Explainable Artificial Intelligence for Patient Safety: A Review of Application in Pharmacovigilance

Machine and deep learning algorithms are using everywhere to predict future values or to classify objects but still some organizations are hesitating to employee this models for their business growth prediction. In such organization Medical Health is the one which is hesitating to employ machine learning models to predict drug side effects or their effectiveness. To avoid hesitance ML organizations has introduce another tool called XAI (explainable AI) which can explain about the model prediction. XAI will explain about features which contribute most for AI or ML models in accurate prediction. Hence XAI is consider as the black box for ML or Deep learning models. XAI will give clear picture about the model prediction to those who has hesitation.

In propose paper author is giving reviews by collecting all medical papers and its study data and then employing different ML and DL algorithms like Tree Model, Neural Network Model and Graph Model.

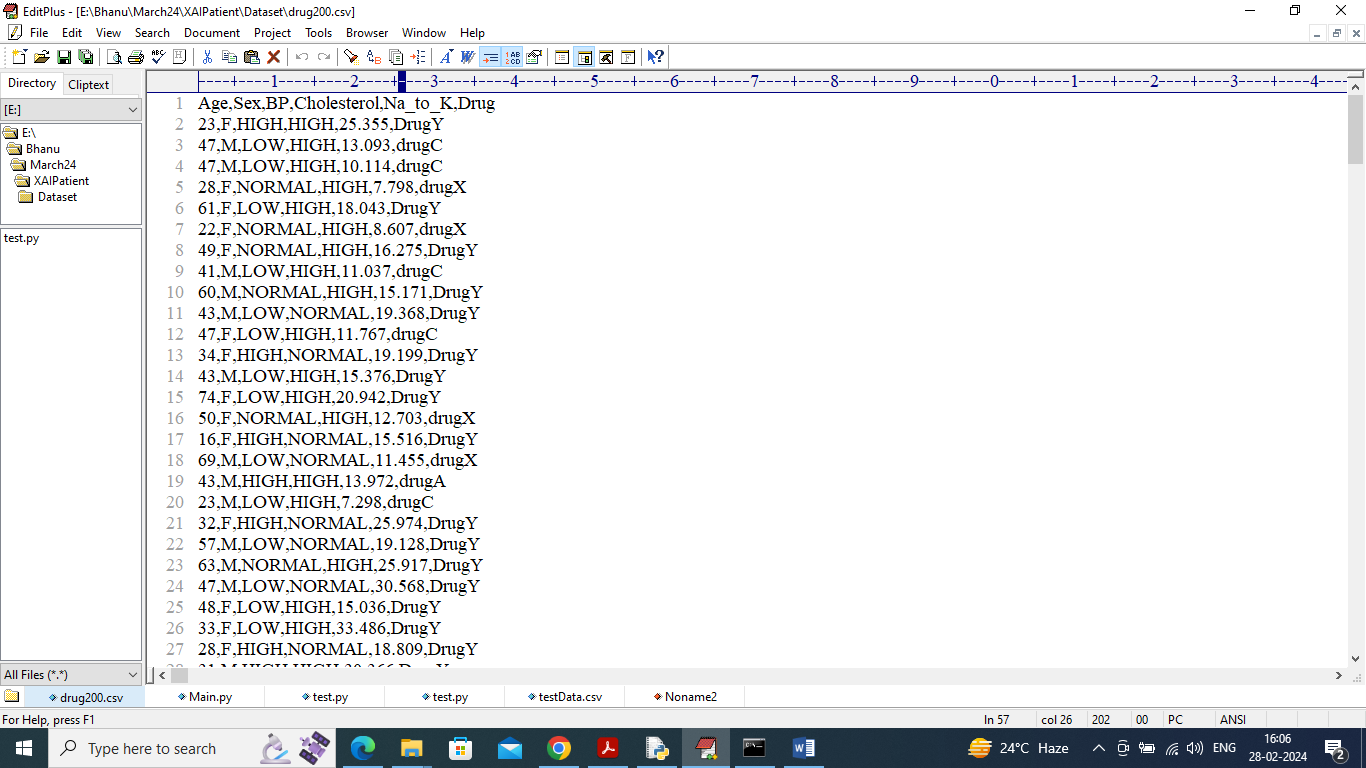
Prediction of above 3 models can be explain with SHAP tool which can be used to explain models. In this paper author has listed out many datasets which can be used to train and test above algorithms. We have chosen ‘Drug Classification’ dataset given in paper from KAGGLE repository.

Drug Classification dataset can be used to train model to predict drug for the patients based on his health conditions like BP, Cholesterol, Gender, Age etc.

**Propose and Extension Algorithms**

In propose paper as Tree model author used XGBOOT algorithm and then as Graph Model author used GRAPHCNN and as Neural Network author used MLP and as extension we have used CNN2D algorithm which will optimized or filtered features using 2 dimension neurons. Optimized features help in getting enhance accuracy.

In below screen showing dataset details



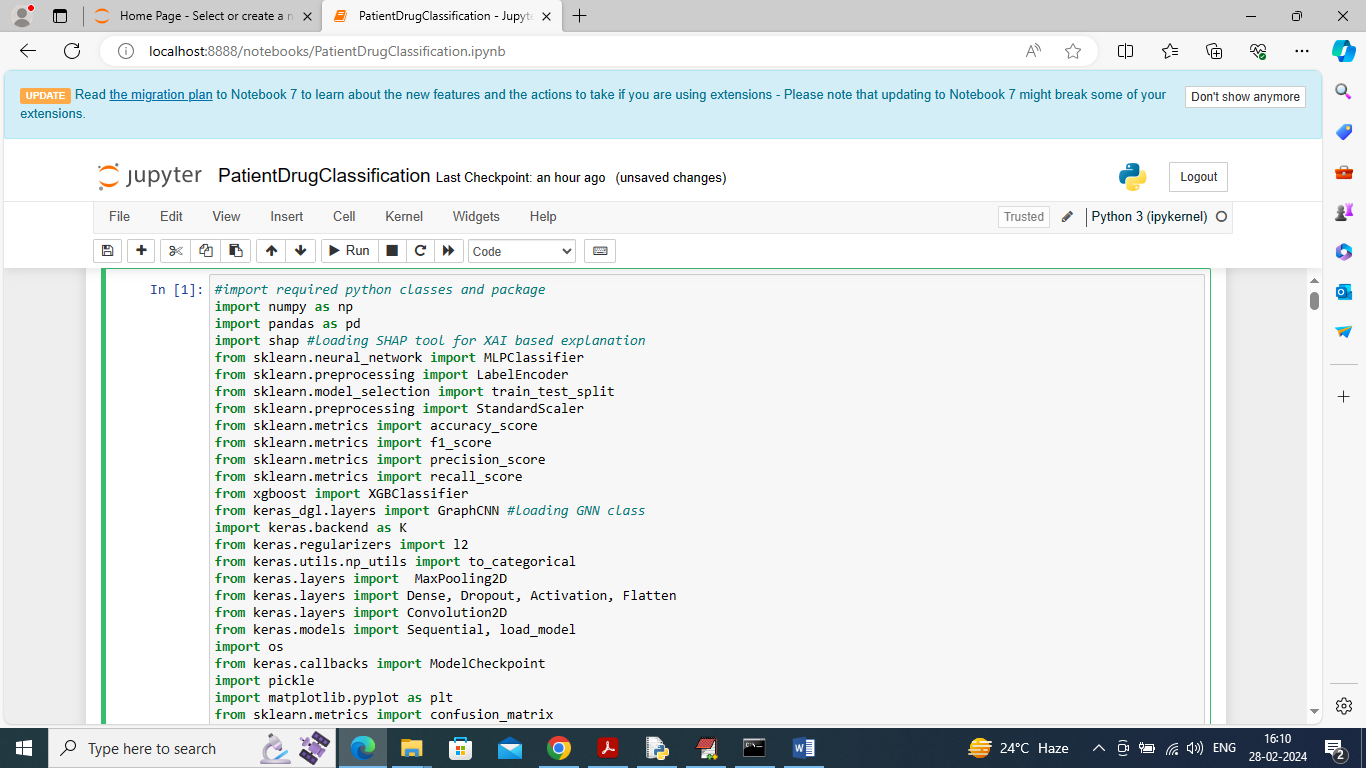
In above dataset screen first row contains dataset column names and remaining rows contains dataset values and in last column we can see class label as drug type based on patient condition. So by using above dataset we will train, test and explain above algorithms.

