Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications

In this paper author is combining multiple algorithms such as intelligence of Reinforcement learning and deep recurrent neural networks (RNN) to forecast crop yield prediction. To predict crop yield many existing machine learning and deep learning algorithms such as CNN and RNN are used but this algorithms are not accurate in prediction so author combining both Deep Reinforcement and Deep learning to form new algorithm called DQN (deep Q learning Network). Proposed algorithms able to achieve high accuracy with less MSE error.

Combining the intelligence of reinforcement learning and deep learning, deep reinforcement learning builds a complete crop yield prediction framework that can map the raw data to the crop prediction values. The proposed work constructs a Deep Recurrent Q-Network model which is a Recurrent Neural Network deep learning algorithm over the Q-Learning reinforcement learning algorithm to forecast the crop yield. The sequentially stacked layers of Recurrent Neural network is fed by the data parameters. The Q- learning network constructs a crop yield prediction environment based on the input parameters. A linear layer maps the Recurrent Neural Network output values to the Q-values. The reinforcement learning agent incorporates a combination of parametric features with the threshold that assist in predicting crop yield. Finally, the agent receives an aggregate score for the actions performed by minimizing the error and maximizing the forecast accuracy.

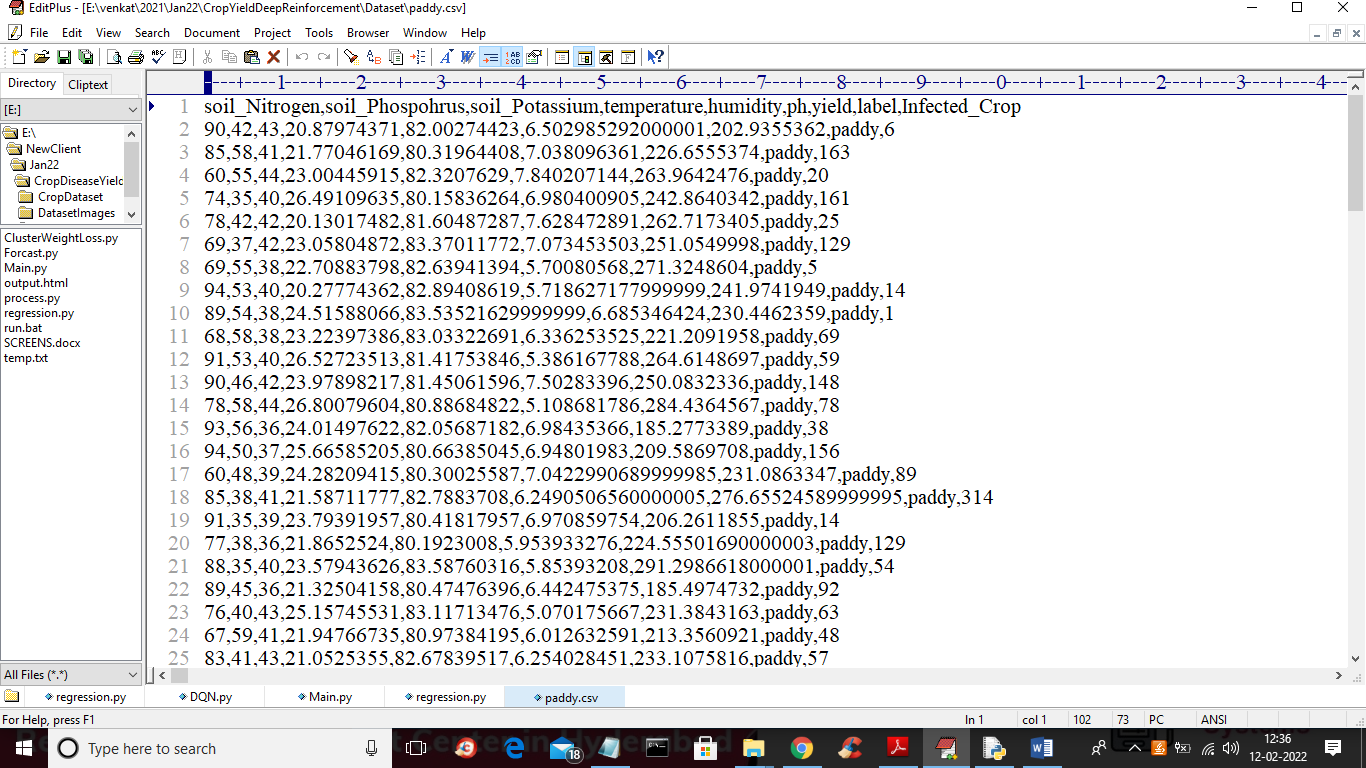
DQN algorithms works like human brain intelligence if brain decision is accurate then we will reward decision else apply penalty and DQN will continuously call RNN model to predict crop yield and if prediction accuracy is high then reward will be applied and if prediction accuracy more then penalty will be applied and this continues till we get best accuracy result.

DQN algorithm works on 3 variables such as ENVIRONMENT or AGENT, STATE and ACTION.

Environment refers to application area and state refers to paddy dataset and then ACTION will perform prediction on paddy dataset and if prediction accuracy then ACTION will raised reward else raised penalty.

In propose DQN algorithm RNN get pre-trained on paddy dataset and then DQN will call RNN model to improve its prediction accuracy by applying random test data and this continues till no more accuracy can be enhanced.

To implement this project we have used paddy dataset which contains soil information like available nitrogen, phosphorus and potassium and then using environment data such as HUMIDTY, TEMPERATURE etc. Below is the dataset screen shots

