Detecting Mental Disorders in Social Media Through Emotional Patterns - The case of Anorexia and Depression

Now-a-days due to stress all ages (elders suffer due to job stress and minor suffer due to education competition) of peoples are suffering from mental disorders and timely detection of such disorder can help victim in recovering and to detect such emotion many algorithms has been introduce such as analysing sentiments or linguistic from user POST or reviews to detect depression or Anorexia but this technique directly detect whether review is positive or negative without identifying emotion sub-domains as user may express his depression emotion in all words compare to normal users and by identifying sub-domains from emotion we can identify actual level of depression.

In propose paper to effectively detect depression author has introduce sub-emotion detection via two modules such as static BOSE (bag of sub-domain emotion) and dynamic BOSE. Static BOSE algorithm combine all reviews as a single document and then detect depression and won’t analyse user history reviews to know emotion variation. Dynamic BOSE similar to static BOSE with single difference of taking all users reviews as HISTORY or CHUNK and then each chunk will be analyse using BOSE algorithm to detect variations in emotion based on review history time.

In propose paper author masking each word in a review with LEXICON emotion words and then this masking will be converted to weighted TF-IDF (term frequency inverse frequency document). TF-IDF will be input to Affinity clustering to cluster each sub domain and then predict depression and control and anorexia. Cluster data will be input to SVM and CNN algorithm to calculate depression detection F1 SCORE. Above process will be called as static BOSE.

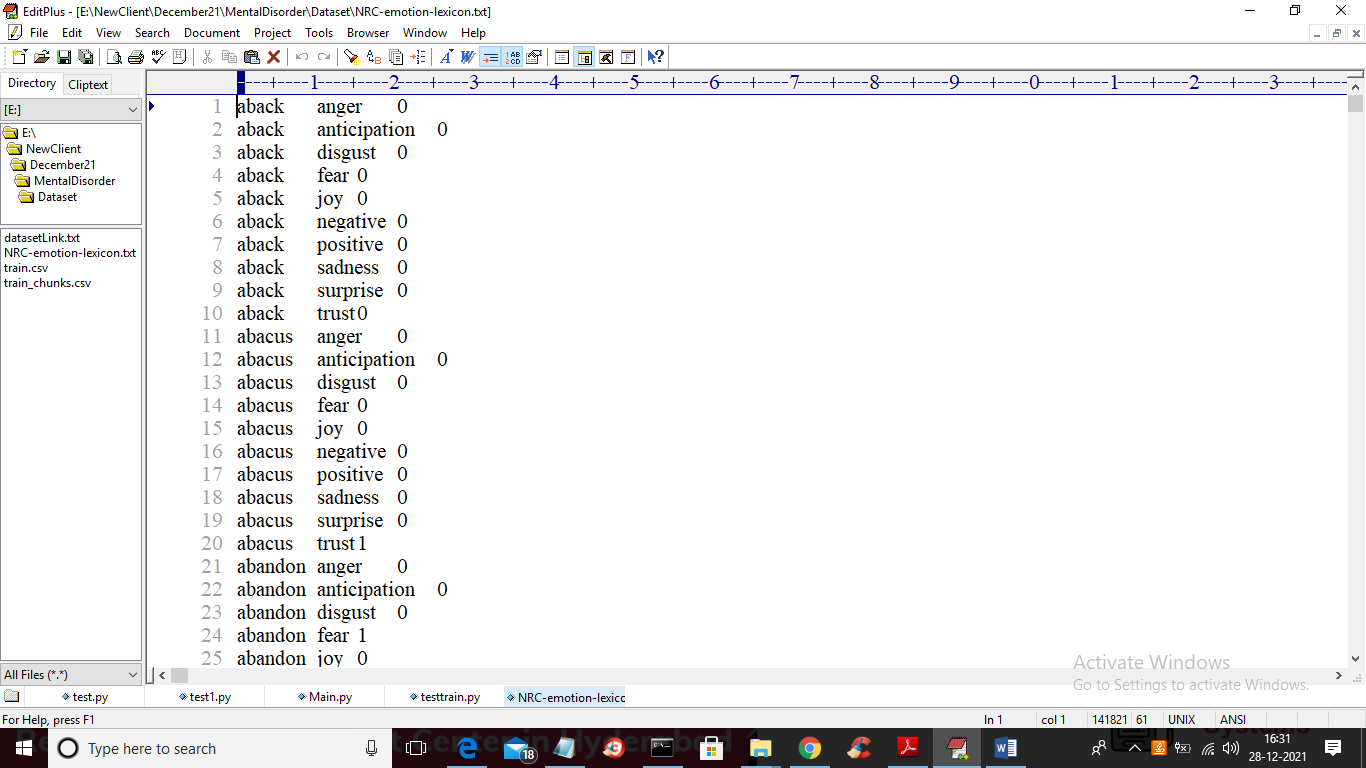
Dynamic BOSE take all history reviews as chunk and then apply above BOSE technique to detect variations in emotions.

Below is the example of Masking review to sub-domain

**The most important thing is to try and inspire people.**

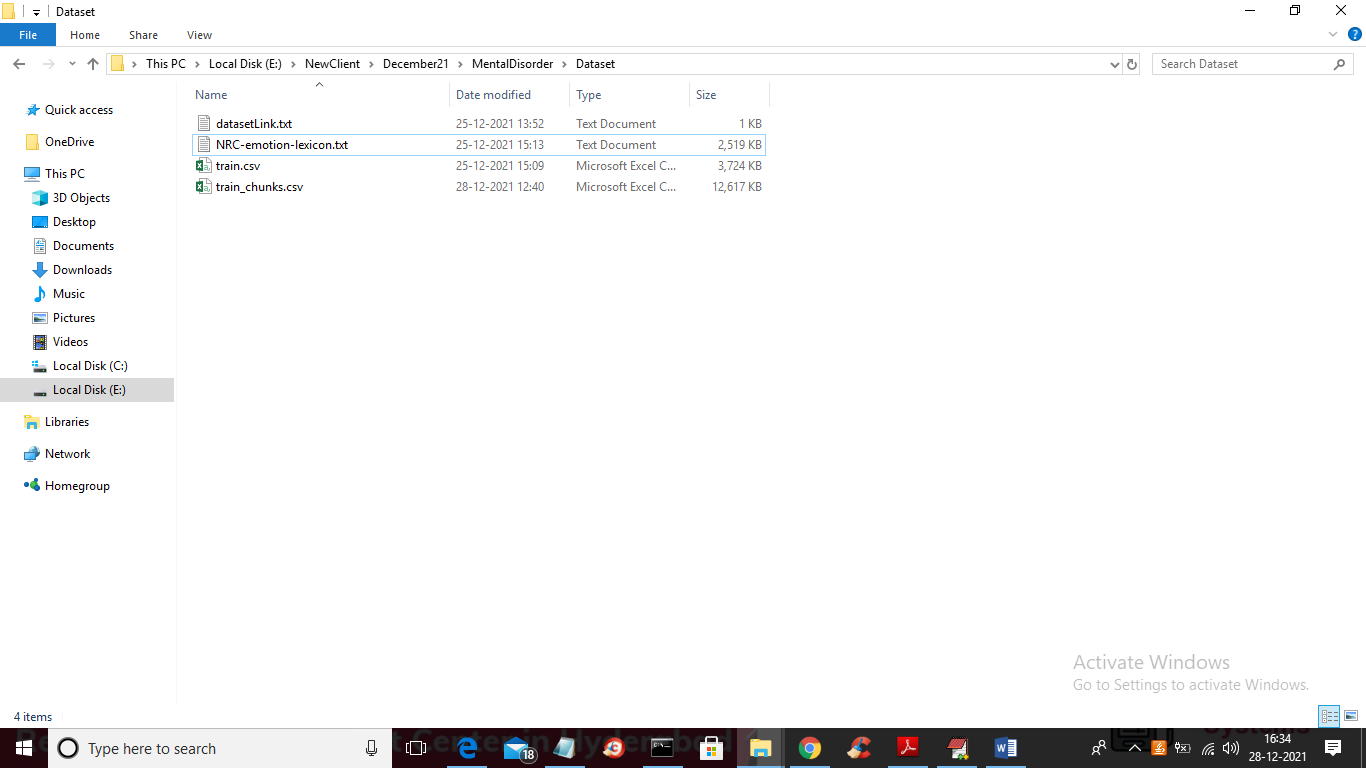
**anticipation27 joy27 positive5 negative62 anticipation10 anticipation29 positive20 negative80 trust23 joy16**

