Predictive Maintenance for Factory Equipment

In industrial environment production is the main factor to balance supply demand and to progress industry revenue but sometime due to machinery failure production will stop which effect supply and industry revenue. In olden days we don’t have any possible way to predict machinery failure before time till it fully break down but now all industries are using sensors to monitor machinery health and by utilizing this sensor data we can predict machine health or its failure and it’s available life and based on life technicians will arrange maintenance. Timely maintenance will make machines to work perfectly and production will continue non-stop.

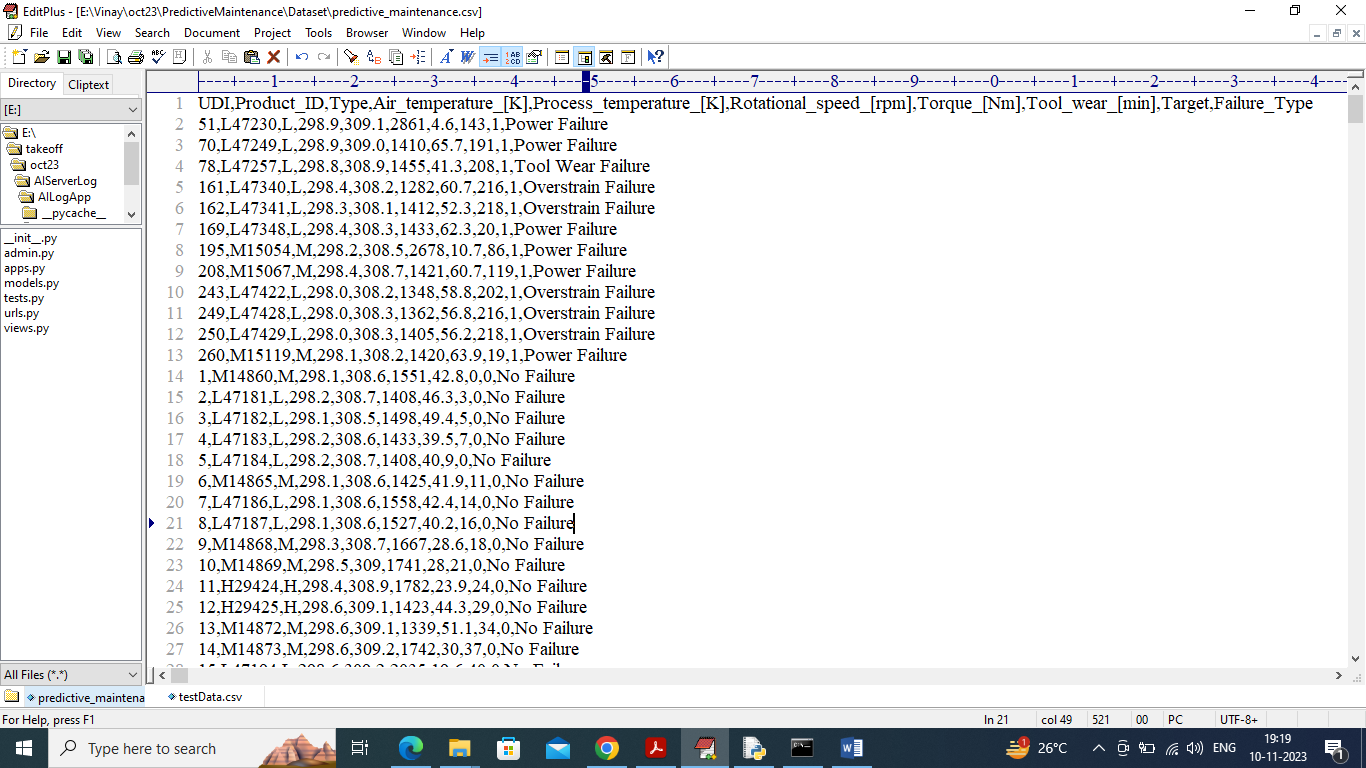
To predict failure we can employ machine or deep learning algorithms which will get trained on past data and can predict future value by taking current input. This trained models can continuously read input from sensor data and then predict machine health or failure.

To make prediction accurate we have experimented with various machine and deep learning algorithms such as SVM, Decision Tree, KNN, CNN (Convolution Neural Network) and Random Forest. Each algorithm performance is evaluated in terms of accuracy, precision, recall, Confusion Matrix, ROC Graph and FCSORE. All algorithms able to achieve accuracy of 90% and CN manage to get an accuracy of 95%.

To train and test above mention algorithm performance we have utilize ‘Predictive Maintenance Sensor Dataset’ from KAGGLE repository and this dataset can be downloaded from below URL

<https://www.kaggle.com/datasets/shivamb/machine-predictive-maintenance-classification>

In below screen we are showing dataset details

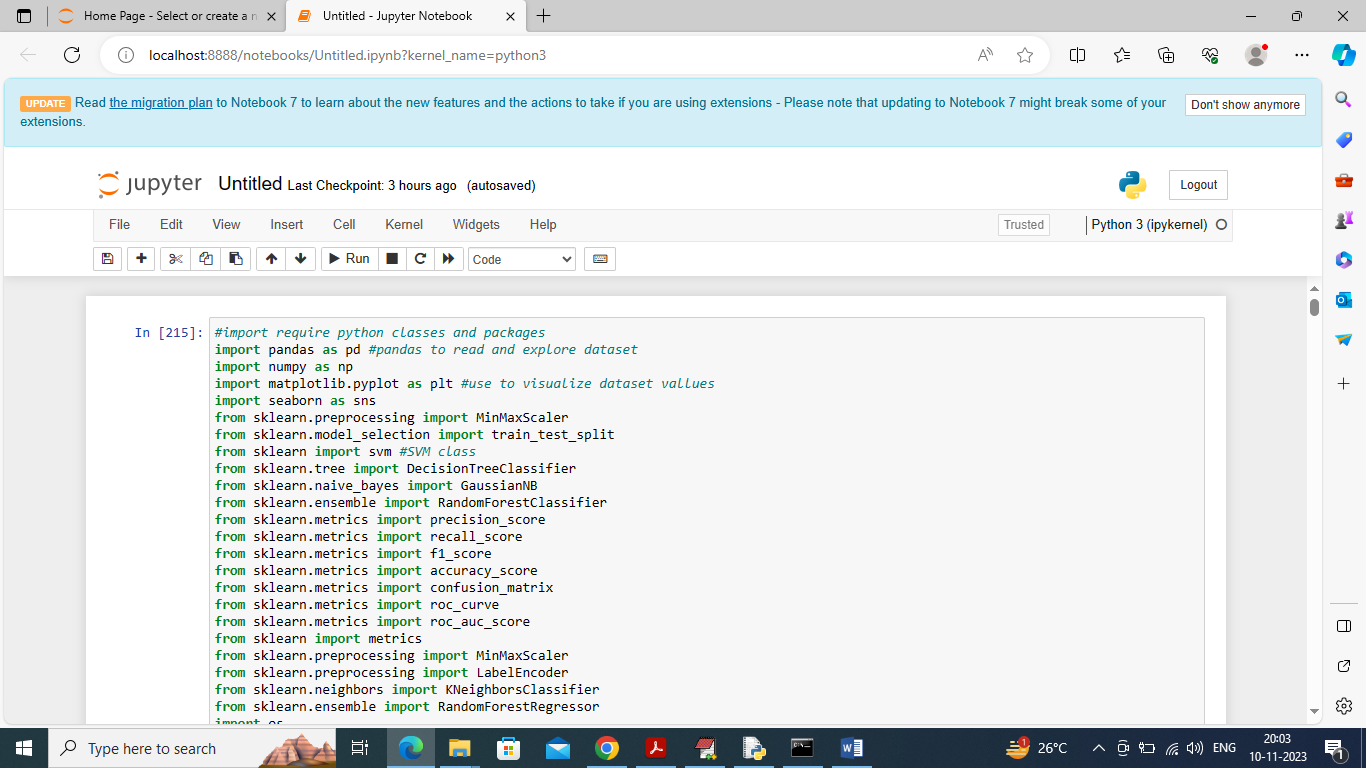


In above dataset screen first row represents Dataset Column Names and remaining rows contains dataset values and in last column we have “Failure Type’ as Machine Failure class labels. So by using above dataset we will train and test all algorithm performance.

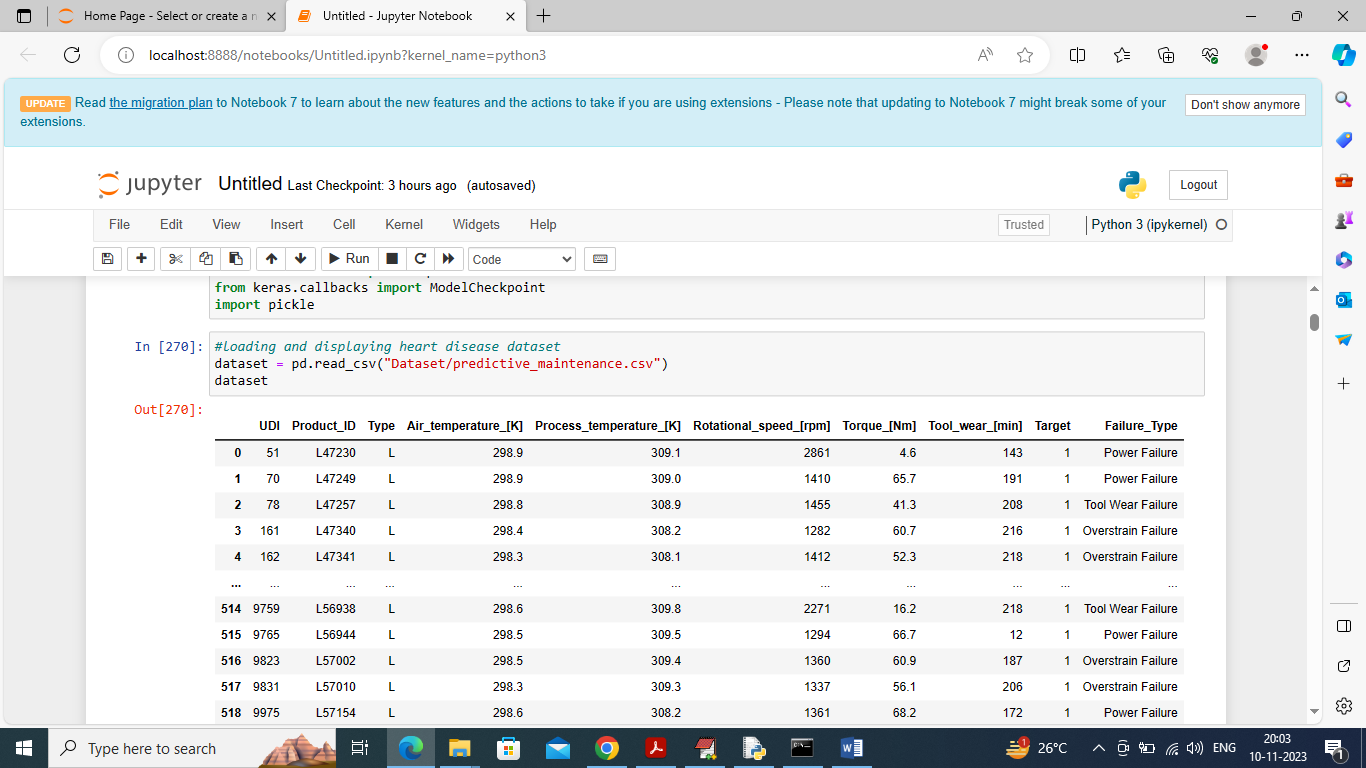
Before training we have applied various data analysis such as Graph Visualization, Features Selection, data shuffling and normalization.

