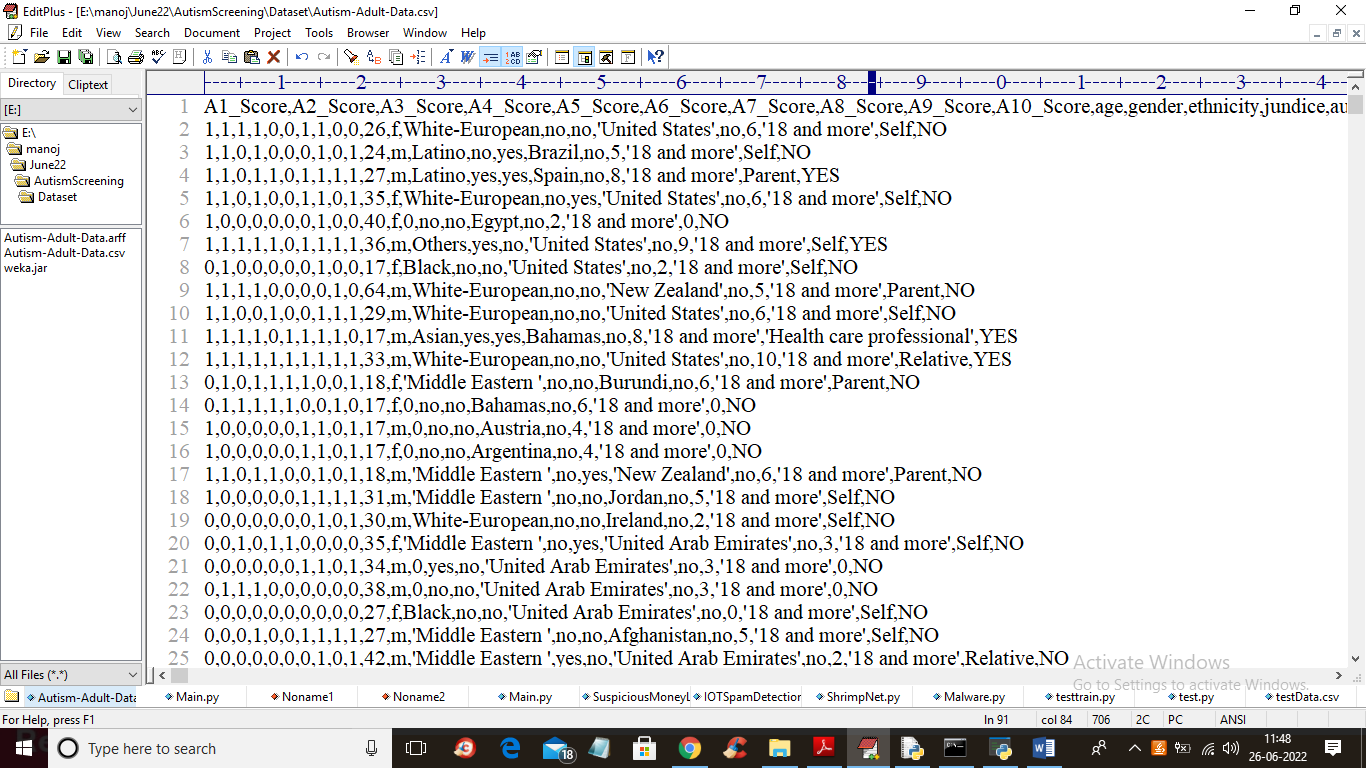
Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques

Autism is a neuro based disorder which effect human brain from childhood to adulthood and this disorder make person to laugh unnecessary, no felling of pain, unable to make eye contact and many more disorder. To detect such disease author is experimenting with various machine learning algorithms such as SVM, KNN, Naïve Bayes, Logistic Regression and deep learning algorithms such as Artificial Neural Networks which trained on single dimensional array and CNN (Convolution Neural Network) which get trained on 2 or multidimensional array. In all algorithms CNN is giving 100% accuracy.

To train above algorithms author has used AUTISM dataset from UCI machine learning and this dataset contains 704 records and 21 columns and each column is associated with class label as NO or YES where yes means autism detected. Below screen showing dataset details



In above dataset screen first row contains dataset column names and remaining rows contains dataset values and each row is associated with class label as ‘NO or YES’. In dataset we have columns such as patient age and some questions and answers score such as patient is responding or not.

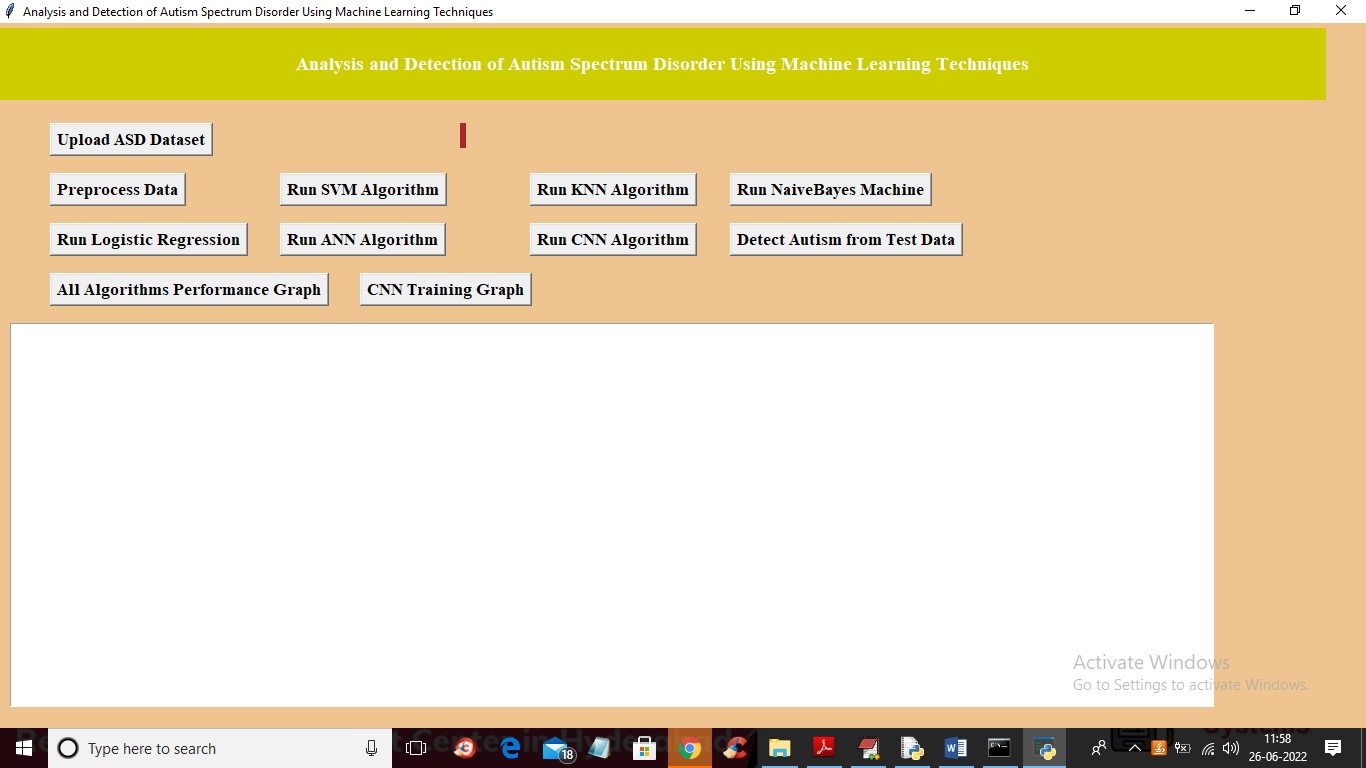
By using above dataset author training all machine and deep learning algorithms and evaluating their performance in terms of accuracy, sensitivity, precision etc.