In above statement first sentence is the original statement and second is mask sentences which we can obtained from LEXICON emotion dictionary which contains words and its emotion. Below is the LEXICON screen shot



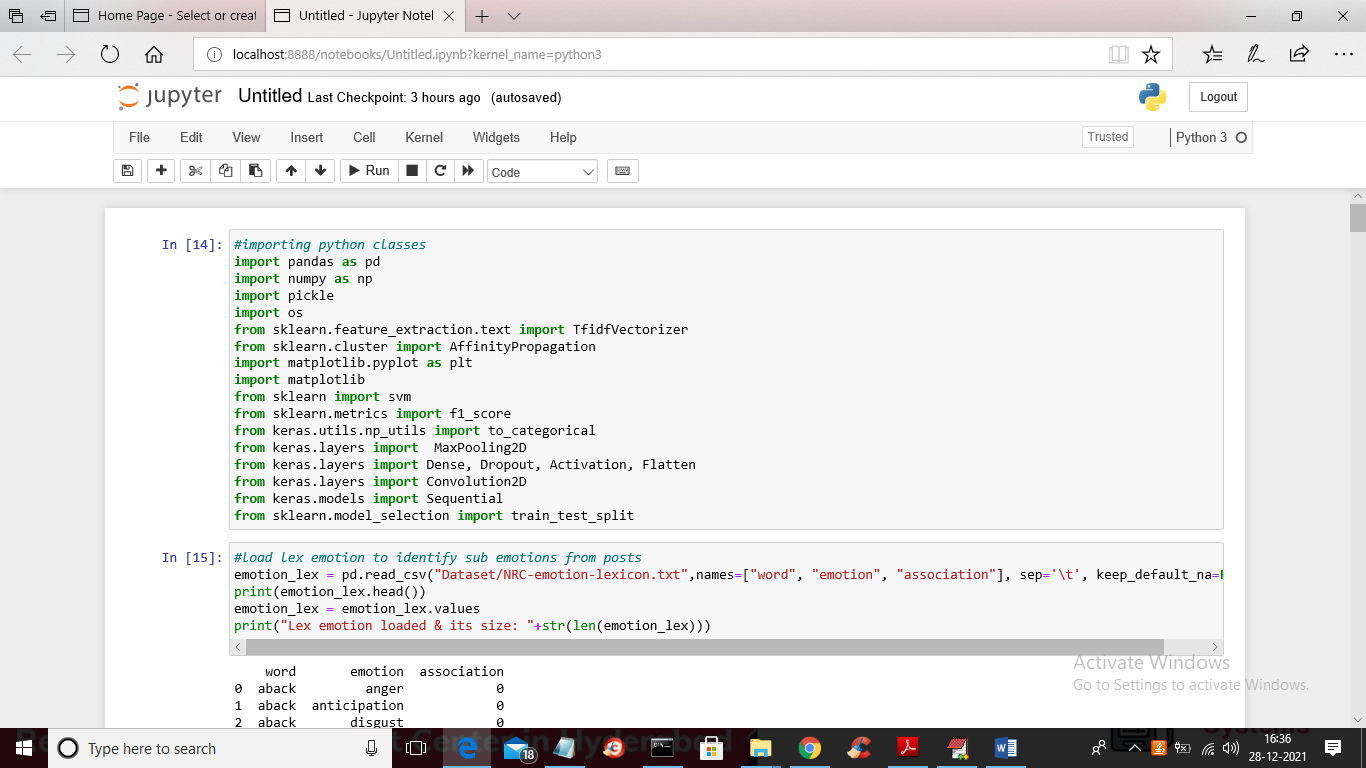
In above LEXICON dictionary we can see each word is associated with emotion and we will use same LEXICON to mask all sentences to emotion.

In propose paper we have taken two dataset one is with all reviews and the other with chunk reviews and using first dataset we will run static BOSE and using second dataset we will run dynamic BOSE. Below is the normal and chunk dataset screen

