

NLP Lab 3 Report

BENYAMINA YACINE LAZREG IASD - G01

`yl.benyamina@esi-sba.dz`

26 avril 2024

1 Results from previous labs

After exporting the cleaned and stemmed text which was realised in Lab 1 and training it using different techniques and models (Lab 2) the following results were obtained :

	Model	Vectorizer	Accuracy	Precision	Recall	F1-score
0	relu	cv	0.751405	0.756136	0.751405	0.750617
1	relu	tfidf	0.763834	0.764323	0.763834	0.763673
2	logistic	cv	0.771667	0.772280	0.771667	0.771728
3	logistic	tfidf	0.758386	0.758884	0.758386	0.758201
4	tanh	cv	0.739826	0.740088	0.739826	0.739812
5	tanh	tfidf	0.738805	0.739987	0.738805	0.738905

FIGURE 1 – Results of using a combination of different activation functions and vectorizations methods

	Model	Vectorizer	Accuracy	Precision	Recall	F1-score
0	relu	cbow	0.490039	0.509813	0.490039	0.472720
1	relu	sg	0.650604	0.651380	0.650604	0.650786
2	relu	ft	0.396561	0.406460	0.396561	0.260886
3	relu	gl	0.560531	0.559977	0.560531	0.559292
4	logistic	cbow	0.404223	0.431555	0.404223	0.262031
5	logistic	sg	0.625575	0.627517	0.625575	0.624826
6	logistic	ft	0.404904	0.335679	0.404904	0.237868
7	logistic	gl	0.570918	0.573939	0.570918	0.571613
8	tanh	cbow	0.509109	0.516222	0.509109	0.503290
9	tanh	sg	0.651967	0.652010	0.651967	0.651492
10	tanh	ft	0.396731	0.262799	0.396731	0.253020
11	tanh	gl	0.584199	0.584849	0.584199	0.579401

FIGURE 2 – Results of using a combination of different activation functions and embedding methods

According to the figures we can clearly see that the model which uses CountVectorizer and Sigmoid activation function has the best results across all metrics, this is going to be the model that we'll try to improve.

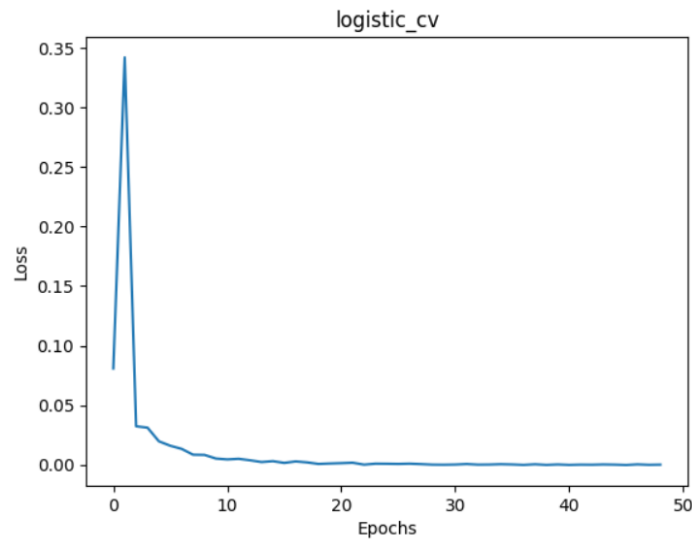


FIGURE 3 – The loss graph of the best model

2 Improving the model

The model was trained on a simple architecture of two hidden layers of size (32, 64) with Sigmoid activation function, a learning rate of 0.01 and Adam Optimizer.

The trivial thing to do is to try Random Search to try and improve the architecture of the neural network, therefore Random Search was realized with 3 target variables;

- The number of neurons in the first layer which is a choice between [8, 16, 32, 64, 128] neurons
- The number of neurons in the second layer which is a choice between [8, 16, 32, 64, 128] neurons
- The learning rate which is a random value between 0.0001 and 0.01 and follows the log sampling

It is also worth noting that adding Dropout as another variable with random value between 0.1 and 0.5 was tried but canceled later because after running more than 20 trials with dropout as a parameter the model could not achieve an accuracy higher than 40%

Running the Random Search for 100 trials, with Adam optimizer and 10 epochs each yields the following results :

```
Trial 100 Complete [00h 01m 02s]
val_accuracy: 0.7217149138450623

Best val_accuracy So Far: 0.8258783519268036
Total elapsed time: 02h 16m 43s
```

FIGURE 4 – Result of 100 trials

```
{'units_1': 64, 'units_2': 32, 'learning_rate': 0.0004503919820457184}
```

FIGURE 5 – The best Hyper parameters

The best obtained Hyper parameters were :

- 64 neurons for the first hidden layer
- 32 neurons for the second hidden layer
- A learning rate of 0.00045039

3 Final results and discussion

After running the model with the best Hyper parameters for 15 epochs the following results were obtained :