To implement this project we have designed following modules

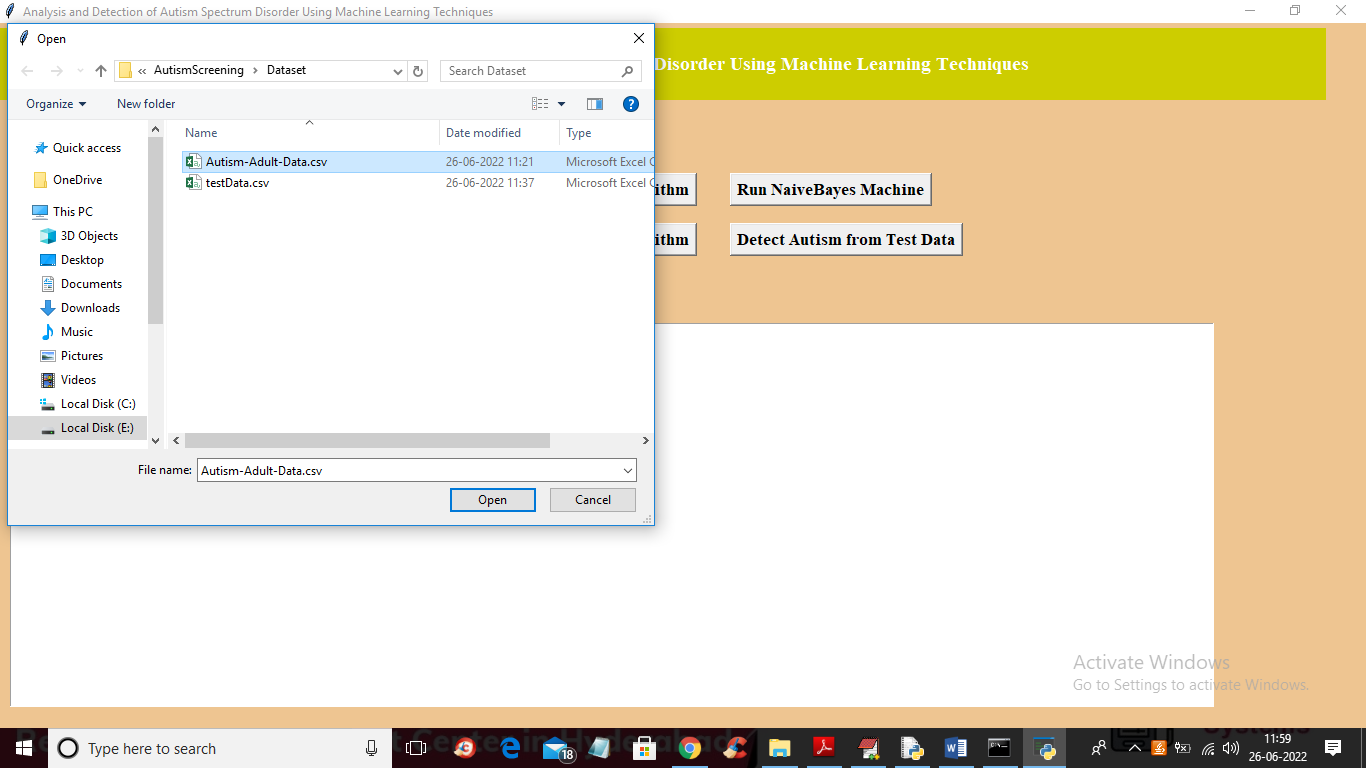
1. Upload ASD Dataset: using this module we will upload dataset to application
2. Preprocess Data: using this module we will read entire dataset and then replace missing values with 0 and then convert all non-numeric values to numeric by using LABEL ENCODING Algorithm as this algorithm will assigned unique integer ID to non-numeric values. After processing we will split dataset into train and test where application used 80% dataset for training and 20% dataset for testing
3. Run SVM Algorithm: now processed train data will be input to SVM algorithm to trained prediction model and this model will be applied on 20% test data to compute SVM prediction accuracy.
4. Run KNN Algorithm: now processed train data will be input to KNN algorithm to trained prediction model and this model will be applied on 20% test data to compute KNN prediction accuracy.
5. Run Naïve Bayes Algorithm: now processed train data will be input to Naïve Bayes algorithm to trained prediction model and this model will be applied on 20% test data to compute Bayes prediction accuracy.
6. Run Logistic Regression Algorithm: now processed train data will be input to LR algorithm to trained prediction model and this model will be applied on 20% test data to compute LR prediction accuracy.
7. Run ANN Algorithm: now processed train data will be input to ANN algorithm to trained prediction model and this model will be applied on 20% test data to compute ANN prediction accuracy.
8. Run CNN Algorithm: now processed train data will be input to CNN algorithm to trained prediction model and this model will be applied on 20% test data to compute CNN prediction accuracy.
9. Detect Autism from Test Data: using this module we will upload test data and then CNN will predict weather test data is normal or contains Autism disorder
10. All Algorithms Performance Graph: using this module we will plot accuracy graph of all algorithms
11. CNN Training Graph: using this module we will plot CNN accuracy and loss graph of training