SCREEN SHOTS

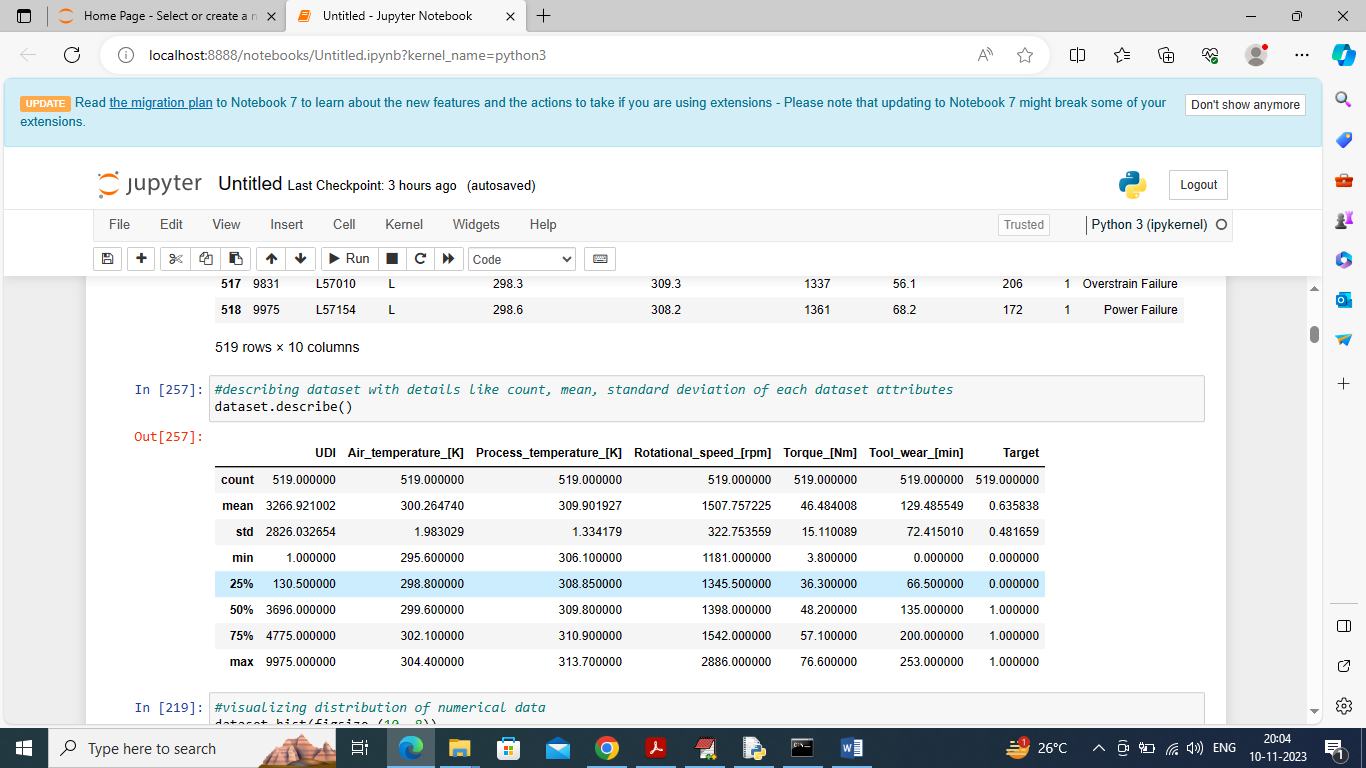
We have coded this project using JUPYTER NOTEBOOK and below are the code and output screens with blue color comments



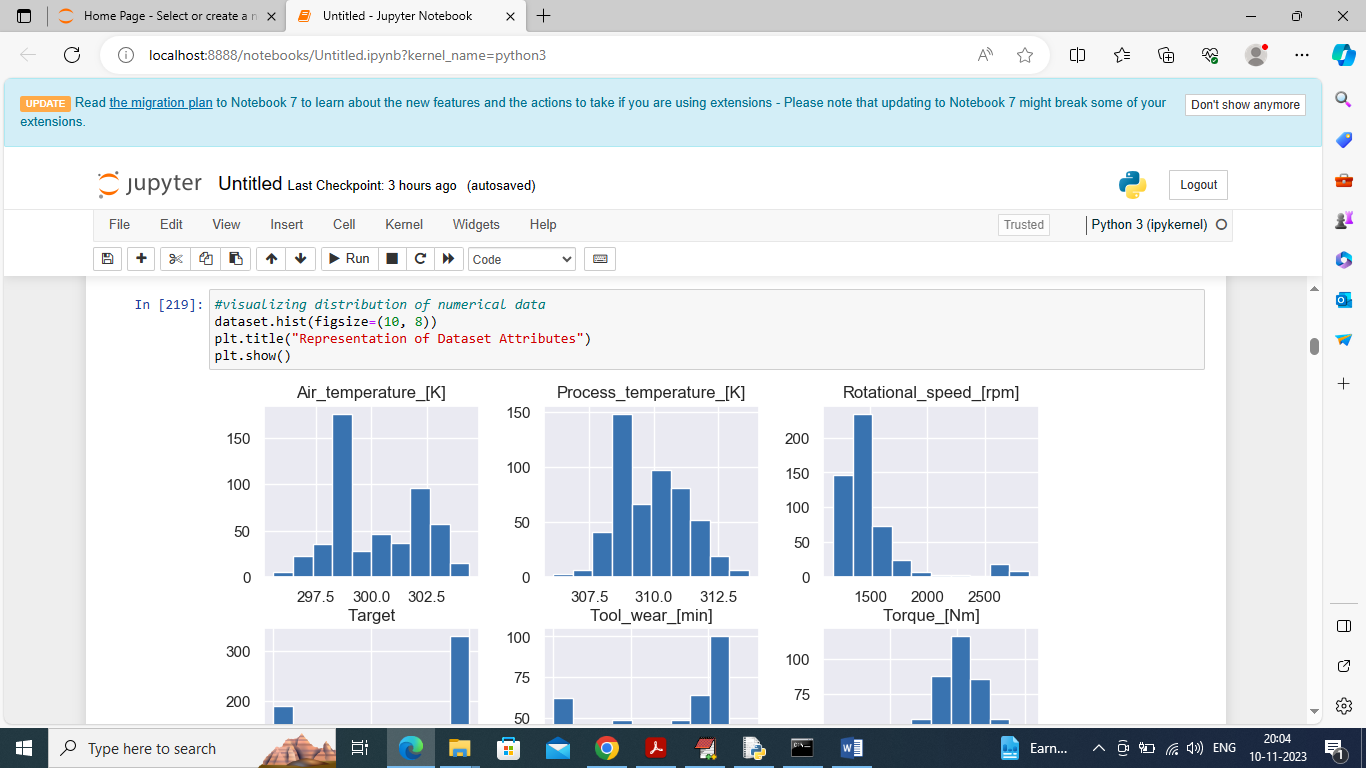
In above screen importing require python classes and packages



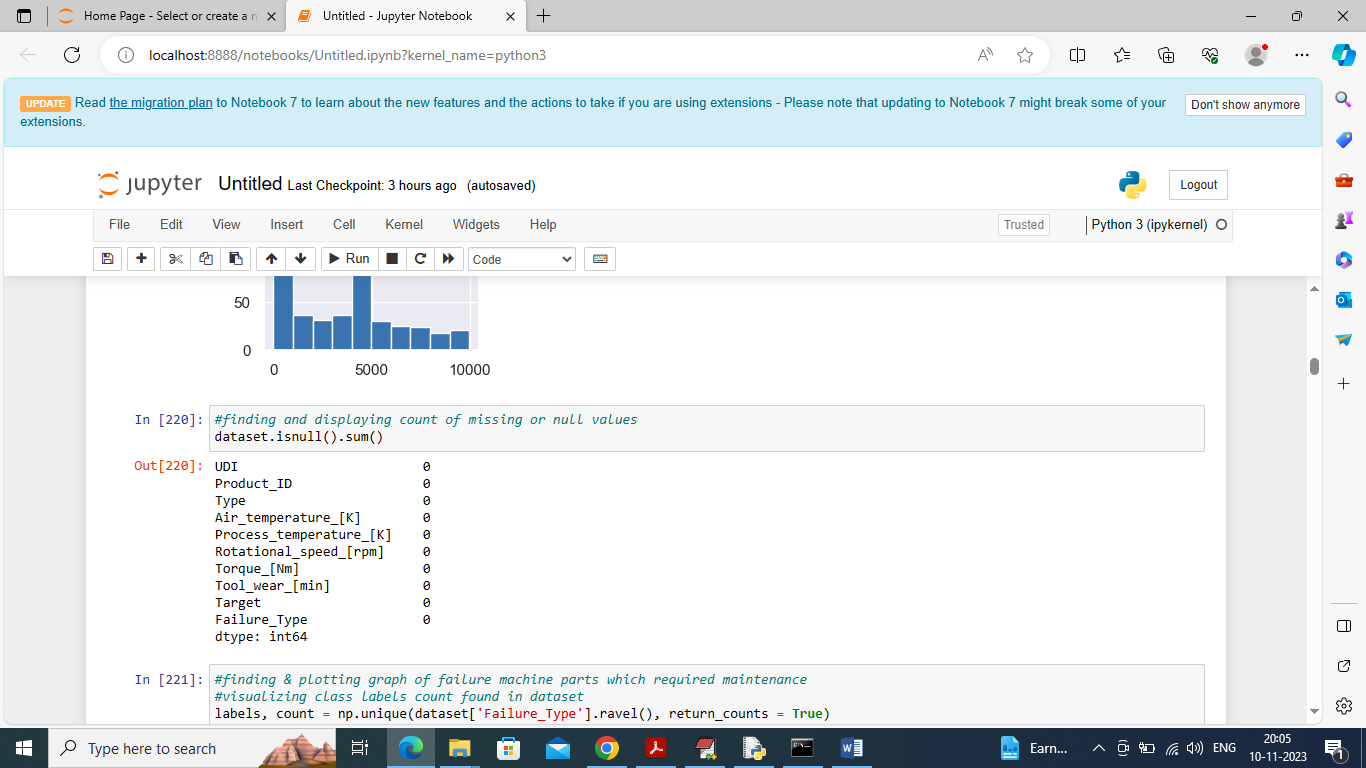
In above screen loading and displaying predictive maintenance dataset