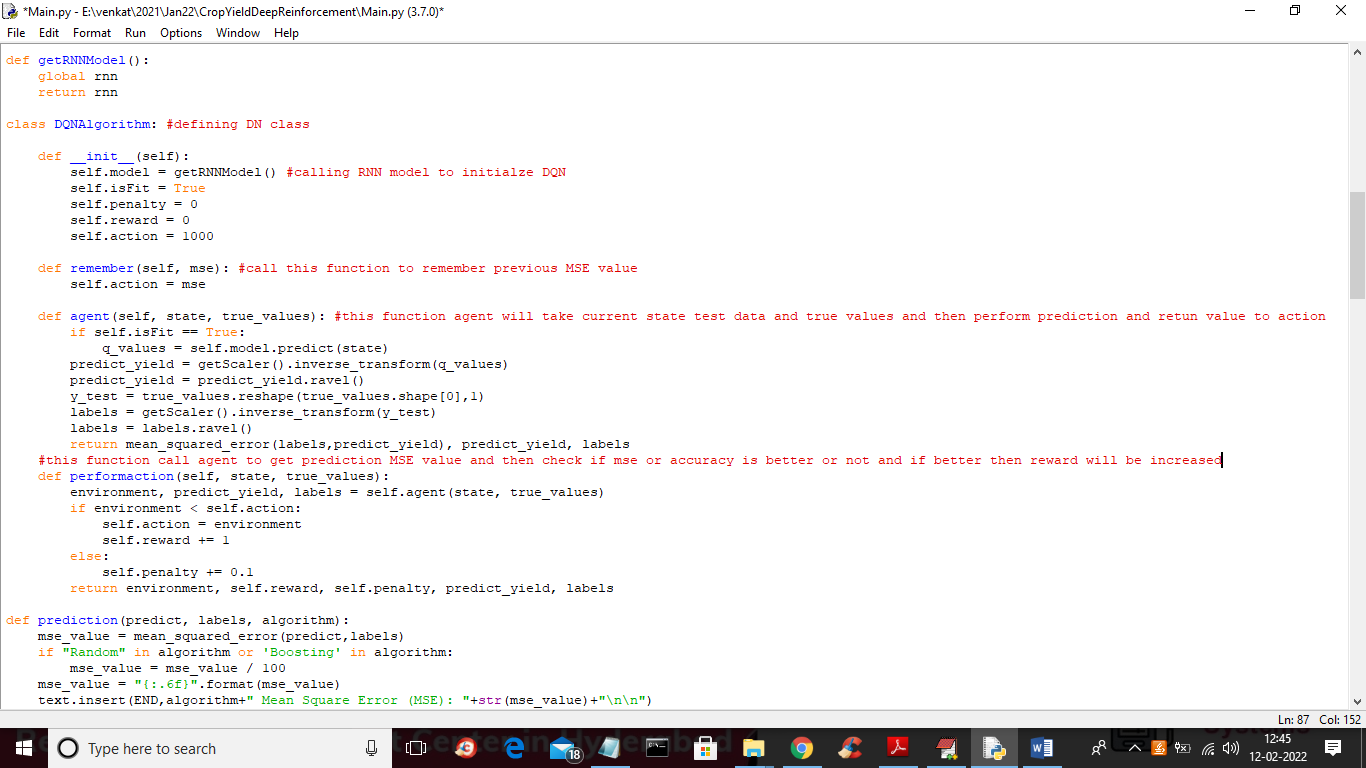


In above dataset first column contains dataset column names and remaining are the dataset values. We will use above dataset to train DQN algorithm

To implement this project we have designed following modules

1. Upload Paddy Crop Dataset: using this module we will upload dataset to application
2. Preprocess Dataset: in above dataset we can see we have numeric and non-numeric values and deep learning algorithms only take numeric data so we need to process dataset to remove missing and non-numeric values so by using this model we will remove such data
3. Train RNN Algorithm: now process data will be input to RNN algorithm to trained crop yield prediction model
4. Run Proposed DQN Model: this algorithm environment/AGENT function will initialize RNN model and then its STATE will read test data and call ACTION to perform prediction and based on predicted accuracy rewards or penalty will be applied
5. Run Random Forest Algorithm: using this module we will trained existing Random Forest algorithm and calculate its accuracy, MSE and compare with propose DQN algorithm
6. Run Gradient Boosting Algorithm: using this module we will trained existing Gradient Boosting algorithm and calculate its accuracy, MSE and compare with propose DQN algorithm
7. MSE Comparison Graph: using this module we will plot MSE graph of all algorithms and then lower the MSE the better is the algorithm

