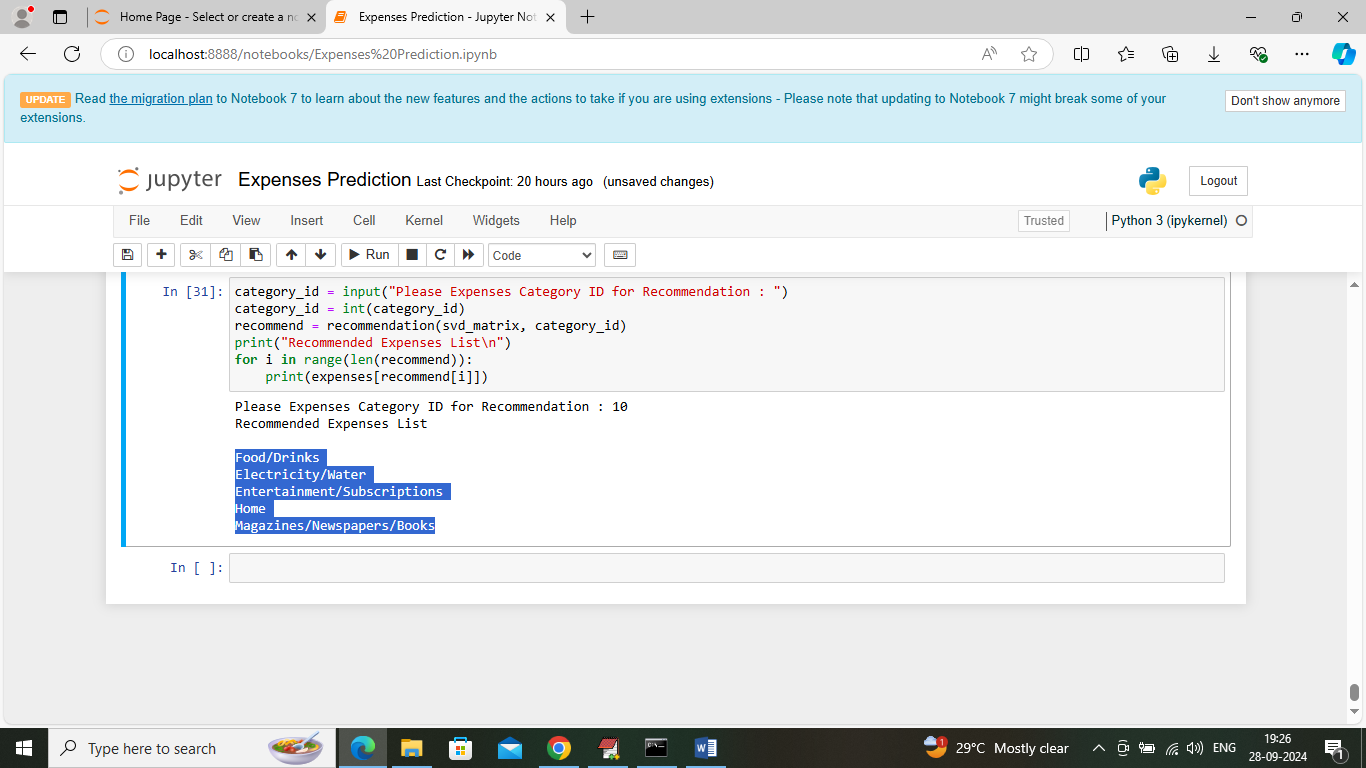
In above screen predicting expenses for different category expenses



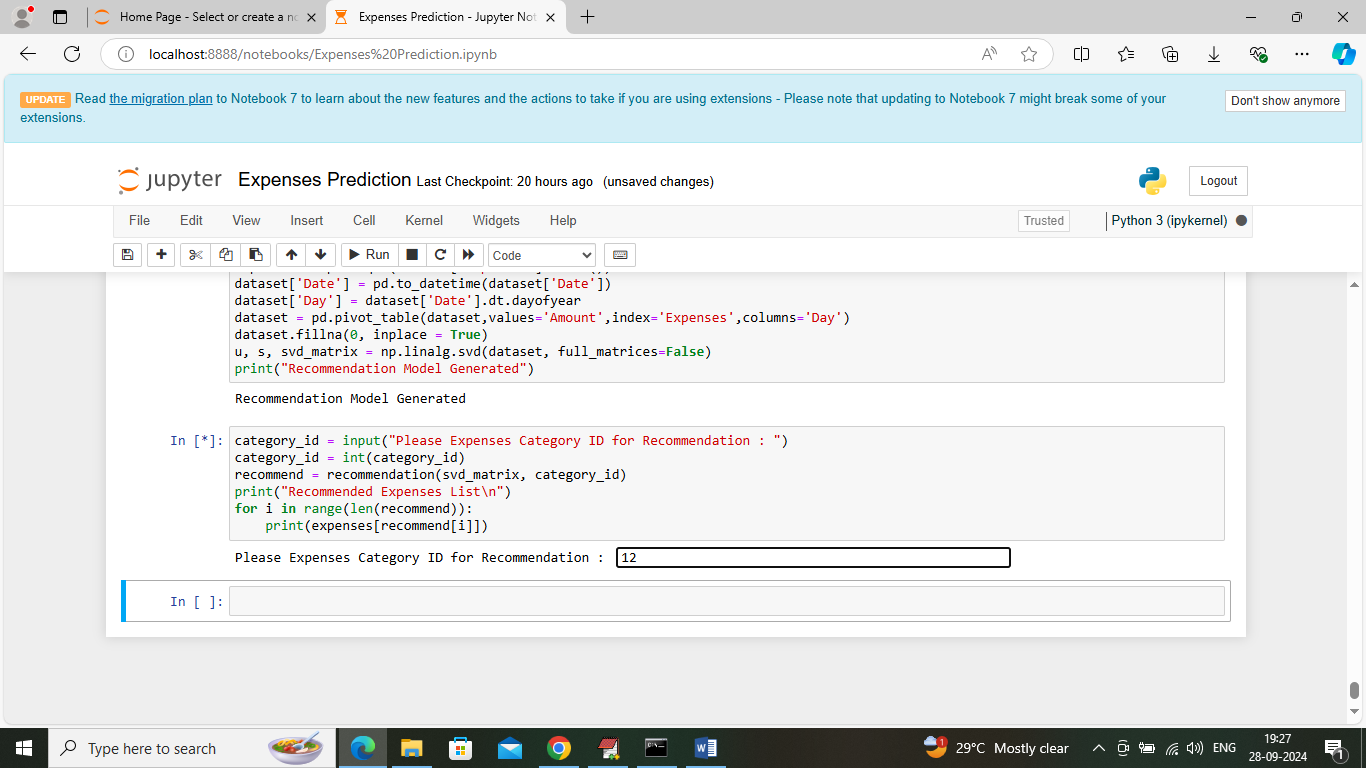
In above screen defining recommendation code and after executing this block then application will ask to input category expenses ID and then system will recommend new category expenses which you can done in future



