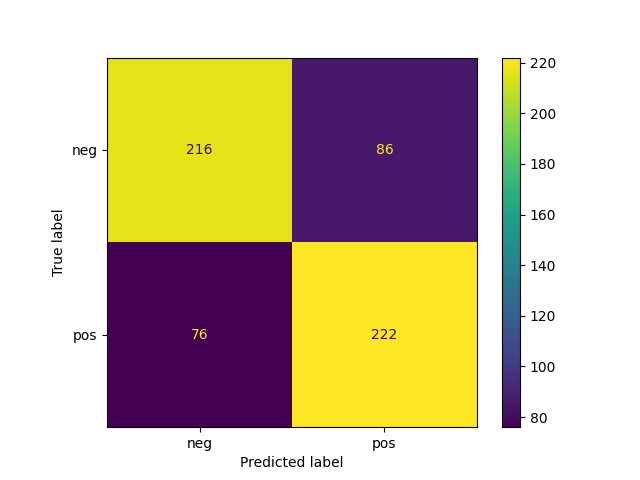
Tanat Arora- 6410381

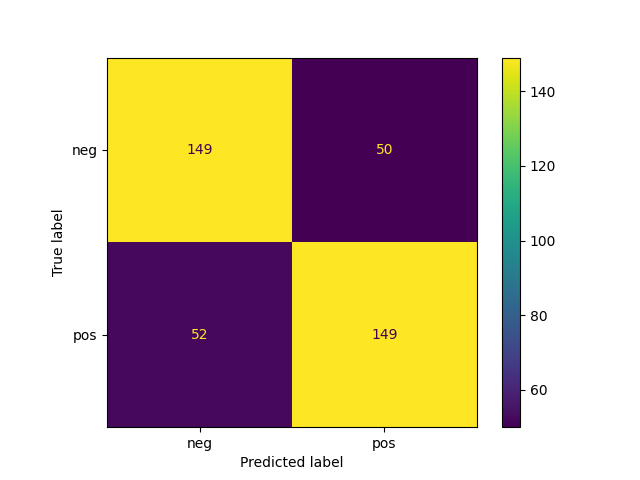
I use TFIDFVectorizer instead of count\_vec to fit transform the data because it seems to be giving a better result.

For Train/Test split ratio: 70:30

A screenshot of a computer

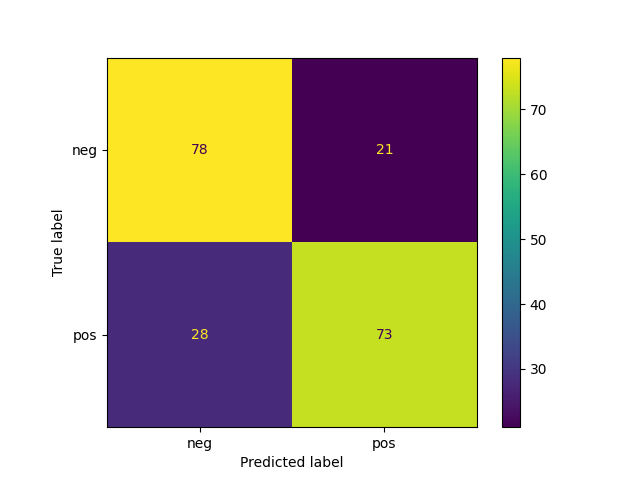
Description automatically generated with medium confidence

For Train/Test split ratio: 80:20

 A screenshot of a computer

Description automatically generated with medium confidence

For Train/Test split ratio: 90:10



A screenshot of a computer program

Description automatically generated with low confidence

If we compare the F1\_score (I am using F1\_score because it is a balanced measure of precision and recall) between the three results we got.   
We can see that the model where we split 09:01 has the F1\_score of 0.75 which is highest when compared to the other two models which are 0.74 and 0.73 respectively.

Therefore, the model with a train/test splitting ratio of 0.9:0.1 gives the best results in terms of performance.