Explanation XAI graph

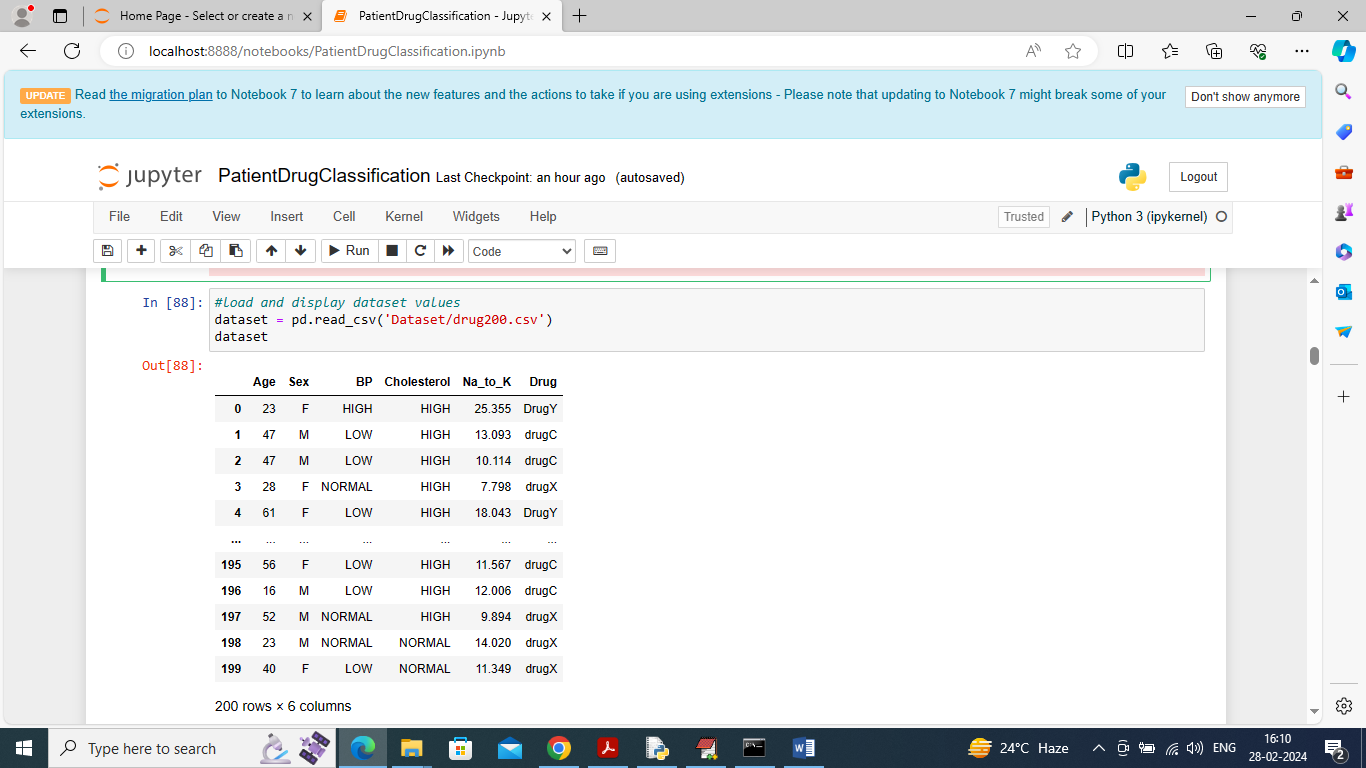
To explain model we have utilized Force Plot, Summary Plot, Violin Plot and Decision Plot. All this plot explain about all features which contribute most to the model for correct prediction.

Here we need to plot multiple graphs using XAI whose graph will plot only in JUPYTER notebook not in web or TKINTER so we have developed this code using JUPYTER notebook.

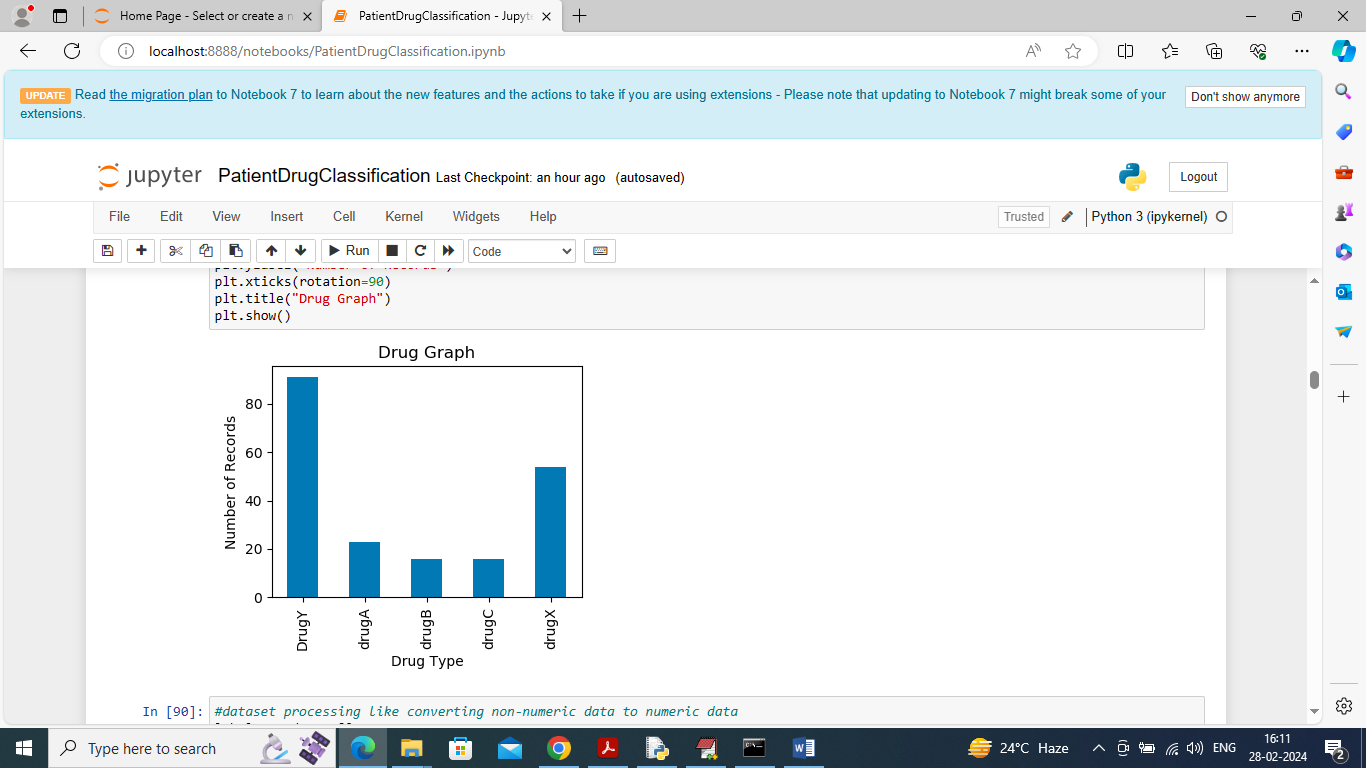
Below are the code and output screens with blue colour comments



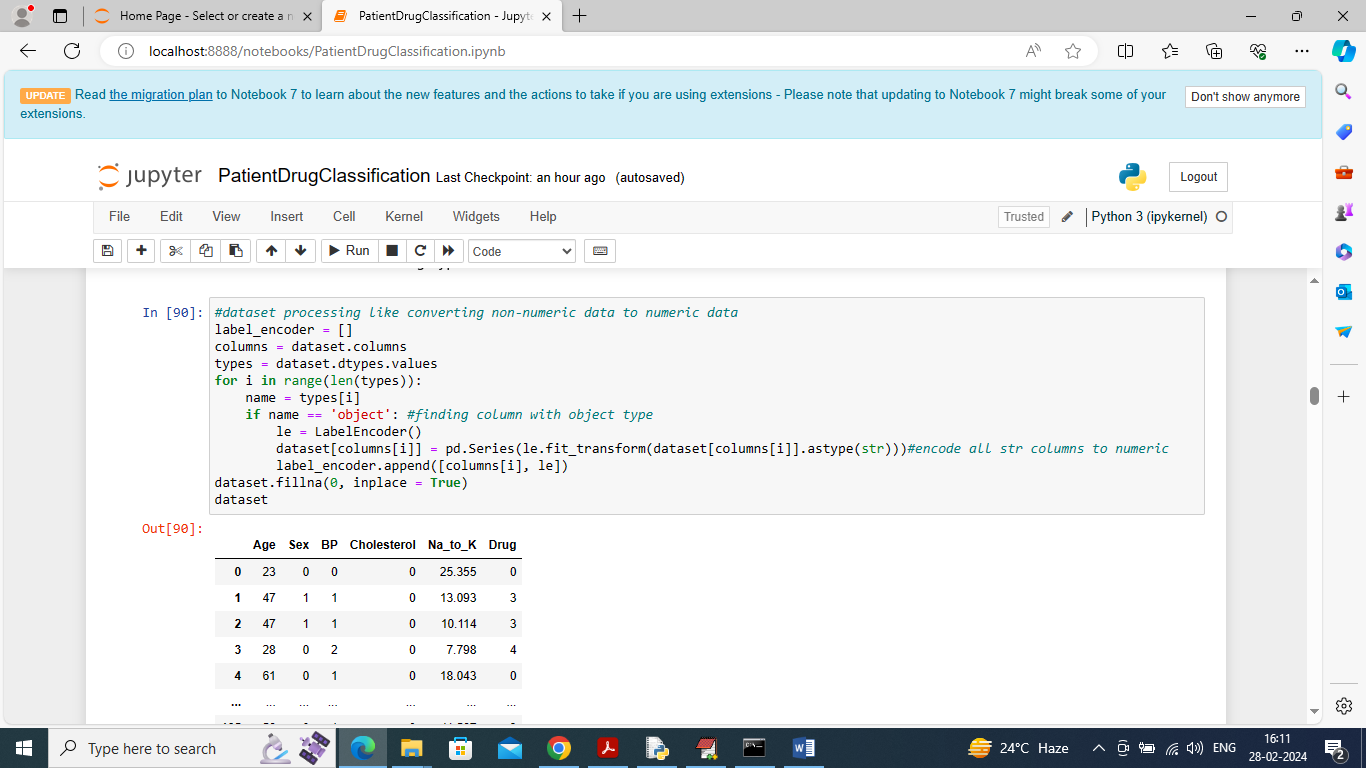
In above screen importing required python classes and packages