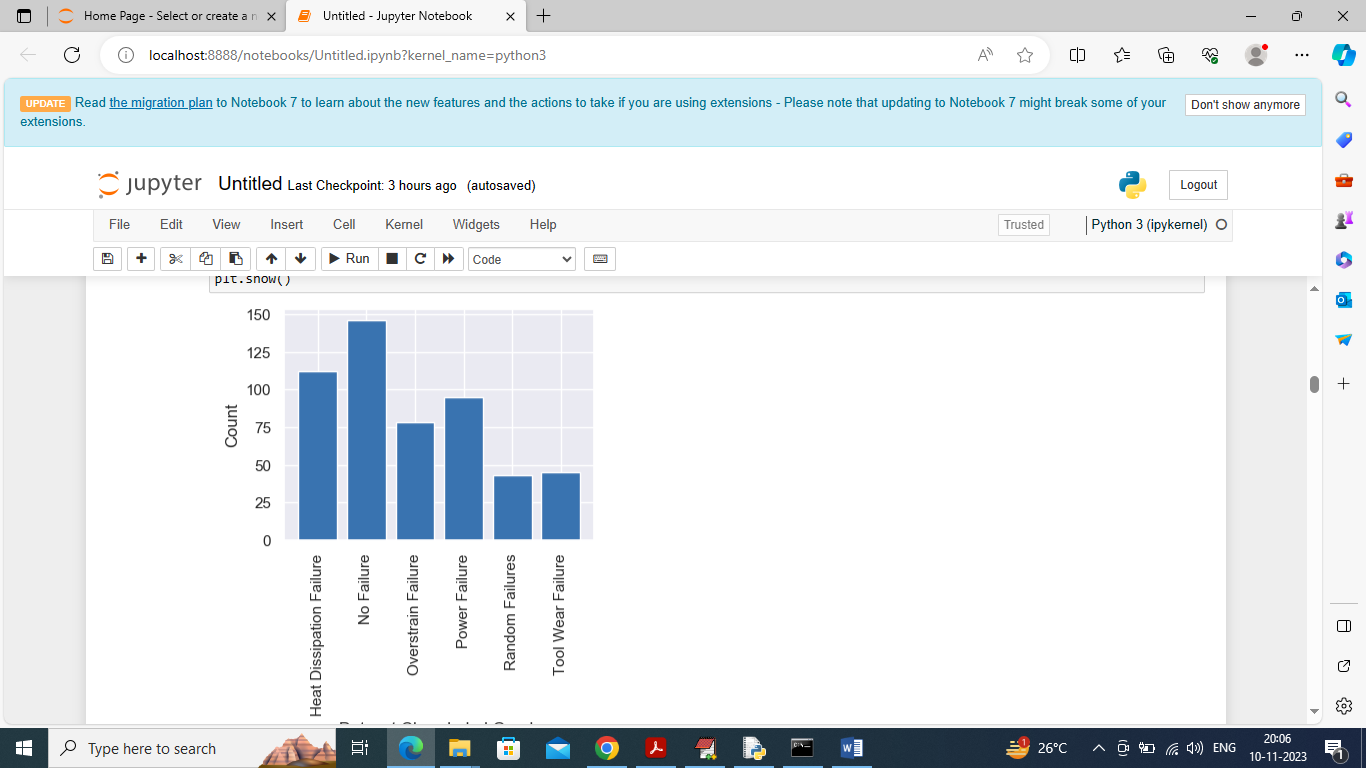
In above screen describing dataset values as ‘Mean, standard deviation, min, max and other percentage of values



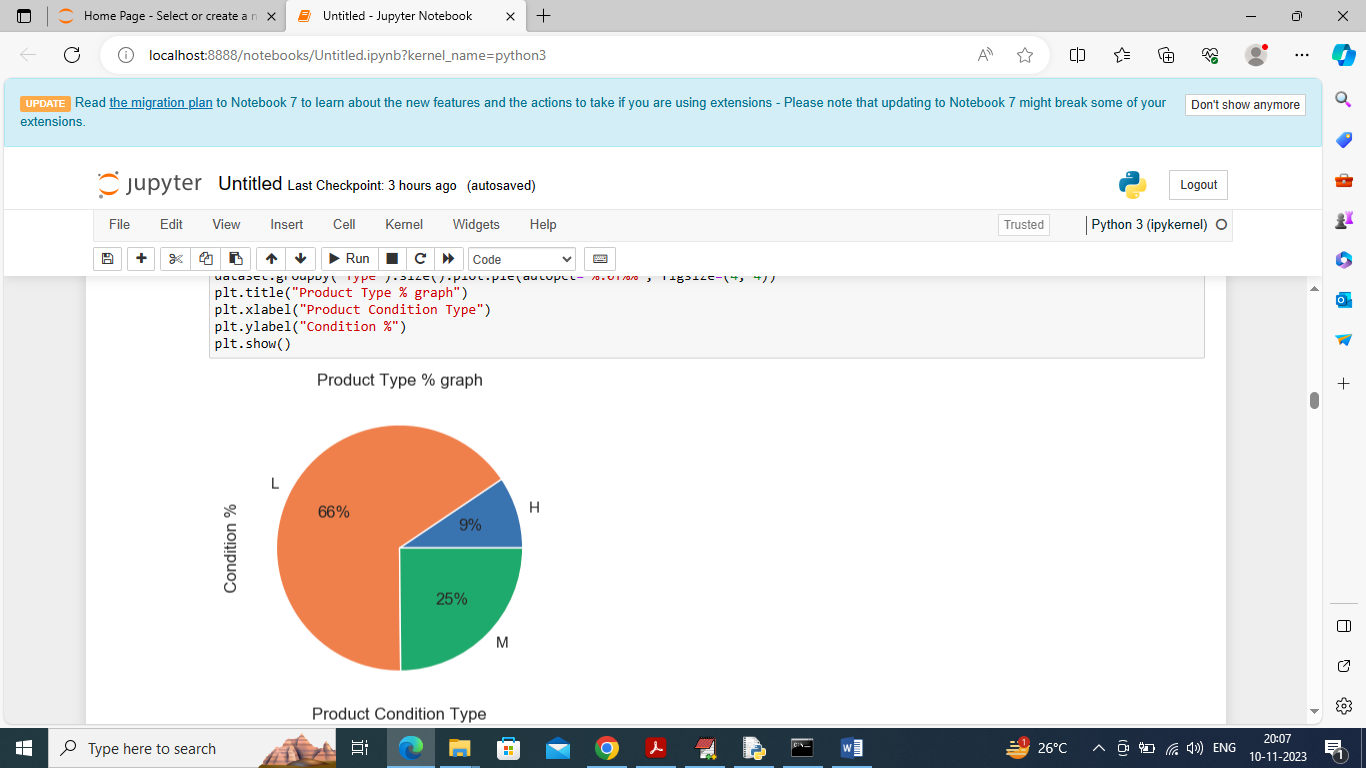
In above screen displaying graph of values distribution for each column values and by seeing above graph we can understand how values of columns distributed from one range to other range



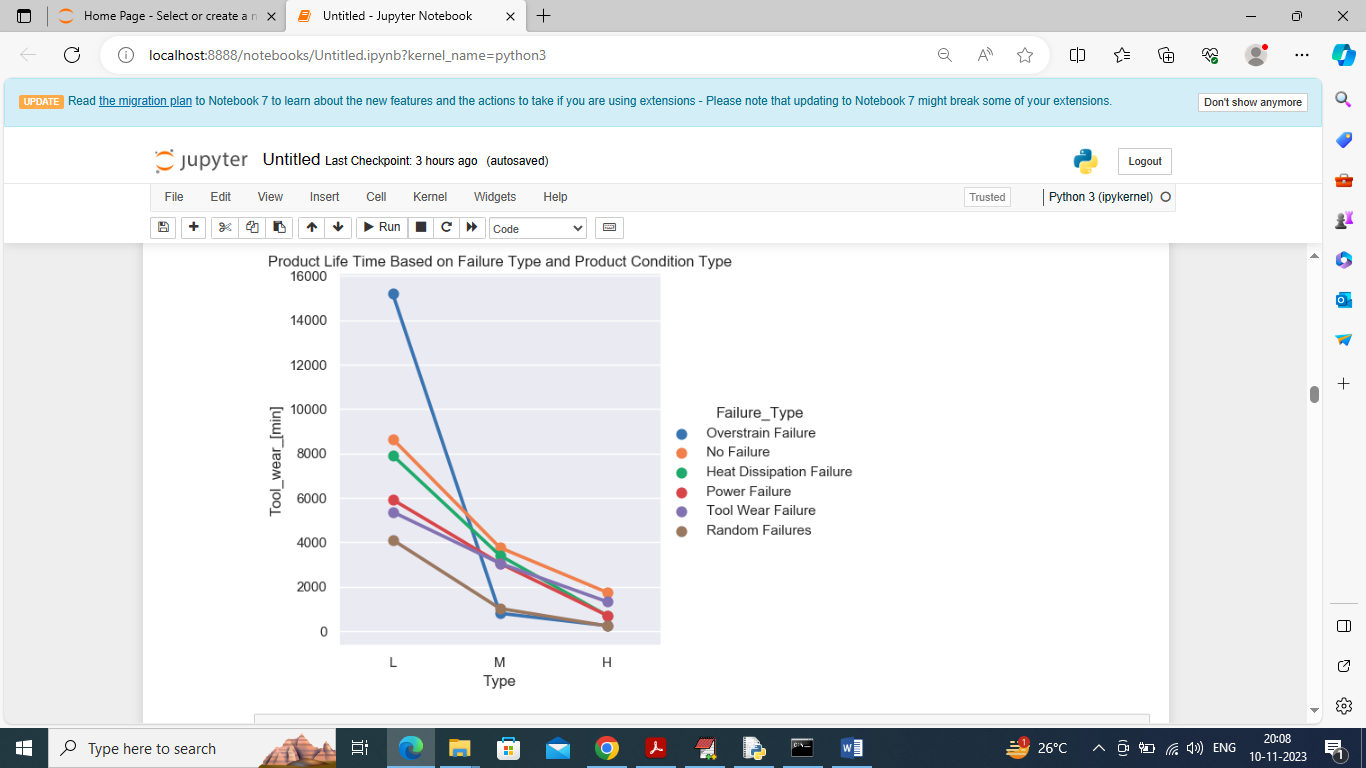
In above screen checking and displaying count of missing values and above dataset contains NO Missing values



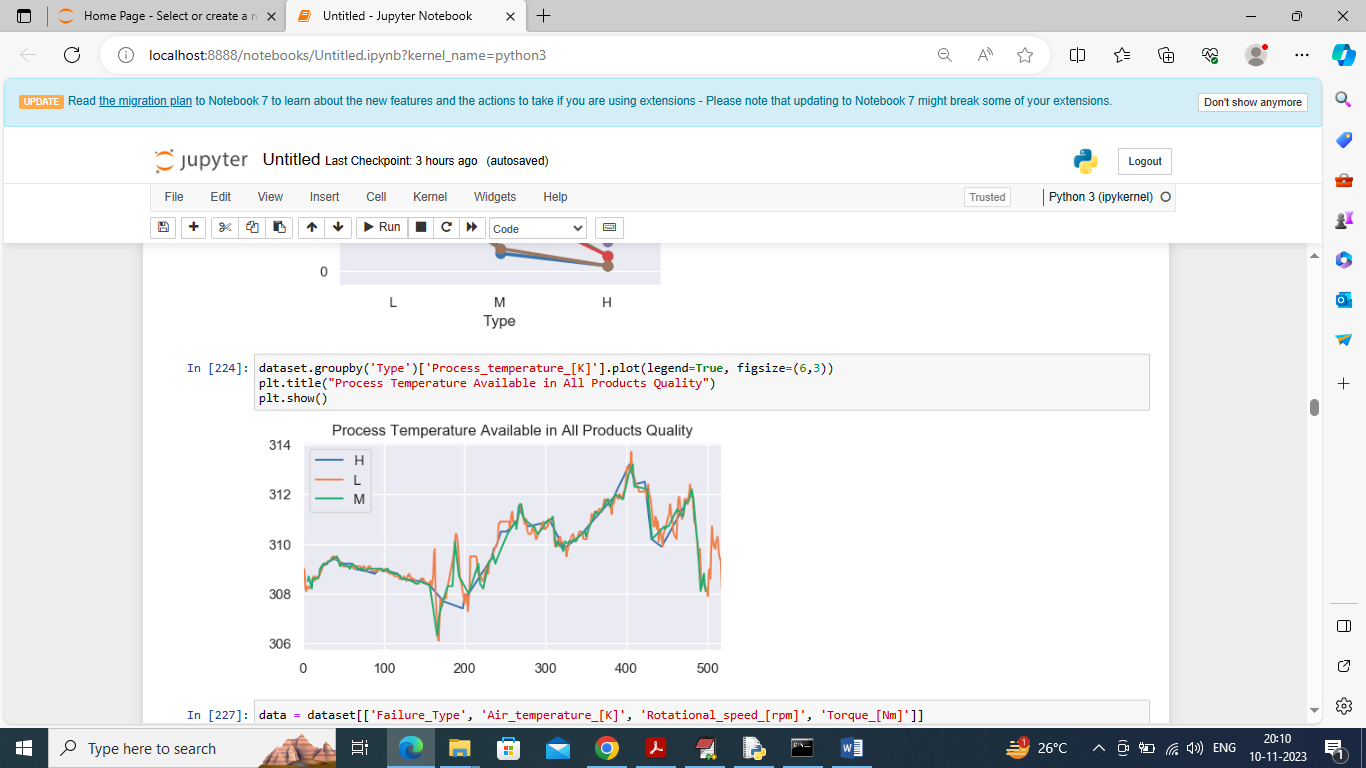
In above graph displaying different Failures found in the dataset where x-axis represents ‘Failure Name’ and y-axis represents Number of instances or samples found under that failure