SCREEN SHOTS

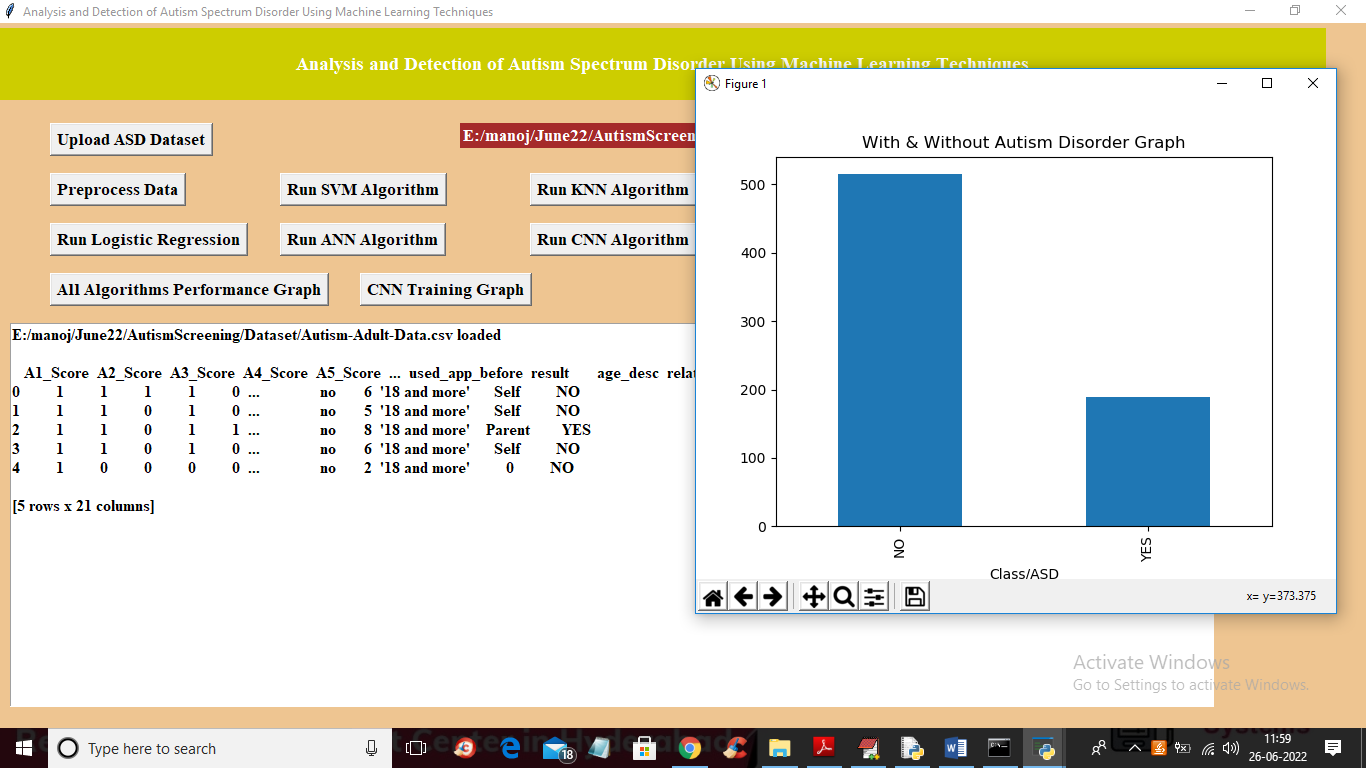
To run project double click on ‘run.bat’ file to get below screen



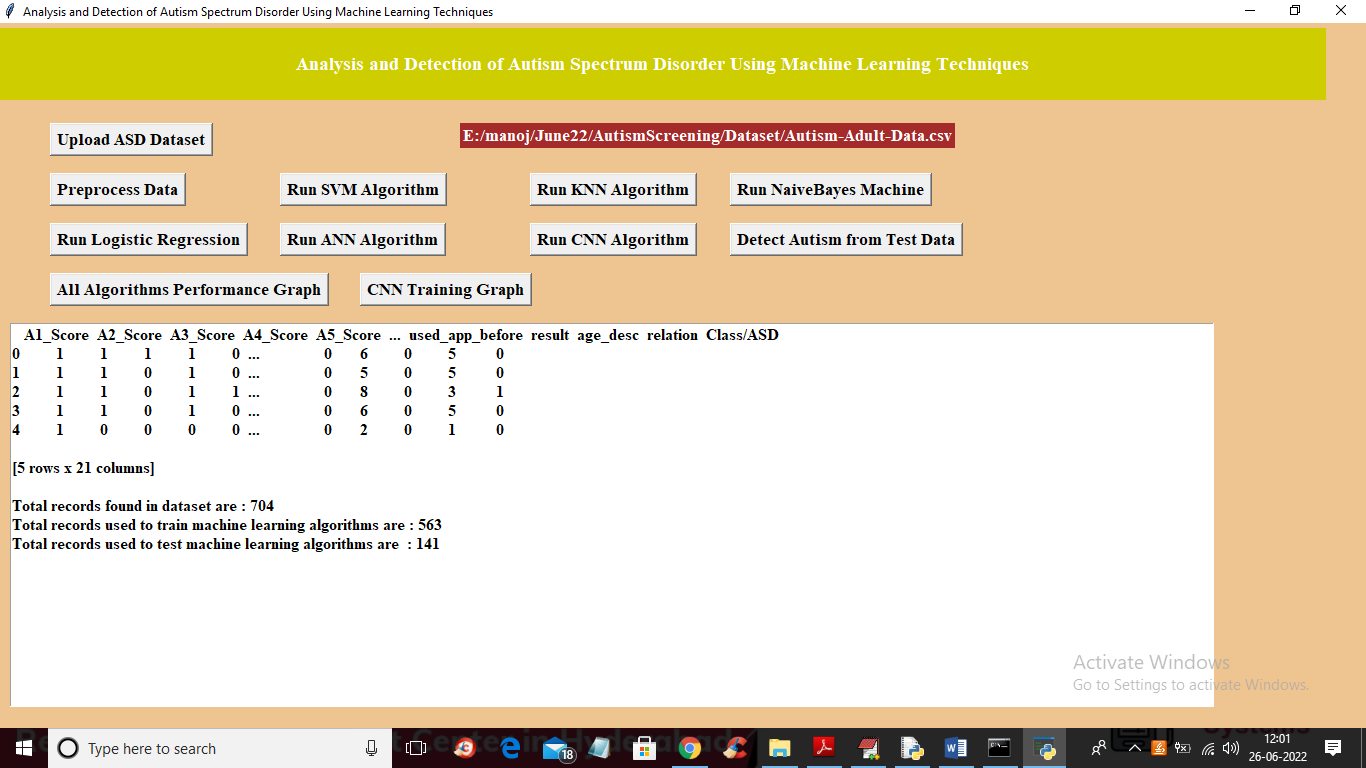
In above screen click on ‘Upload ASD Dataset’ button to upload dataset and get below output