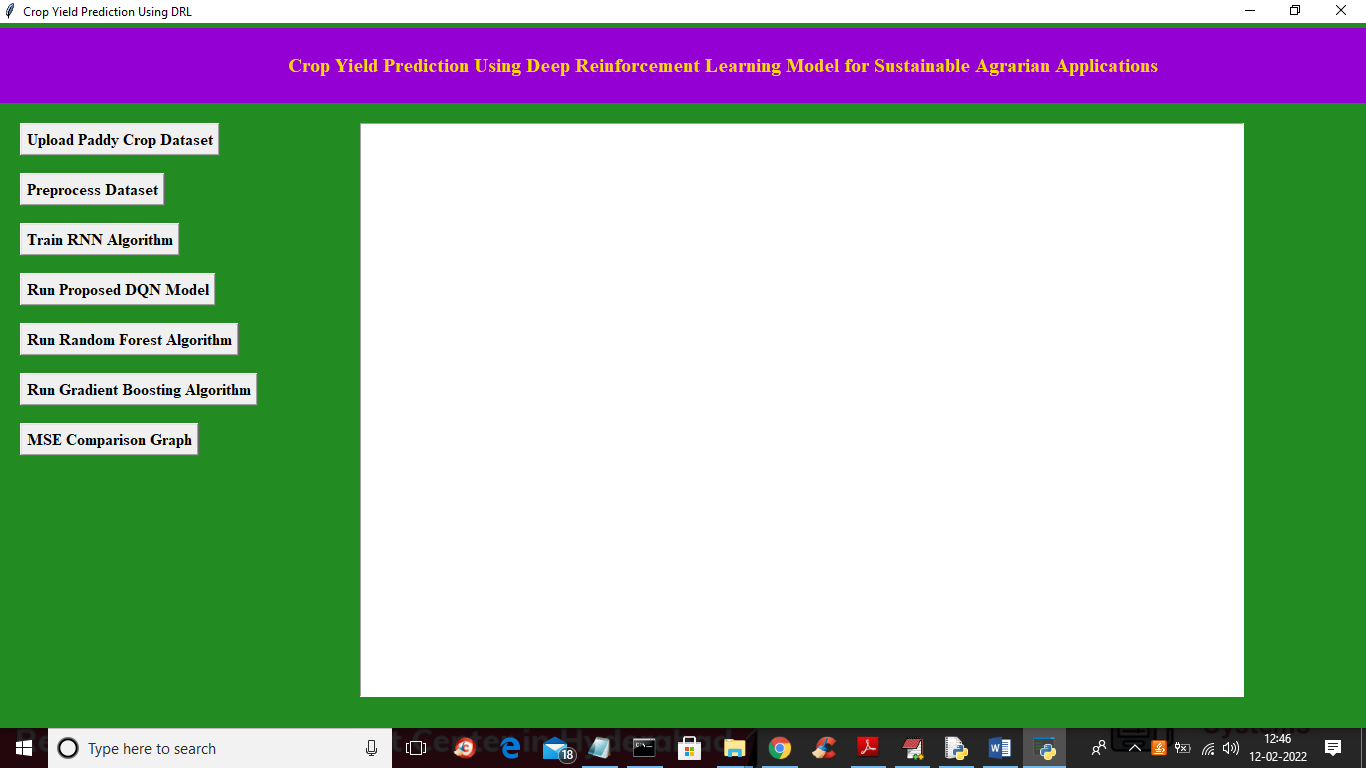
In below screen I am showing DQN code



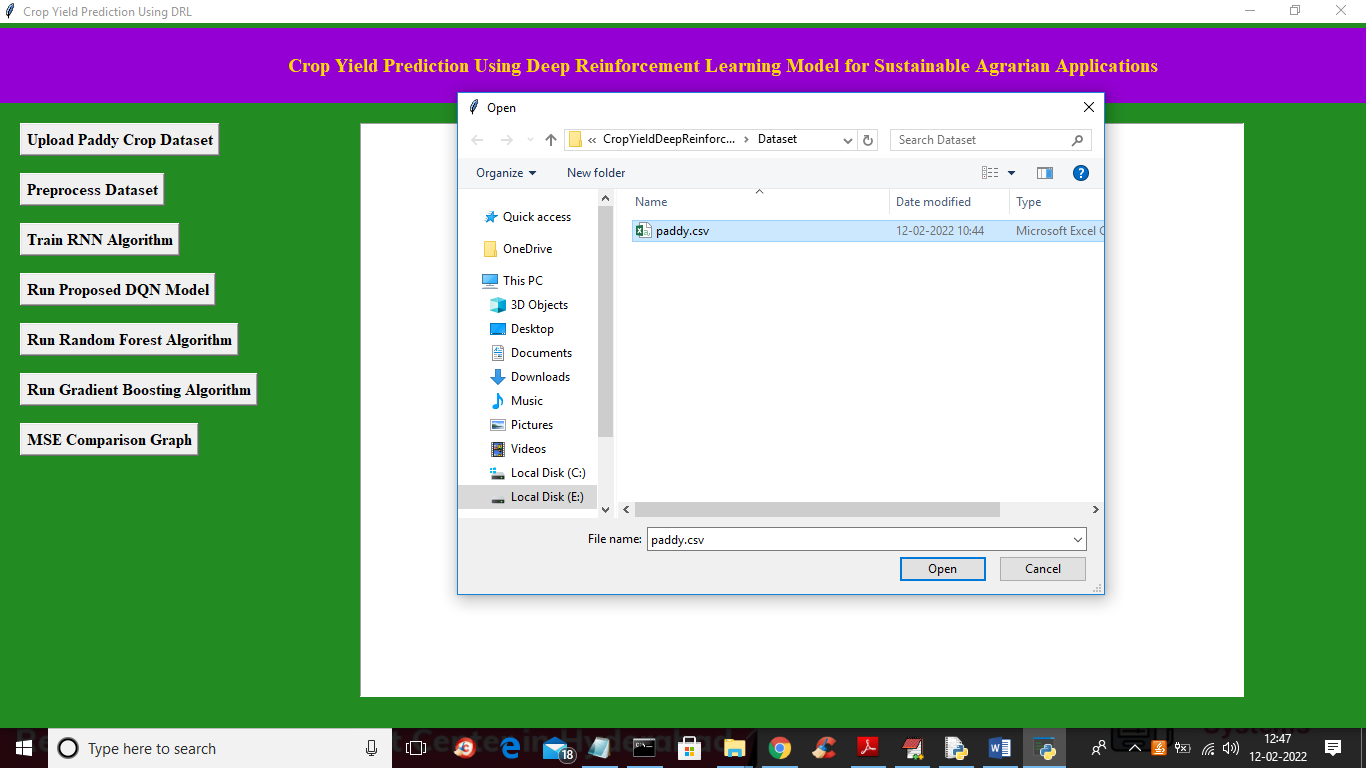
In above screen read red colour comments to know about DQN algorithm

SCREEN SHOTS

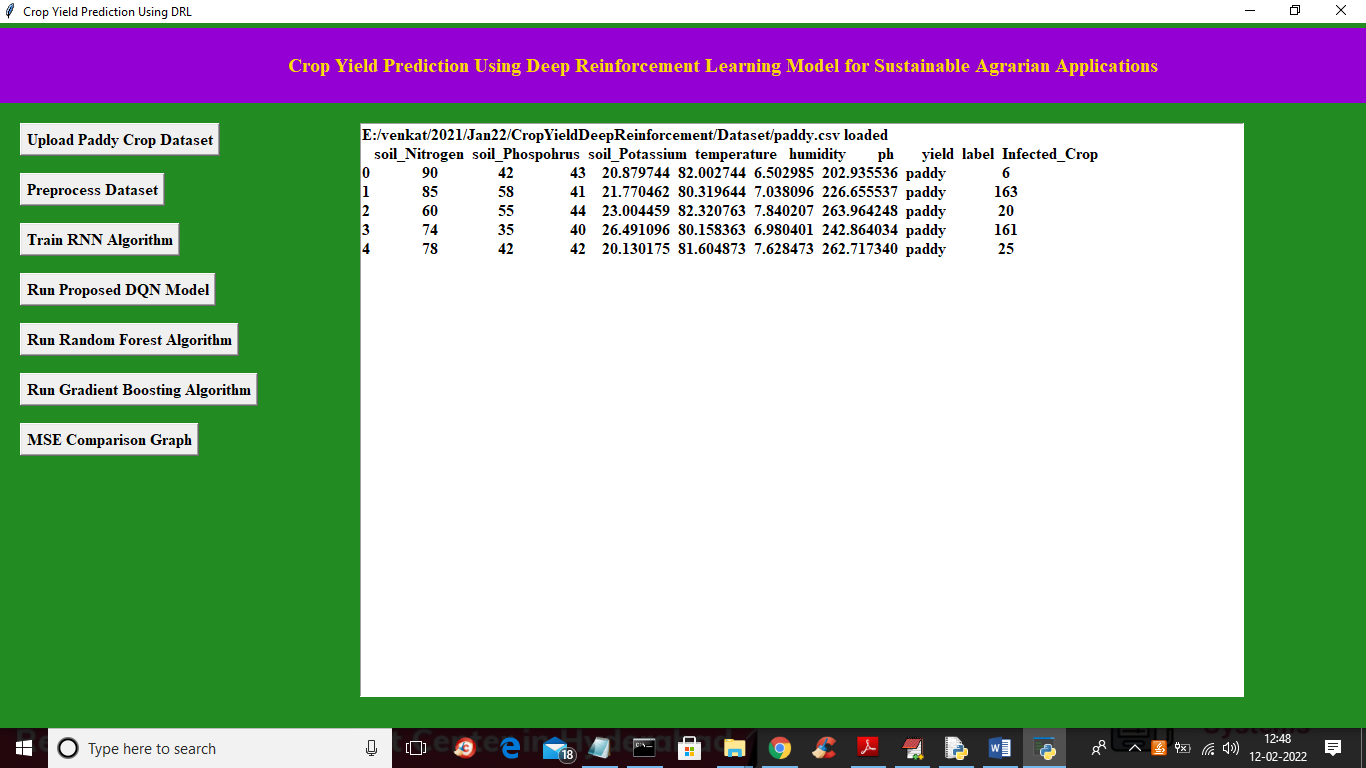
To run project double click on ‘run.bat’ file to get below output



