Fine-Tuning BERT for Multi-Class Text Classification

IR4
Rachelli Adler 213836687
Shiffy Scheiner 328949714
Github: https://github.com/shiffy01/IR4.git
Course number: 157125.5785
Date: 29/01/2025

Given files:
The sentences and models sentiment from previous home work. - IR3

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	ument Num	tence Num	Sentence	I/P	del1 sentim	model1 scor	del2 sentim	model2 scor	del3 sentim	model3 scor	del4 sentim	model4 scor	del 5 sentim	model 5 scor	del 6 sentim	model 6 scor	del 7 sentim	model 7 scor	majority	verage scor	final
2	1	1	pope rene	P	NEU	0.791487	NEG	0.468124	NEG	0.724178	POS	0.998213	NEG	0.472699	NEU	0.920949	NEU	0.99832	NEG	0.76771	I
3	2	1	biden is sti	I	POS	0.968796	POS	0.715602	POS	0.577471	POS	0.99867	POS	0.81125	POS	0.987807	POS	0.999273	POS	0.865553	I
4	2	2	united stat	I	NEU	0.811672	POS	0.314722	NEG	0.733376	POS	0.988878	NEU	0.485027	NEU	0.845714	NEU	0.995922	NEU	0.73933	
5	7	1	relatives o	P	NEG	0.670948	NEG	0.366417	NEG	0.775636	NEG	0.995543	NEG	0.692221	NEU	0.694596	NEU	0.937242	NEG	0.733229	I
6	8	1	icj rules isr	I	NEG	0.798723	NEG	0.511674	NEG	0.781512	NEG	0.998236	NEU	0.494758	NEU	0.528842	NEU	0.999164	NEG	0.730416	P

Step 1:
Based on the I/P column and the max of the sentiments(7 columns)
We put labels into the sentences.

```
    "anti-i": 0,
    "pro-p": 1,
    "neutral": 2,
    "anti-p": 3,
    "pro-i": 4
```

Step 2:
We summed up the categories and noticed that together we had:
neutral: 3322
anti-i: 3184
anti-p: 1273
pro-i: 469
pro-p: 59

We needed more pro palestinien sentences. So we changed the qualification to this:

If 7 or 6 of the 7 models agreed on the same sentiment then we took the sentence witht the sentiment. If 5 of the models agreed that the sentence is negative and the other two said its neutral then we took it as negative. If 4 of the models said positive and the other 3 said neutral then we took the sentence as positive.

```python
if neg>5:
    return True
if pos>4:
    return True
if neu>4:
    return True
if neg==5 and neu==2:
    return True
if pos==4 and neu==3:
    return True
return False
```

The statistics of this after the change was:
anti-i: 3184
anti-p: 1273
pro-p: 127
neutral: 693
pro-i: 712

Inorder for the data set to be balanced we took 125 of each category.

Step3:
shuffled the data set and took 125 sentences from each of the 5 categories. (shuffle so we don't take all 125 from the same magazine)
Now we have are dataset. Ready for fine tuning.

Step4:
We split the data set to train 82%, test 18%
And then the train dataset we split to train 85% and validate 15%.
```
Training set size: 435
Validation set size: 77
Testing set size: 113.
```

Step5:
Imported a tokenizer bert-base-uncased and embedded the train test and validate dateframes.

The `SentimentDataset` class is used to wrap the dataset in a format that is compatible with PyTorch's `DataLoader`, which is essential for efficient training. PyTorch requires data to be in a specific format to properly handle batching, shuffling, and parallel loading during model training.

Step6:
The model used to classifier. That we did fine tuning on `'bert-base-uncased'`
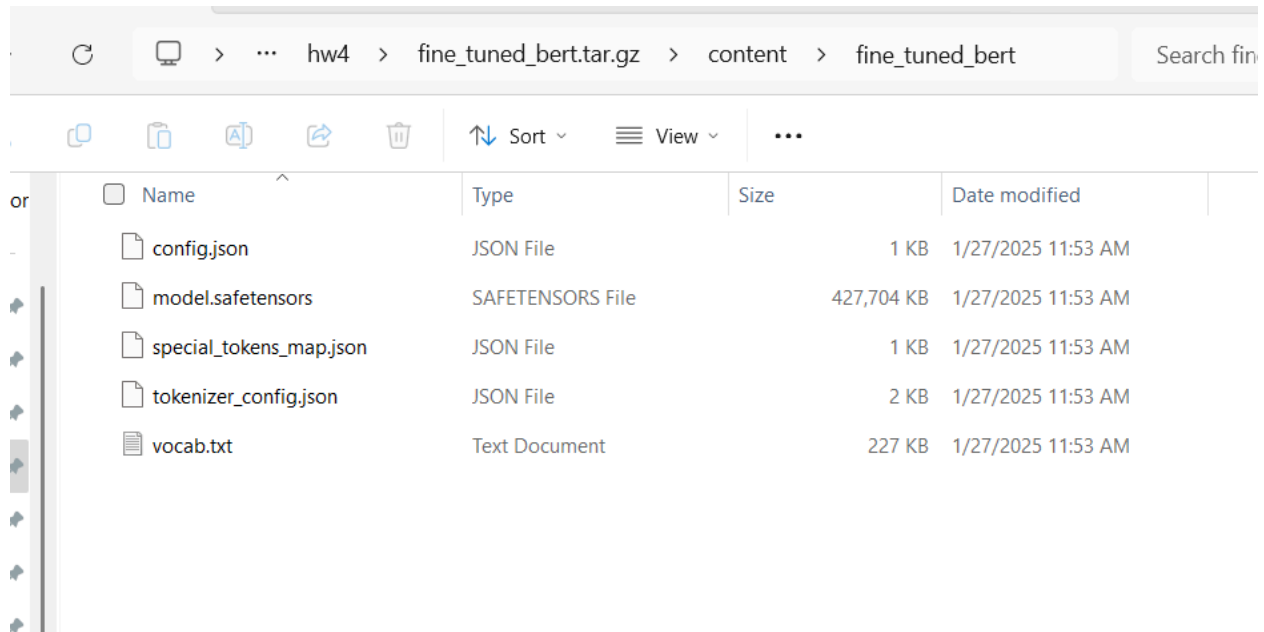We ran the model on colab: Link:
https://colab.research.google.com/drive/1IE9D3lCm4p_m4hBsHaeBZB39RFR1Ab_9?usp=sharing

```
trainer = Trainer(
wandb: WARNING The `run_name` is currently set to the same value as `TrainingArguments.output_dir`. If this was not intended, please specify
                                    [280/280 04:32, Epoch 10/10]
```