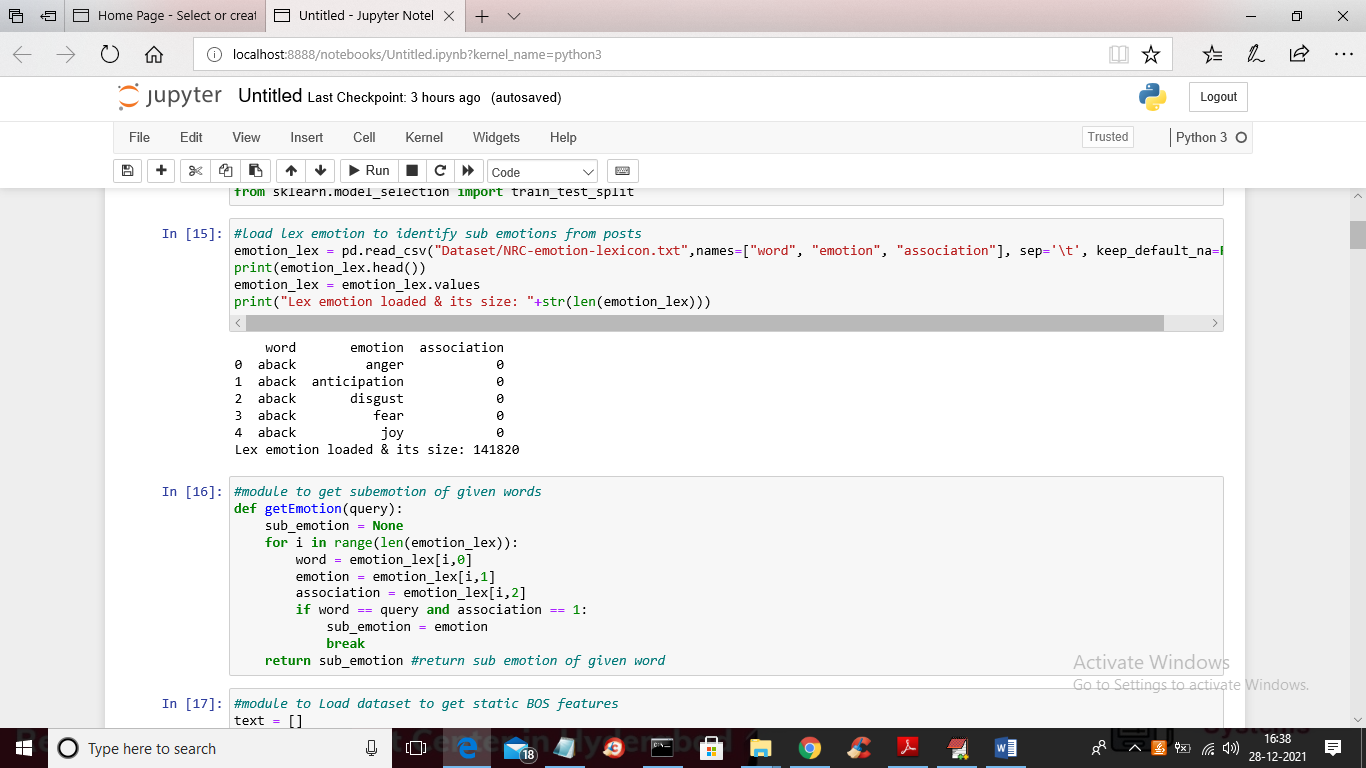


In above screen train.csv contains all single reviews for each user and train\_chunks.csv contains multiple reviews for each user as chunk history reviews. By using both datasets we will run static and dynamic BOSE and below are the output screens