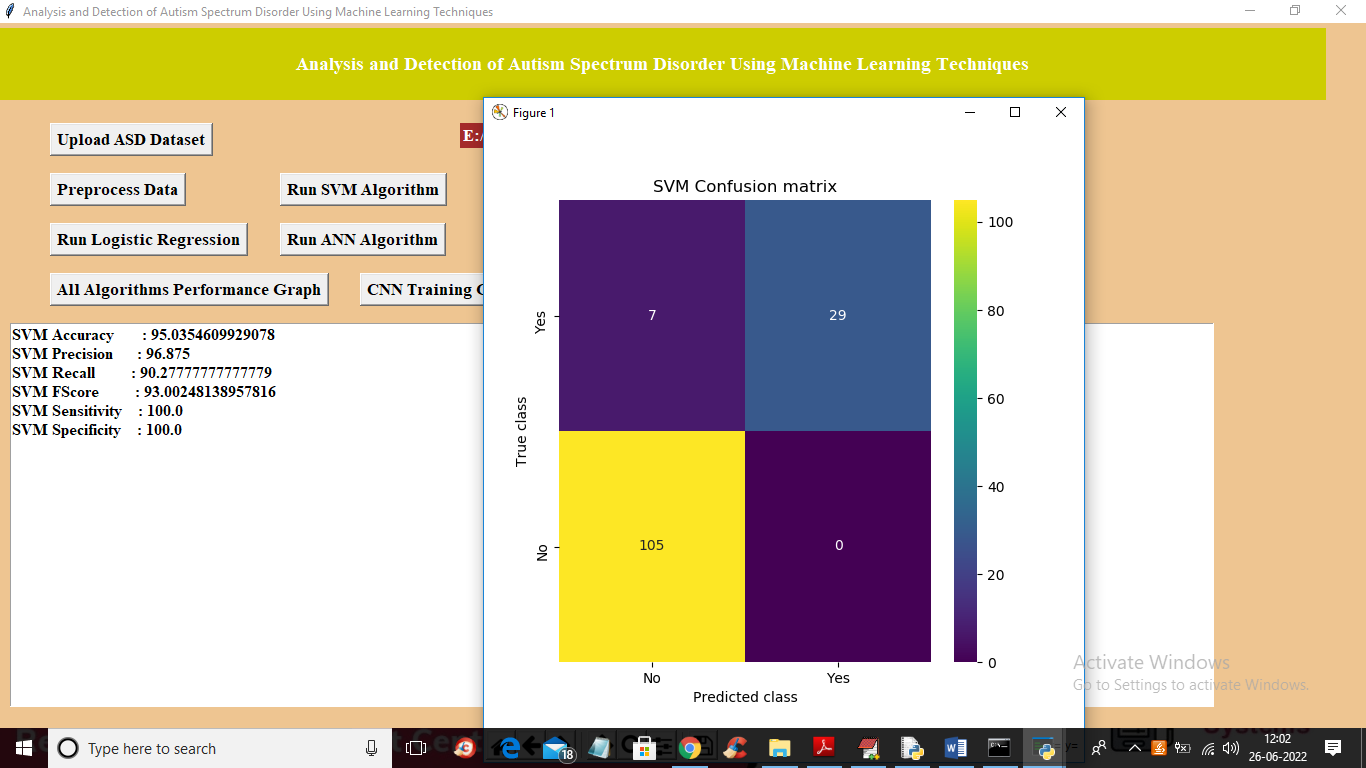
In above screen selecting and uploading ‘Autism’ dataset and then click on ‘Open’ button to get below output



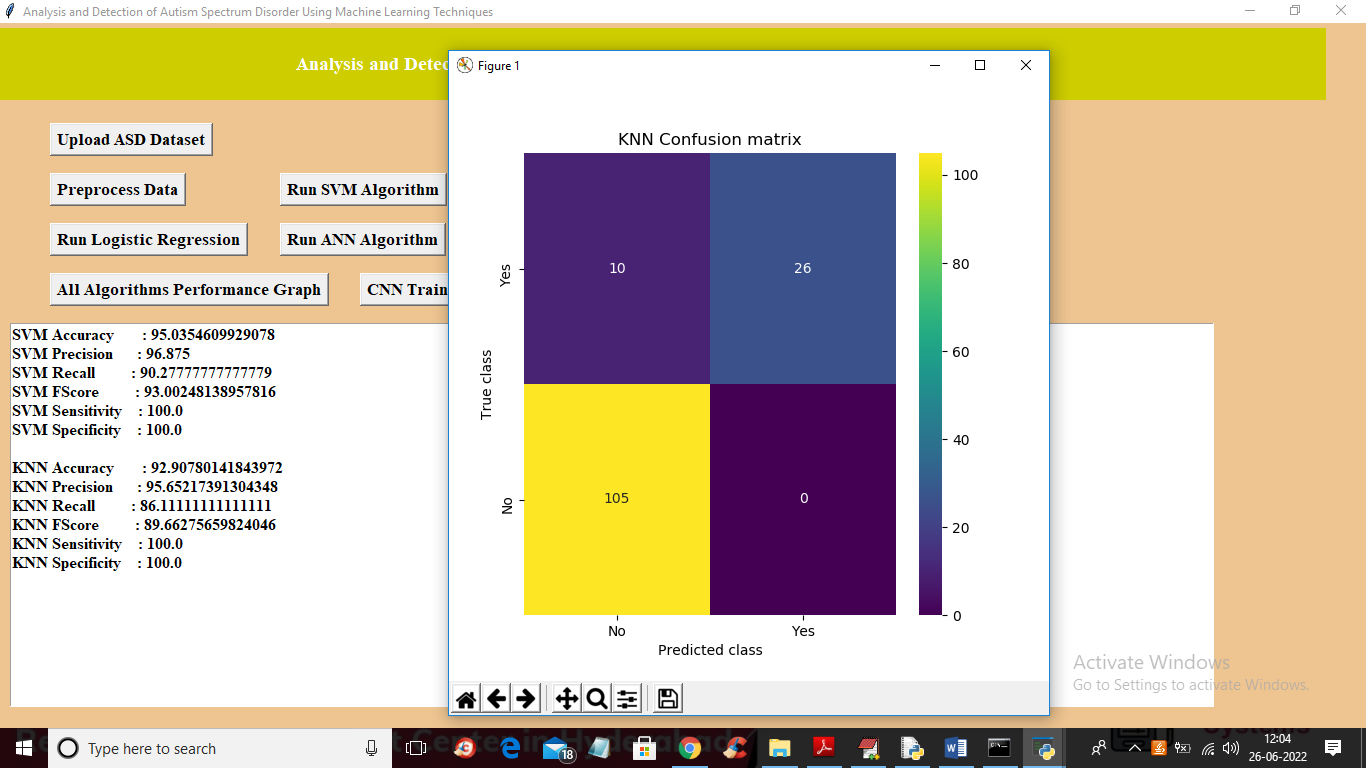
In above screen dataset loaded and we can see dataset contains some missing and non-numeric values so we need to process dataset to replace missing values and non-numeric values and in above graph we are showing number of ‘YES and NO’ patients found in dataset where x-axis contains labels and y-axis contains count and now close above graph and then click on ‘Preprocess Data’ button to get below output