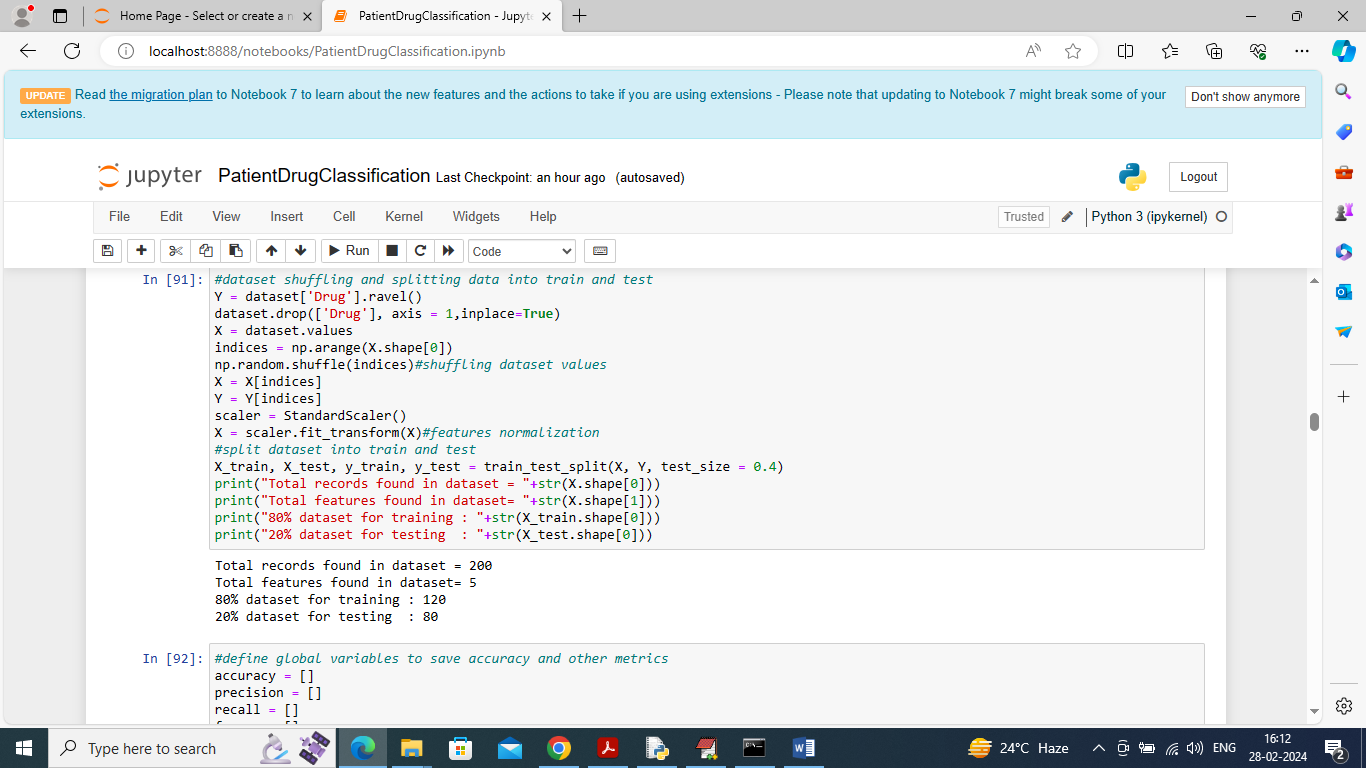
In above screen loading and displaying dataset values



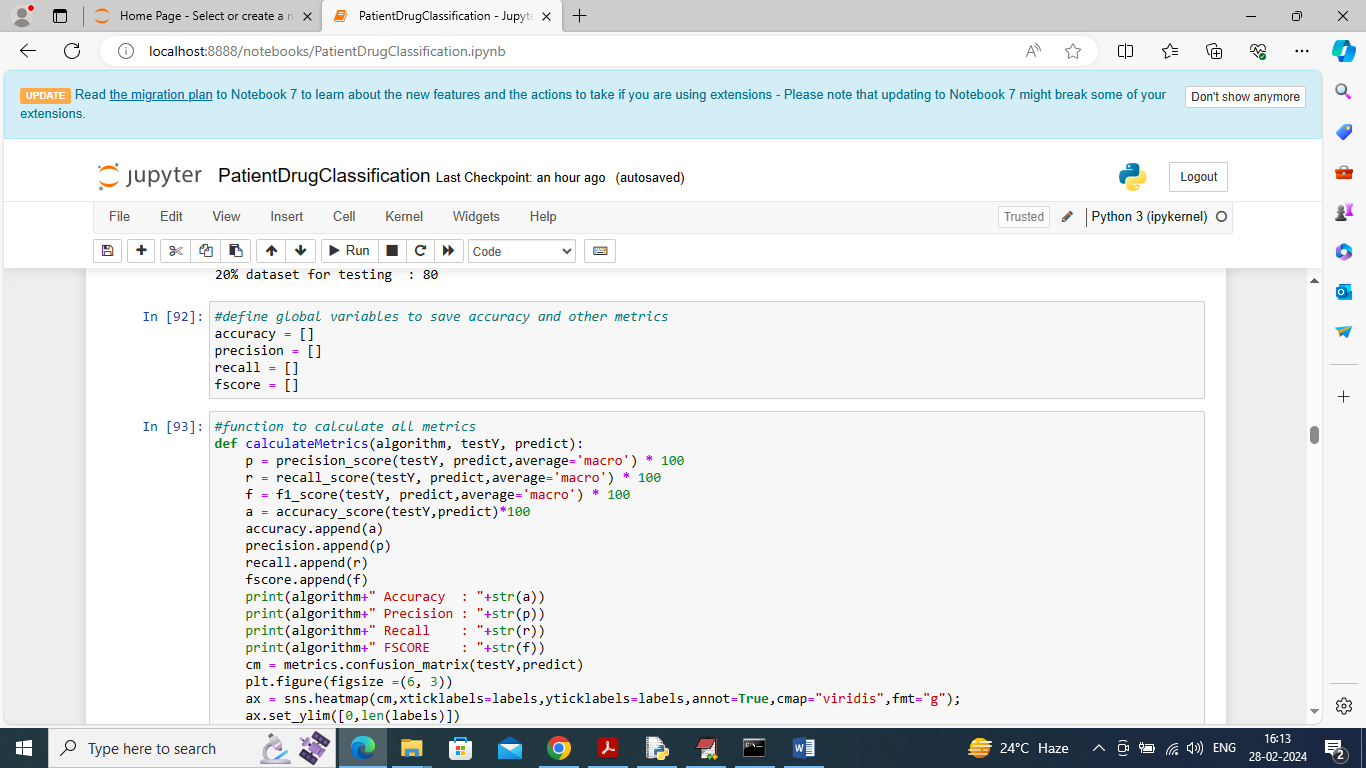
In above screen displaying types of Drugs and Number of records found in dataset as graph format



