



Twitter Bot Detection – Help Page

Made by – Or Cohen & Eliran Menashe

If is your first time at our application, WELCOME!

Important Note-

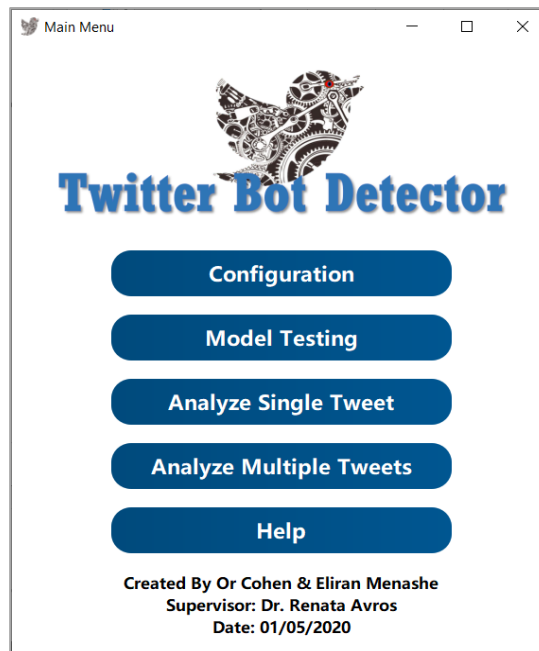
The main folder contains the following sub-folders:

- **Data:**
bots_tweets.txt, human_tweets.txt – These files are the tweets dataset used for training.
glove.twitter.27B.200d.txt, wiki-news-300d-1M.vec – These files contains the word embedding vectors.
COVID.csv – This file contains unsupervised tweets about covid-19 for bot detection testing.
- **Output:**
This folder contains all the results that generated from our software.
For example: all models that have been created from configuration window, and all the excel results that we get from predicting phase.
This folder already contains the default models that are already trained and ready for use.

Note: all models must be saved on the 'output' folder, otherwise you can not see them in models list on predicting phase.

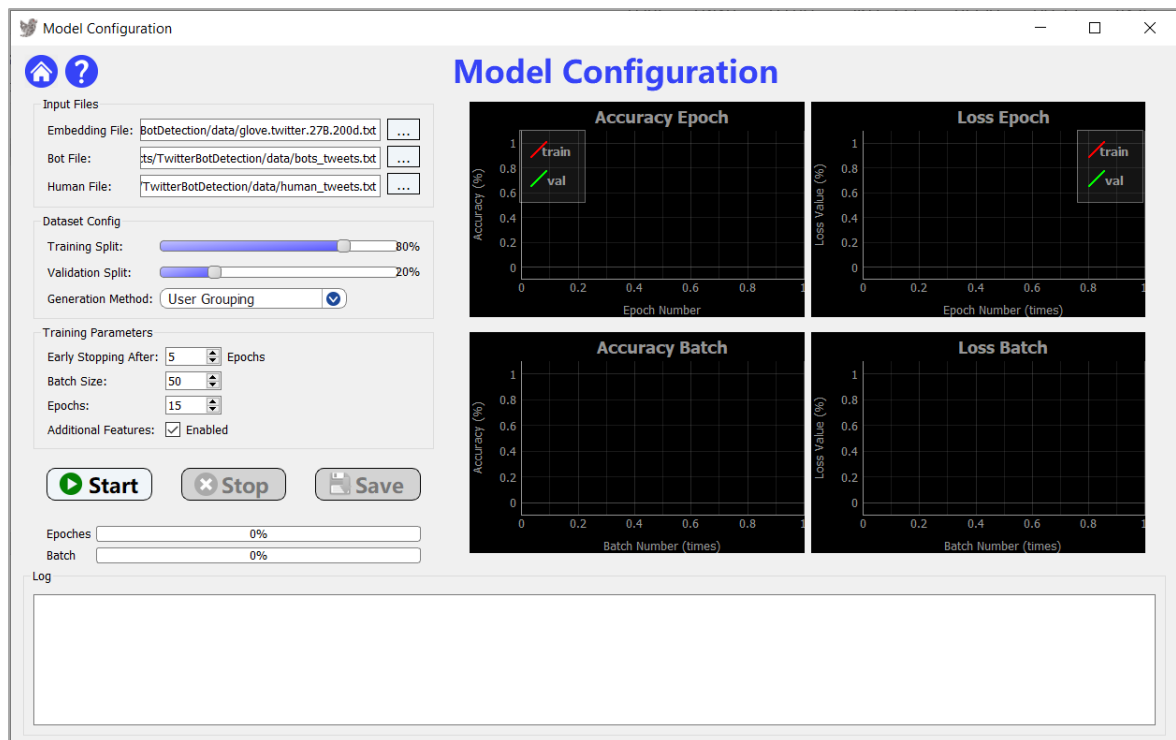
Explanations About the Application-

Main Window-



- 'Configuration'-
This button opens the "Configuration" window for creating a new model.
- 'Model Testing'-
This button opens the "Model Testing" window for testing the accuracy of the existing model.
- 'Analyze Single Tweet'-
This button opens the "Analyze Single Tweet" window for predicting a single tweet.
- 'Analyze Multiple Tweets'-
This button opens the "Analyze Multiple Tweets" window for predicting multiple tweets from a dataset file.
- 'Help'-
This button opens the instructions about the system that contains explanations about each button in each window in the system.