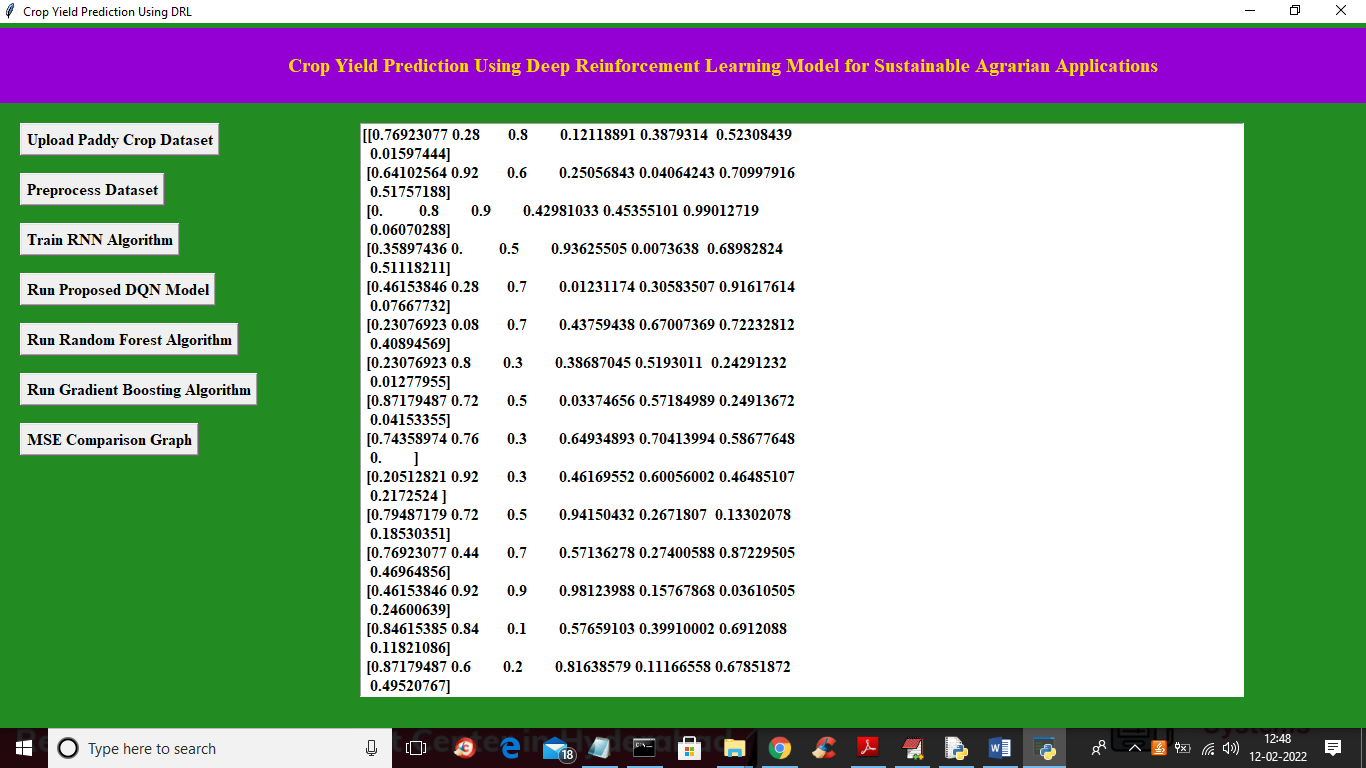
In above screen click on ‘Upload Paddy Crop Dataset’ button to upload dataset and to get below screen



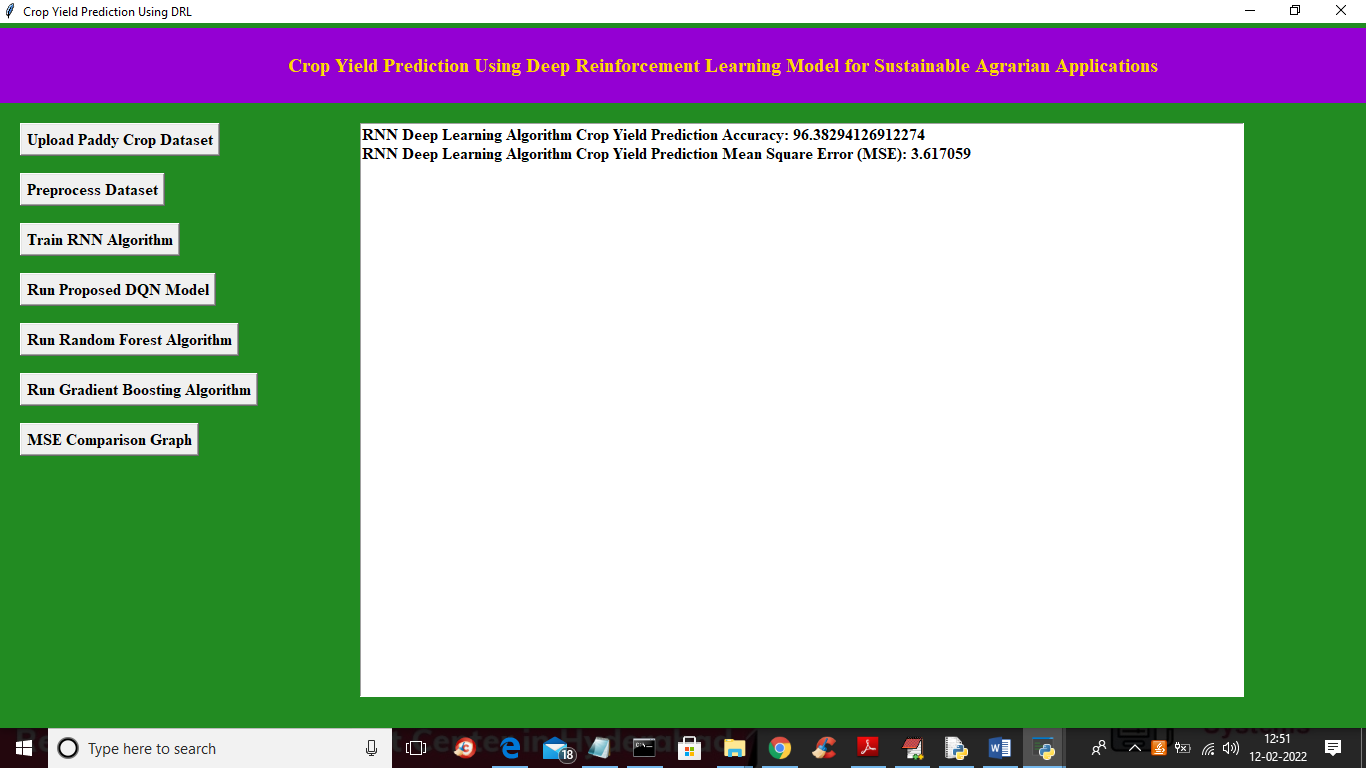
In above screen selecting and uploading ‘paddy.csv’ dataset file and then click on ‘Open’ button to load dataset and to get below output



