**Code Explanation of Query Search**

1. Files “QuerySearch\_WITH\_PRE\_PROCESSING.py” and “QuerySearch\_WITHOUT\_PRE\_PROCESSING.py”
2. From file “input\_test\_file.csv” get WORD1 and WORD2 for each row.
3. Get the intersection of both the words for each row and create a list of intersection for each row of 2 words.
4. Then create an output file “output\_test\_file\_with\_pre\_processed.csv” or “output\_test\_file\_without\_pre\_processed.csv” where the last column “OUTPUT\_INTERSECTION” will contain 1 if intersection list is greater than 0, ie. The 2 words where identified in at least one document. Otherwise update this column with 0.
5. Now we have actual “OUTPUT\_INTERSECTION” from “input\_test\_file.csv” and Predicted “OUTPUT\_INTERSECTION” from “output\_test\_file\_with\_pre\_processed.csv” or “output\_test\_file\_without\_pre\_processed.csv”
6. From this I am creating confusion matrix to get the following parameters (accuracy, precision, recall, fmeasure) for evaluation and then generation evaluation file – “output\_test\_evaluation\_with\_pre\_processed.csv” or “output\_test\_evaluation\_without\_pre\_processed.csv”