TEXT SENTIMENT FROM TWITTER

NAME	STUDENT IDENTIFICATION NUMBER
Andrianakou Stavroula	2022202404018
Kapelou-Lampazevits Ioustina	2022202404003
Kyriakidis Giorgos	2022202404022

- PROFESSOR: THEODOROS GIANNAKOPOULOS

DATA PRESENTATION AND PREPROCESSING

The task is to judge the sentiments from different messages(tweets/texts).

We decided to keep only the instances with the label Positive and Negative

First of all emojis were spotted in the text so we converted the emoji face to text because the emojis are crucial and reflect the sentiment

Next we removed from the text the https, @, www characters , string punctuation converted all the words to lowercase removed the stop words

We had to convert each word to vector using glove-wiki-gigaword-100. We did it strategically because we wanted our neural network to generalize good that why we choose a general language(Wikipedia) and not specific

For the simple forward neural network we just created the average for each sentence (tweet)

For the CNN and RNN (simple RNN, bidirectional RNN, bidirectional LSTM) a different technique was used. For each word in each tweet, we created a vector we stacked them together and created a table(matrix) for each tweet.

Network Architectures

SIMPLE NN 256NEURONS -> 64NEURONS-> 1 OUTPUT

Validation dataset results

Classification Report:

	precision	recall	f1-score	support
0	0.87	0.88	0.88	4382
1	0.87	0.85	0.86	3900
accuracy			0.87	8282
macro avg	0.87	0.87	0.87	8282
weighted avg	0.87	0.87	0.87	8282

Test dataset results

	precision	recall	f1-score	support
Negative	0.91	0.94	0.92	266
Positive	0.94	0.91	0.92	277
accuracy			0.92	543
macro avg	0.92	0.92	0.92	543
weighted avg	0.92	0.92	0.92	543

CNN 1 CONV LAYER 128 FILTERS-> 64 NEURONS -> 1 OUTPUT

CNN 1 CONV LAYERS 128 FILTERS-> CNN 1 CONV LAYERS 128 FILTERS-> 64 NEURONS ->1 OUTPUT

CNN1 CONV LAYER 128 FILTERS -> CNN 1 CONC LAYERS 128 FILTERS -> CNN 1 CONV LAYERS 128 FILTERS -> 64 NEURONS -> 1 OUTPUT

Validation dataset results(CNN 3 CONV)

Classification Report:

	precision	recall	f1-score	support
0	0.95	0.87	0.90	4382
1	0.86	0.94	0.90	3900
accuracy			0.90	8282
macro avg	0.90	0.91	0.90	8282
weighted avg	0.91	0.90	0.90	8282

Test dataset results(CNN 3 CONV)

Classification Report:

	precision	recall	f1-score	support
0	0.98	0.96	0.97	266
1	0.96	0.98	0.97	277
accuracy			0.97	543
macro avg	0.97	0.97	0.97	543
weighted avg	0.97	0.97	0.97	543

Network architectures

SIMPLE RNN 128 UNITS -> 64 NEURONS-> 1 OUTPUT

BIDIRECTIONAL RNN 128 UNITS X 2=256 -> 64 NEURONS -> 1 OUTPUT

BIDIRECTIONAL LSTM 128 UNITS X2 = 256 -> 64 NEURONS -> 1 OUTPUT

Validation dataset results(LSTM)

Classification Report:

	precision	recall	f1-score	support
0	0.94	0.90	0.92	4382
1	0.90	0.93	0.91	3900
accuracy			0.92	8282
macro avg	0.92	0.92	0.92	8282
weighted avg	0.92	0.92	0.92	8282

Test dataset results(LSTM)

Classification Report:

	precision	recall	f1-score	support
0	0.98	0.98	0.98	266
1	0.98	0.98	0.98	277
accuracy			0.98	543
macro avg	0.98	0.98	0.98	543
weighted avg	0.98	0.98	0.98	543

Combination of tweets and movies

We tried the zero-shot technique where we used our 3 conv layer CNN on the test data of the second dataset movies reviews.

PERCENTAGE OF MOVIES USING IN TWITTER	COMBINE TWITTER AND PART OF MOVIES ACCURACY
0% ZERO SHOT	0.63
10%	0.68
20%	0.69
30%	0.73
100%	0.75

Transfer Learning

- We experimented also with transfer learning. We took our saved model CNN that was trained on the dataset twitter and learned the weights of that sentiment prediction and used it on the different the percentage of the movies dataset to retrain the Dense part of it

percentage	transfer model twitter to movies (3 conv frozen)	2 conv frozen	1 conv frozen
0%	-	-	-
10%	0.63	0.63	0.64
20%	0.61	0.63	0.64
30%	0.6	0.66	0.66
100%	0.6	0.72	0.75

Only use B (movies dataset)

Last but not least we tried to train the CNN on the dataset b (each time used different percentage) to see what will happen and if this specific architecture is appropriate for the movies dataset

percentage	Only use movies
0%	-
10%	0.6
20%	0.59
30%	0.65
100%	0.78

Conclusion

We can conclude that the movies and twitter are two different datasets that have very different languages and tones that are used. An assumption can be made that the movie’s dataset uses a more formal tone. Thats why during the transfer learning only 1 con layer must be kept frozen to achieve good accuracy. Still when the CNN was trained on the dataset movies it proved that more tailored models or architecture could potentially achieve even higher performance. Overall, these findings illustrate the trade-offs between generalization and domain-specific learning and emphasize the value of controlled fine-tuning in cross-domain sentiment analysis

THANK YOU FOR YOUR ATTENTION!!!