Epoch	Training Loss	Validation Loss	Accuracy	Precision	Recall	F1
1	1.638600	1.608250	0.272727	0.155795	0.272727	0.193048
2	1.553500	1.540659	0.389610	0.312002	0.389610	0.336658
3	1.488000	1.444760	0.415584	0.368031	0.415584	0.377071
4	1.317500	1.301246	0.519481	0.530177	0.519481	0.484953
5	1.080700	1.062752	0.623377	0.617292	0.623377	0.604298
6	0.783200	0.843139	0.701299	0.741337	0.701299	0.685788
7	0.498000	0.750494	0.753247	0.789336	0.753247	0.746812
8	0.255500	0.664977	0.779221	0.809536	0.779221	0.776940
9	0.101600	0.517174	0.857143	0.862748	0.857143	0.857234
10	0.046600	0.683523	0.818182	0.825294	0.818182	0.814240

```
/usr/local/lib/python3.11/dist-packages/sklearn/metrics/_classification.py:1565: UndefinedMetricWarning: Precision is ill-defined and being
  _warn_prf(average, modifier, f"{metric.capitalize()} is", len(result))
/usr/local/lib/python3.11/dist-packages/sklearn/metrics/_classification.py:1565: UndefinedMetricWarning: Precision is ill-defined and being
  _warn_prf(average, modifier, f"{metric.capitalize()} is", len(result))
                                    [8/8 00:00]
Evaluation Results on Test Set: {'eval_loss': 0.5774326920509338, 'eval_accuracy': 0.8053097345132744, 'eval_precision': 0.8163648319132031,
Updated Excel file saved to /content/with prediction.xlsx
```

The model:

Examples of sentences with a correct classification:

Anti-Israeli:
BBC article 166 sentence 17:
These inaccuracies and errors may violate israel's obligation under international law to provide "effective advanced warnings", says janina dill, co-director of the oxford institute of ethics, law and armed conflict.
Classification score: 0.964014113

BBC article 177 sentence 15:
we asked the idf how many tunnels, and what proportion of the total tunnel network, they had destroyed.
Classification score: 0.958807826042175

BBC article 151 sentence 28:
israel does not want that war either.
Classification score: 0.962320148944854

Pro-Palestinian:

J-Post article 53 sentence 10:
In recent years, Iranian training and knowledge helped Hamas upgrade its missiles and rockets, with the results being shown in this latest round.
Classification score: 0.954369664192199

BBC article 323 sentence 22:
gaza is my life.
Classification score: 0.94982773065567

J-Post article 192 sentence 8:
They launched a peaceful, popular resistance.
Classification score: 0.942691922187805

Neutral:
BBC article 206 sentence 9:
the pentagon said the visit had been scheduled previously but had been brought forward "due to recent developments".
Classification score: 0.964078724384307

BBC article 216 sentence 10:
it is thought that arrow dealt with most of the ballistic missiles aimed at israel.
Classification score: 0.950028717517852

NY Times article 454 sentence 9:
"i don't know how much time i have left," the mother said last year.
Classification score: 0.95287948846817

Anti-Palestinian:
J-Post article 347 sentence 1:
gaza factions threaten escalation, lash out at egypt over reconstruction delays
Classification score: 0.975630640983581

J-Post article 228 sentence 4:
And I call on social media organizations to immediately rid their platforms of vile posts that promote violence against Jews with such vile declarations as, "Hitler was right" he concluded.
Classification score: 0.938451528549194

J-Post article 358 sentence 3:
Hamas is under pressure, and it's getting worse.

Classification score: 0.978379011154174

Pro-Israeli:
J-Post article 193 sentence 47:
Societal Perspective: Iron Dome, as an effective defense system, significantly contributes to Israel's sense of community and national pride.
Classification score: 0.963795065879821

J-Post article 486 sentence 15:
Following her return, Netanyahu spoke with the young woman's mother who thanked him for his efforts, according to a PMO statement.
Classification score: 0.872908890247344

J-Post article 193 sentence 48:
It contributes to Israel's degree of resilience by reducing overall levels of fear and anxiety, creating a unified home front with all its civilians feeling protected.
Classification score: 0.952621102333068

Examples of sentences with an incorrect classification:

Anti-Israeli:
J-Post article 563 sentence 21:
Arabs there are under complete Israeli sovereignty, yet lack the same services that are available to Israelis.
Classification: Neutral

BBC article 158 sentence 13:
three of the killed aid workers were british nationals.
Classification: Anti-Palestinian

BBC article 375 sentence 4:
israel's ambassador to the un, gilad erdan, said the body had welcomed a "terror state" into its ranks.
Classification:Neutral

BBC article 156 sentence 14:
dr bashar murad, the executive director of the palestinian red crescent, said many of its paramedics were refusing to evacuate wounded and sick people from the north to the south in their ambulances due to the lack of safety guarantees, despite their protection under international law and co-ordination efforts by international organisations.