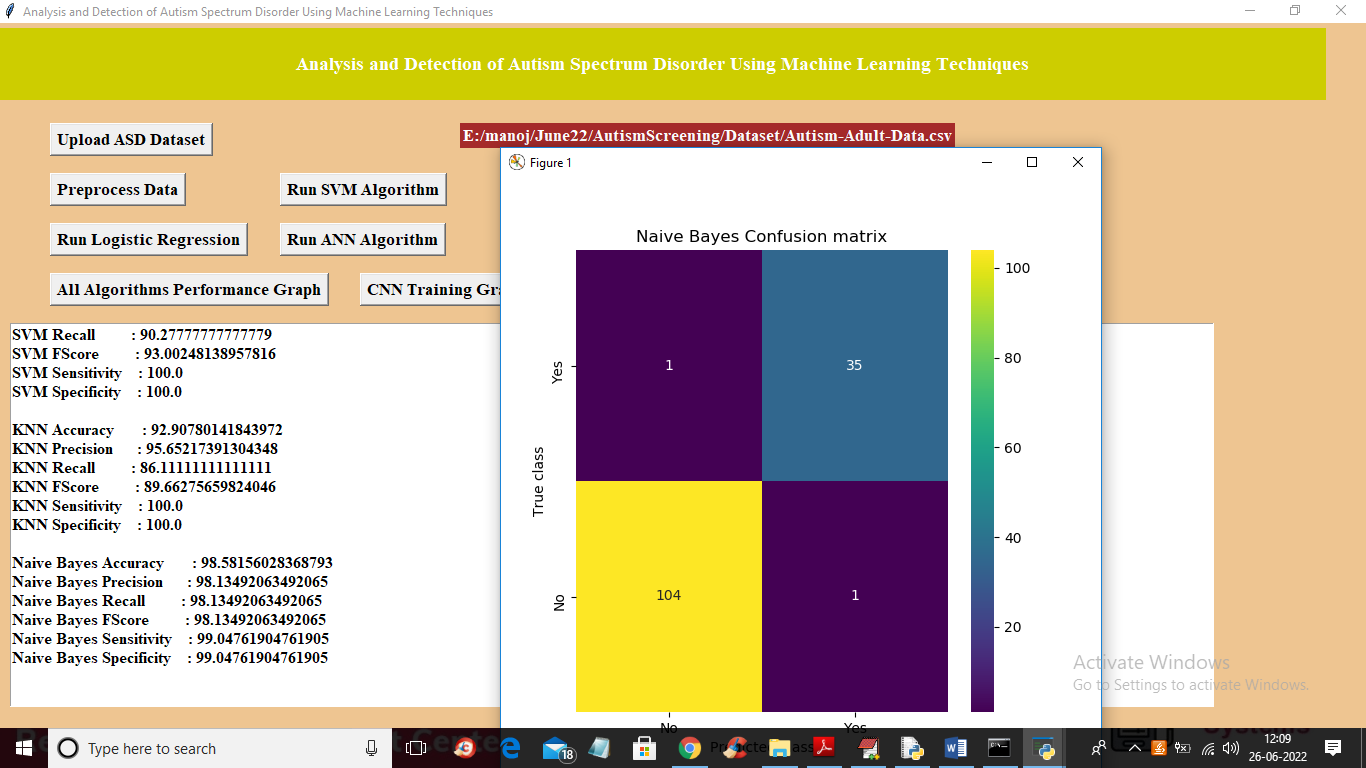
In above screen we can see all data is changed to numeric values and then we can see total dataset size with train and test split details and now dataset is ready with train and test part so click on ‘Run SVM Algorithm’ button to train SVM and get below output



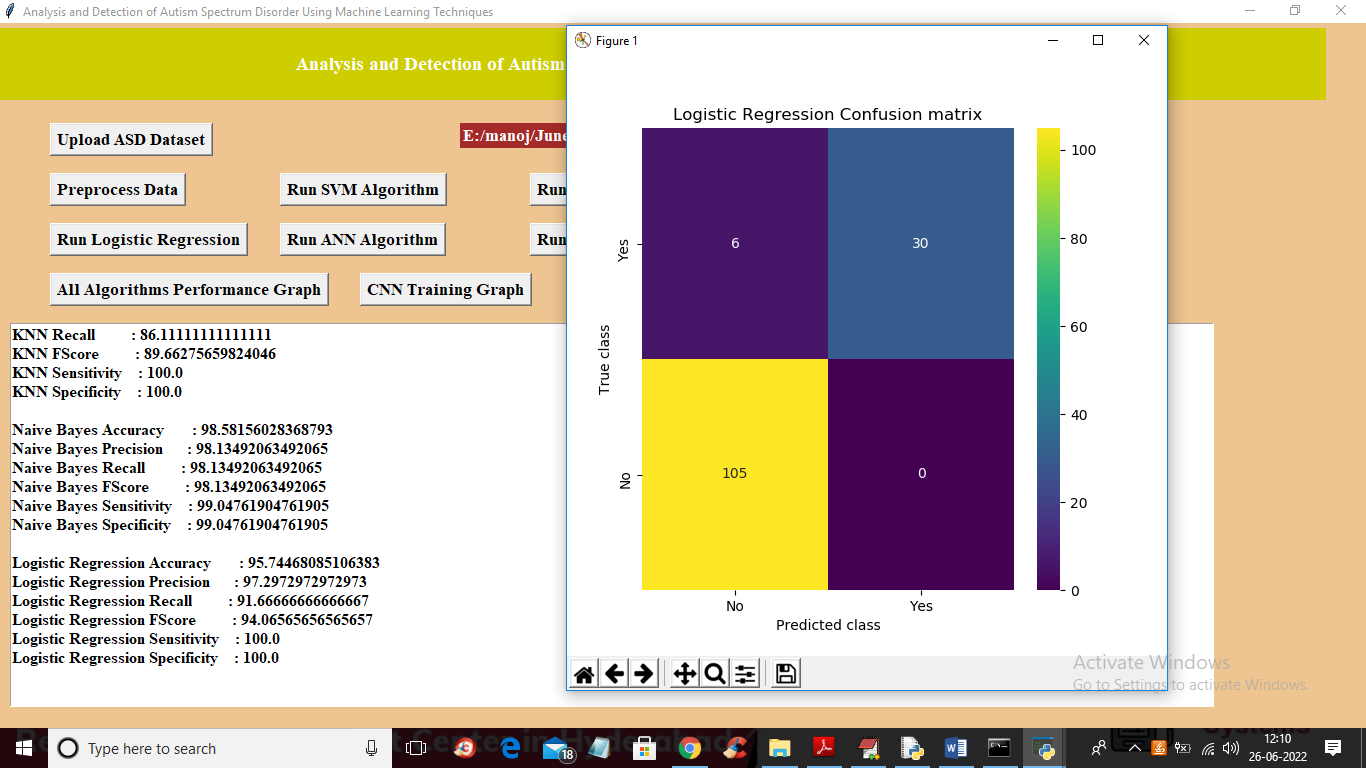
In above screen with SVM we got 95% accuracy and in SVM confusion matrix graph we can see x-axis represents Predicted labels and y-axis represents Test label and we can see only 7 records are incorrectly predicted and 105 and 29 records are correctly predicted. Now close above graph and then click on ‘Run KNN Algorithm’ button to get below output