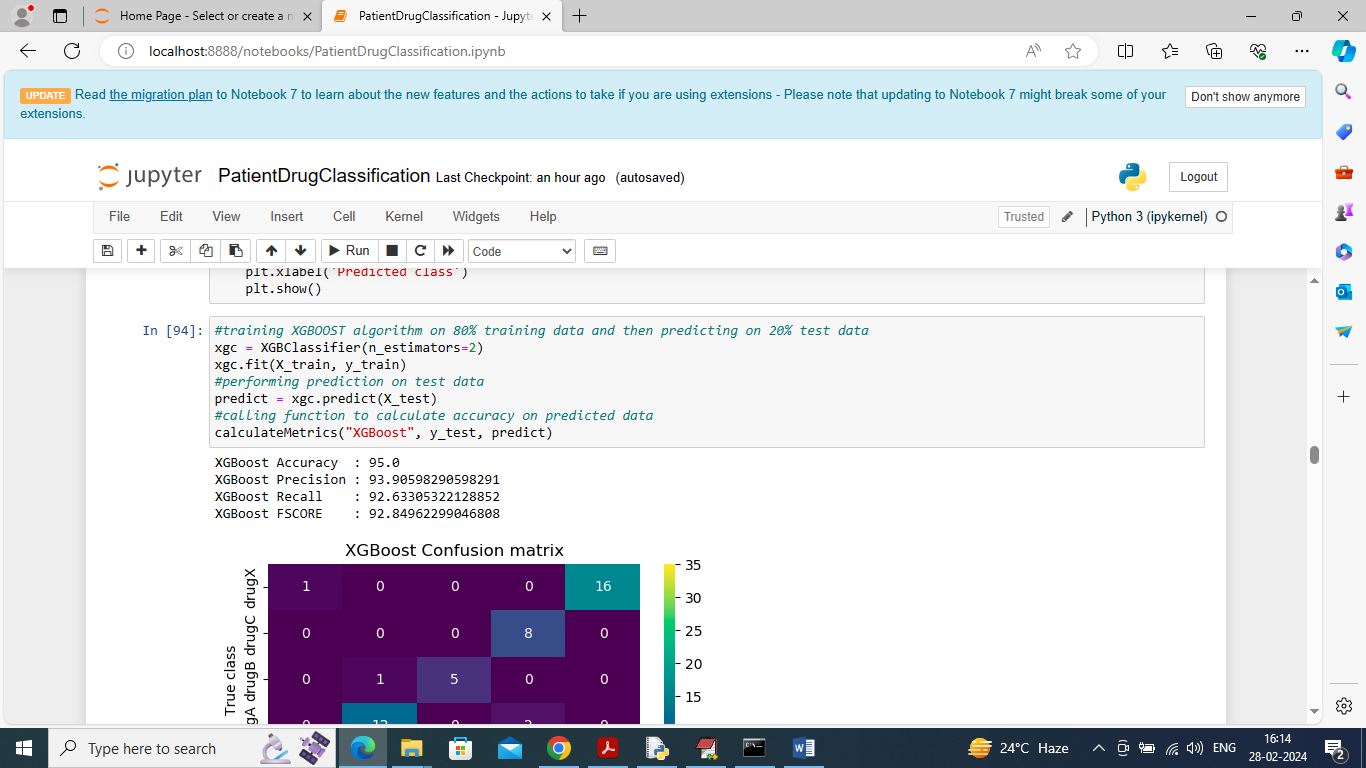
Dataset contains both numeric and non-numeric values so in ab0ve screen employing label encoder class to convert non-numeric data to numeric data and after conversion can see all values are in numeric format



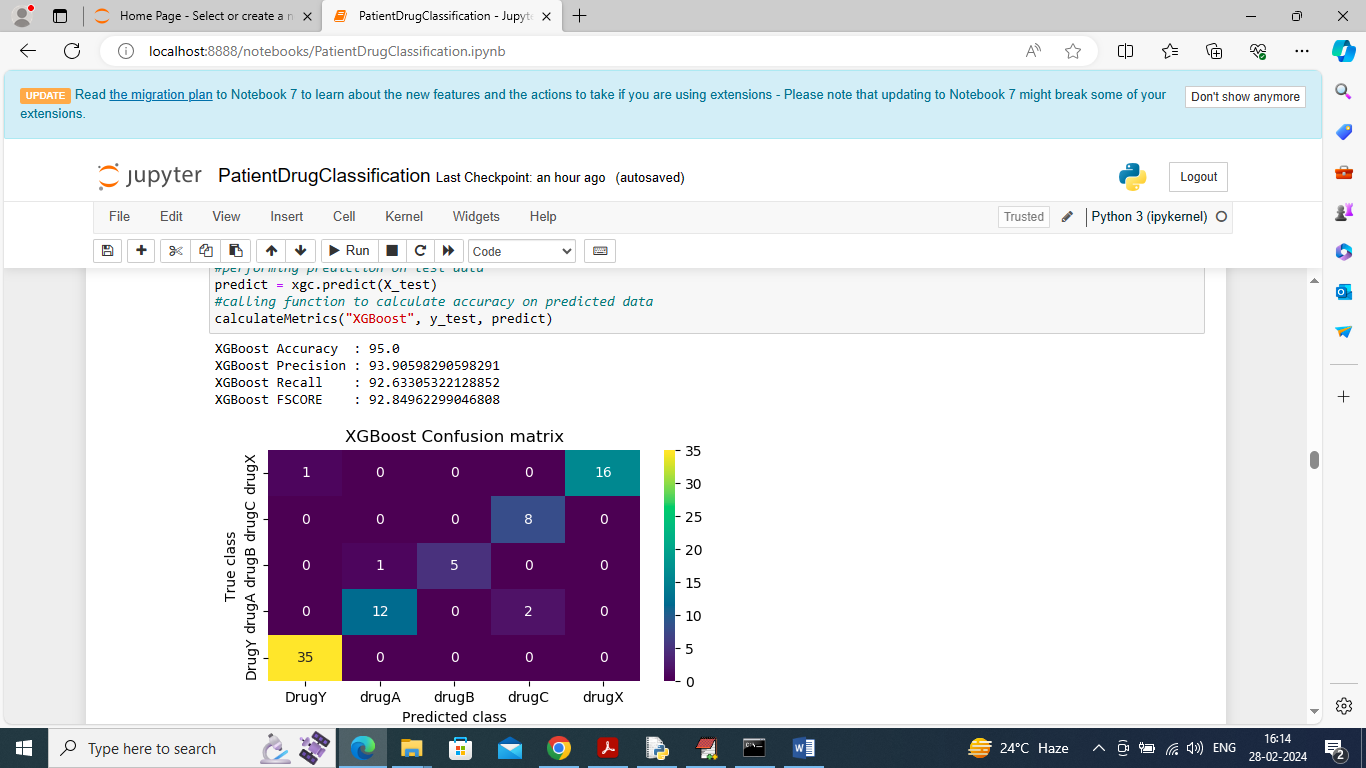
In above screen applying processing like data normalization, shuffling and splitting into train and test part with a ratio of 80:20. 80% for training and 20% for testing



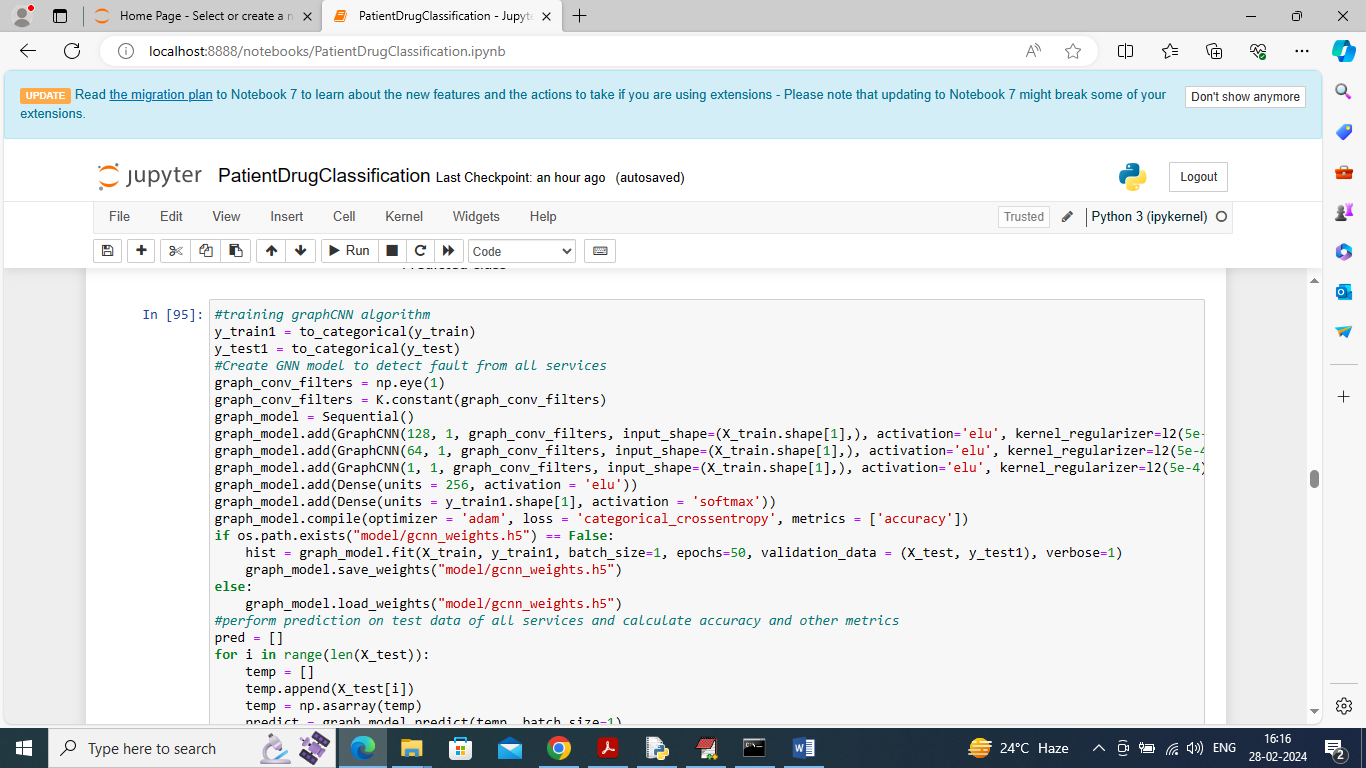
In above screen defining function to calculate accuracy and other metrics