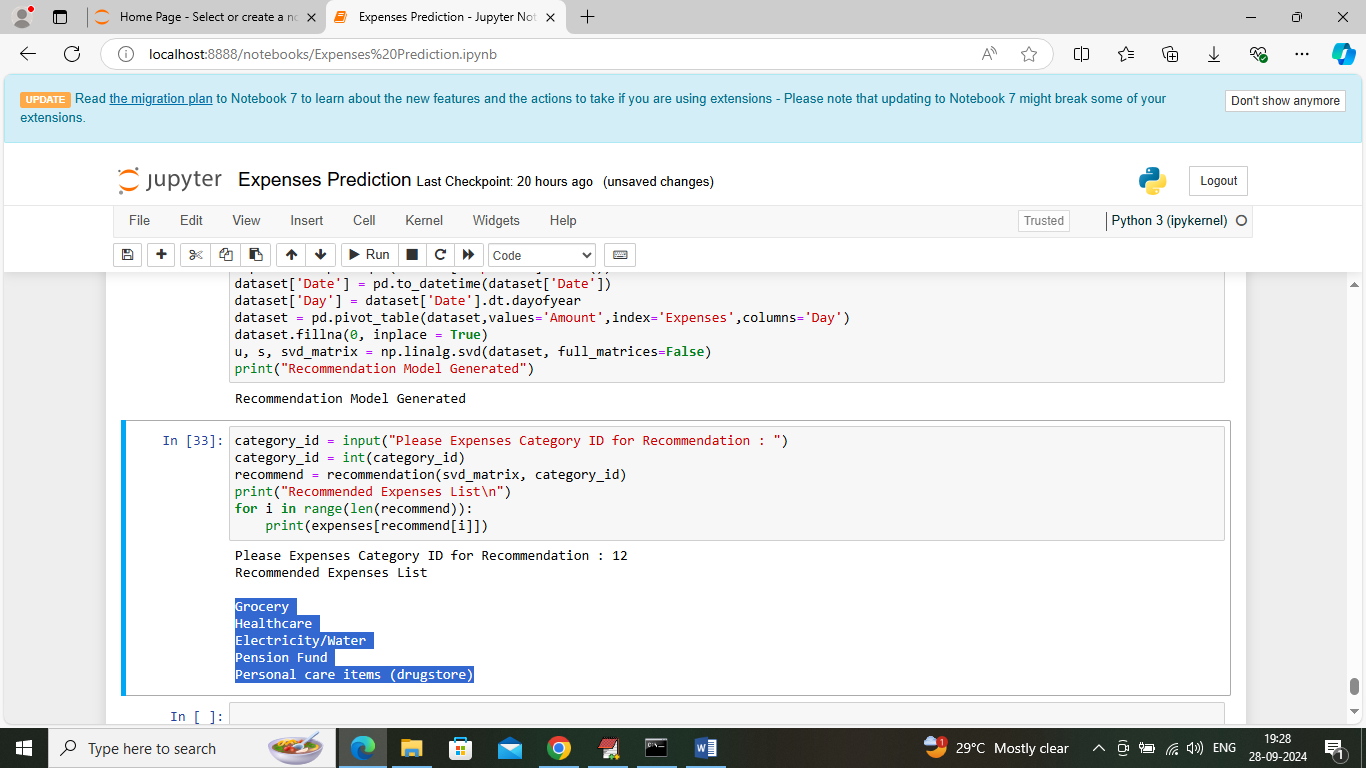
In above screen I entered category ID as 10 and then press enter key to get below output



In above screen can see top 5 category names as recommendation and in below screen I am giving another category id which you can enter just by running block again



In above screen I am giving category id as 12 and then press button to get below recommendation



In above screen can see top 5 recommendation names for given 12 category ID. You can category id from 1 to 24.