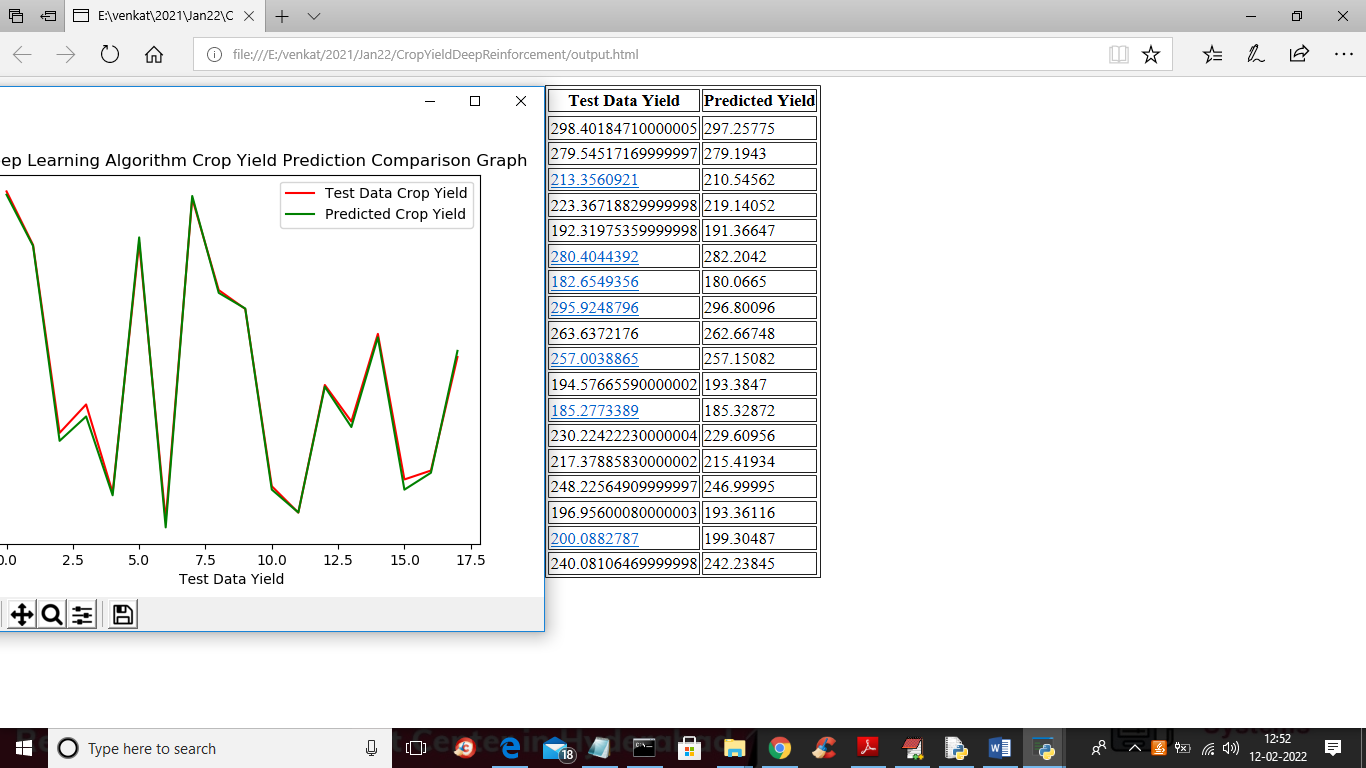
In above screen we can see dataset loaded and it contains some non-numeric values so we need to process and now click on ‘Preprocess Dataset’ button to process the dataset and to get below screen



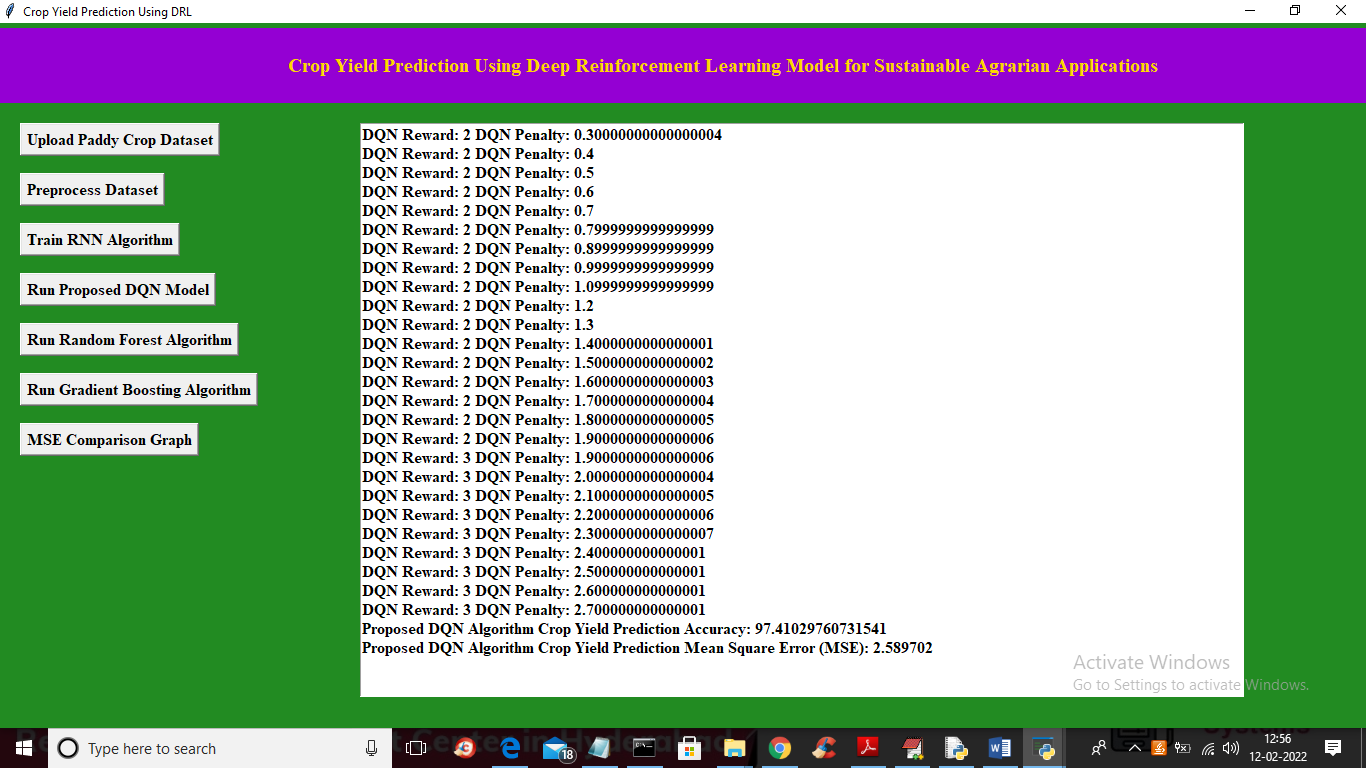
In above screen we can see dataset process and all non-numeric values are removed and now click on ‘Train RNN Algorithm’ button to train RNN and get below output