Classification: Anti-Palestiinian

Pro-Palestinian:

BBC article 9 sentence 27:
but he felt hopeful he was on the verge of finally getting out of gaza.
Classification: Neutral

BBC article 443 sentence 4:
"he just embodied what it meant to be a royal marine commando - courage,
determination, unselfishness" andrew campbell was also deployed to afghanistan with
mr henderson.
Classification: Pro-Israeli

J-Post article 239 sentence 17:
The success of those initiatives was helped, in part, by the calm that existed in Gaza,
which Mladenov had helped maintain.
Classification: Neutral

NY-Times article 427 sentence 4:
now, this green, a variety of mallow, is making up an outsize portion of many gazans'
diets by providing an inexpensive way to blunt hunger.
Classification: Anti-Palestinian

Neutral:
BBC article 76 sentence 2:
but our group brings together women from these three different faiths," ms mirza said.
Classification score: Pro-Palestinian

BBC article 458 sentence 2:
yet there is an increasing focus on what could and should happen if and when the guns
fall silent in gaza.
Classification: Anti-Palestinian

BBC article 96 sentence 9:
he later visited a hospital where injured palestinians are being treated.
Classification: Anti-Palestinian

BBC article 356 sentence 19:

the israeli leader would have to choose between his government's survival and the vital support the us president has given him in recent months.
Classification score: Pro-Israeli

Anti-Palestinian:
BBC article 9 sentence 22:
ahmad, who lives in gaza but has a friend in canada fundraising for him, says he has to leave because of chronic knee inflammation, the medication for which is no longer available in gaza.
Classification: Neutral

J-Post article 563 sentence 10:
I decided not to turn, and a few seconds later a huge explosion went off.
Classification: Anti-Israeli

BBC article 237 sentence 2:
if, for some reason, the brent crude prices spike, or if there is some kind of freedom of navigation issue in the gulf.
Classification: Anti-Israeli

BBC article 185 sentence 17:
i have too many friends there, in gaza" no-one has been able to show me around the ruins of khan younis or gaza city, or the tents of around 1.4 million displaced civilians in rafah, in the way ron bahat showed me nir oz.
Classification: Pro-Palestinian

Pro-Israeli:
BBC article 97 sentence 25:
preparing for elections is, sabri saidam believes, the biggest feasible step to restore faith in palestinian politics.
Classification: Pro-Palestinian

J-Post article 132 sentence 8:
PA Security Forces, the US-trained Palestinian security forces, which cooperate with the IDF, are considered an essential pillar of stability in the West Bank.
Classification: Pro-Palestinian

BBC article 274 sentence 11:
the force said: "we have reflected on the strength of the response to our previous statement."

Classification: Neutral

J-Post article 193 sentence 31:
In the decade following that decision, the US significantly increased its financial contribution to equipping additional Iron Dome systems and launchers.
Classification: Pro-Palestinian

Insights on the classification of whole articles:

Another interesting thing that can be done with this model is to classify entire articles (rather than just sentences) by taking all the sentences in the article and finding the average or most common label. For example, in our data we have a few sentences from article 9 of the BBC Newspaper, and our sentence classifications were all predicted as neutral. So we can conclude that the whole article is neutral.

BBC	9	1	however,	neutral	2	2	0.960327446
BBC	9	9	some peo	neutral	2	2	0.963774681
BBC	9	27	but he felt	pro-p	1	2	0.959007025
BBC	9	22	ahmad, w	anti-p	3	2	0.607828319

Another example is article 70 in The Jerusalem Post. The most common predicted label is Anti-Israel:

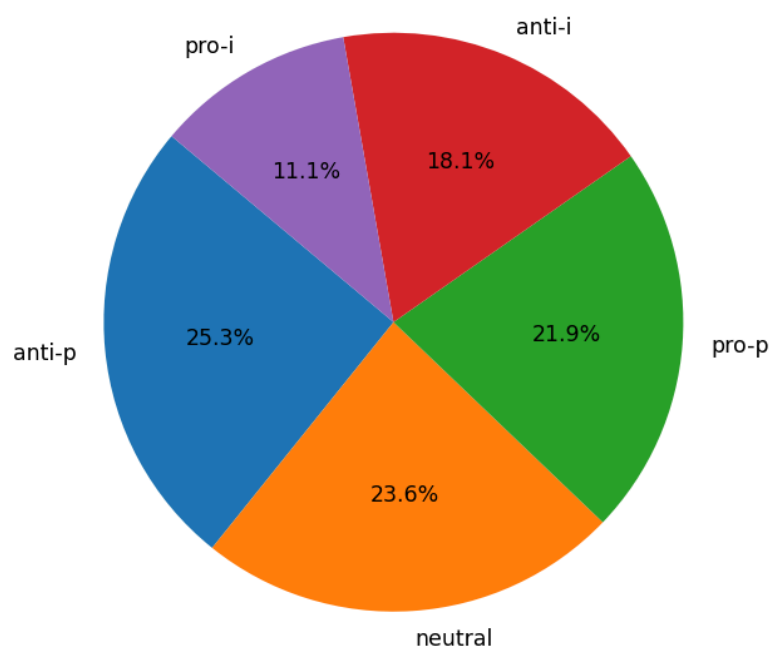| jpost	70	17	Rahat is a	neutral	2	2	0.956145942
jpost	70	25	Anyone w	anti-i	0	0	0.929527998
jpost	70	24	According	neutral	2	2	0.96256727
jpost	70	7	According	anti-i	0	0	0.96960777
jpost	70	1	israel's be	anti-i	0	0	0.963752747

Another example:
This show that this article is pro israeli.