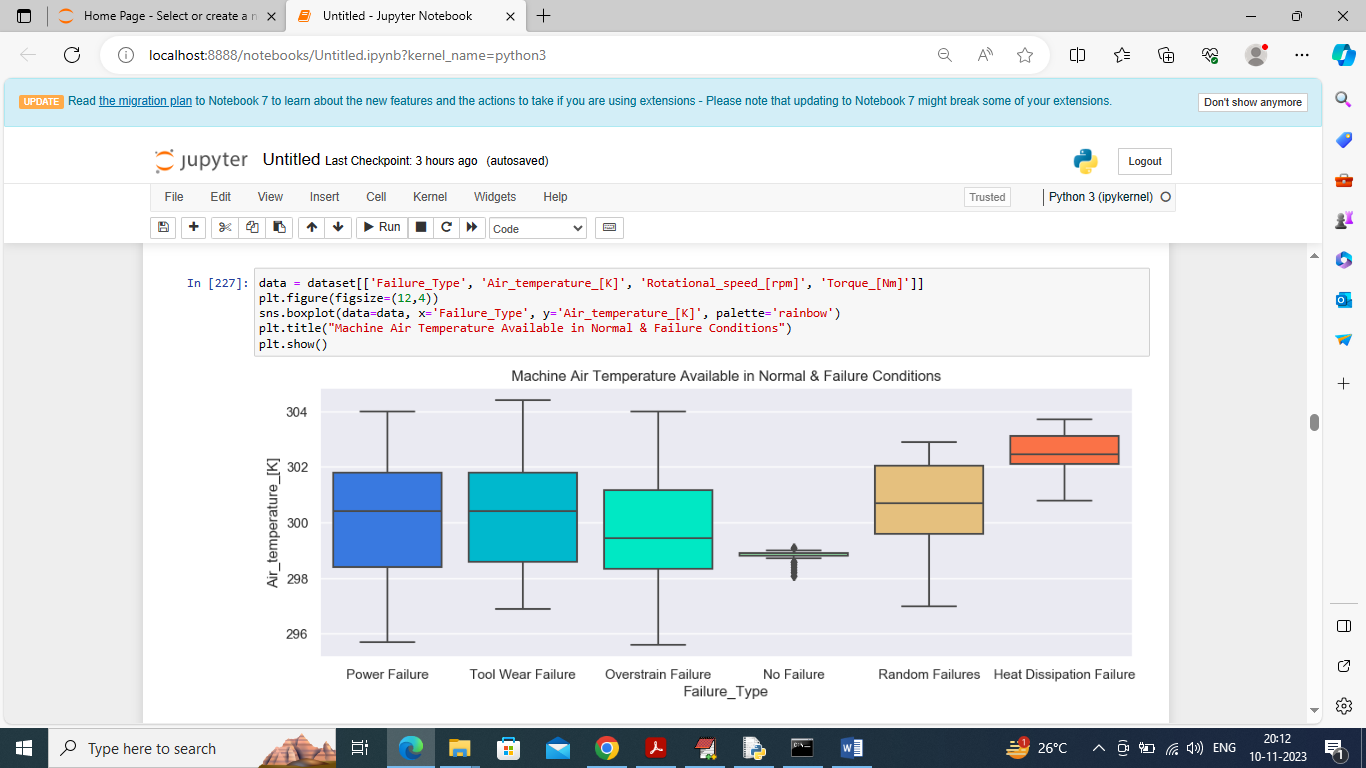
In above graph displaying product maintenance status where L represents Low Quality, M represents Medium and H represents High and in above graph we can see % of product quality in machine



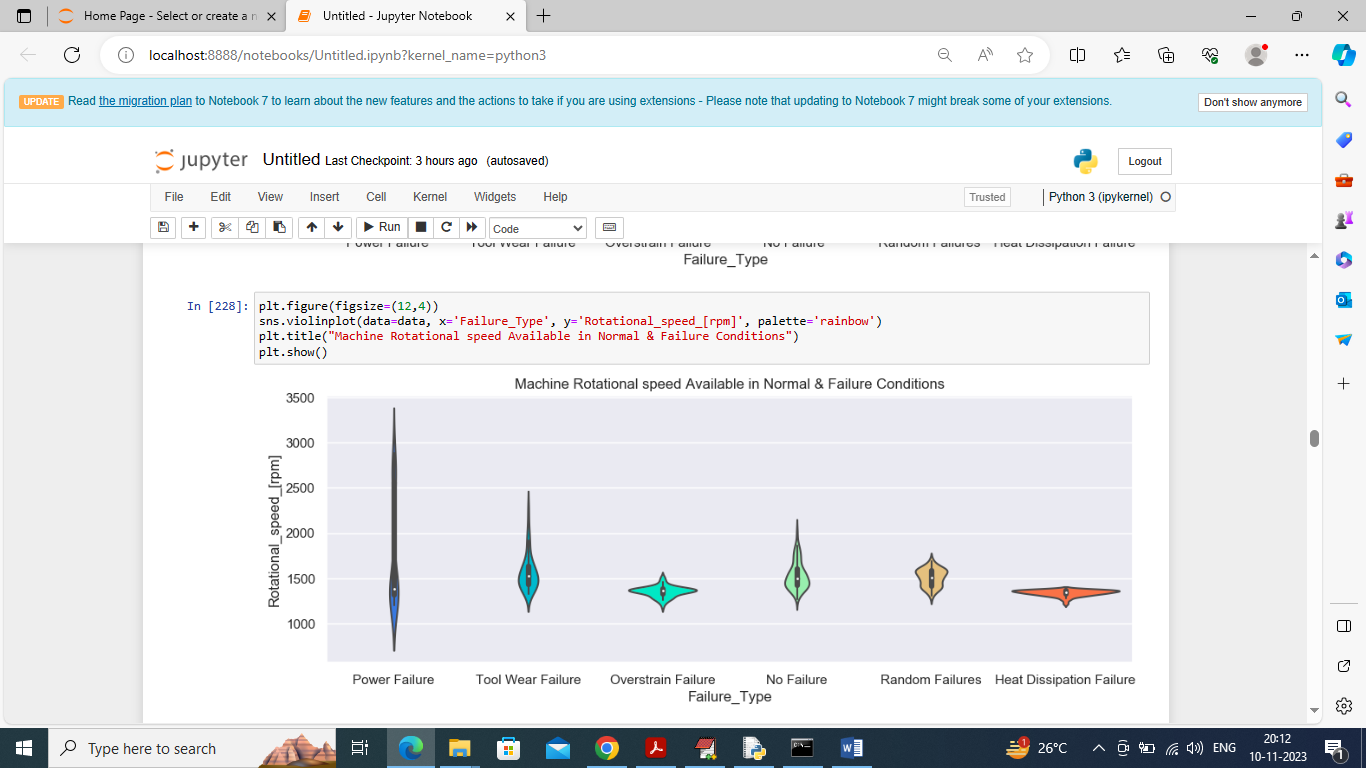
In above graph displaying % of maintenance required for each machinery product and in graph x-axis represents Low, High and Medium product condition and y-axis represents available life and each different line represents Failure for which maintenance required



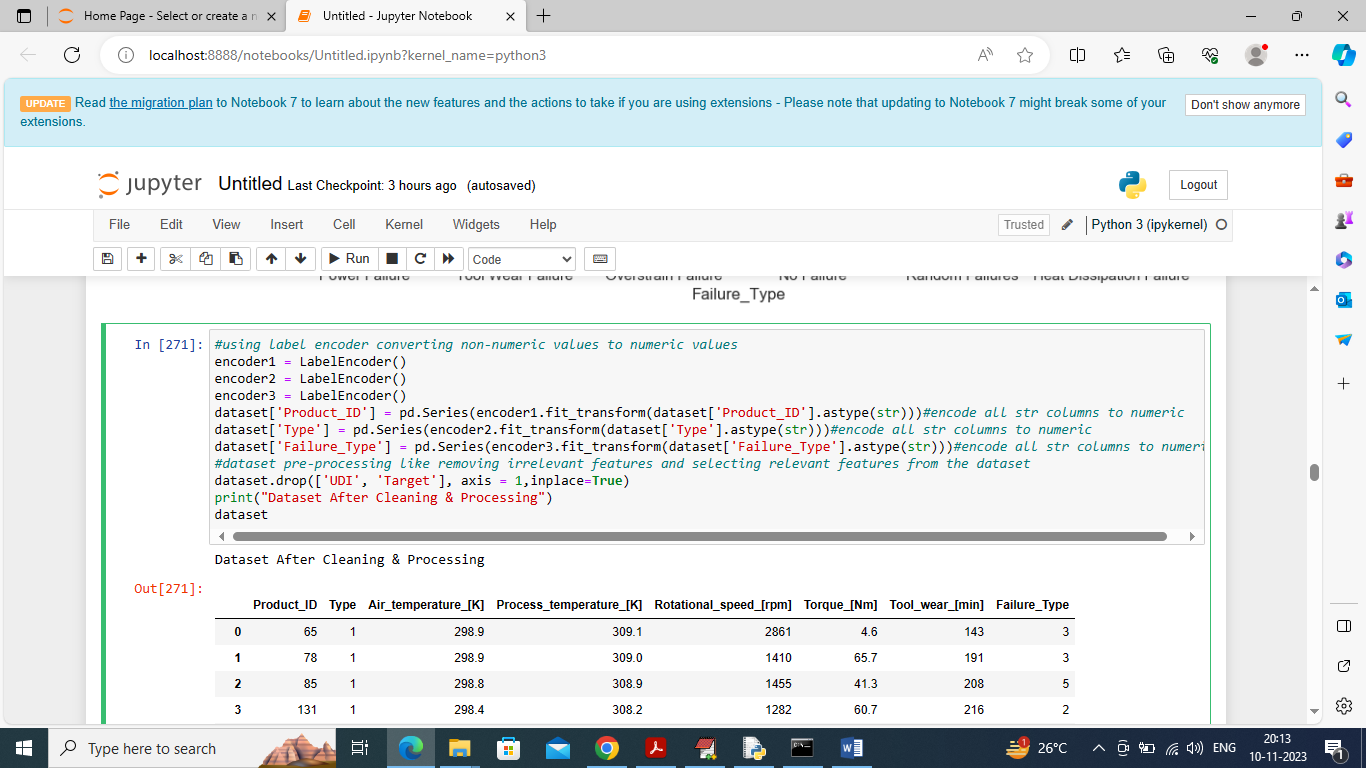
In above graph displaying Machine Process Temperature for various condition of product where x-axis represents Number of Records and y-axis represents Process Temperature and different color lines represents High, Low and Medium product condition