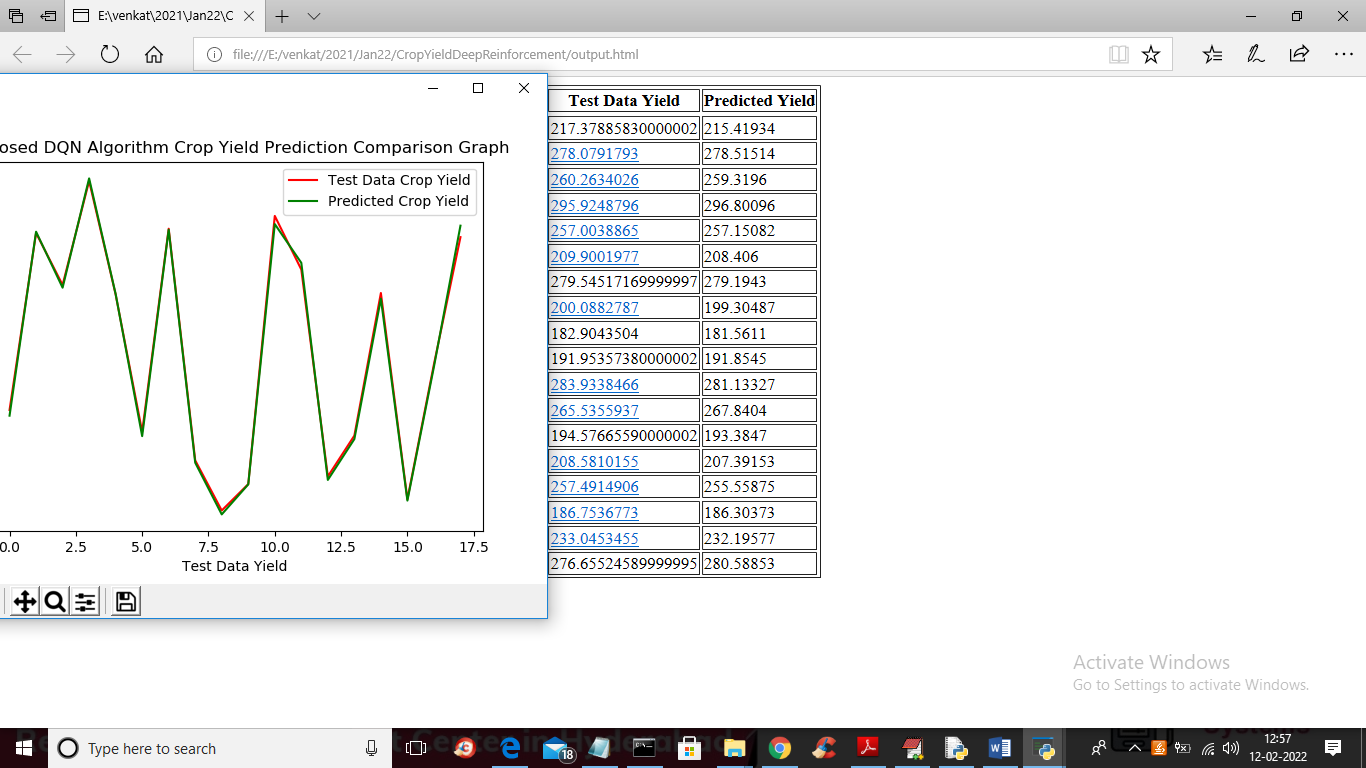
In above screen with RNN algorithm we got accuracy as 96% and MSE as 3.61 and below is the RNN crop yield prediction graph



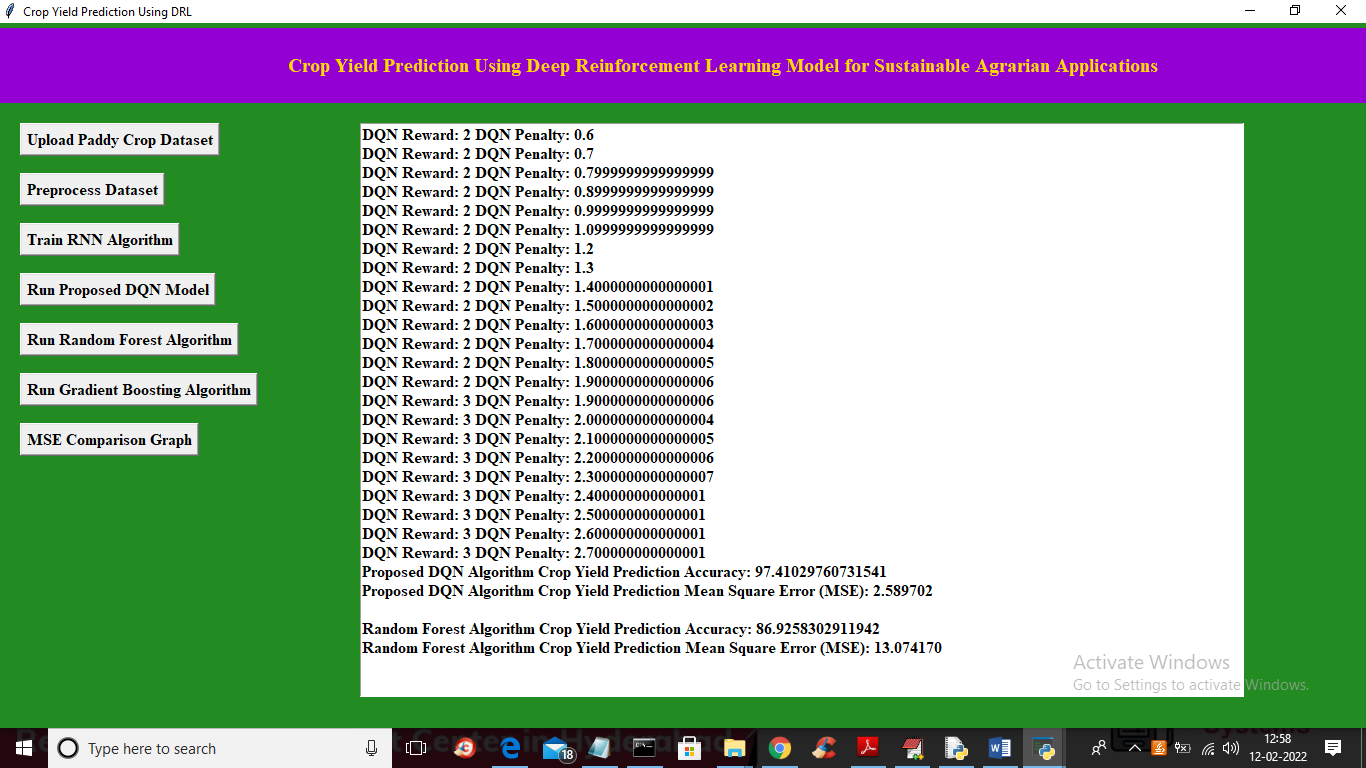
In above graph red line represents crop TEST DATA and green line represents crop yield PREDICTED DATA and in graph we can see both lines are fully overlap so its prediction is little accurate and we can see original TEST DATA YIELD and Predicted Yield also and you can see both values are closed and now click on ‘Run Proposed DQN Model’ button to train propose DQN and get below output