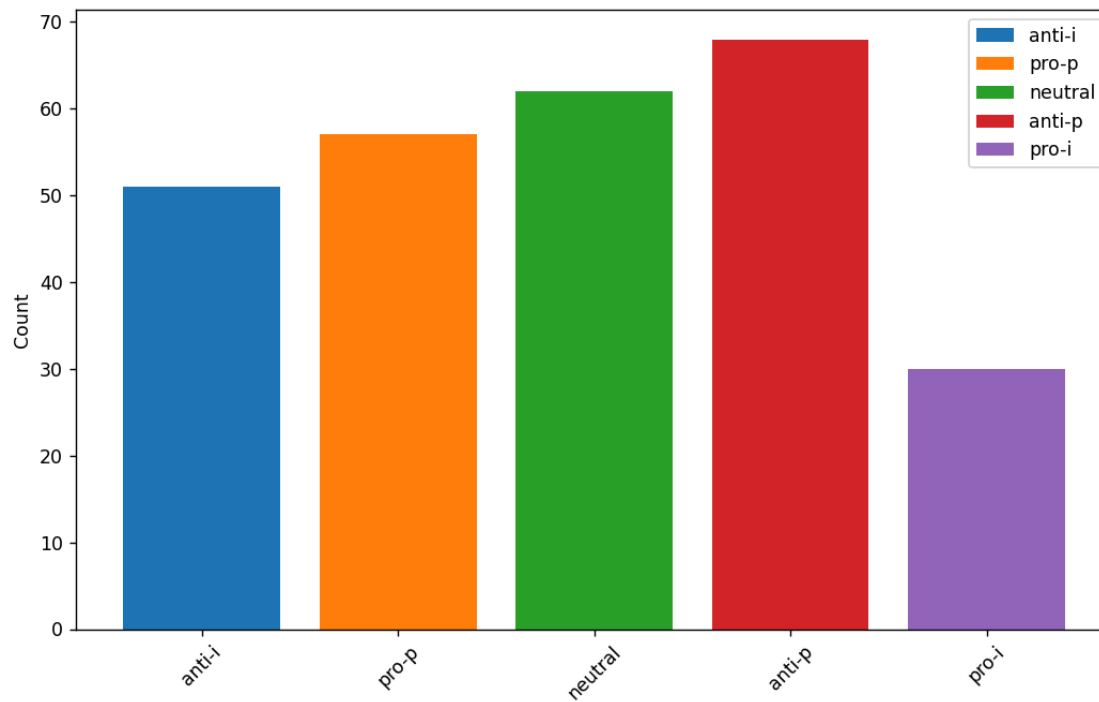| 50	jpost	107	1	museum of tolerance jerusalem spotligh	pro-i	4	4	0.77139
51	jpost	107	6	Each clip was created to ignite a light c	pro-i	4	4	0.956641
52	jpost	107	7	In this harsh reality, this important exhib	pro-i	4	4	0.930724
53	jpost	107	9	The volunteers, the fighters, the rescue	pro-p	1	1	0.941316

In theory this can be applied to all the articles.