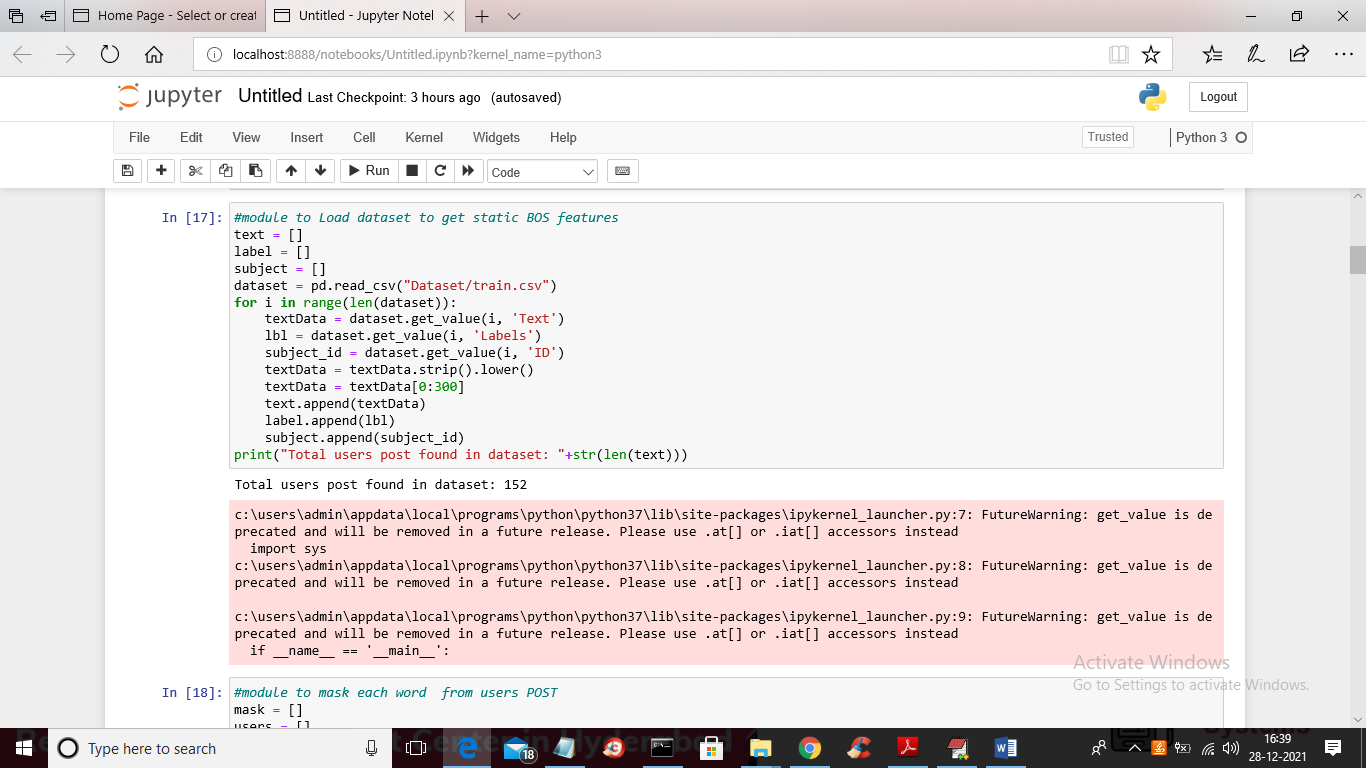
OUTPUT SCREENS



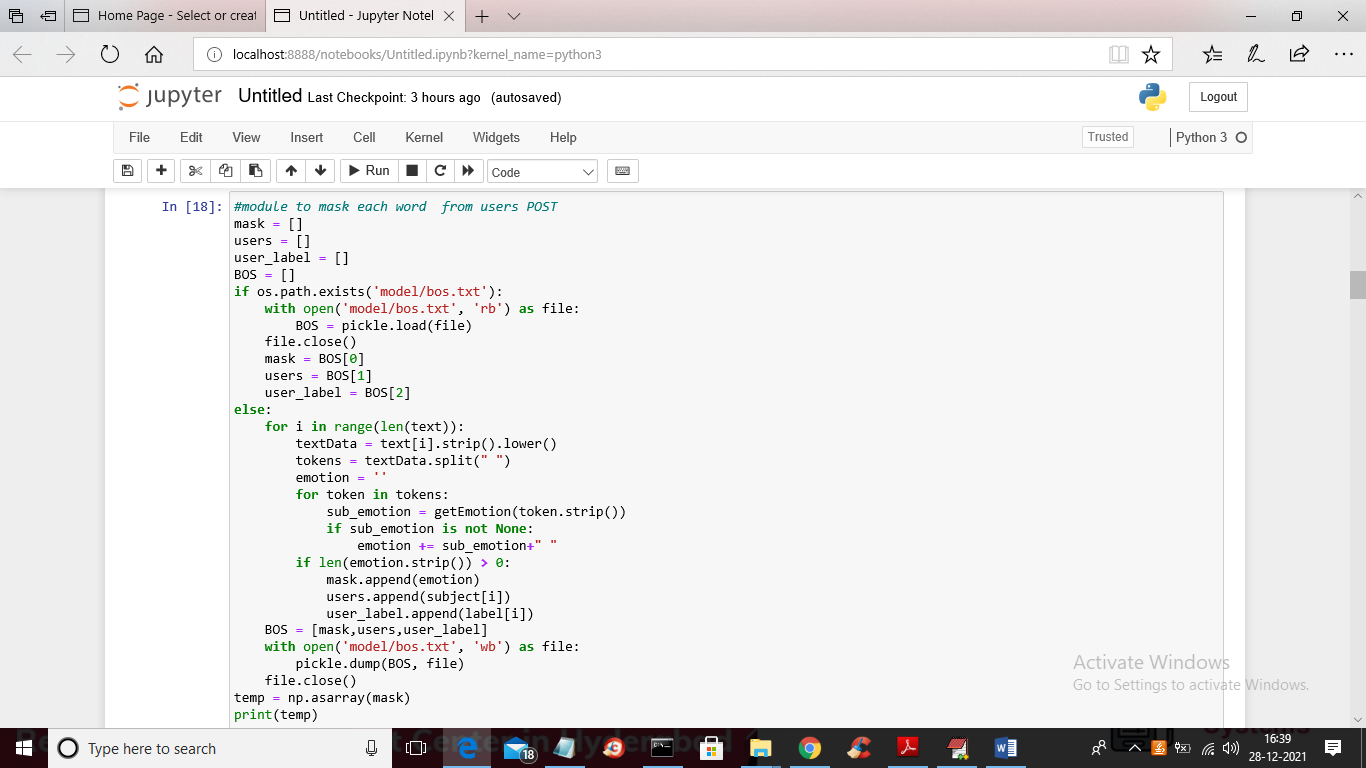
In above screen we are loading python packages and in below screen you can see we are reading and storing LEXICON dictionary



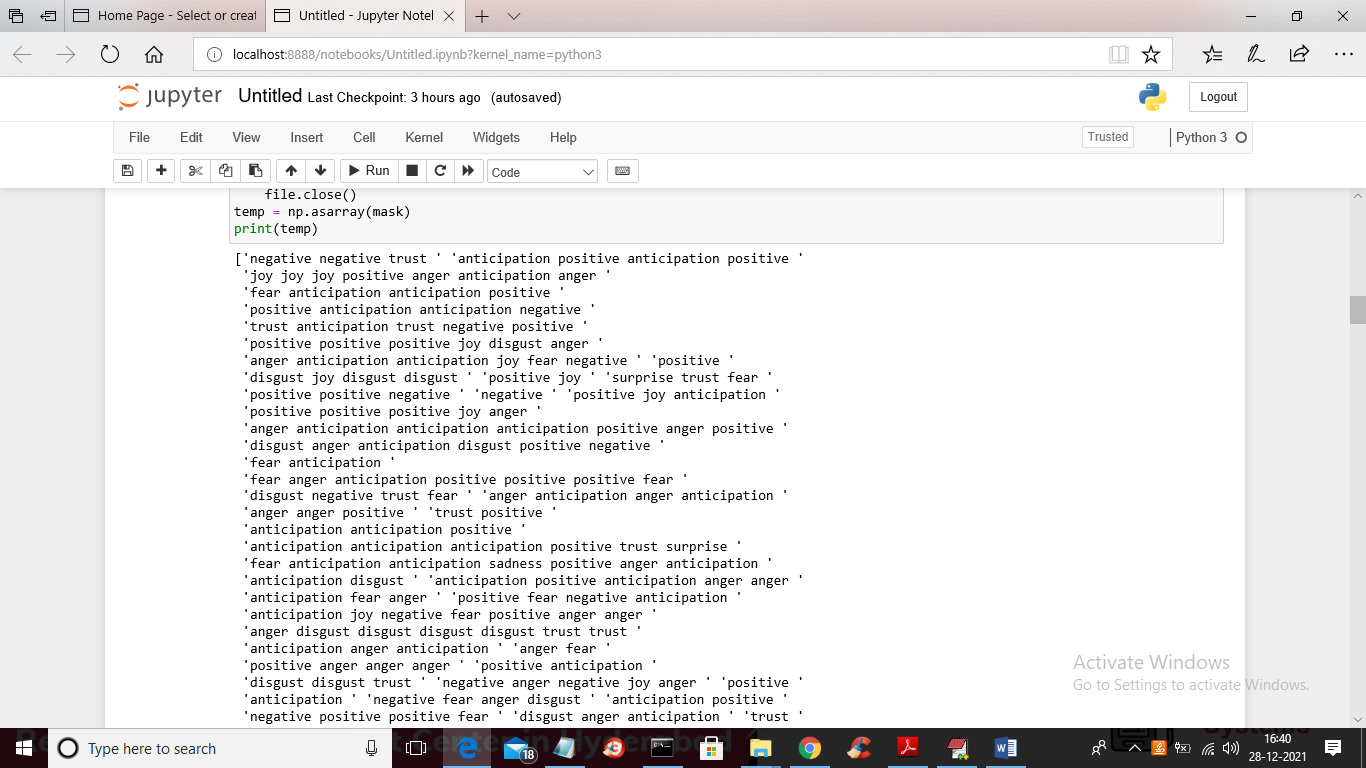
In above two modules we are displaying few words from LEXICON and then defining function to get emotion from given words.