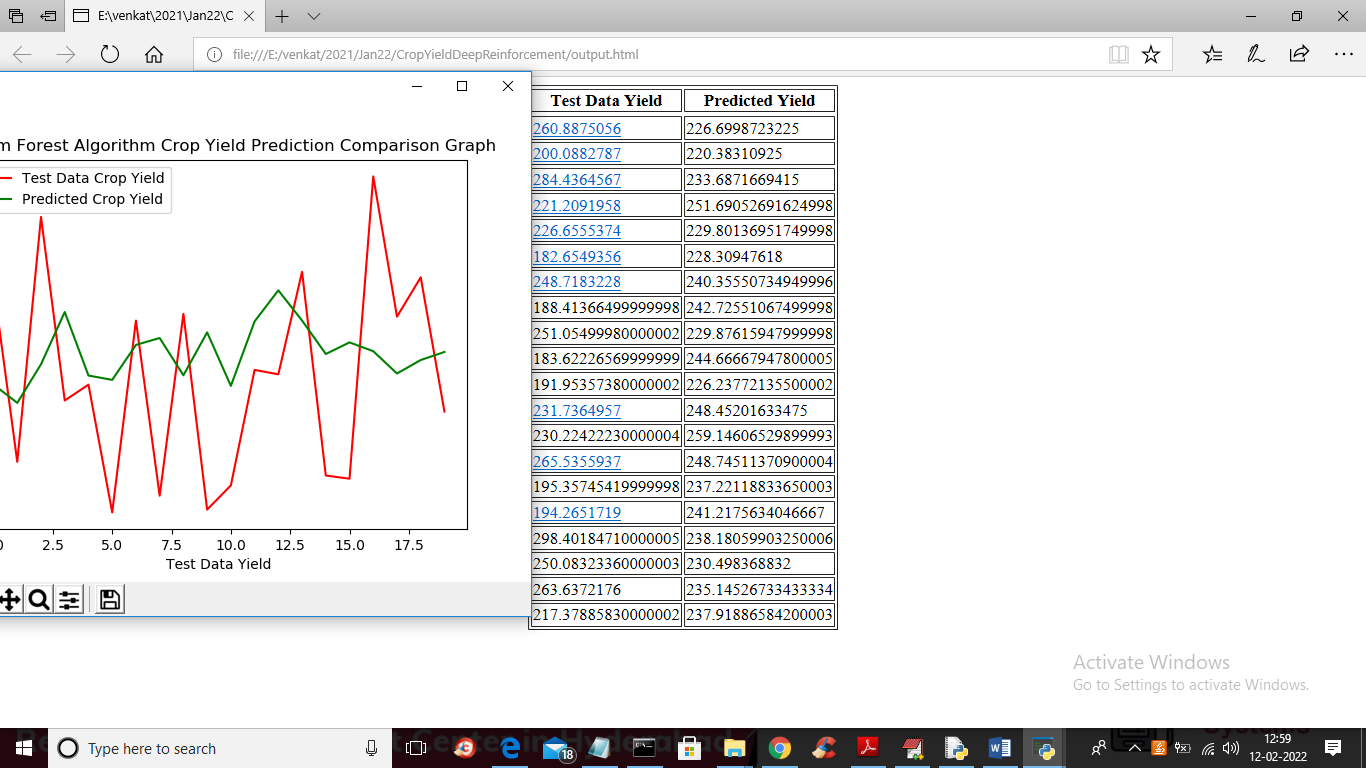
In above screen we can see DQN reward and penalty values and we got accuracy as 97.41% and MSE as 2.58% and below is the prediction graph of DQN



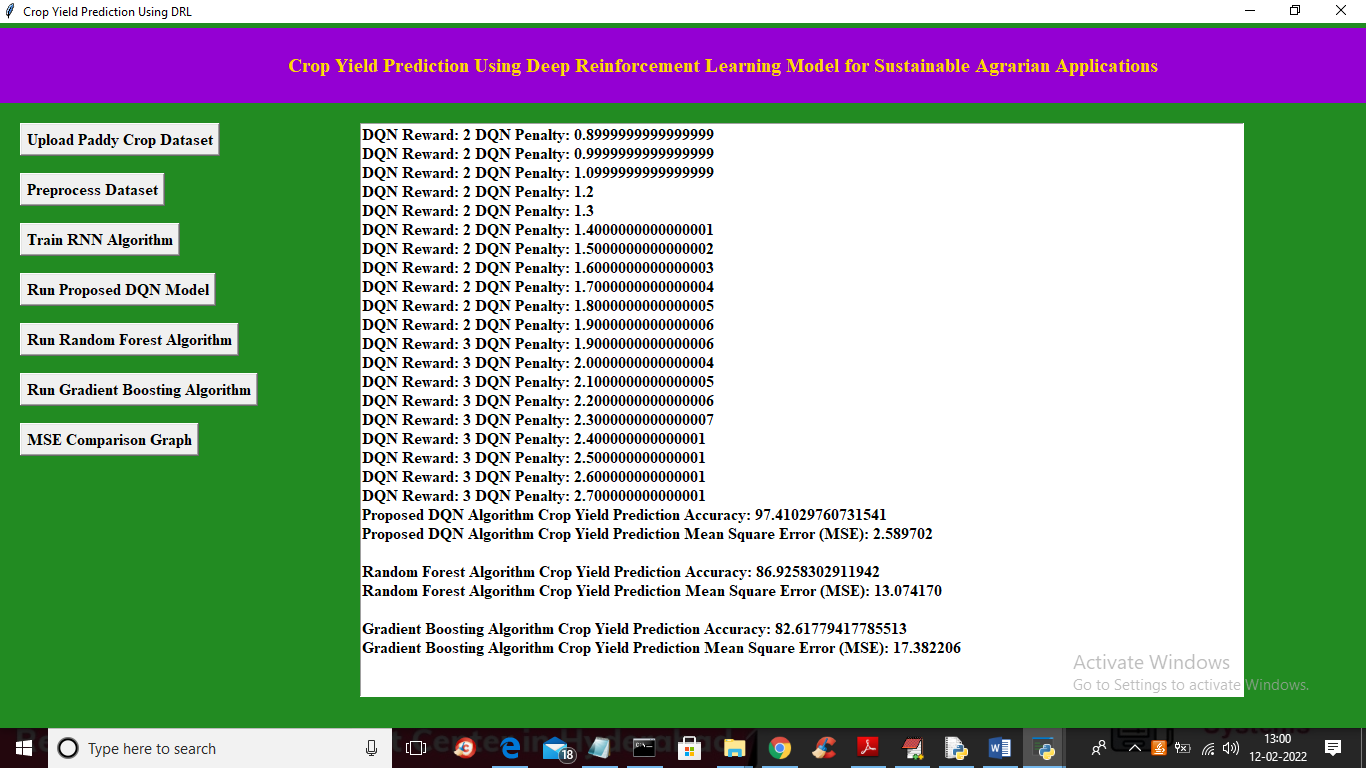
In above proposed DQN graph we can see both red and green lines are fully overlapped so its prediction if accurate and now close above graph and then click on ‘Run Random Forest Algorithm’ button to get below output