Statistics newspaper
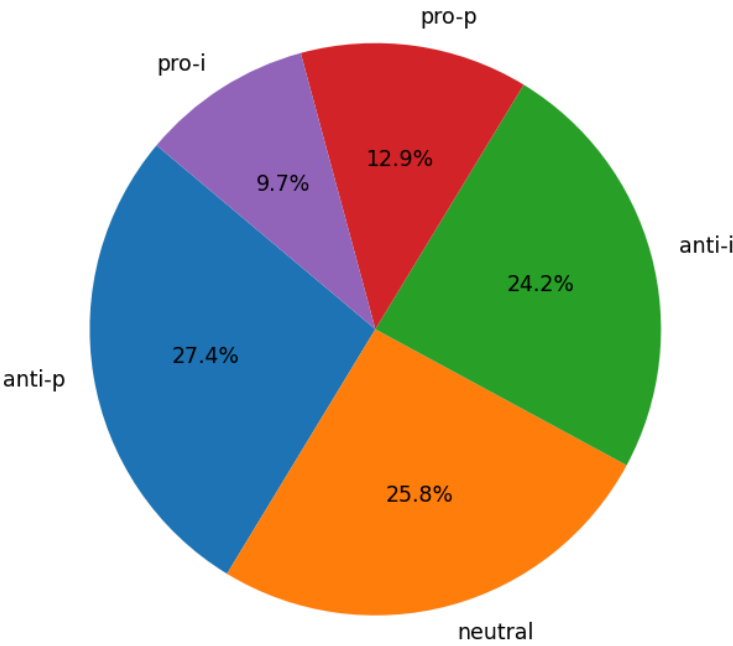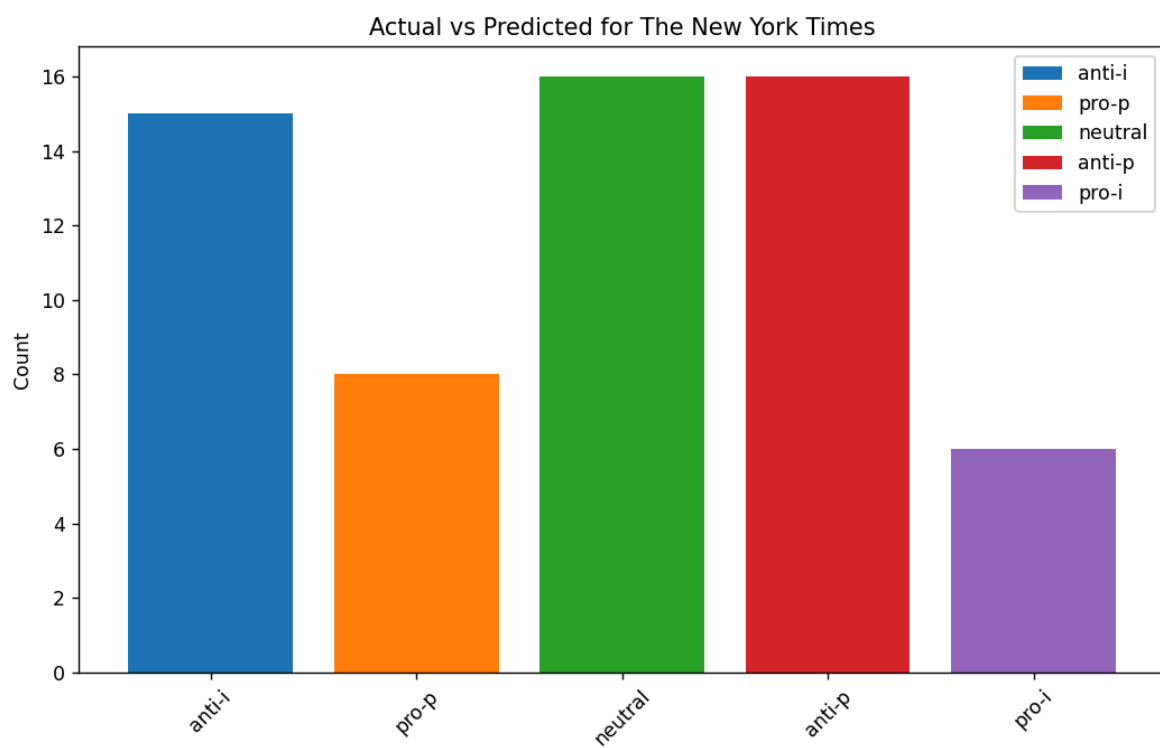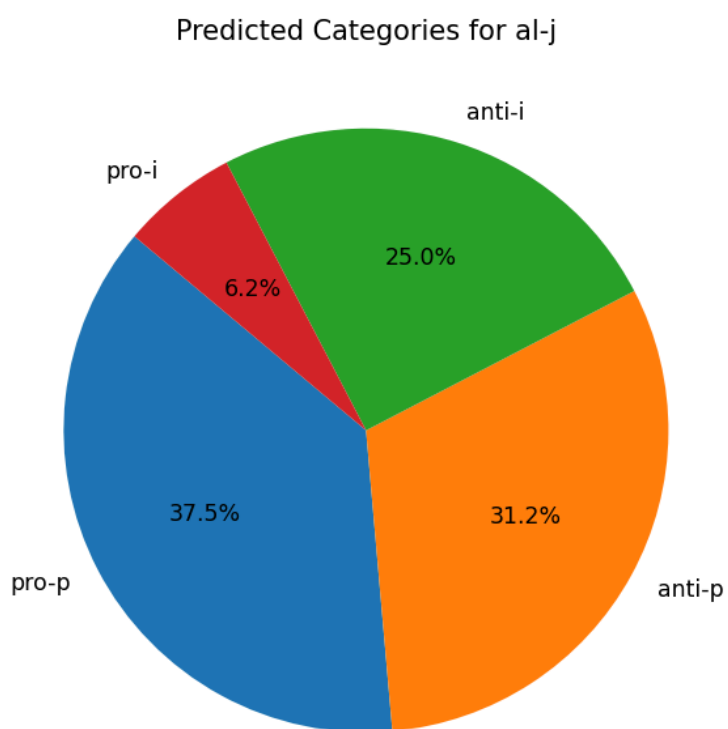
Predicted Categories for BBC