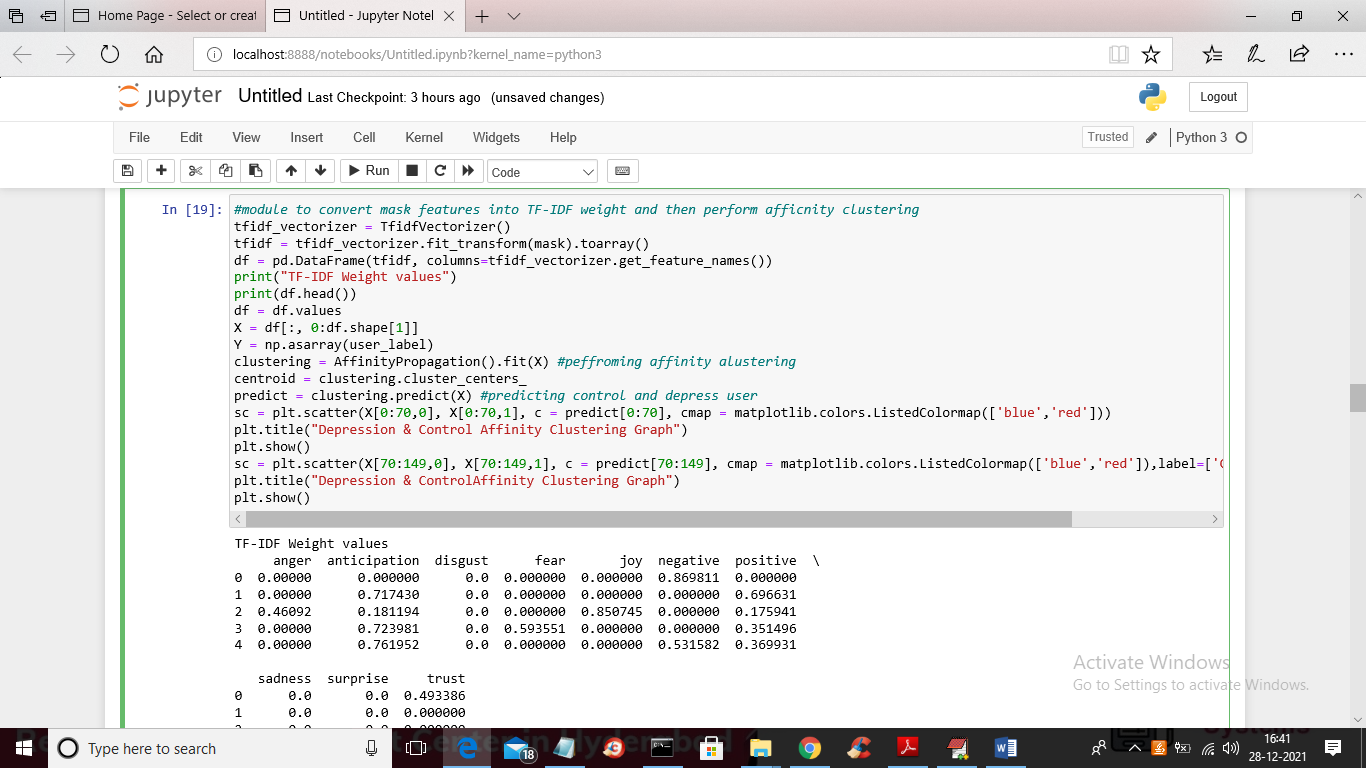
In above screen we are reading dataset and then storing in array variables and in below screen we are masking all dataset reviews