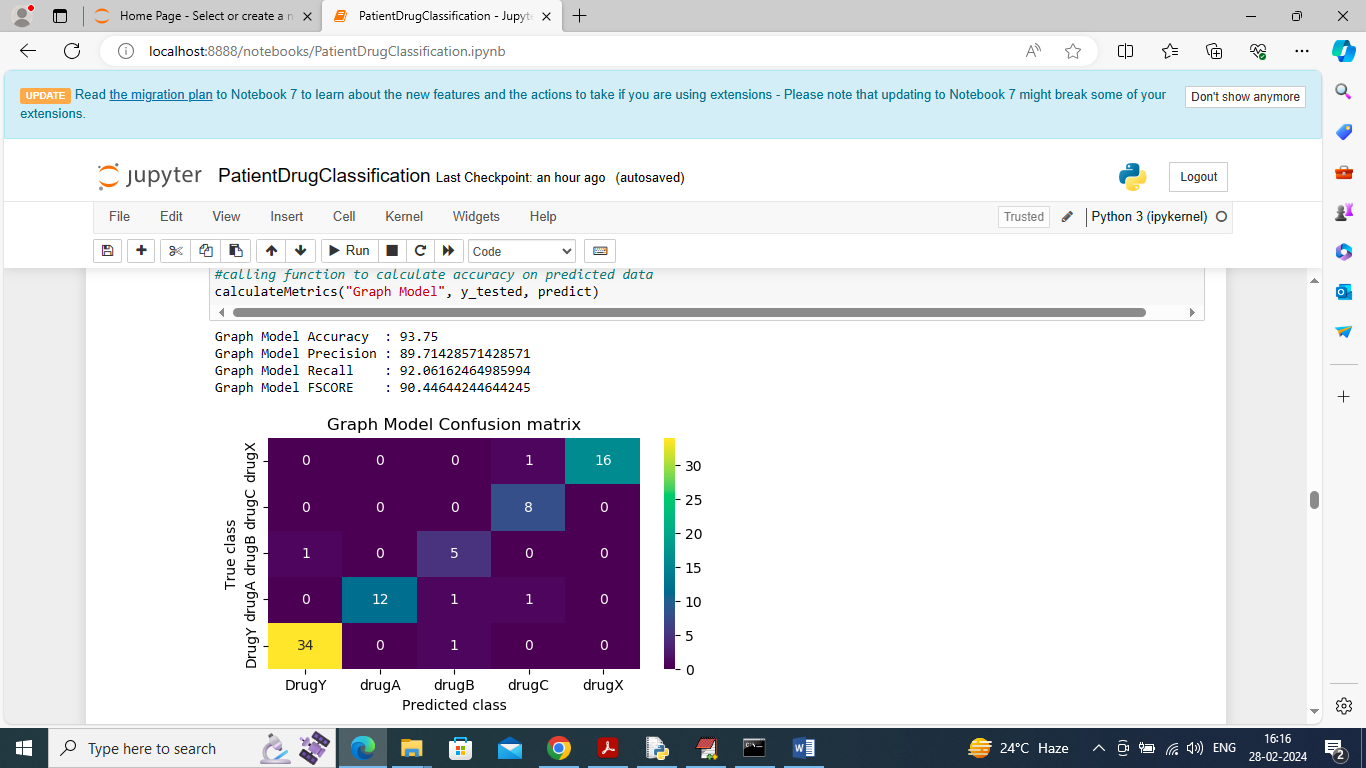
In above screen training XGBOOST algorithm as tree model and it got 95% accuracy and can see other metrics like precision, recall and FSCORE.



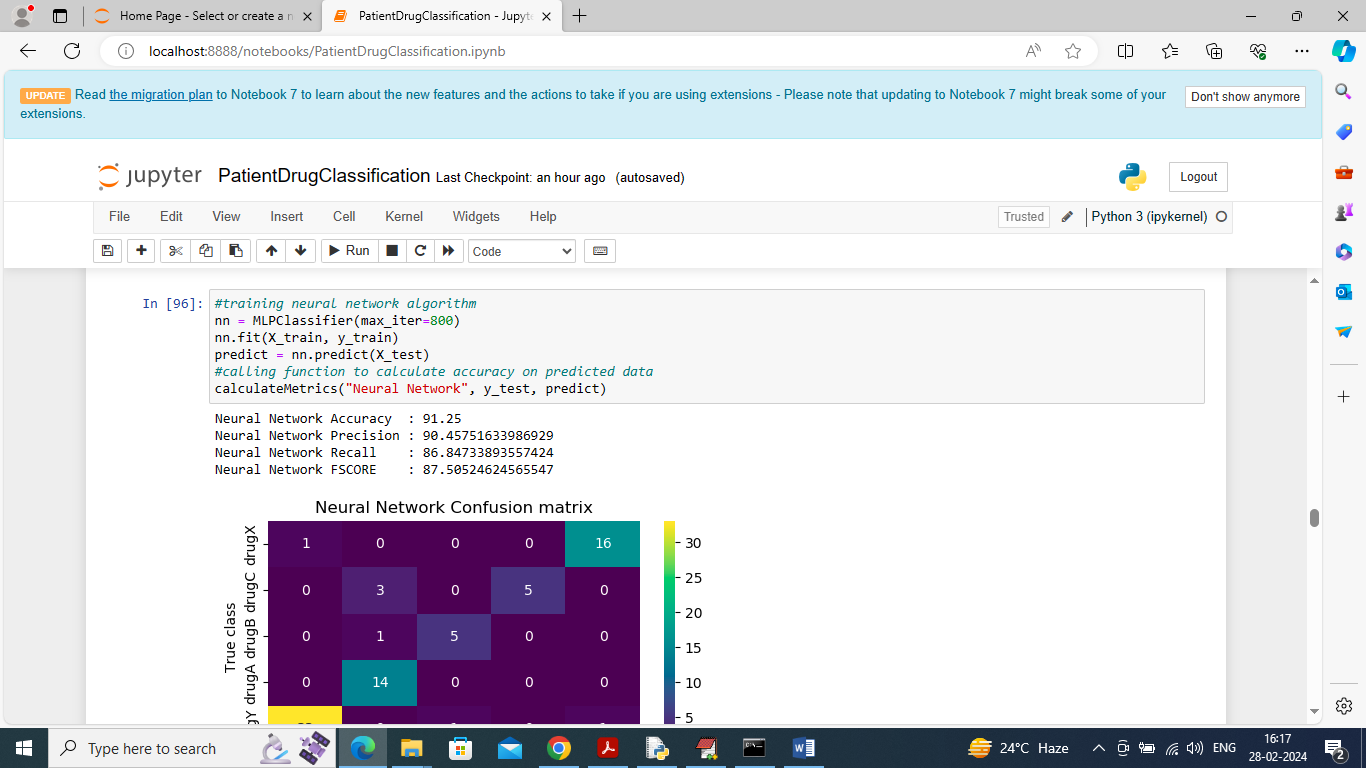
In above XGBOOST confusion matrix graph x-axis represents ‘Predicted Labels’ and y-axis represents True labels and then all different colour boxes in diagnol represents correct prediction count and all blue boxes represents incorrect prediction count which are very few