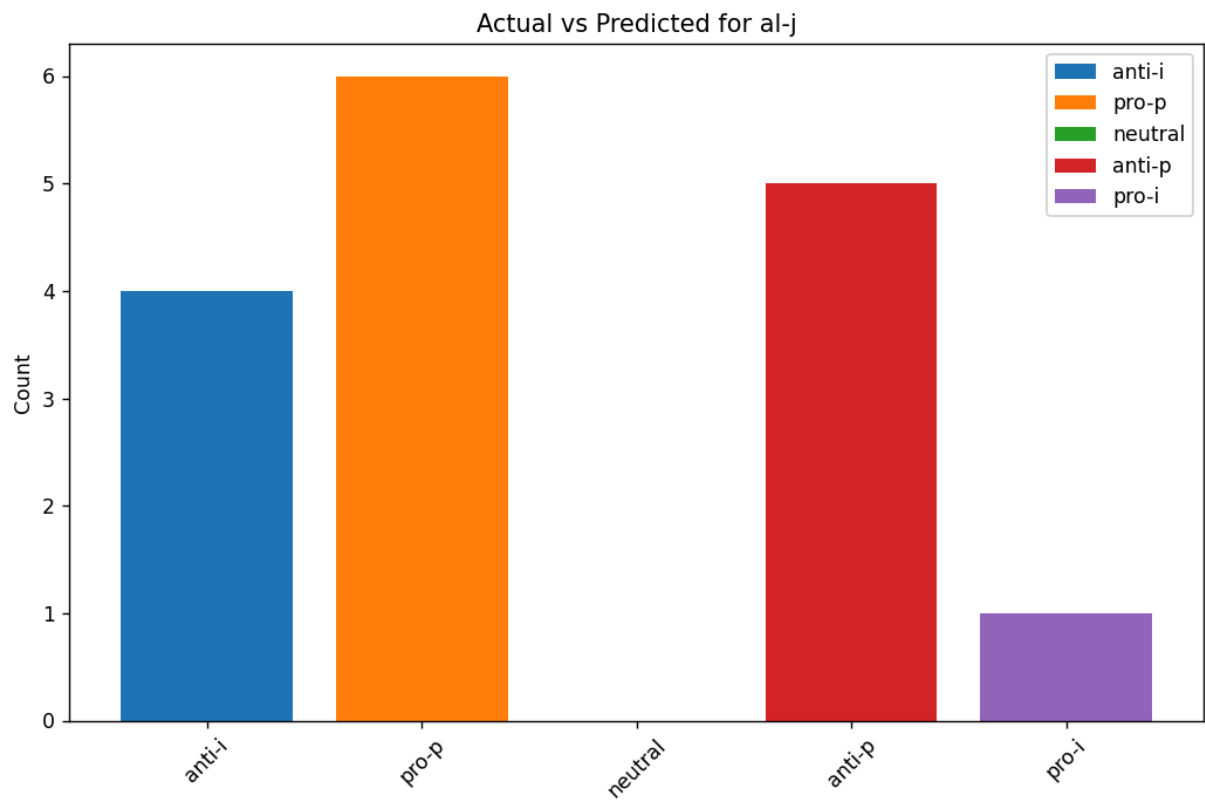Actual vs Predicted for BBC
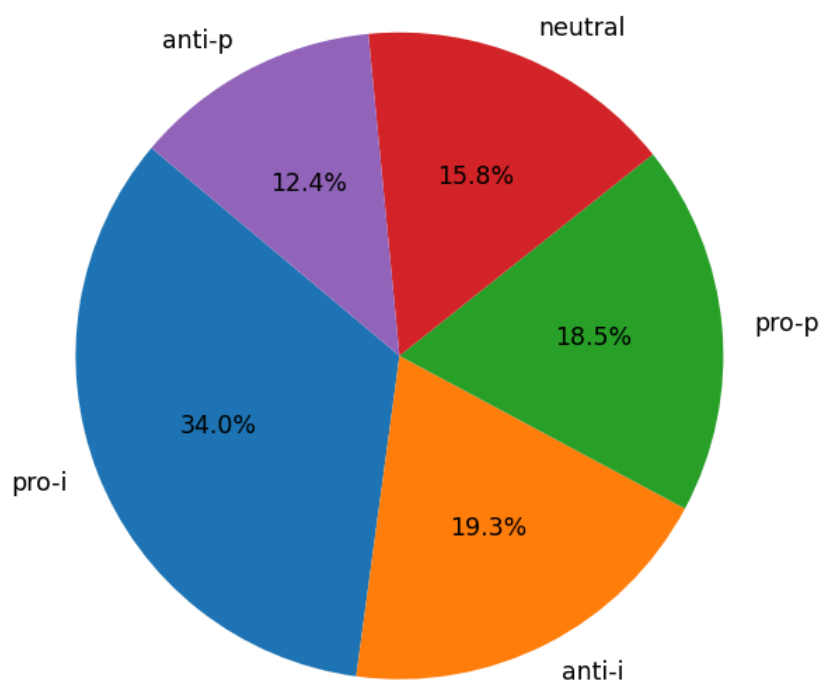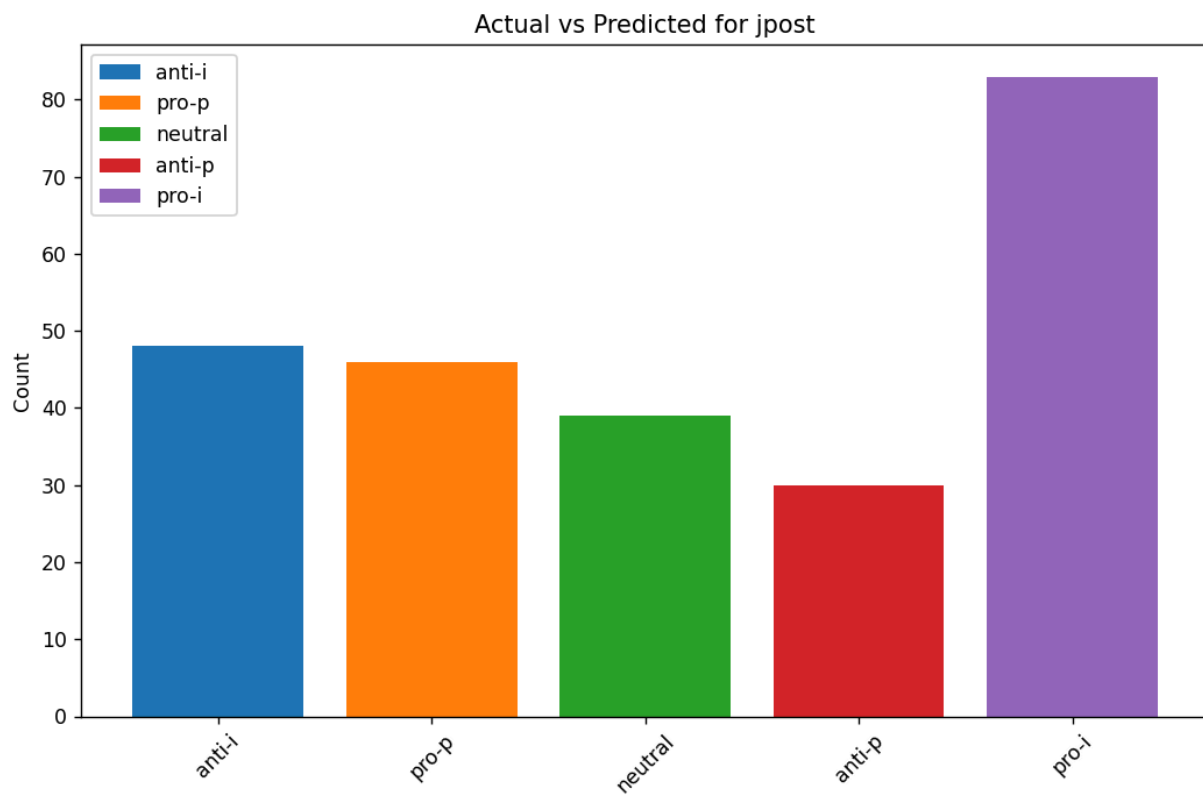
NY TIMES