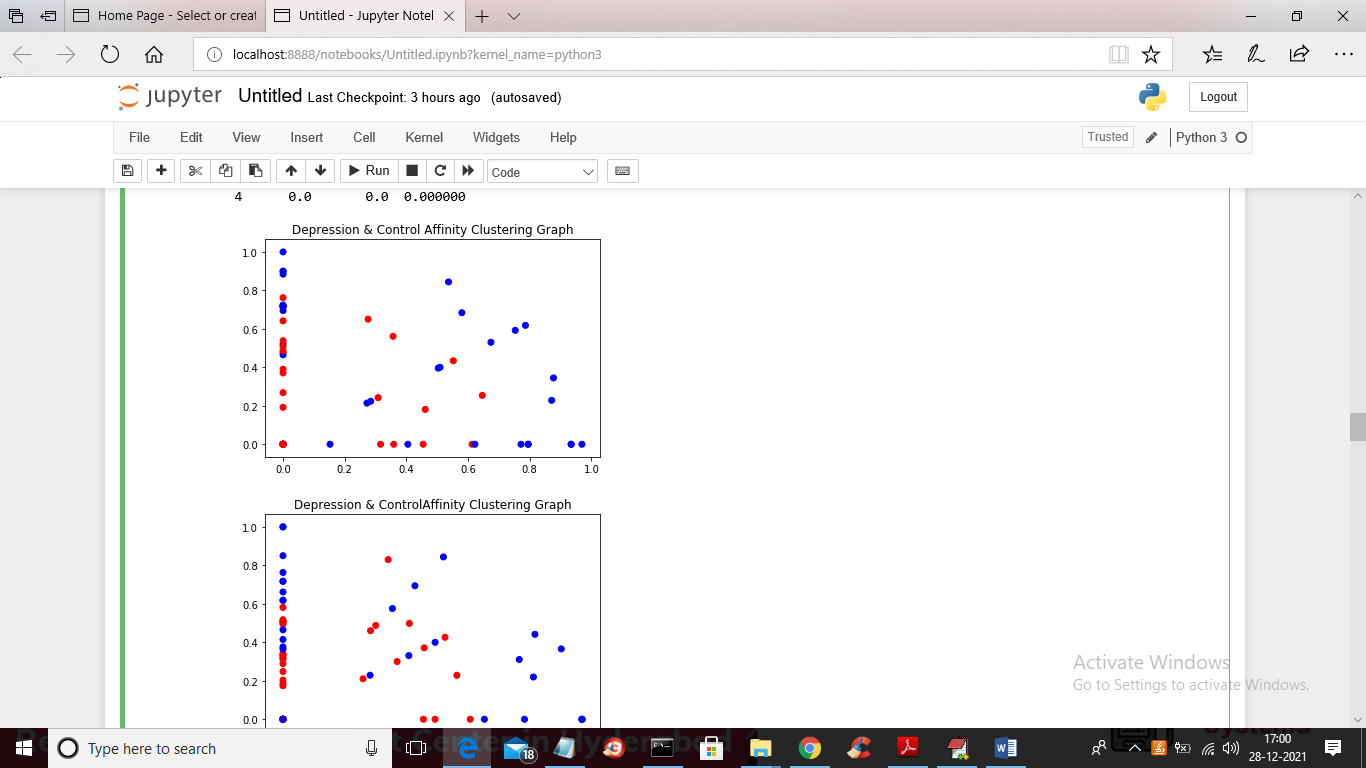
Below is the mask sentences



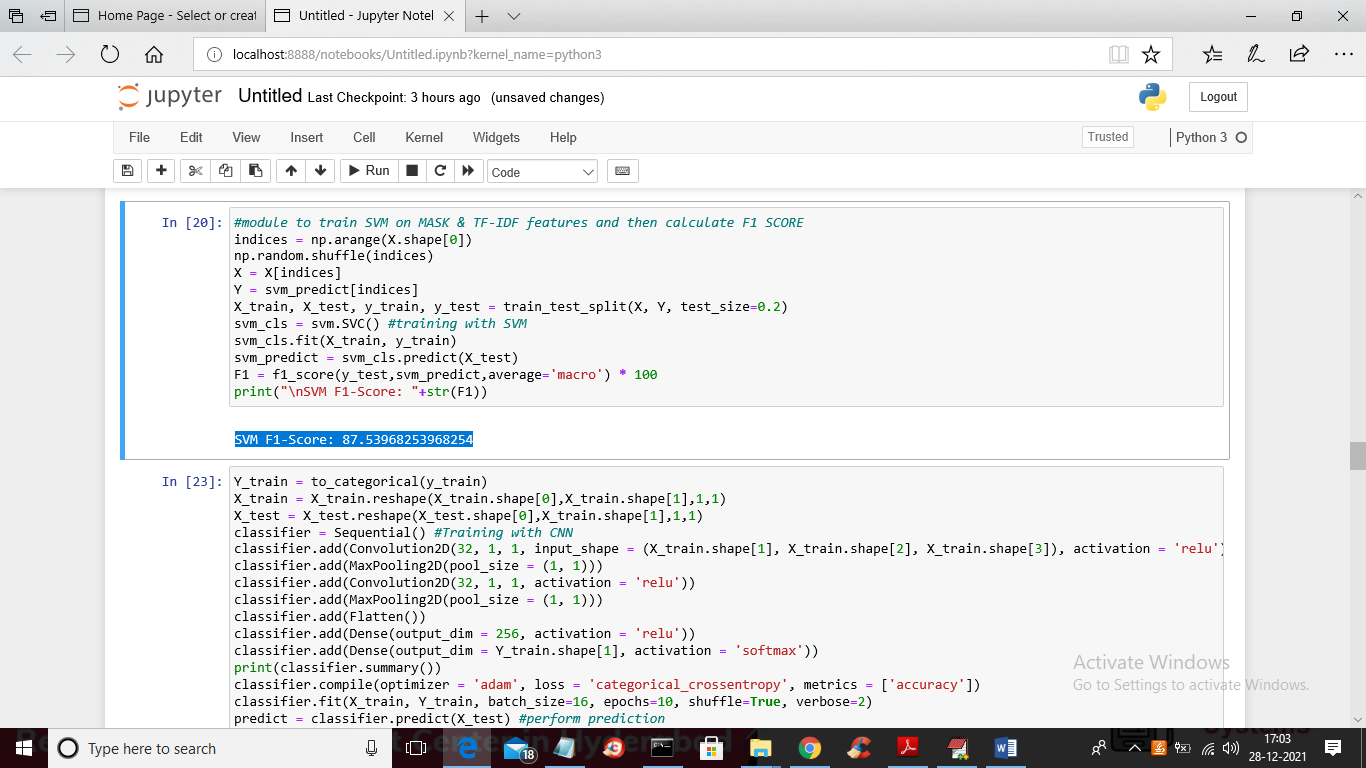
In above screen we can see all sentences or reviews converted to emotion masking and in below screen we are converting mask data into TF-IDF and then applying Affinity Clustering



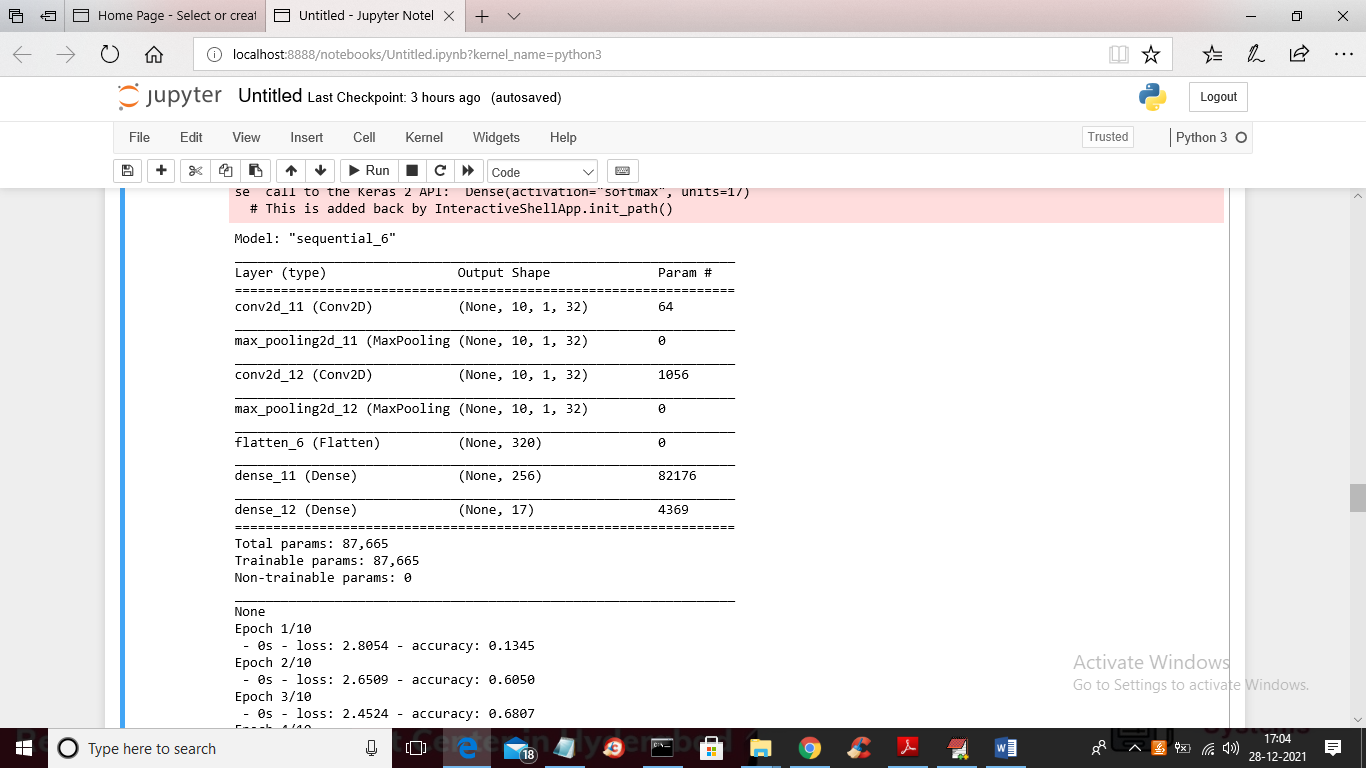
In above screen we are displaying TF-IDF values based on sub-domain emotions and then apply Affinity Clustering to get below graph and we got depression and anorexia user from affinity cluster prediction result.