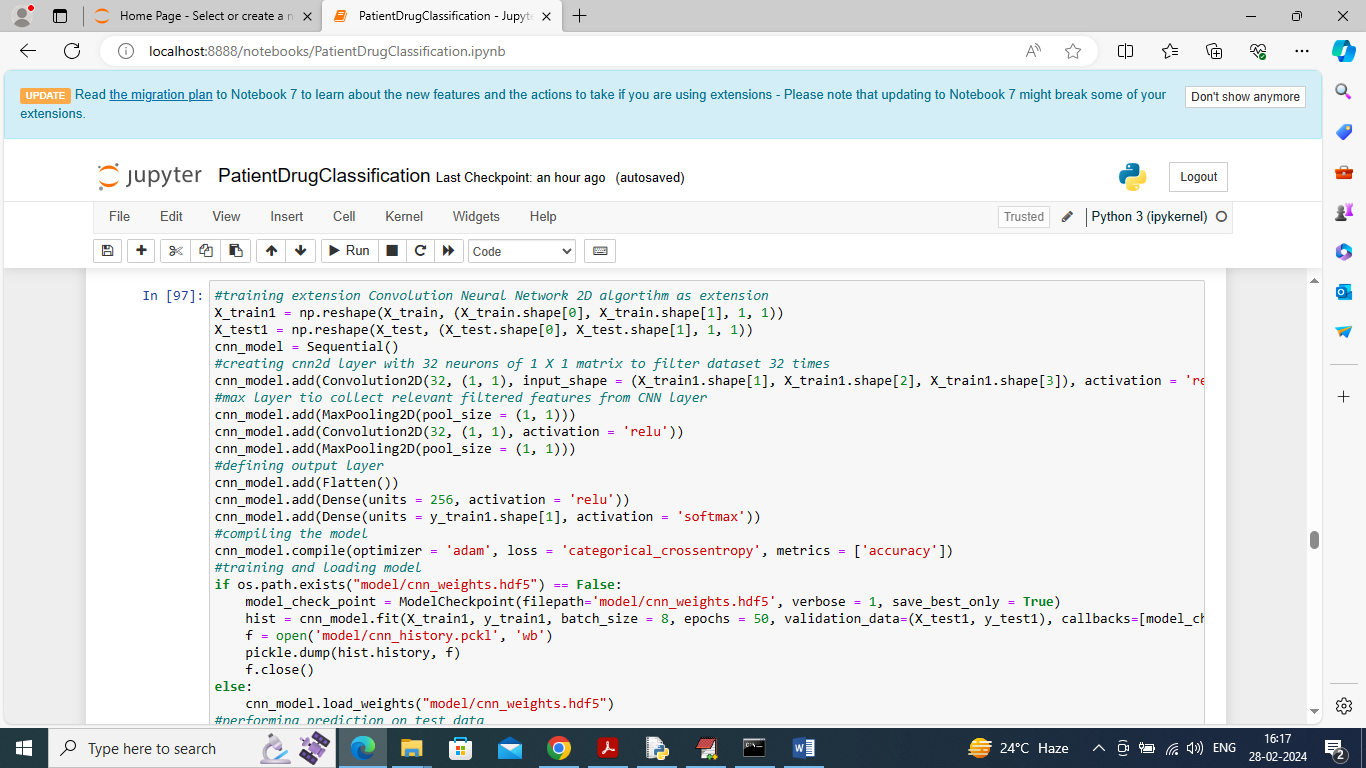
In above screen training GRAPHCNN algorithm and after execution will get below output



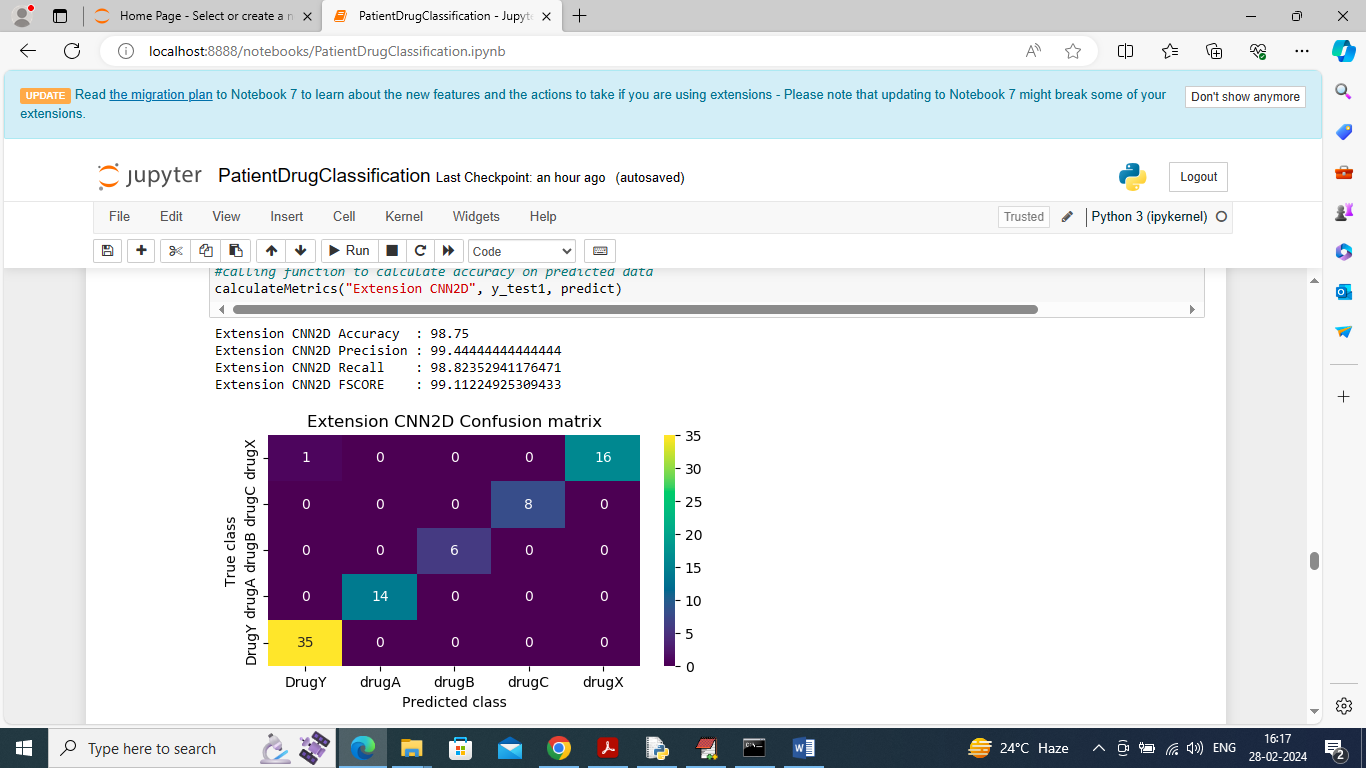
In above screen GRAPHCNN got 93% accuracy and can see all other metrics output



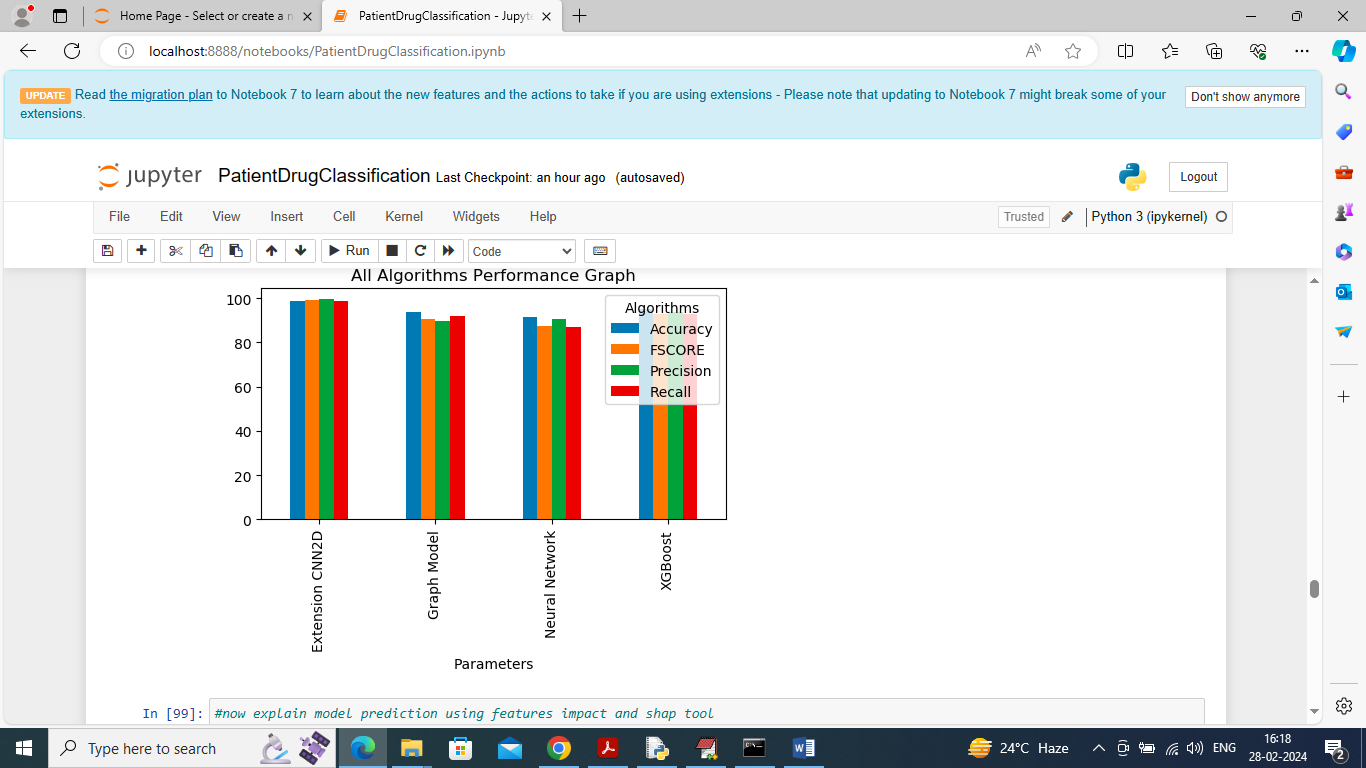
In above screen neural network got 91% accuracy