Model Configuration-



- **'Back Button'**-
This button will back to the previous window - "Main Window".
- **'Help Button'**-
This button opens the instructions about the system that contains explanations about each button in each window in the system.
- **'Input Files'**-
 - **'Embedding File'**- You need to choose the embedding file for training.
 - **'Bot File'**- You need to choose the bot dataset file that contains all the bot tweets.
 - **'Human File'**- You need to choose the human dataset file that contains all the human tweets.
- **'Dataset Config'**-
 - **'Training Split'**- You need to choose the percentage of training split.
 - **'Validation Split'**- You need to choose the percentage of validation split.
 - **'Generation Method'**-
You need to select between two methods for dataset generation:
User Grouping: This method generates datasets by user resolution to get better accuracy.
Random Pairing: This method generates datasets by random.

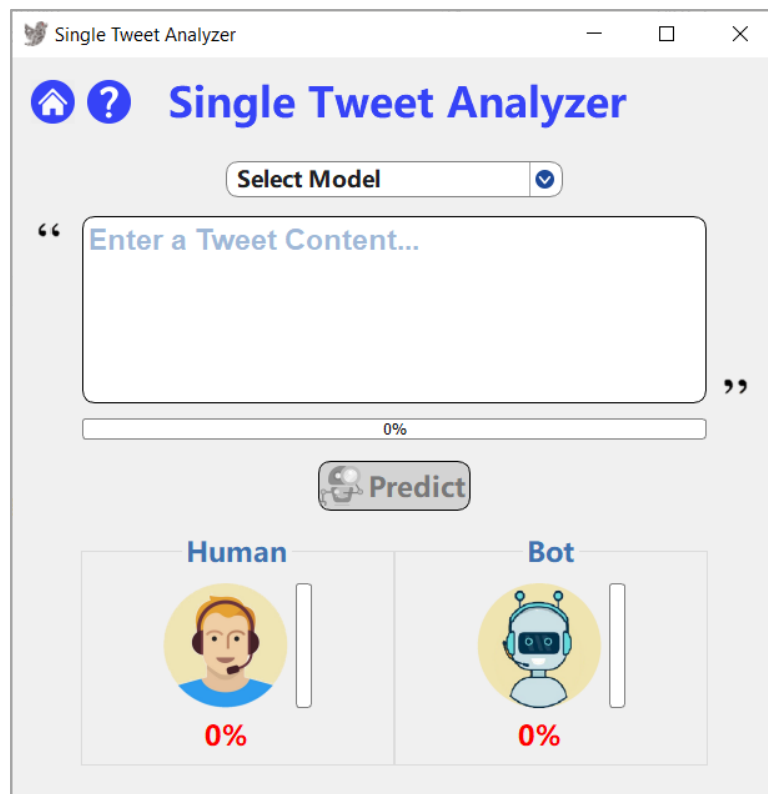
- 'Training Parameters'-
 - 'Early Stopping After'- Early stopping is a method that allows you to specify an arbitrarily large number of training epochs and stop training once the model performance stops improving on a validation dataset.
 - 'Batch Size'- Select the size of the batch.
 - 'Epochs'- Select the number of the epochs.
 - 'Additional Features'- If this feature is enabled, you can get additional info between pairs during training and predicting phase for getting better accuracy.
- 'Start Button'- Start training process.
- 'Stop Button'- This button enabled during the training process and cause stopping this process.
- 'Save Button'- This button enabled at the end of the training process and save the final model.
- 'Epochs Progress Bar'- This progress bar represents the percentage of the current epoch number relative to the total epochs number.
- 'Batch Progress Bar'- This progress bar resets every epoch and represents the percentage of the current batch size relative to the total batches size.
- 'Graphs'-
 - Accuracy Epoch: This graph shows the accuracy during the training process after every epoch.
 - Loss Epoch: This graph shows the loss during the training process after every epoch.
 - Accuracy Batch: This graph shows the accuracy during the training process in the current batch and resets every epoch.
 - Loss Batch: This graph shows the loss during the training process in the current batch and resets every epoch.
- 'Log'- This text box describes the progress throughout the training process.

Model Testing-

- 'Back Button'-
This button will back to the previous window - "Main Window".
- 'Help Button'-
This button opens the instructions about the system that contains explanations about each button in each window in the system.
- 'Select Model'- This combo box allows you to select an existing trained model.
- 'Input Configuration'-
 - 'Bot File'- You need to choose the bot dataset file that contains all the bot tweets.
 - 'Human File'- You need to choose the human dataset file that contains all the human tweets.
 - 'Bot Tweets'- Select the number of tweets for random selection from the bot file.
 - 'Human Tweets'- Select the number of tweets for random selection from the human file.