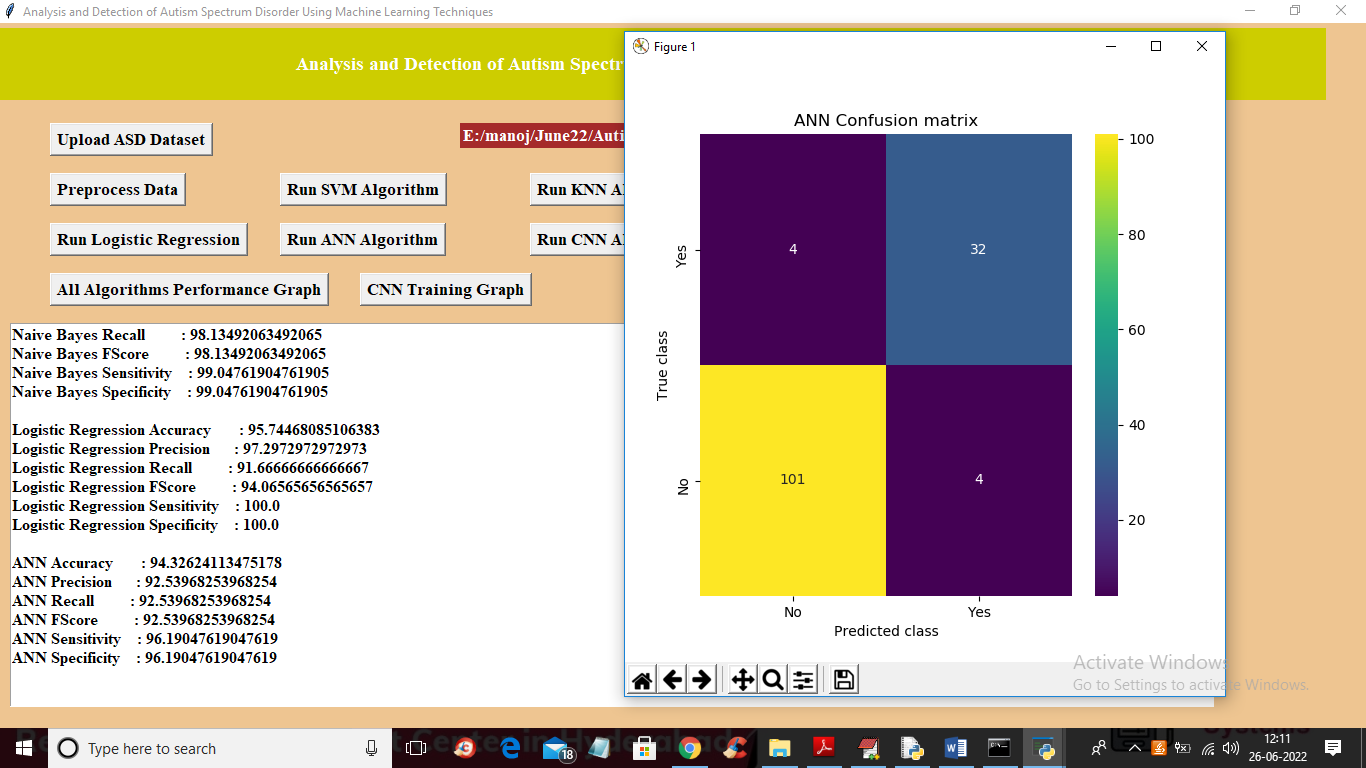
In above screen with KNN we got 92% accuracy and in confusion matrix graph we can see KNN predicting 10 records incorrectly. Now close above graph and then click on ‘Run Naïve Bayes’ to train Naïve Bayes and get below output



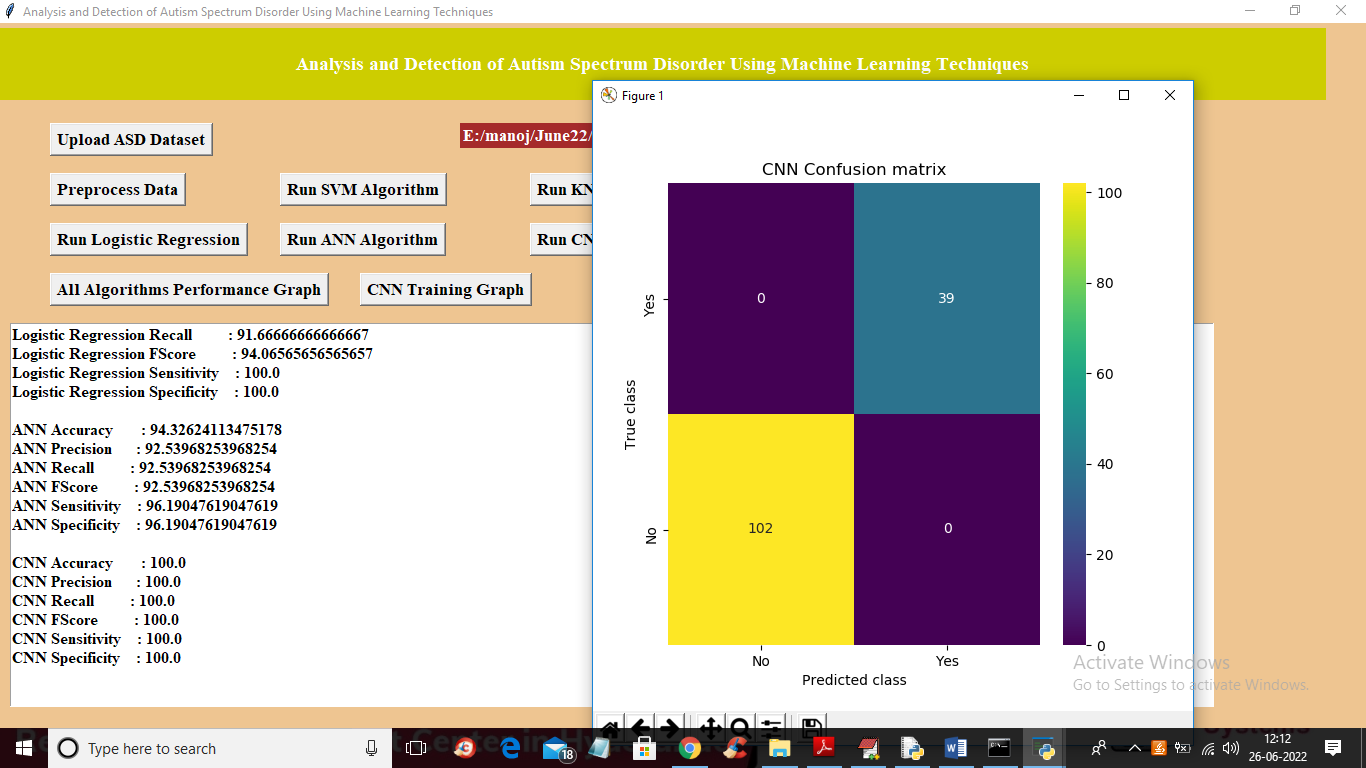
In above screen with Naïve Bayes we got 98% accuracy and in confusion matrix Naïve Bayes predicted only 1 record incorrectly in NO and 1 in YES so 2 records are wrongly predicted. Now close above graph and then click on ‘Run Logistic Regression’ button to train it