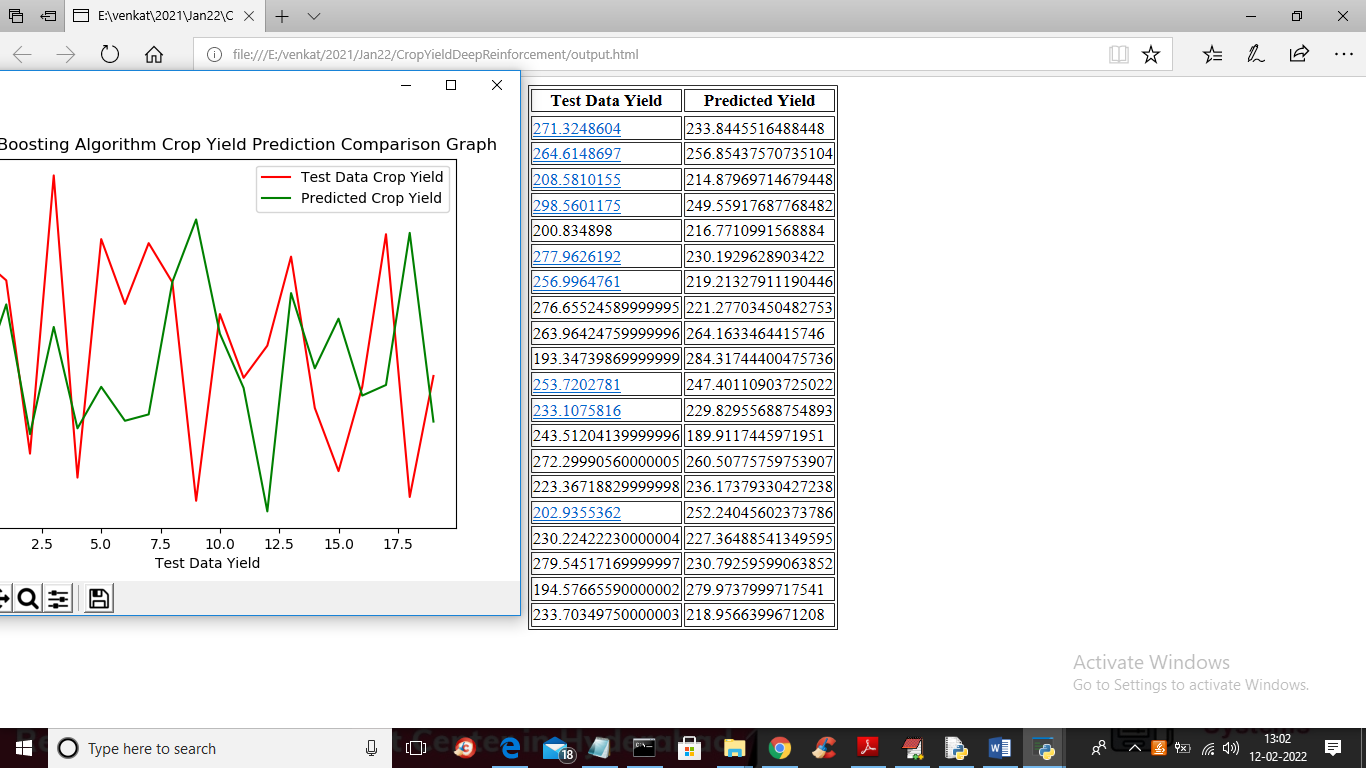
In above screen with Random forest we got 86% accuracy and below is the prediction graph



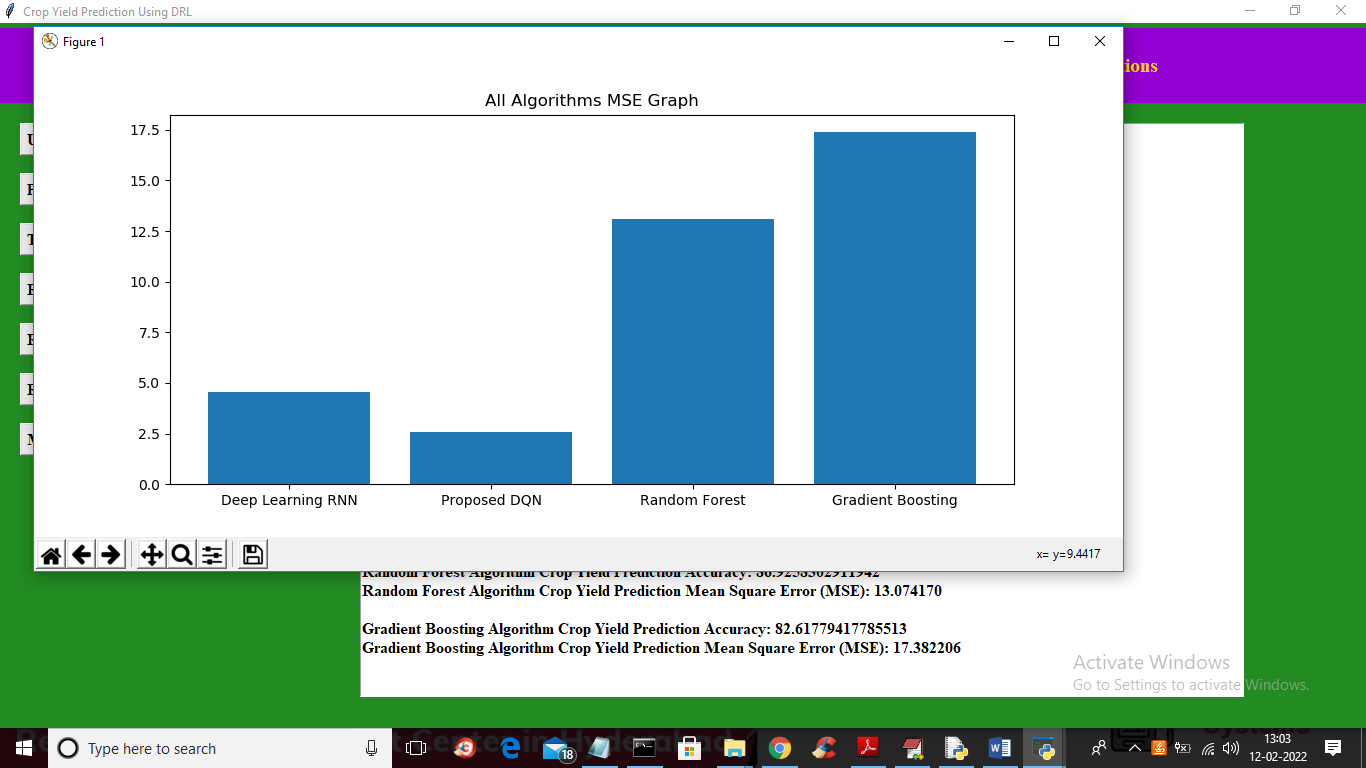
In above graph we can see there is not much overlapping between green and red lines so its prediction is not good and now click on ‘Run Gradient Boosting Algorithm’ button to get below output



In above screen for gradient boosting we got 82% accuracy and below is the prediction graph



In above graph of gradient boosting both lines are not overlapped so its prediction also not accurate and propose DQN only gave 98% accuracy which is better than other algorithms. Now click on ‘MSE Comparison Graph’ button to get below graph



In above MSE graph x-axis represents algorithm names and y-axis represents MSE values and in all algorithms propose DQN only gave less MSE and high accuracy