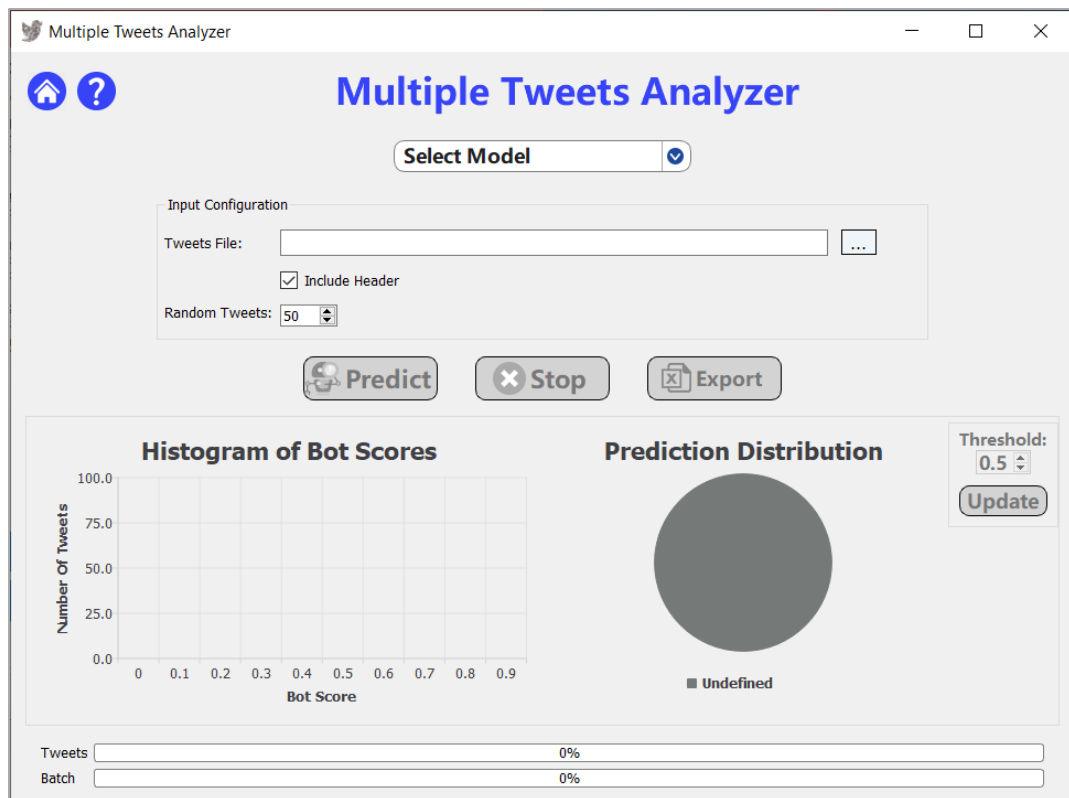
- 'Predict Button'- Start predicting process.
- 'Stop'- This button enabled during the predicting process and cause stopping this process.
- 'Export'- This button enabled at the end of the predicting process and save the results as excel file.
- 'Model Accuracy'- This label shows the final accuracy percentage of the model at the end of the predicting.
- 'Bar chart'-
This chart shows the correct/incorrect prediction percentage for bots and human.
 - 'Threshold'- Select the threshold for classification.
 - 'Update Button'- This button updates the chart based on the threshold.
- 'Tweets'- This progress bar represents the percentage of the current tweet relative to the total tweets.
- 'Batch'- This progress bar resets every tweet and represents the percentage of the current batch size relative to the total batch size.

Single Tweet Analyzer-



- 'Back Button'- This button will back to the previous window - "Main Window".
- 'Help Button'- This button opens the instructions about the system that contains explanations about each button in each window in the system.
- 'Select Model'- This combo box allows you to select an existing trained model.
- 'Enter a Tweet Content'- In this text box you need to write or paste the tweet that you want to predict.
- 'Predict Button'- Start the predicting process.
- 'Human'- Shows the score of human classification in percentage.
- 'Bot'- Shows the score of bot classification in percentage.

Multiple Tweets Analyzer-



- **'Back Button'**- This button will back to the previous window - "Main Window".
- **'Help Button'**- This button opens the instructions about the system that contains explanations about each button in each window in the system.
- **'Select Model'**- This combo box allows you to select an existing trained model.
- **'Input Configuration'**-
 - **'Tweets File'**- You need to choose the dataset file that contains unsupervised tweets. Allowed file types: csv, txt.
 - **'Include Header'**- Check this option if your dataset file contains a header.
 - **'Random Tweets'**- Select the number of tweets for random selection from the dataset file.
- **'Predict Button'**- Start predicting process.
- **'Stop'**- This button enabled during the predicting process and cause stopping this process.
- **'Export'**- This button enabled at the end of the predicting process and save the results as excel file.

- 'Histogram of Bot Scores'- Shows the prediction scores of bots as a histogram.
- 'Prediction Distribution Pie Chart'- Show the classification as pie chart.
- 'Threshold'- Select the threshold for classification.
- 'Update Button'- This button updates the charts based on the threshold.
- 'Tweets'- This progress bar represents the percentage of the current tweet relative to the total tweets.
- 'Batch'- This progress bar resets every tweet and represents the percentage of the current batch size relative to the total batch size.