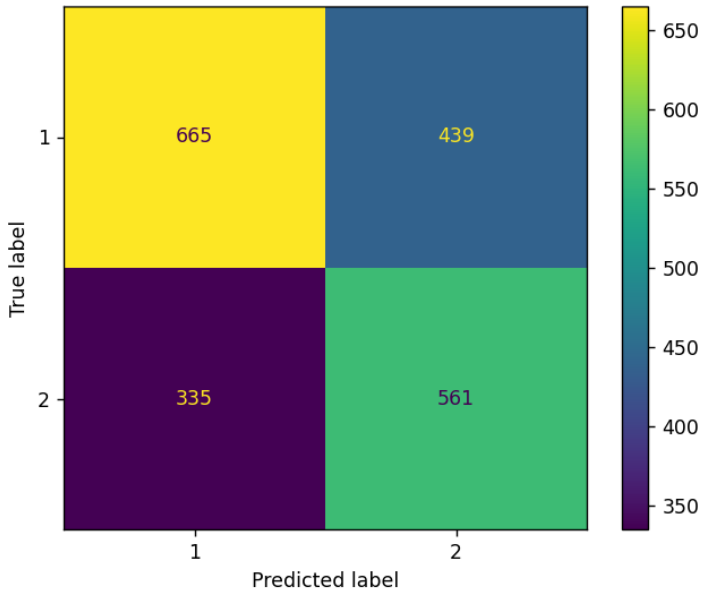
Report

1)SkLearn:

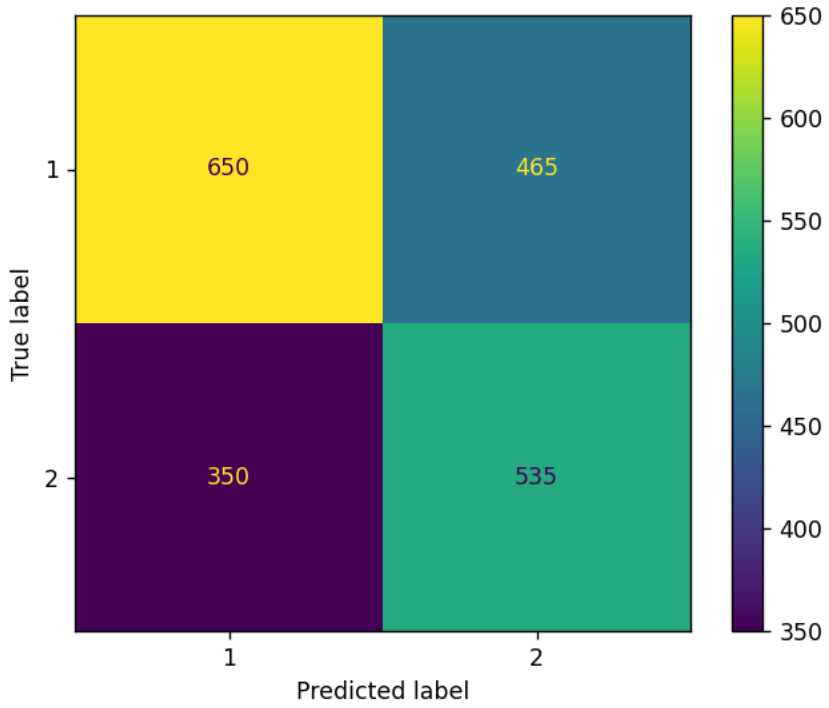
A)By using text classification using gaussianNB and countvectorizer (BOW) and the accuracy of program on sample dataset (10000 from training set &2000 from test)



0.613

	precision	recall	f1-score	support
1	0.60	0.67	0.63	1000
2	0.63	0.56	0.59	1000
accuracy			0.61	2000
macro avg	0.61	0.61	0.61	2000
weighted avg	0.61	0.61	0.61	2000

B) By using text classification using gaussianNB and tfidfvectorizer(TF-IDF) and the accuracy of program on sample dataset (10000 from training set &2000 from test)




```
0.5925

              precision    recall  f1-score   support

           1       0.58      0.65      0.61      1000
           2       0.60      0.54      0.57      1000

    accuracy                           0.59      2000
   macro avg       0.59      0.59      0.59      2000
weighted avg       0.59      0.59      0.59      2000
```

2)By using NLTK we make train test automatically because it's a training model already

	precision	recall	f1-score	support
1	0.79	0.59	0.67	1000
2	0.67	0.84	0.75	1000
accuracy			0.72	2000
macro avg	0.73	0.72	0.71	2000
weighted avg	0.73	0.72	0.71	2000

3)By using spacytextblob we make train test automatically because it's a training model already

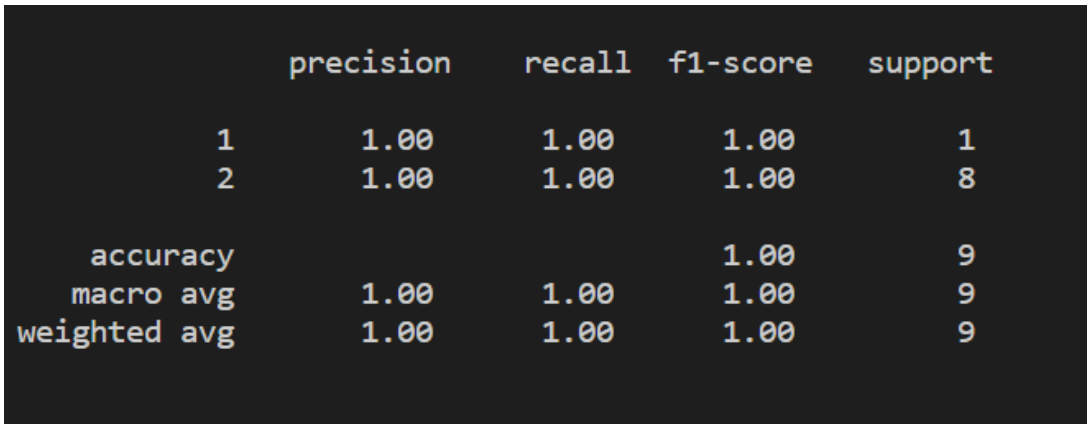
	precision	recall	f1-score	support
1	0.70	0.78	0.74	1000
2	0.75	0.67	0.71	1000
accuracy			0.72	2000
macro avg	0.73	0.72	0.72	2000
weighted avg	0.73	0.72	0.72	2000

(nlp_env)

TABLE OF ACCURACY

	Using Naïve Bayes		SIA	spacytextblob
	BOW	**TF-IDF**	**NLTK**	**Spacy**
Accuracy	**0.61**	**0.59**	**0.72**	**0.72**

After testing on a real data from Amazon using scraping using beautifulsoup on 9 records using SIA model we reach to an accuracy 100% as in the fig.



	precision	recall	f1-score	support
1	1.00	1.00	1.00	1
2	1.00	1.00	1.00	8
accuracy			1.00	9
macro avg	1.00	1.00	1.00	9
weighted avg	1.00	1.00	1.00	9