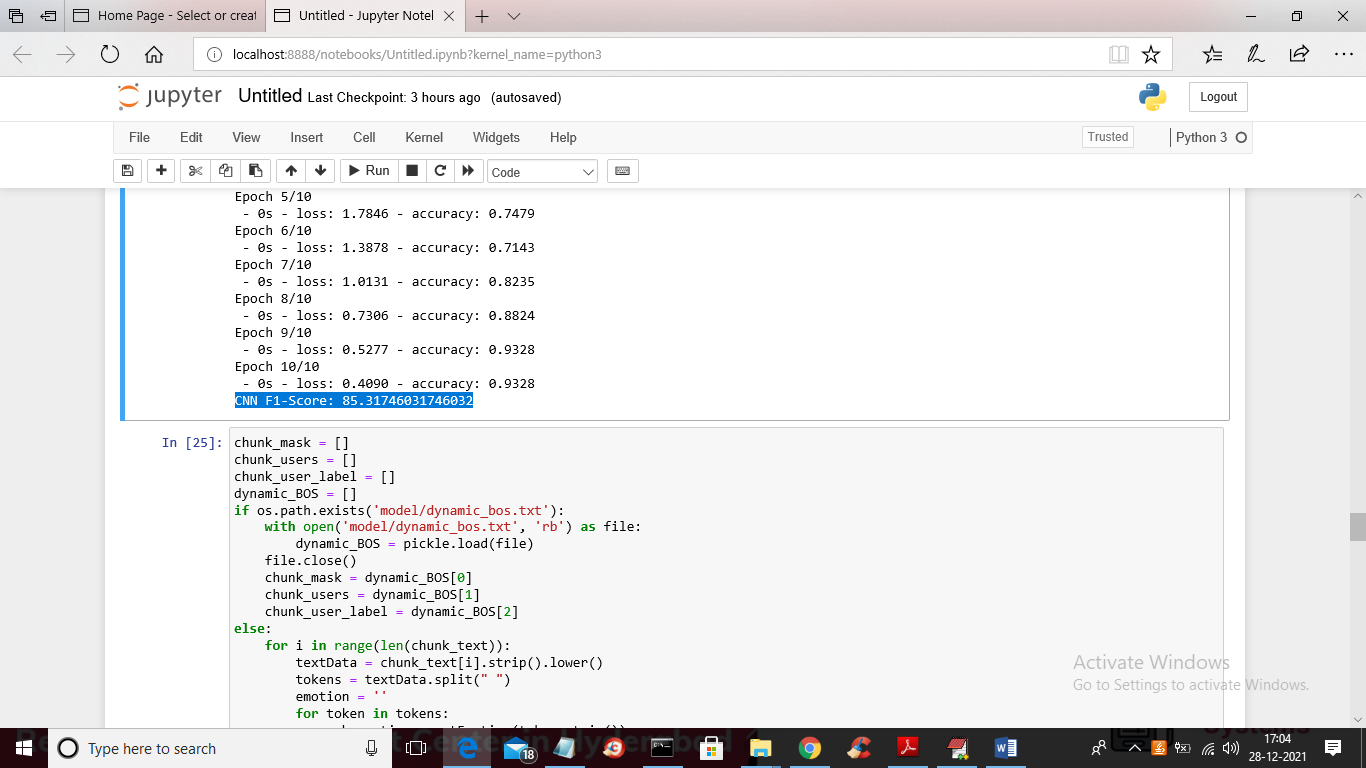


In above graph blue colour dots are then control user and red colour for depression and anorexia users. This TF-IDF cluster data can be trained with SVM and CNN to get depression prediction F1-Score