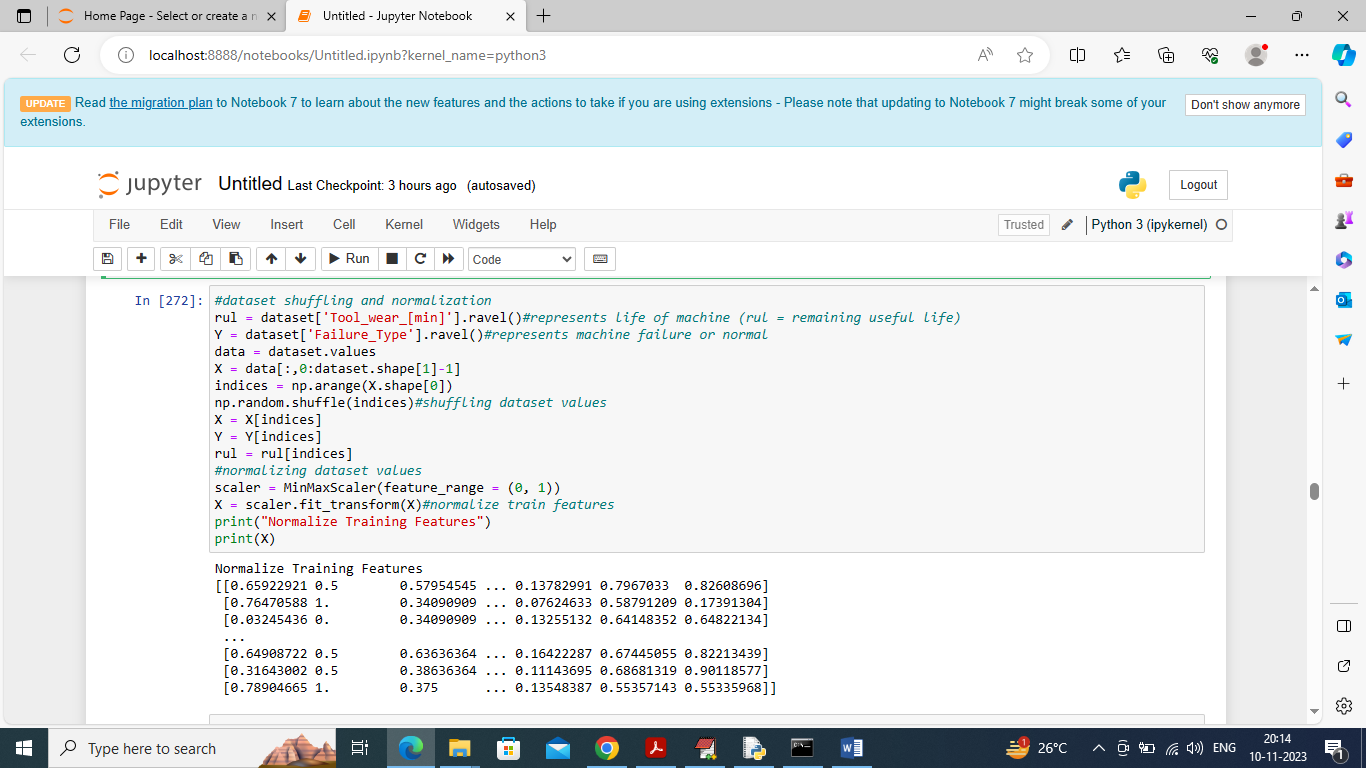
In above graph displaying Air Temperature for different Failure



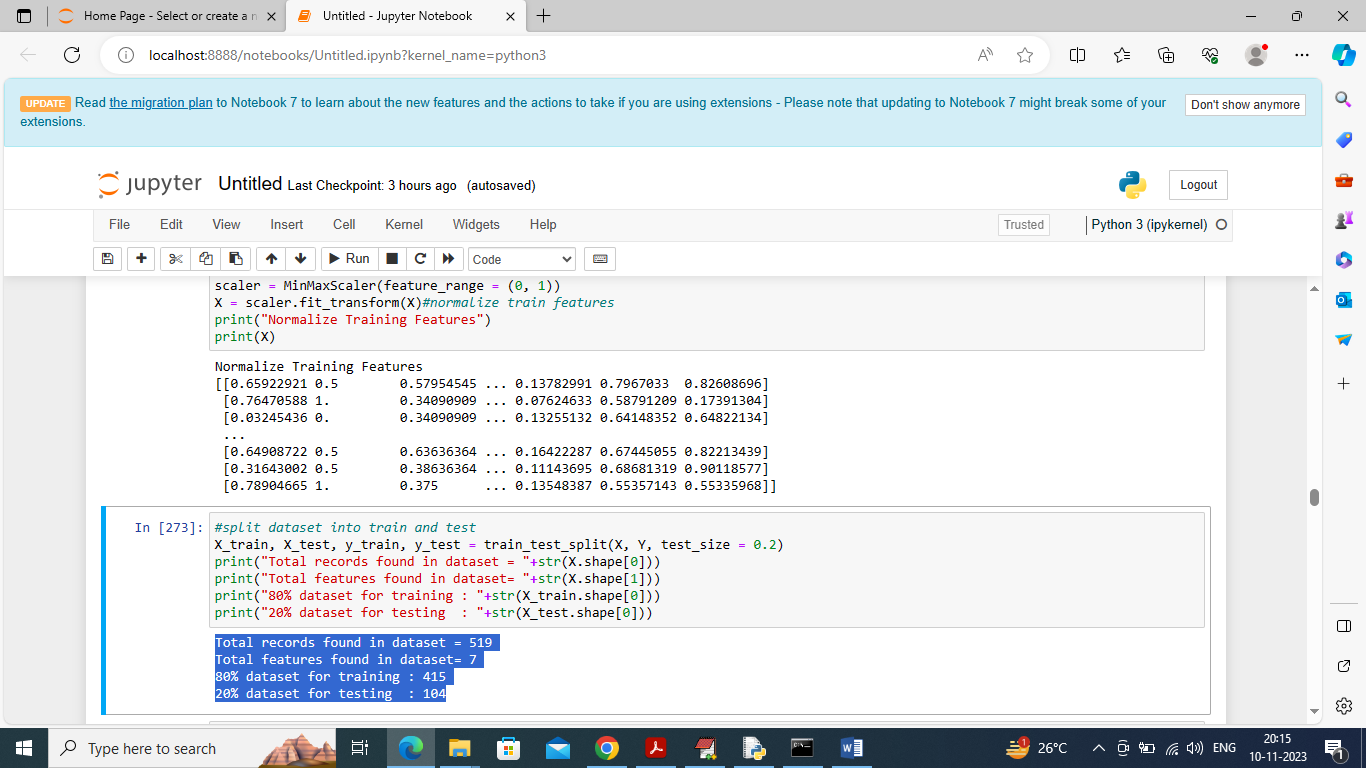
In above graph displaying Machine Rotation Speed for different Failure Condition



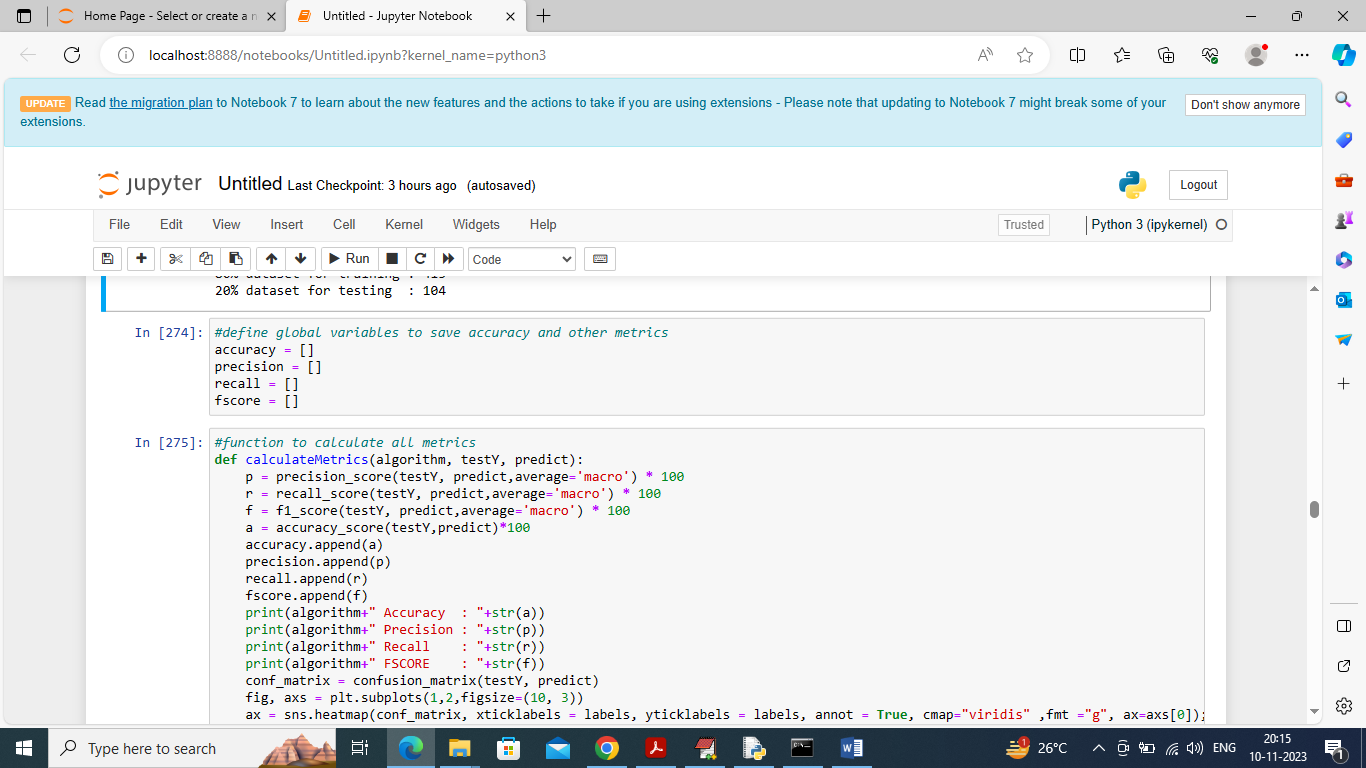
Using above code we are applying data processing to convert non-numeric values to numeric values and after conversion we can see all values are in numeric format as all ML algorithms take input as numeric format so we have converted