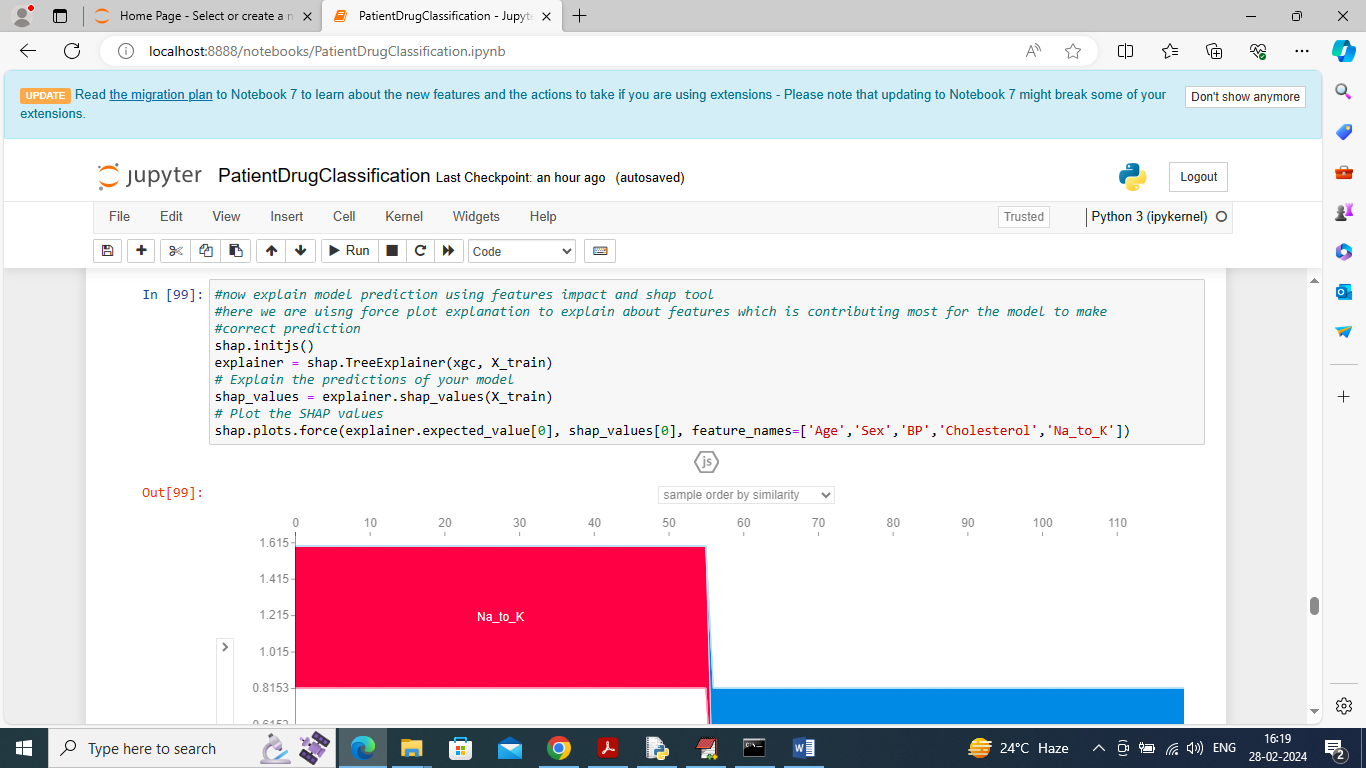
In above screen training extension CNN2d algorithm and below is the output



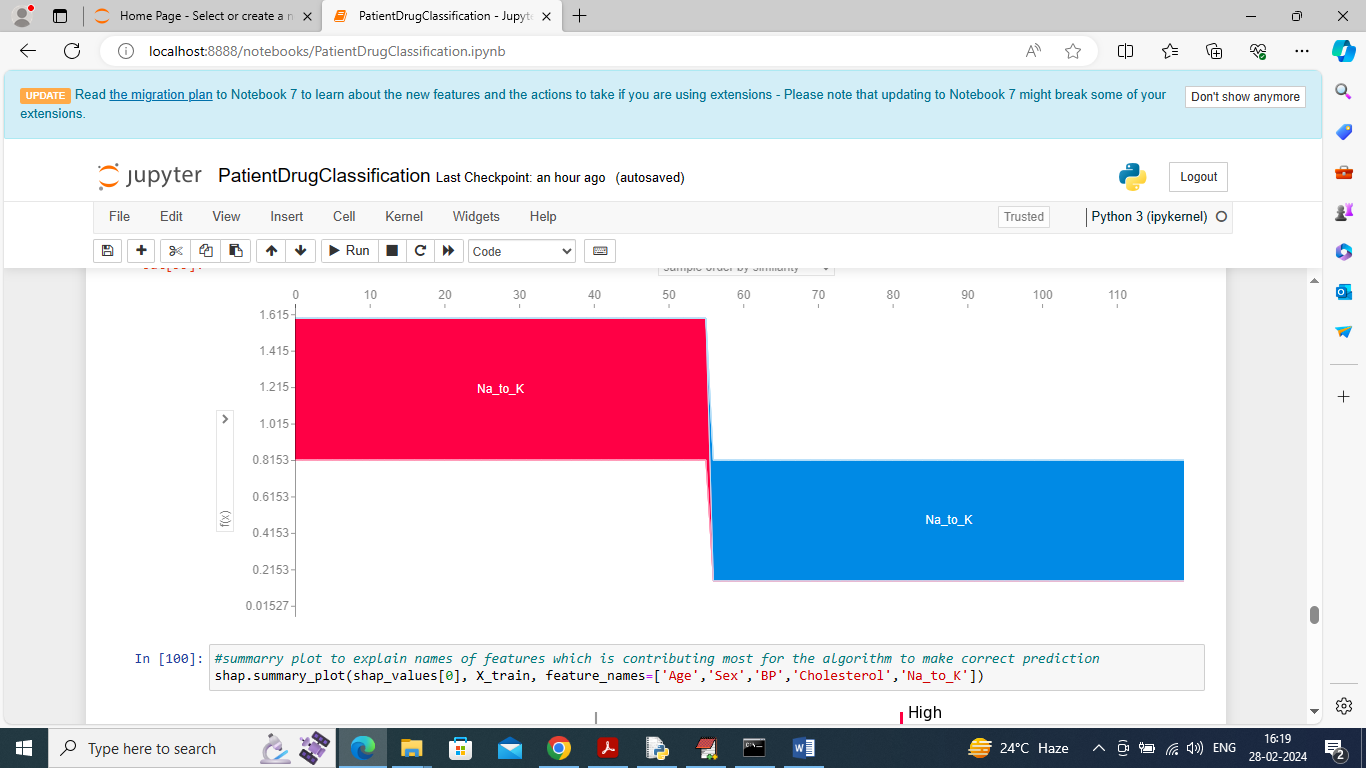
In above screen extension CNN2D got 98% accuracy



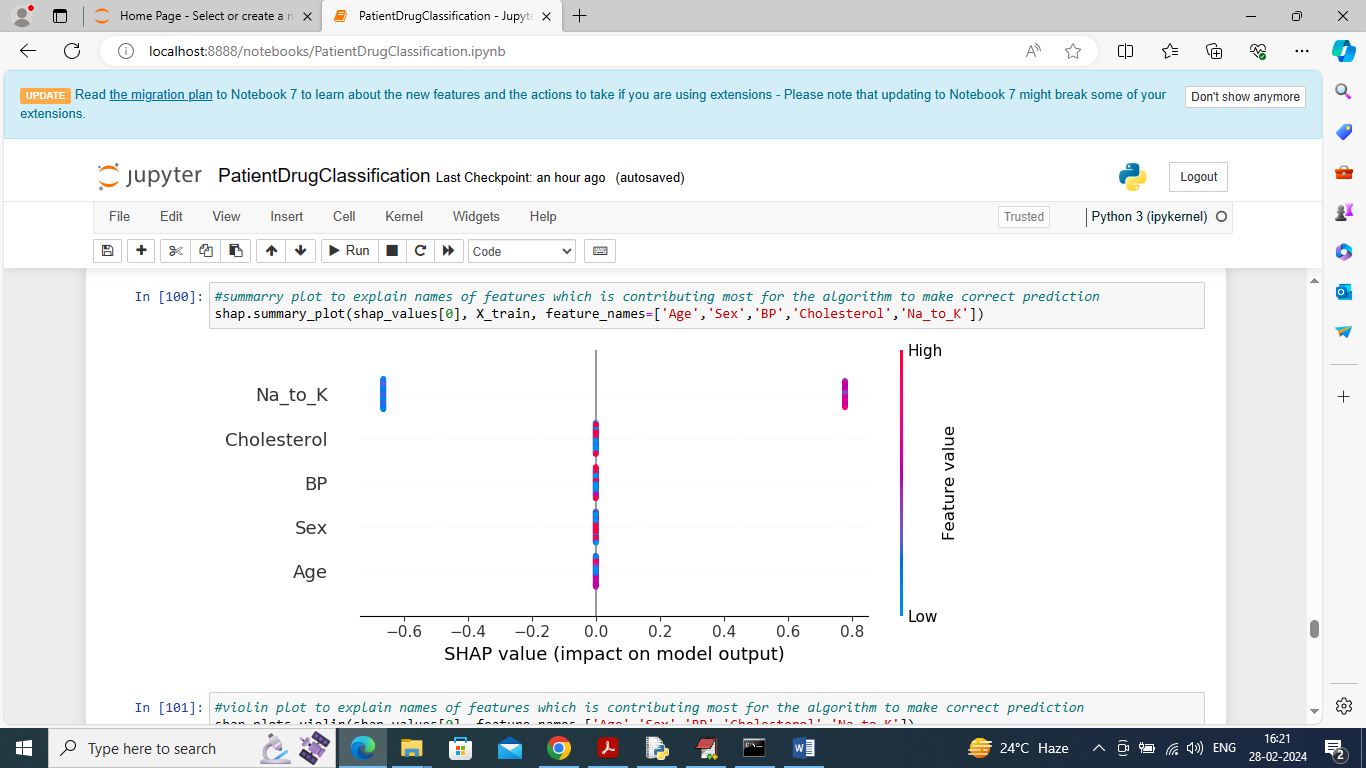
In above screen displaying comparison between all algorithms where x-axis represents algorithm names and y-axis represents accuracy and other metrics in different colour bars and in all algorithms extension CNN2D got high performance