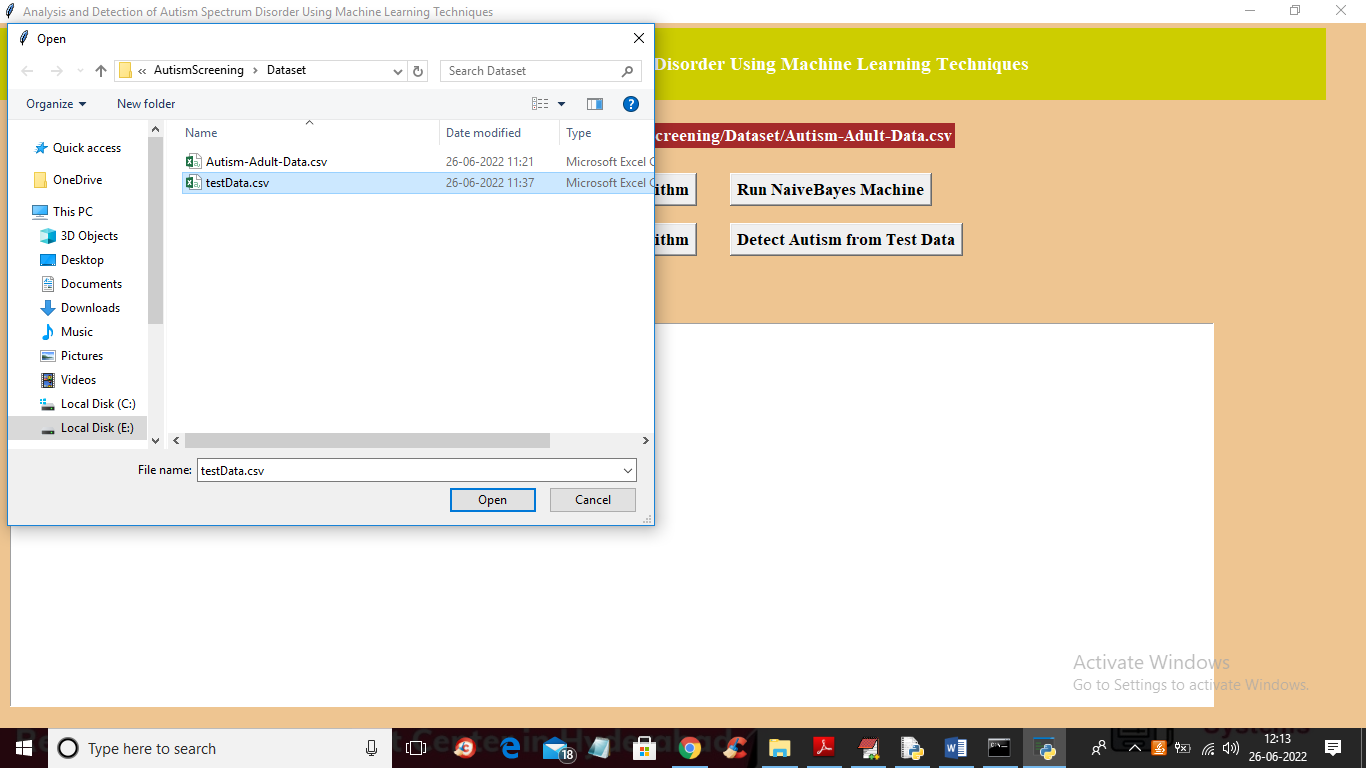
In above screen with logistic regression we got 95% accuracy and in graph we can see its predicted 6 records wrong and now close above graph and then click on ‘Run ANN Algorithm’ to get below output



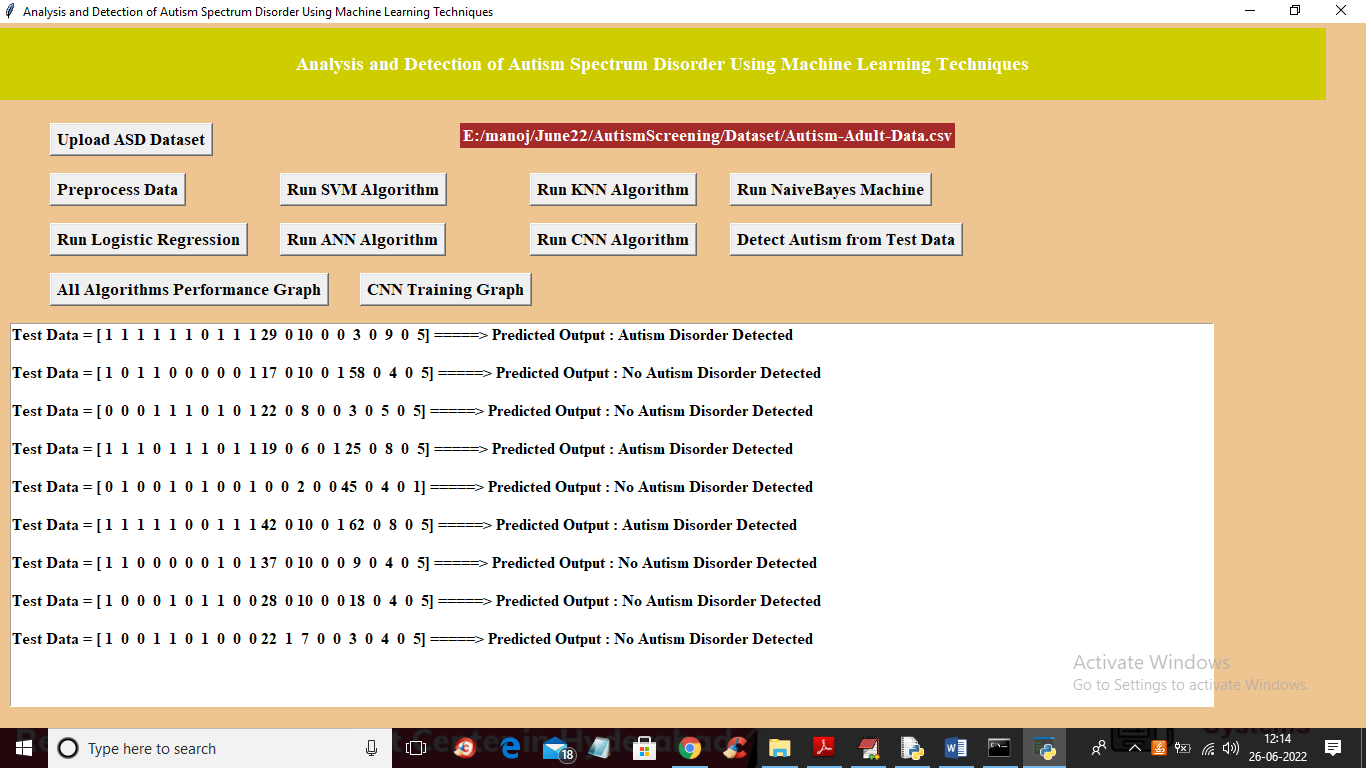
In above screen with ANN we got 94% accuracy and its predicted 8 records incorrectly for both YES and No 4 and 4 and now close above graph and then click on ‘Run CNN Algorithm’ button to get below output