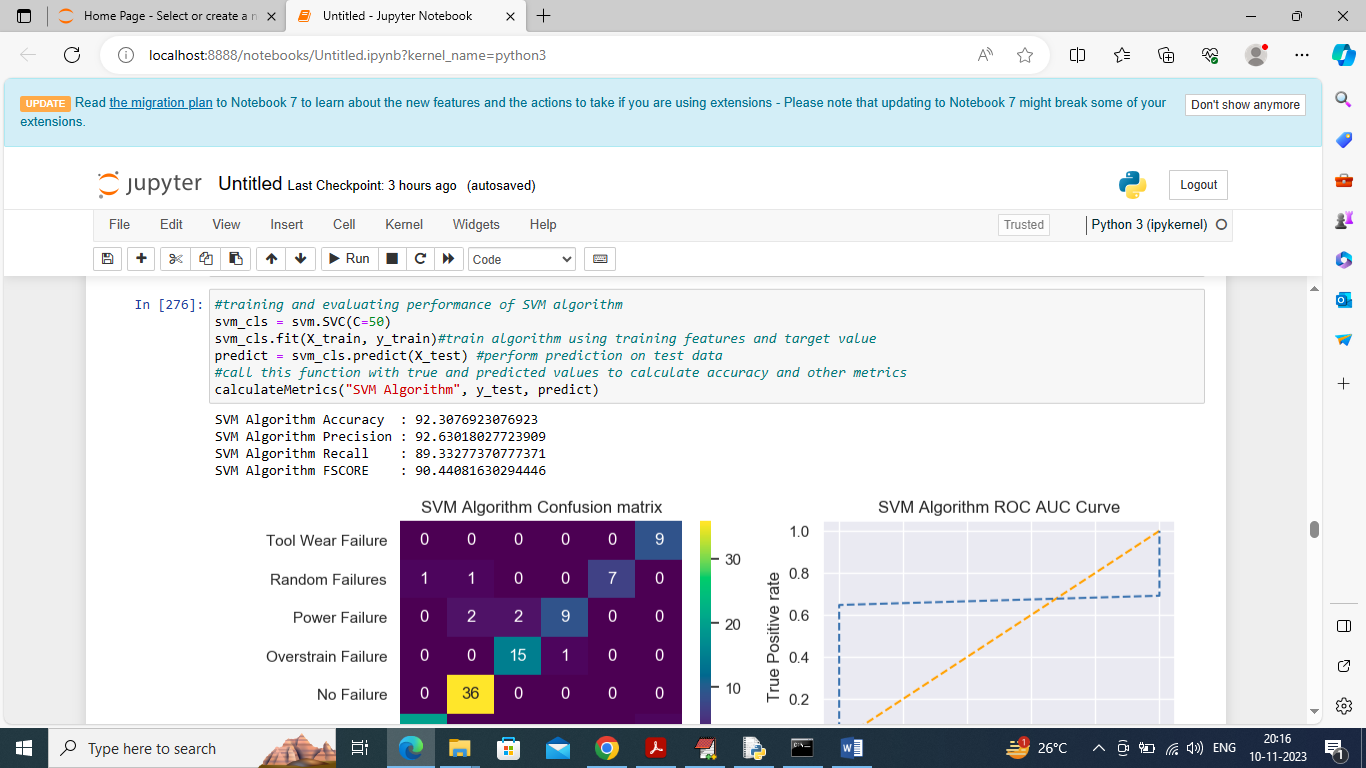
In above screen applying various features processing like Features Selection, shuffling and normalization and after normalization we can see normalized values



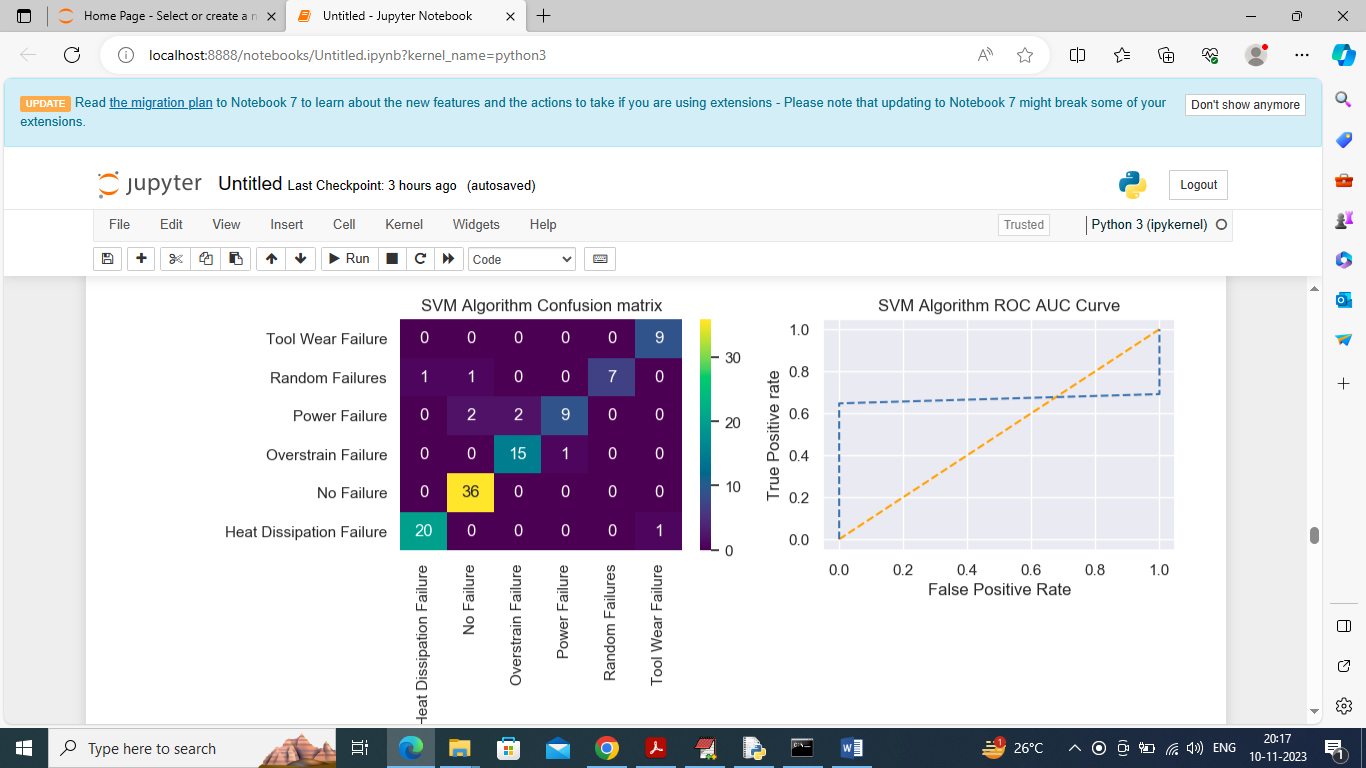
In above screen splitting dataset into train and test where application use 80% dataset size for training and 20% for testing



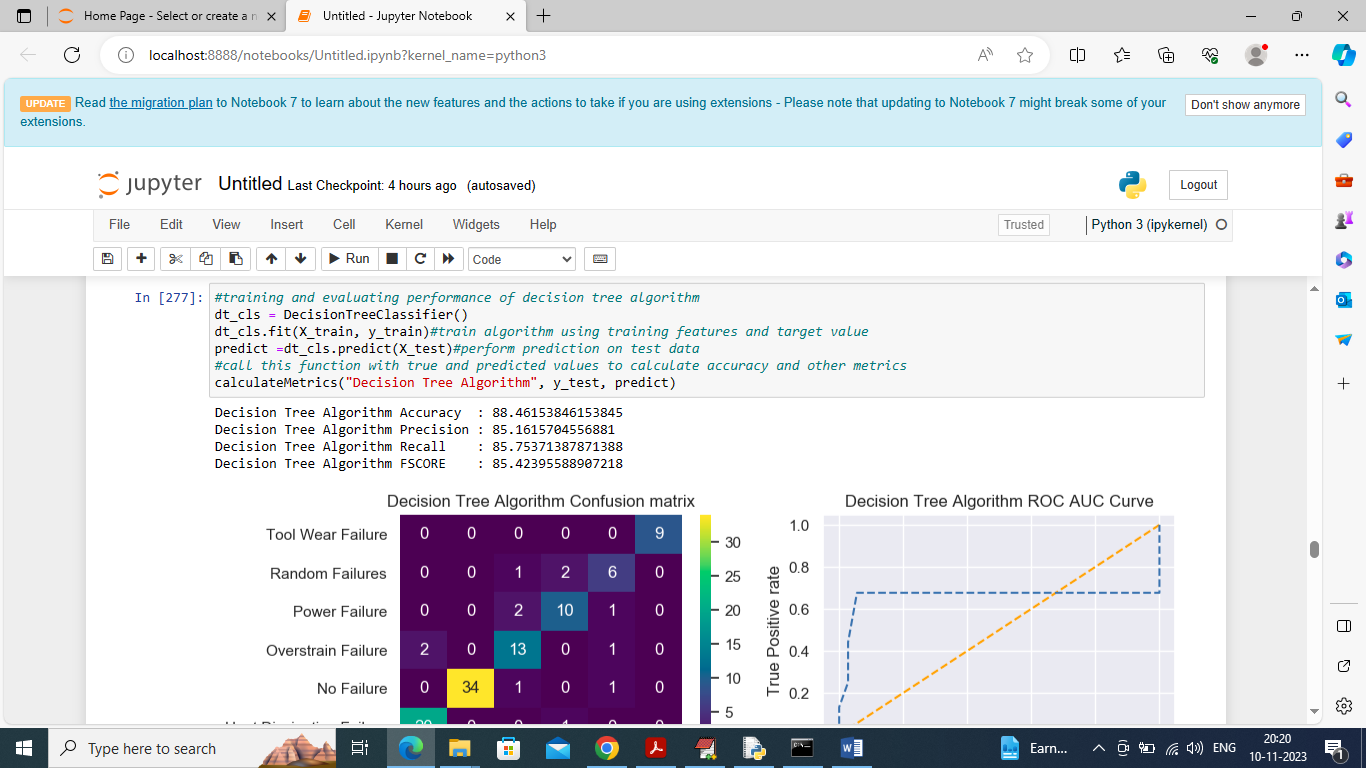
In above screen defining function to calculate accuracy and other metrics



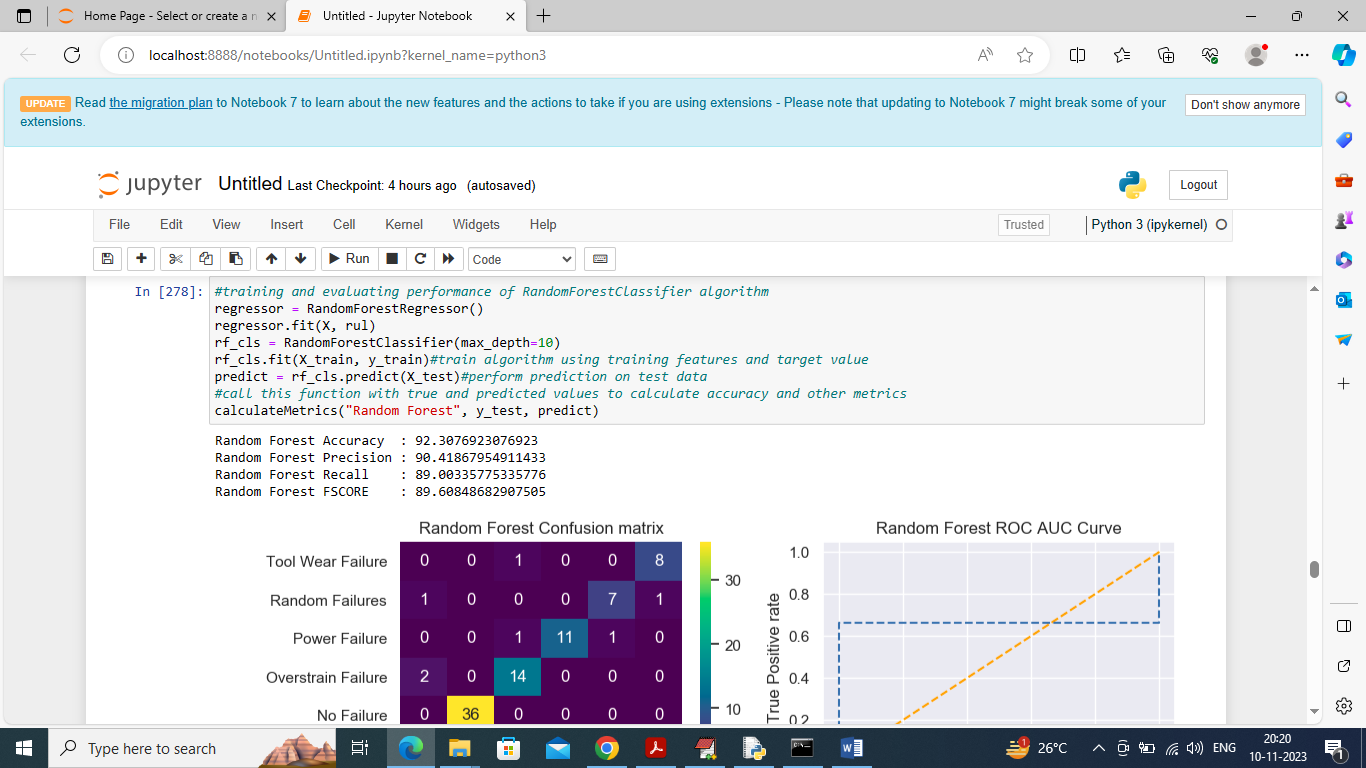
In above screen training SVM algorithm and then performing prediction on test data and after prediction SVM got accuracy as 92% and can see other metrics also and below are the SVM performance graph