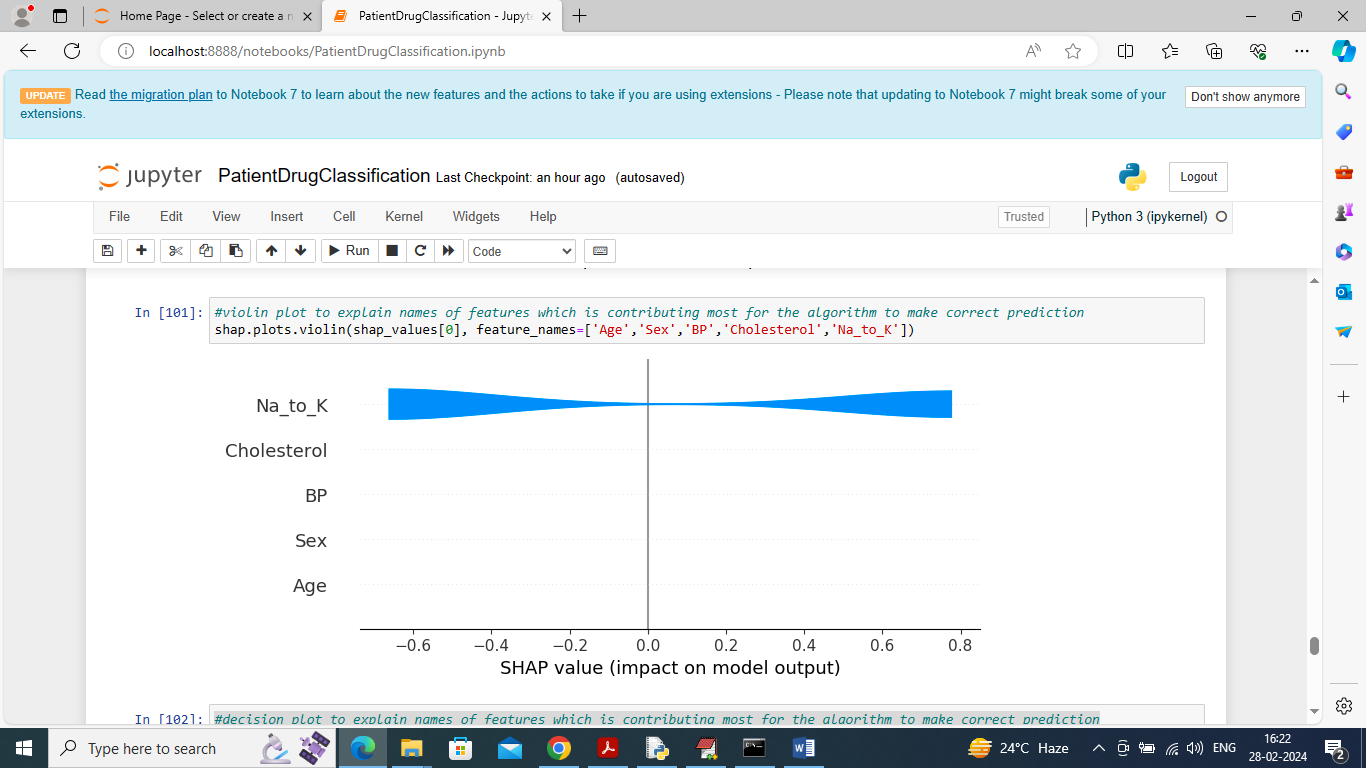
In above screen applying SHAP tool to explain about features which contribute most and after executing of this block will get below output. This plot is displaying using Force Plot function



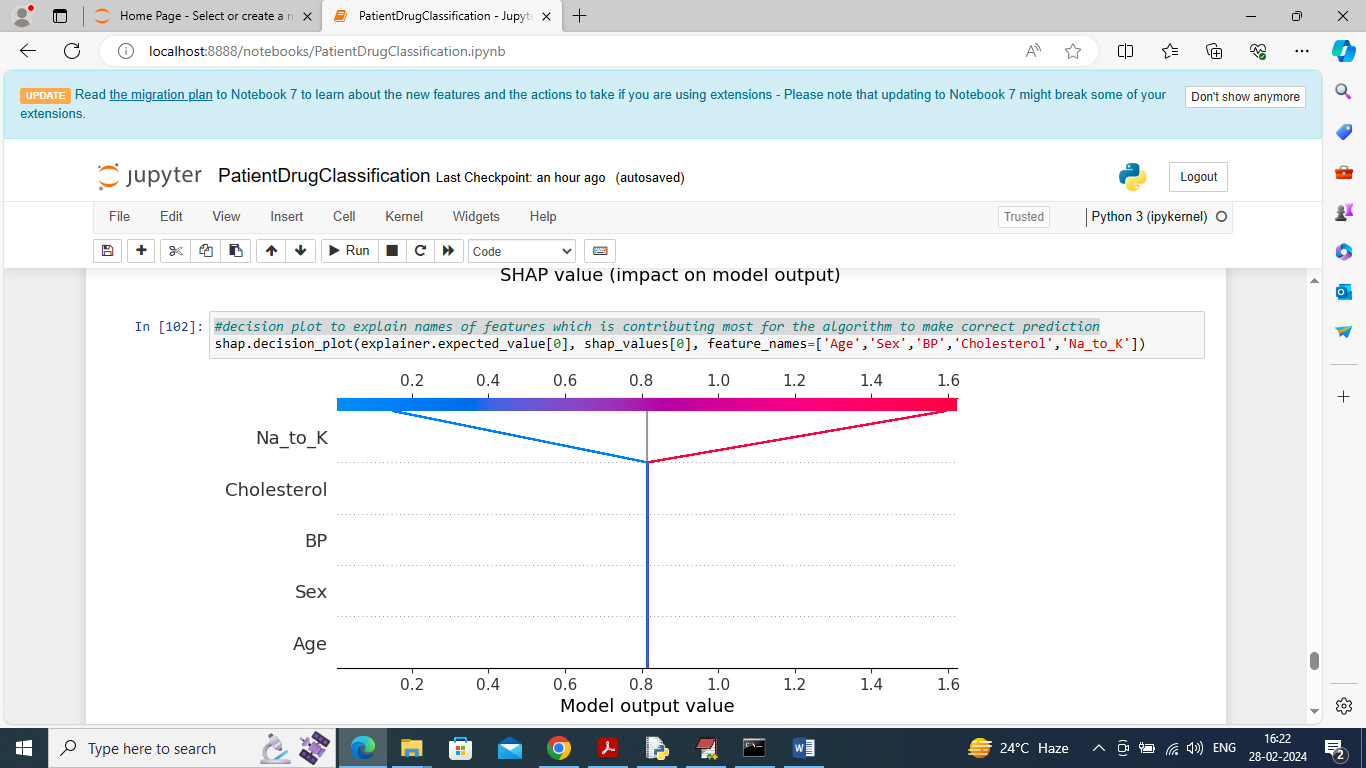
In above graph red bar represents negative features value and blue bar represent positive value and in all features NA\_to\_K feature is contributing most.



In above summary plot also can see XAI saying NA\_to\_K features is best



In above violin plot also XAI saying NA\_to\_K features is best



In above decision plot also XAI saying same ‘NA\_to\_K’ feature contributing most to the model for correct prediction.

So by using above implementation we have trained models and then explain about all features which is contributing or impacting model most for correct prediction.