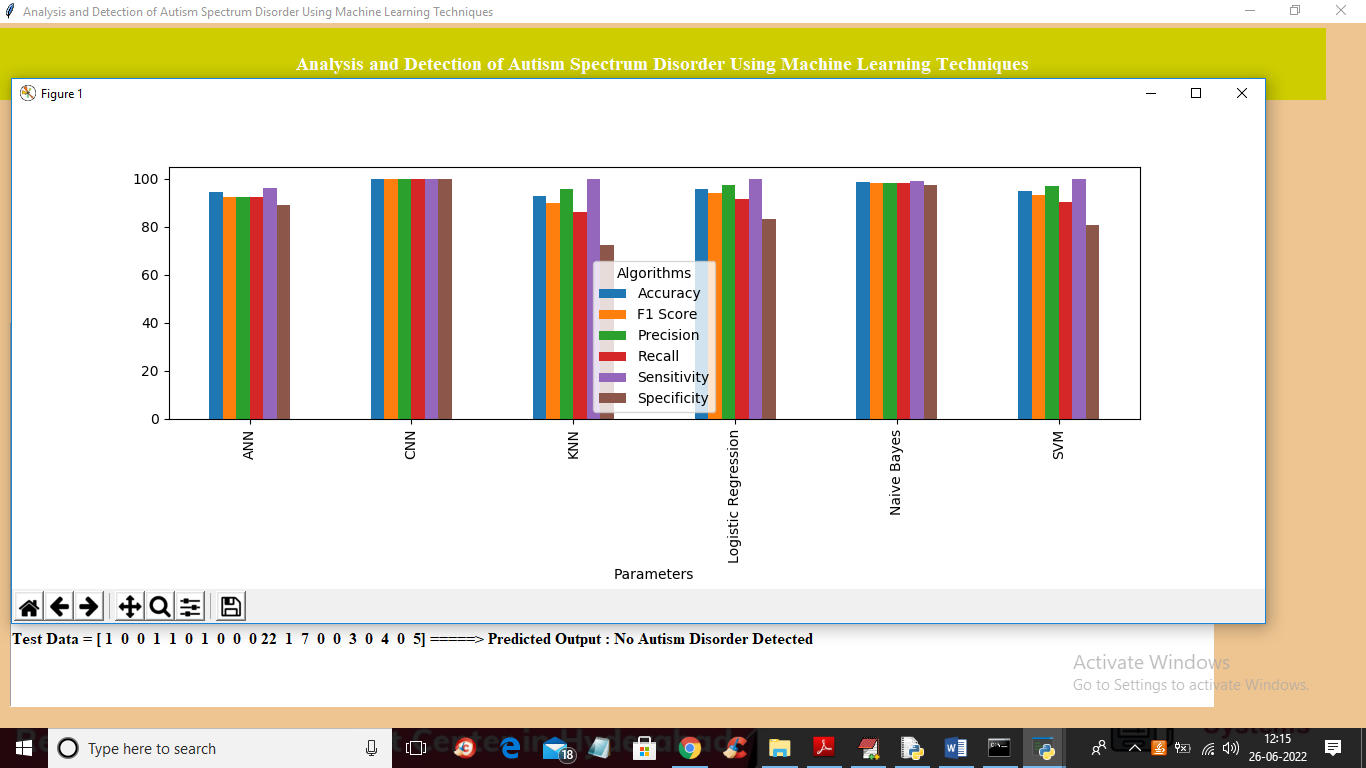
In above screen with CNN we got 100% accuracy and in graph we can see 0 records are incorrectly predicted and now close above graph and then click on ‘Detect Autism from Test Data’ button to upload test data and then predict autism



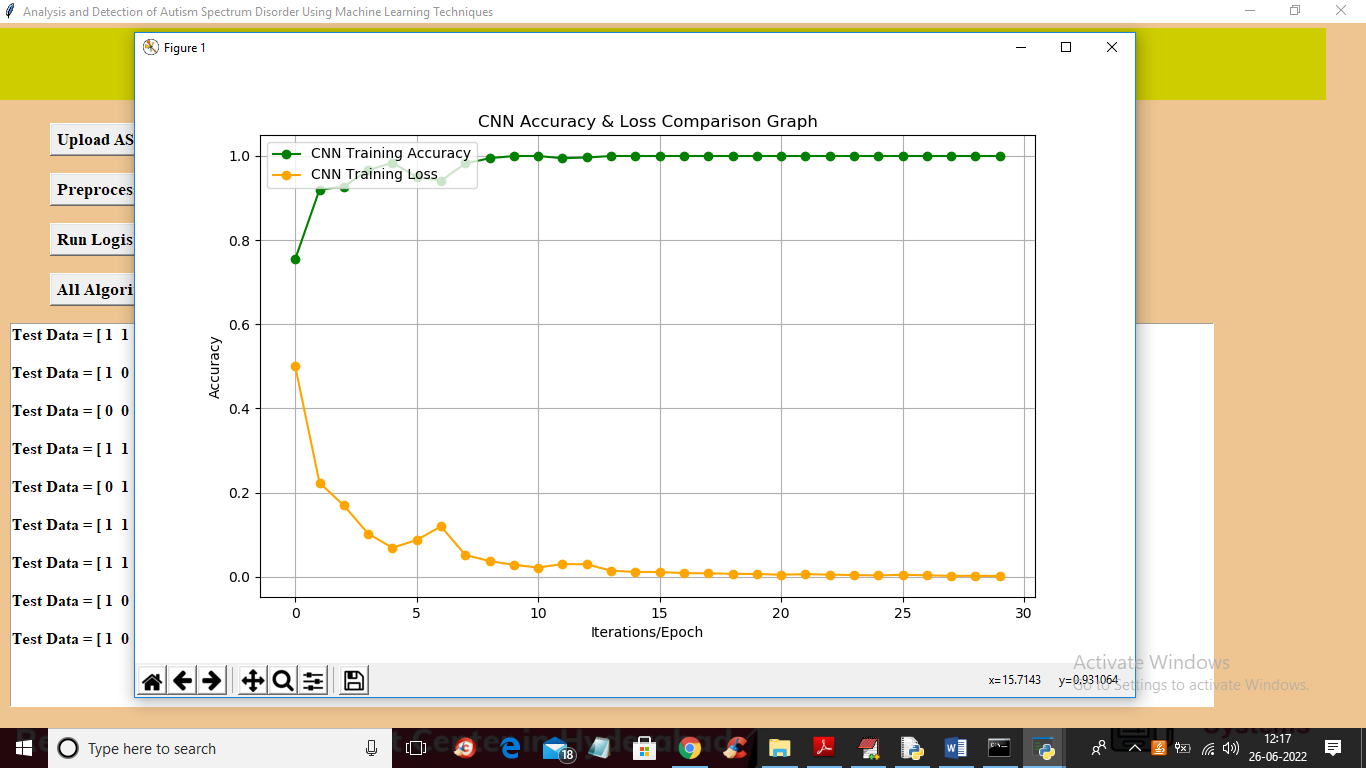
In above screen selecting and uploading ‘testData.csv’ and then click on ‘Open’ button to get below output



In above screen in square bracket we can see TEST data values and after =🡺 arrow symbol we can see predicted output as ‘Autism detected or not’ and now click on ‘All Algorithms Performance Graph’ button to get below graph



In above graph x-axis represents algorithms names and y-axis represents accuracy and other metric values where different colour bar represents different metrics such as precision, accuracy, sensitivity etc. in above graph we can see in all algorithms CNN is giving high performance and now close above graph and then click on ‘CNN Training Graph’ button to get below graph



In above graph x-axis represents training epoch and y-axis represents accuracy and loss values and in above graph green line represents ACCURACY and yellow line represents loss and we can see with each increasing epoch accuracy got increase and loss got decrease. Any model with increasing accuracy and decreasing loss will be consider as best algorithm or model