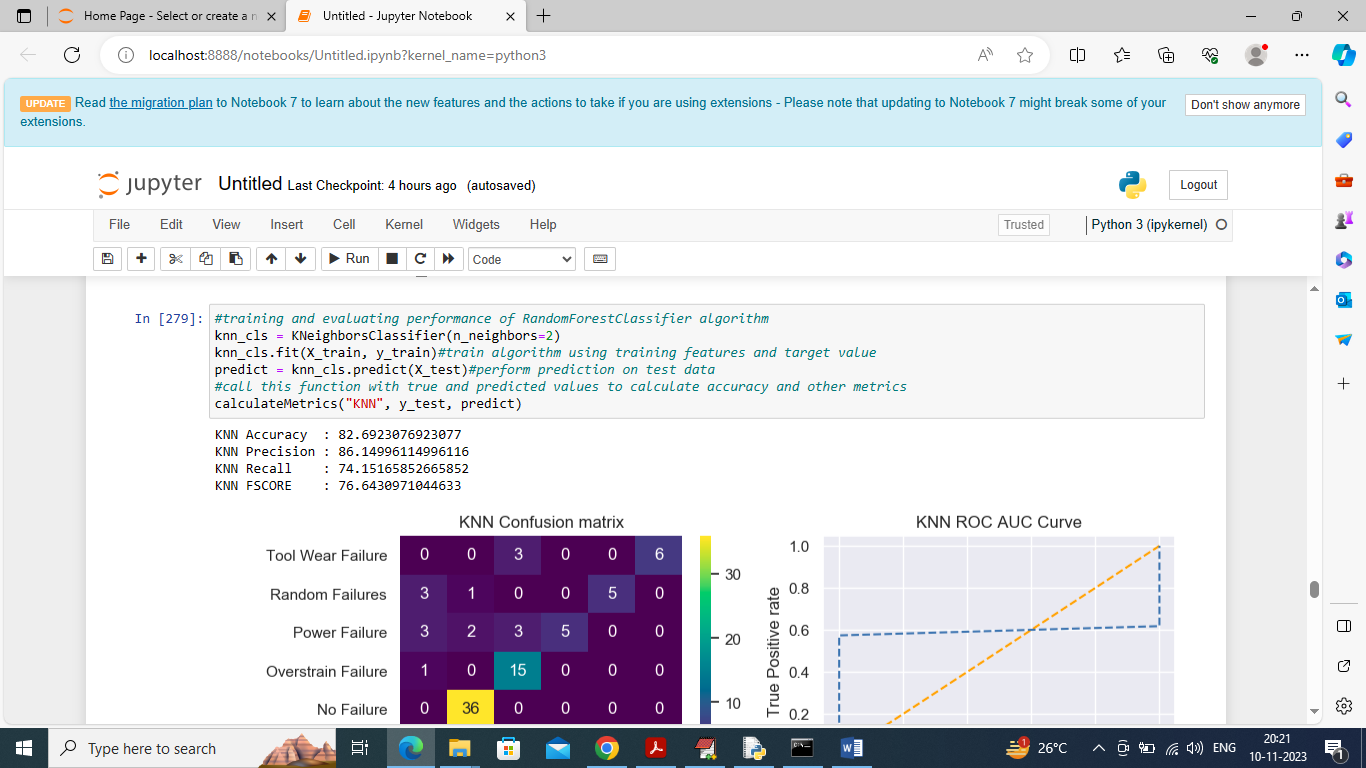
In above confusion matrix graph x-axis represents “Predicted Labels” and y-axis represents “True Labels” and all different color boxes in diagnol represents correct prediction count and remaining all blue boxes contains incorrect prediction count which are very few. In Roc curve graph x-axis represents False Positive Rate and y-axis represents True Positive Rate and if blue line goes below orange line then all predictions are incorrect and if goes above orange line then all predictions are correct and in above ROC graph we can see only few predictions are incorrect



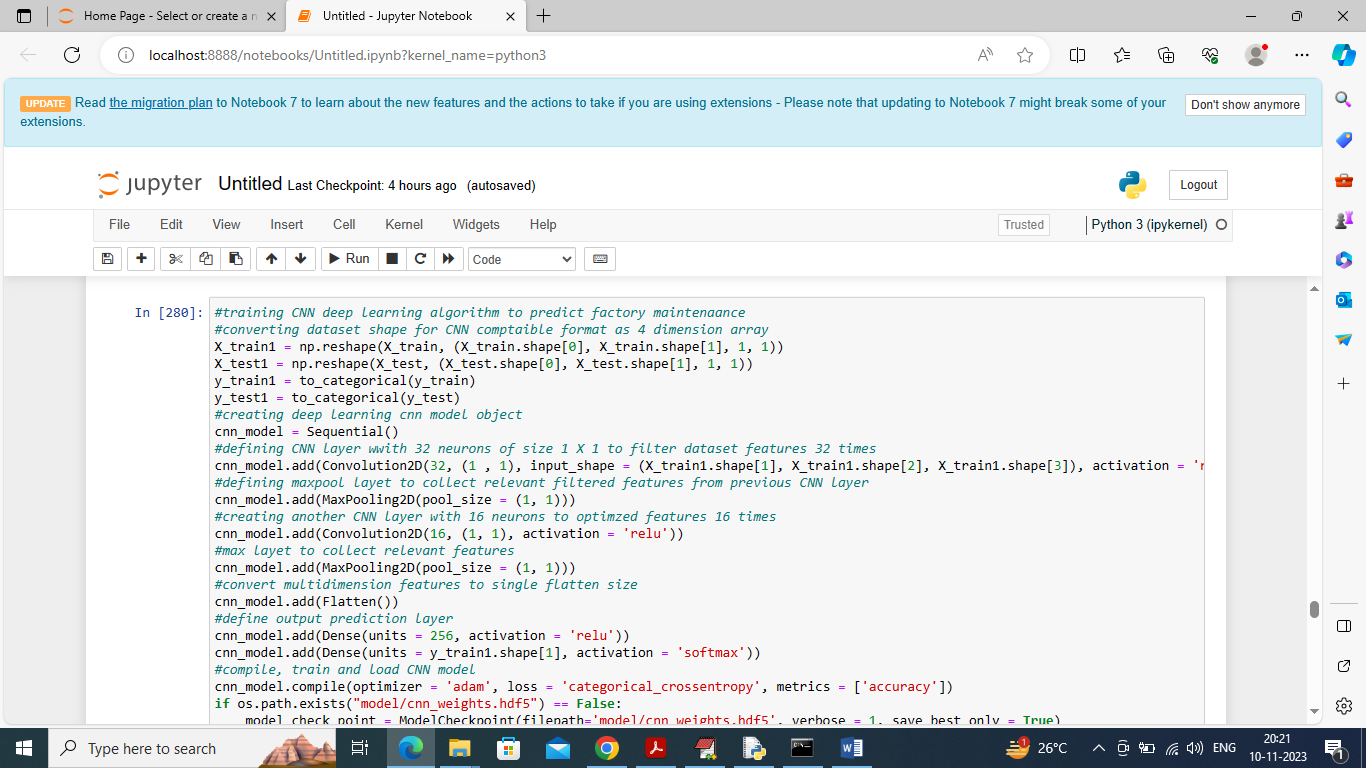
In above screen training Decision Tree algorithm and then it got 88% accuracy and can see other metrics also



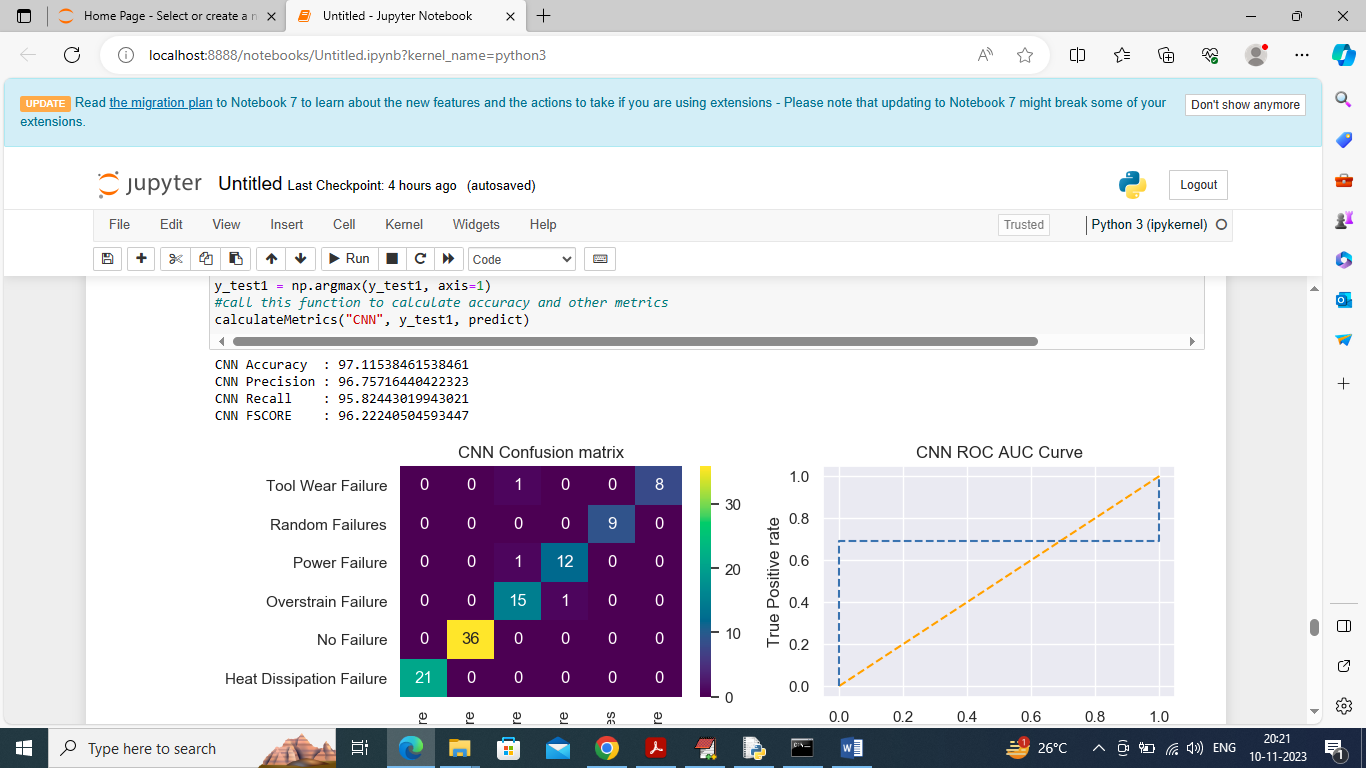
In above screen random forest got 92% accuracy