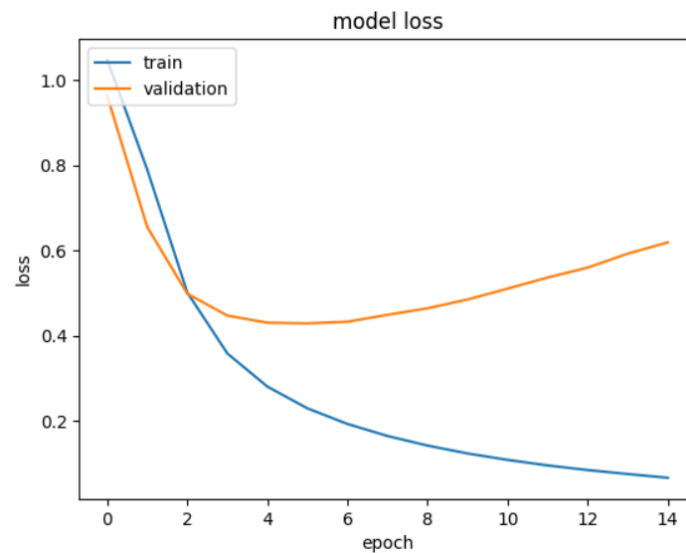


FIGURE 6 – Graph of model loss

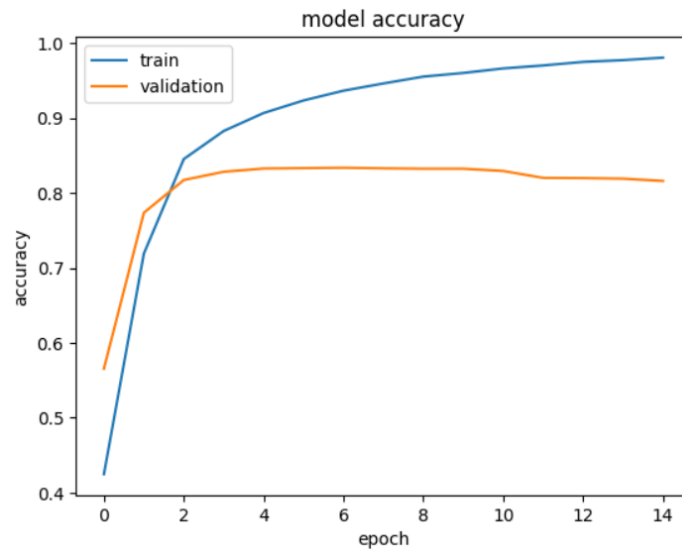


FIGURE 7 – Graph of model accuracy

The model achieves a validation accuracy of 82% but as we can see from the graph the validation loss starts to go up after 5 epochs, which indicates that model is overfitting, As a solution L2 Regularization could be used.

After running the model for 100 epochs with L2 Regularization coefficient of 0.01, the final obtained results are :

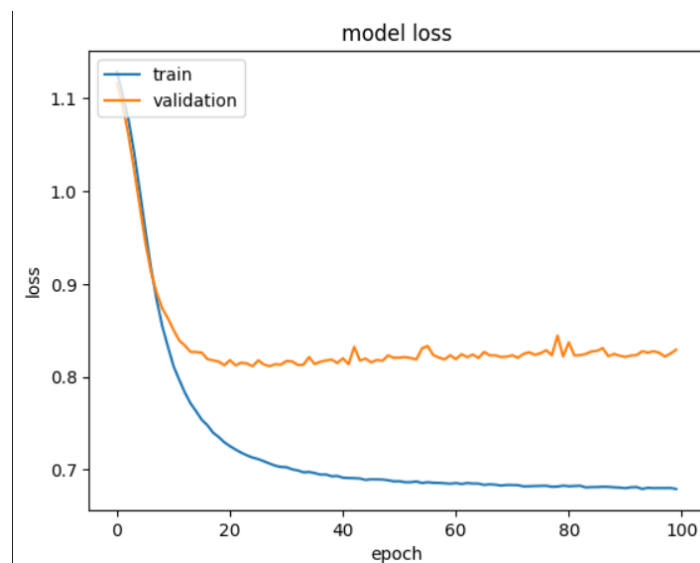


FIGURE 8 – Loss after adding L2 Regularization

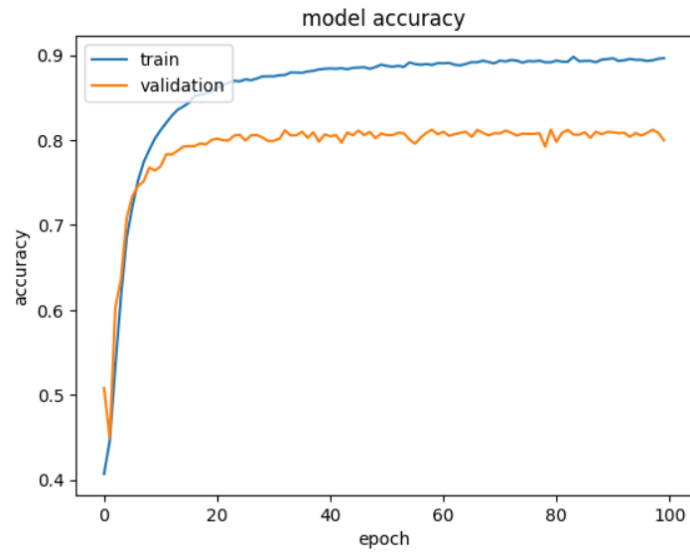


FIGURE 9 – Accuracy after adding L2 Regularization

3.1 Metrics results :

- Validation accuracy : 0.8123084902763367
- Validation precision : 0.8542508161067963
- Validation recall : 0.7597037553787231
- Validation F1 score : 0.8042079549144123

As we can see we have improved all metrics except for recall.