Predicted Categories for The New York Times

Actual vs Predicted for The New York Times

AL-J

Predicted Categories for al-j

Actual vs Predicted for al-j

JPOST

Predicted Categories for jpost

Actual vs Predicted for jpost

Prediction accuracy per newspaper

Statistics for BBC:
{'total_sentences': 288, 'correct_predictions': 268, 'accuracy': 93.05555555555556,
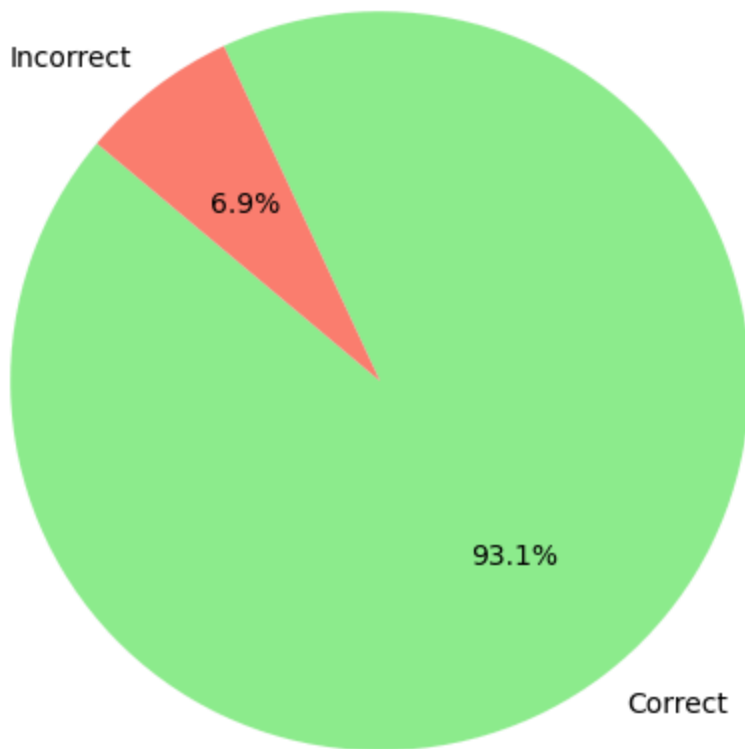'label_counts': Predicted Label
3   73
2   68
1   63
0   52
4   32
Name: count, dtype: int64}

Prediction Accuracy for BBC

Statistics for The New York Times:
{'total_sentences': 62, 'correct_predictions': 61, 'accuracy': 98.38709677419355,
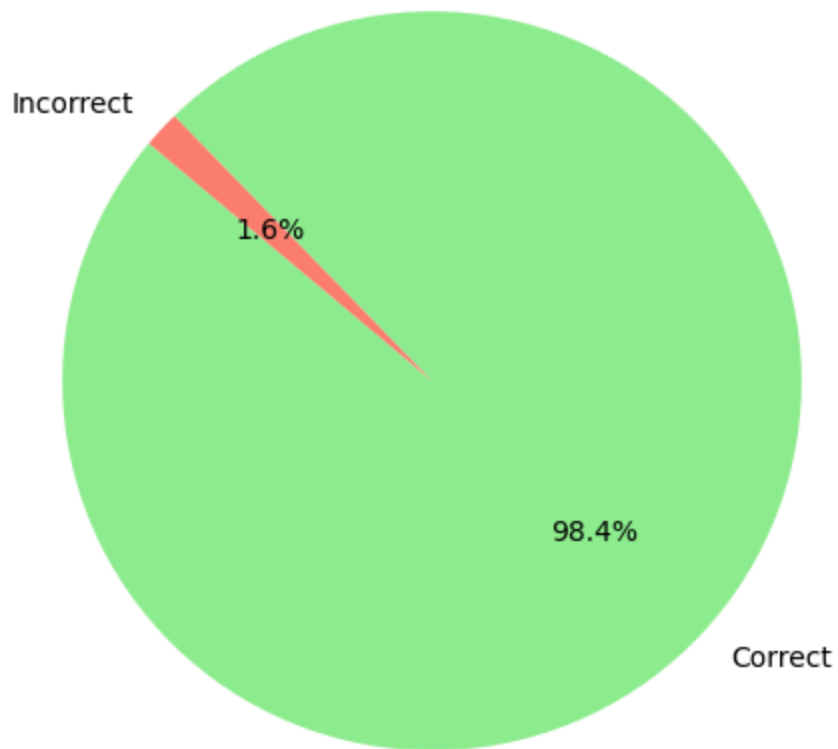'label_counts': Predicted Label
3   17
2   16
0   15
1   8
4   6
Name: count, dtype: int64}