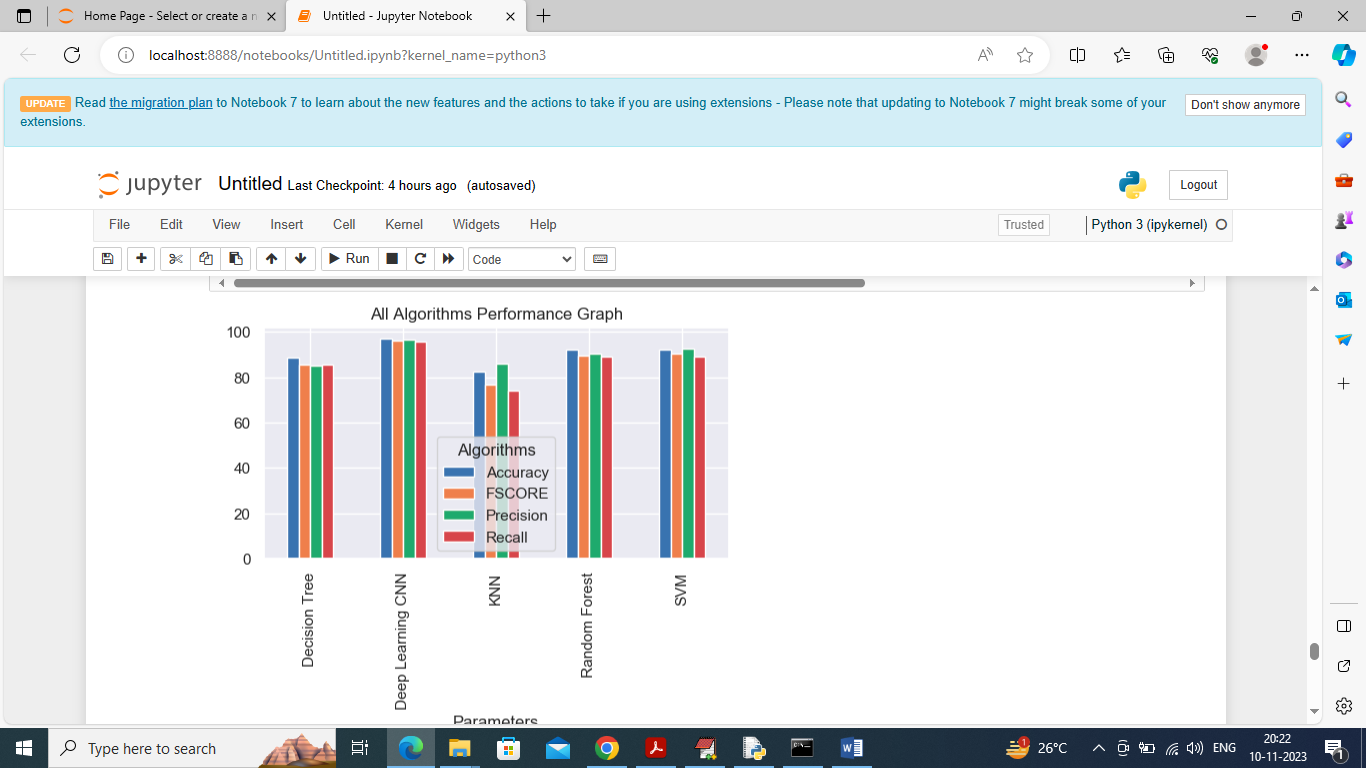
In above screen KNN got 82% accuracy



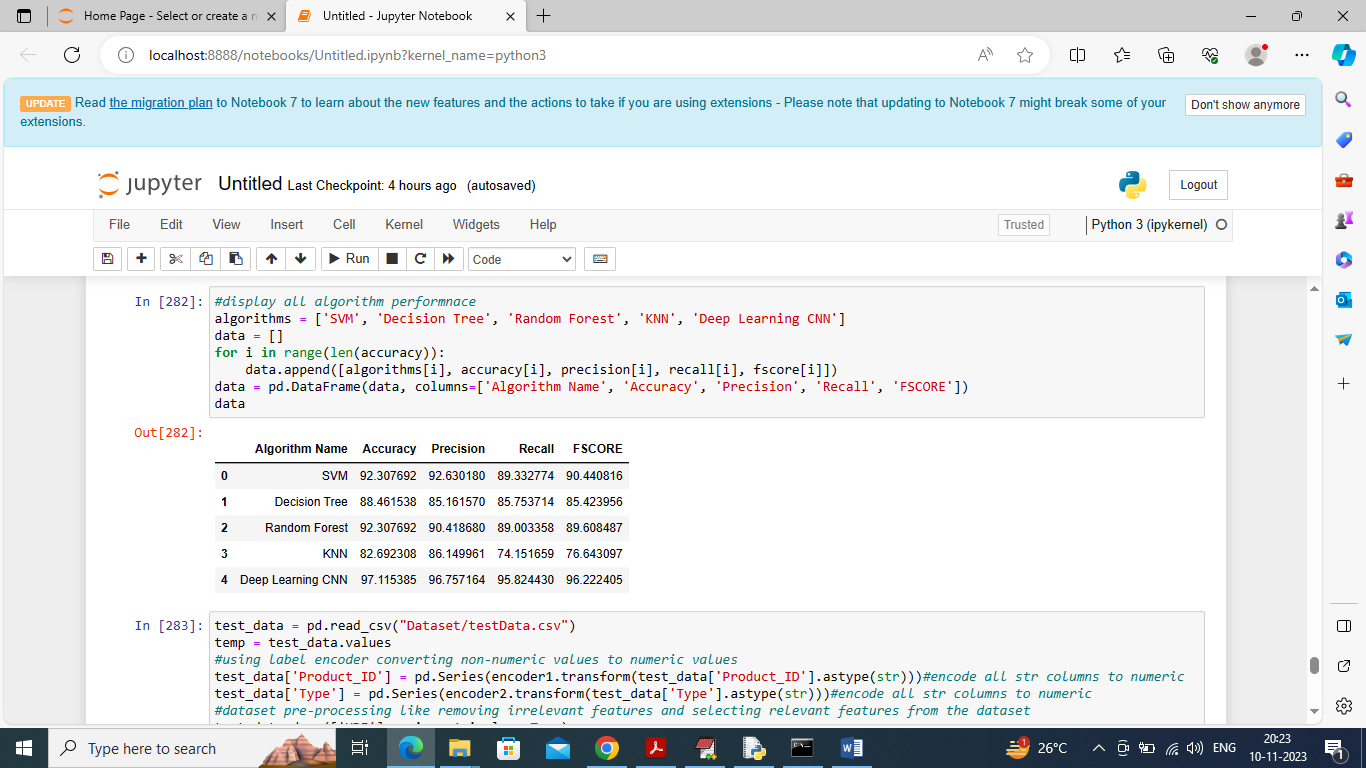
In above screen defining and training deep learning CNN algorithm and after executing above block will get below output



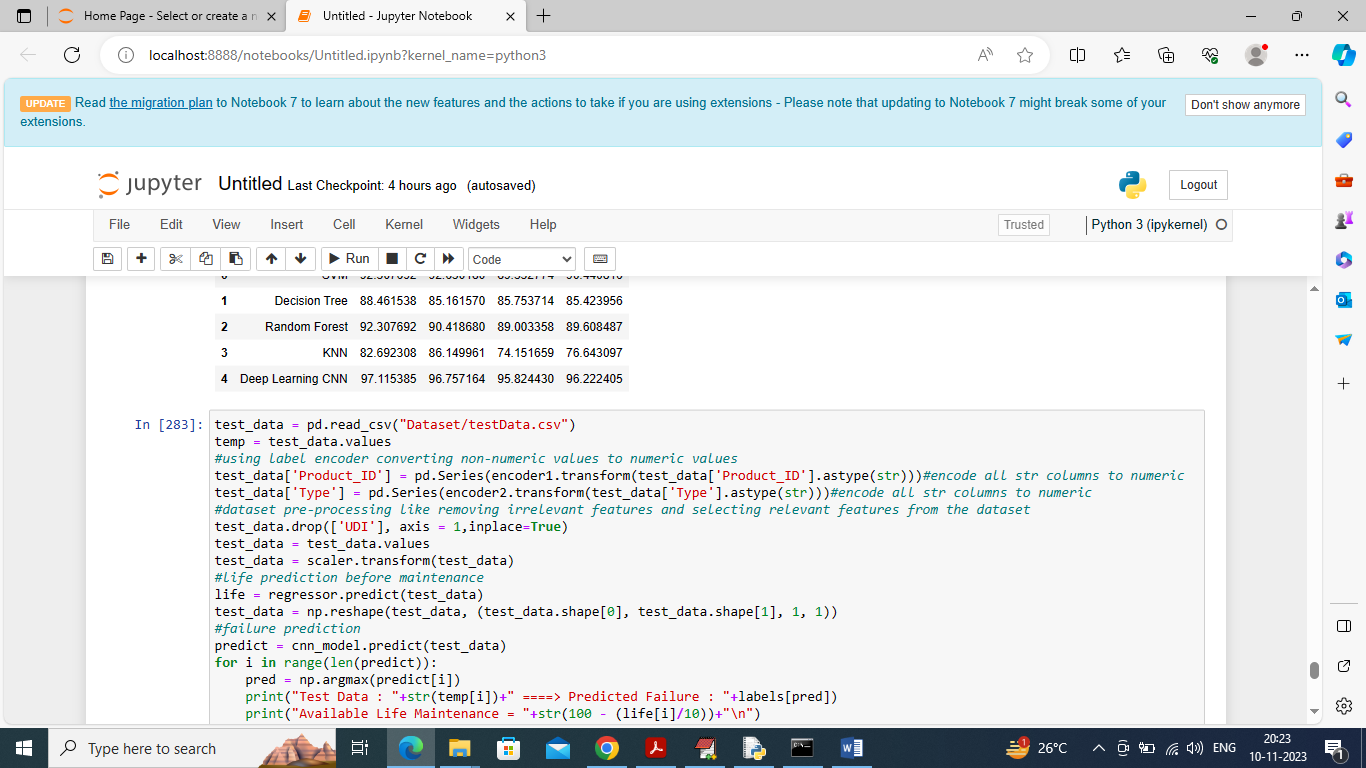
In above screen CNN got 97% accuracy and can see other metrics also



In above graph displaying all algorithm performance where x-axis represents algorithm names and y-axis represents accuracy and other metrics in different color bars and in all algorithms CNN got high performance



In above screen displaying all algorithm performance in tabular format



In above screen defining test Data prediction function which will read test data and then predict Failure Type for maintenance and then suggest available machine life



In above screen in square bracket we can see Sensor Test data and after arrow 🡺 symbol can see predicted Failure and its available life and if life % is less then it maintenance should be schedule sooner. In above prediction for ‘No Failure’ we can see available life is 95% and for other failure we can see life is less and based on life maintenance will be scheduled