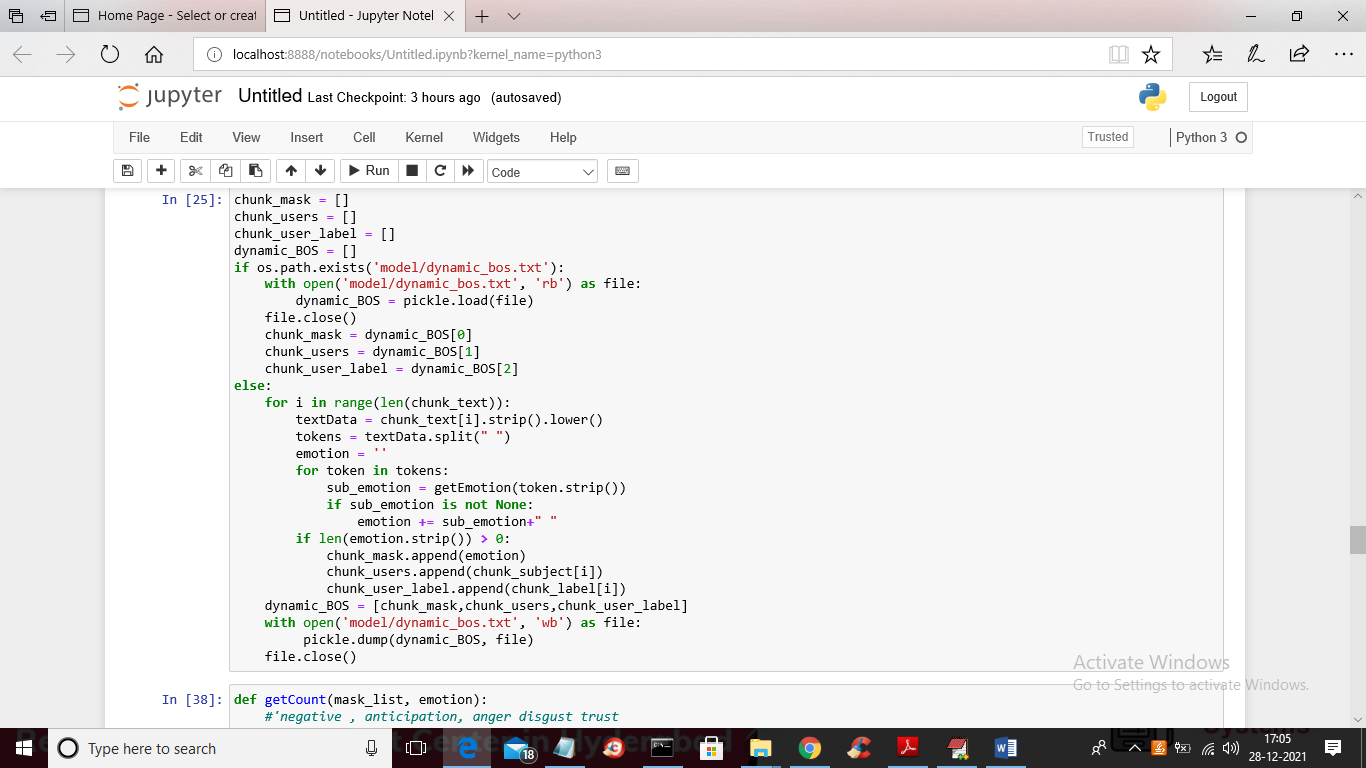


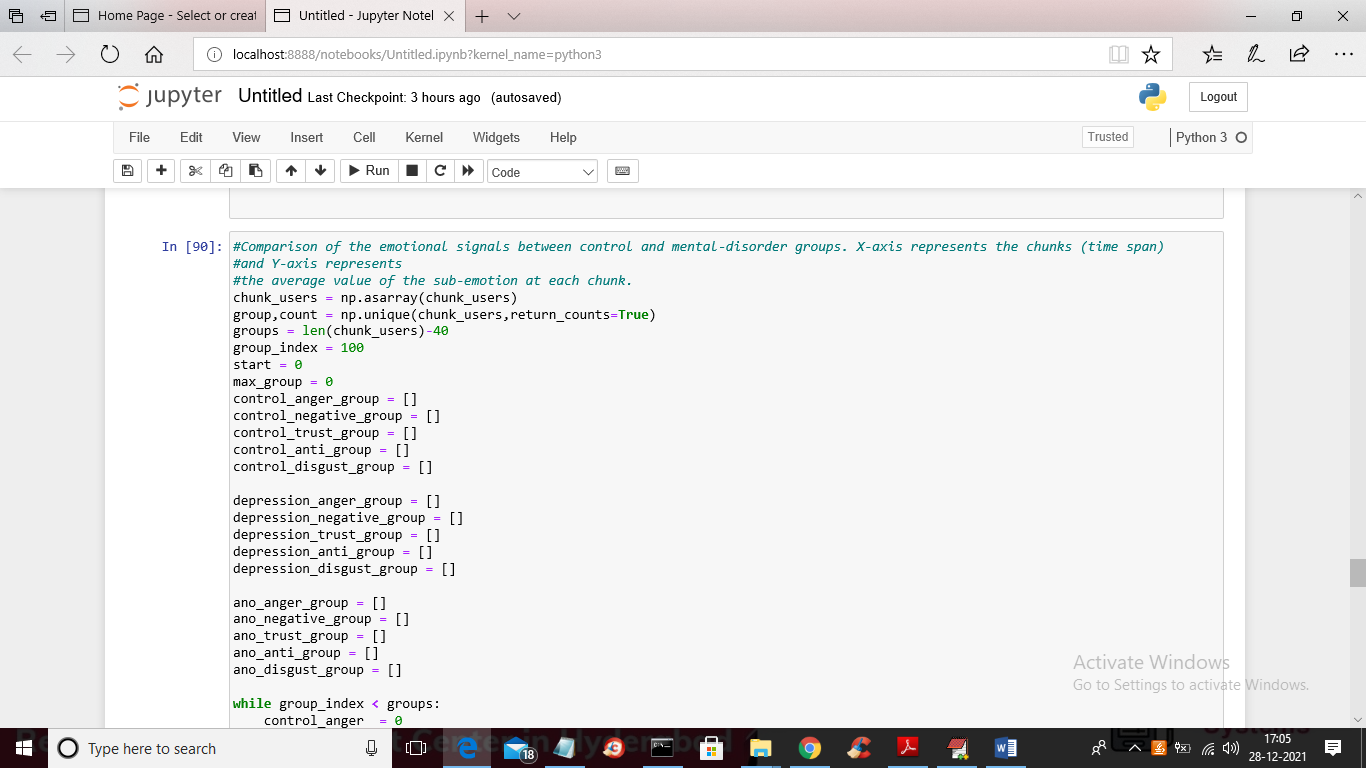
In above screen we are training TF-IDF values with SVM and CNN and then calculating F1 score and in above screen in blue colour text we got F1 score of SVM as 87 and below is the CNN result

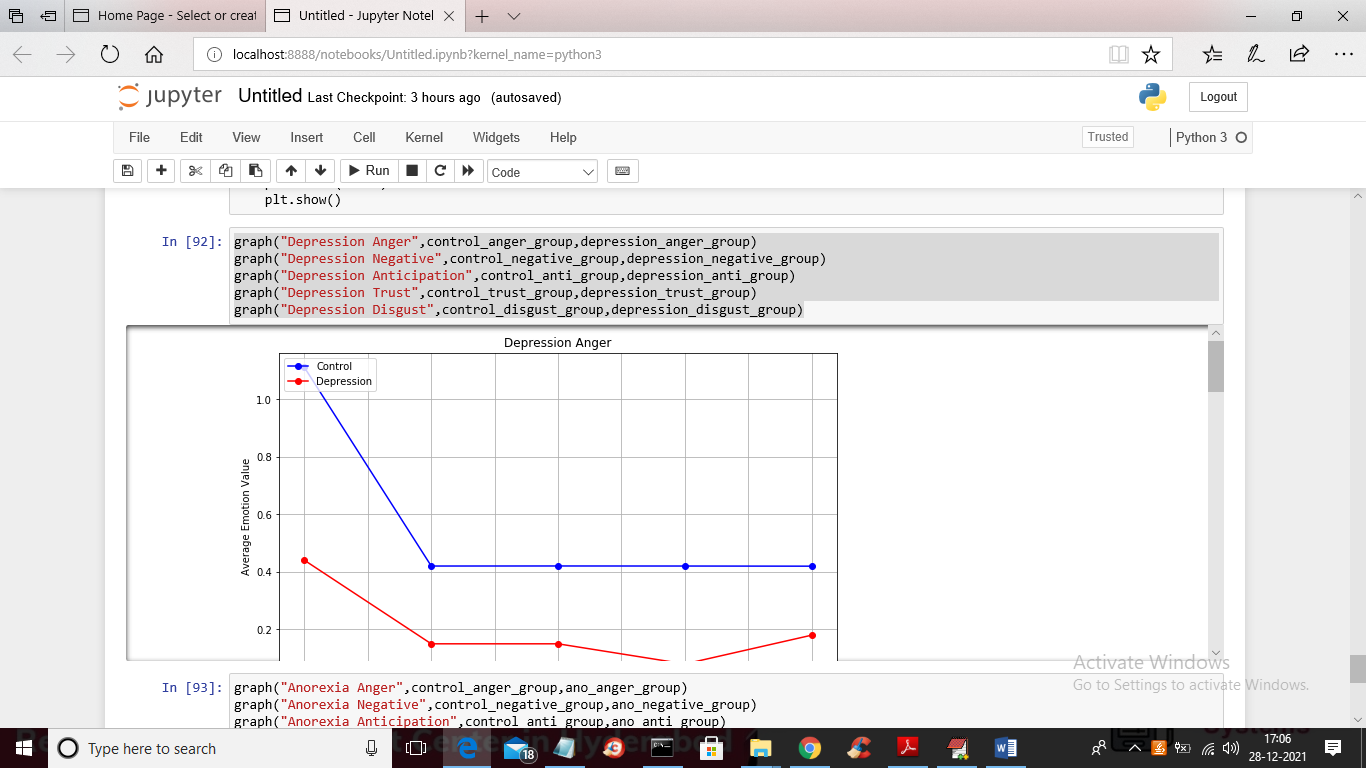




In above screen CNN F1-score is 85% and below is the dynamic BOSE output for group emotions of multiple history reviews from same user







In above screen we are getting group emotion for depression anger and similarly we calculated control, anorexia and depression group emotion for anger, disgust, anticipation, negative etc.