Prediction Accuracy for The New York Times

Statistics for al-j:
{'total_sentences': 16, 'correct_predictions': 16, 'accuracy': 100.0, 'label_counts': Predicted
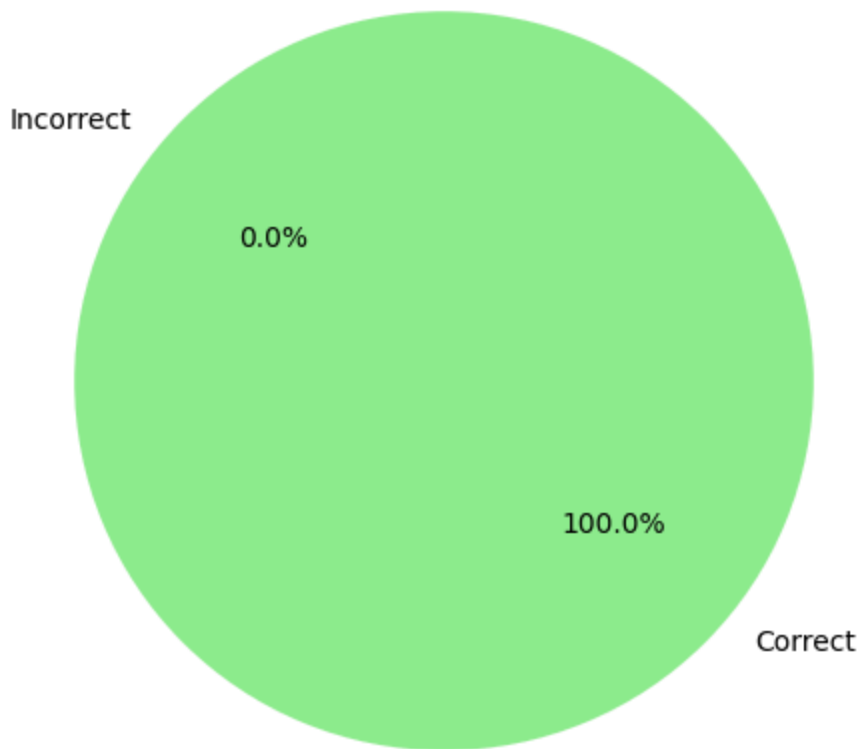Label
1   6
3   5
0   4
4   1
Name: count, dtype: int64}

Prediction Accuracy for al-j

There were only 16 article in al-j so it so happens that those 16 are categorized correctly. Might be overfitting.

When training and testing the model it took random sentences without looking at the newspaper. al- j 's sentences were very clear cut so that what it didnt have any debate where to categorize the sentence.

Statistics for jpost:

{'total_sentences': 259, 'correct_predictions': 246, 'accuracy': 94.98069498069498, 'label_counts': Predicted Label
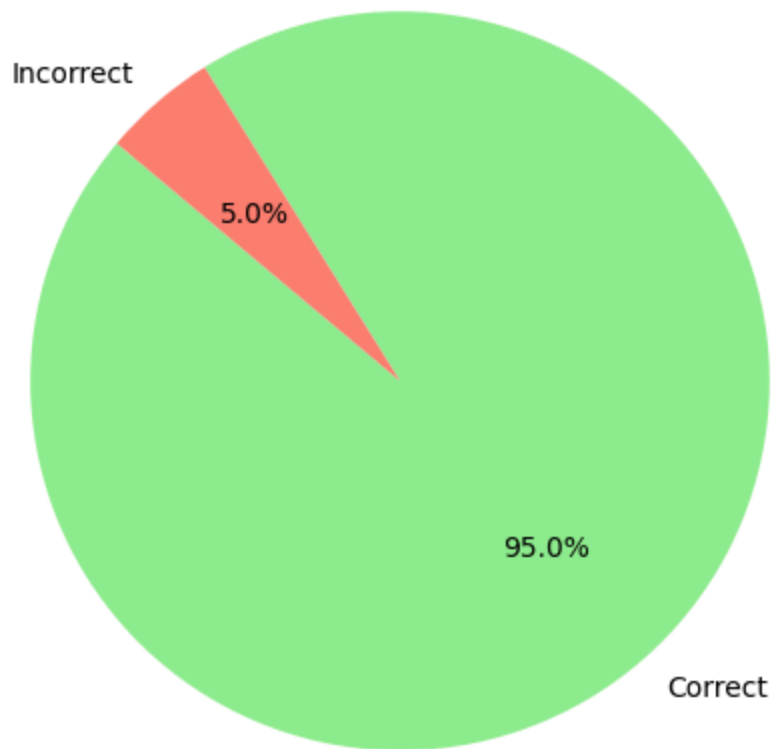
4	88
0	50
1	48
2	41
3	32

Name: count, dtype: int64}

Prediction Accuracy for jpost

Conclusion:

The model is quite good at predicting the sentiment of the sentence.
Weaknesses: it would have been better to train the model on more sentences. Or even on articles. But from research the model works pretty good.