UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

LABORATORY EXCERCISE

Deep Learning 1, laboratory exercise 3

Antonio Lukić
Mentor:

Zagreb, May 2024.

CONTENTS

1. Analysis

1.1. Task 2

Baseline Model Results

Parameters:

Learning rate: 1×10^{-4}

Batch size: 32

Epochs: 5

Test Results:

- **Average Test Accuracy:**

$$\frac{0.7764 + 0.7110 + 0.6709 + 0.7580 + 0.7810}{5} = 0.73946$$

- **Average F1 Score:**

$$\frac{0.7816 + 0.6519 + 0.5480 + 0.7781 + 0.7750}{5} = 0.70692$$

- **Standard Deviation of Test Accuracy:** 0.0424
- **Standard Deviation of F1 Score:** 0.0933
- **Average Confusion Matrix:**

$$\frac{1}{5}\left(\begin{bmatrix}328 & 116\\ 79 & 349\end{bmatrix} + \begin{bmatrix}384 & 60\\ 192 & 236\end{bmatrix} + \begin{bmatrix}411 & 33\\ 254 & 174\end{bmatrix} + \begin{bmatrix}291 & 153\\ 58 & 370\end{bmatrix} + \begin{bmatrix}352 & 92\\ 99 & 329\end{bmatrix}\right)$$

$$= \frac{1}{5}\begin{bmatrix}1766 & 454\\ 682 & 1458\end{bmatrix} = \begin{bmatrix}353.2 & 90.8\\ 136.4 & 291.6\end{bmatrix}$$

1.2. Task 3

Comparative Analysis of RNN Cells

Parameters:

Learning rate: 1×10^{-4}

Batch size: 32

Epochs: 5

Table 1.1: Comparative Analysis of Different RNN Cells

Cell Type	Average Test Accuracy	Average F1 Score	Average Confusion Matrix
GRU	**0.7748**	0.7698	$\begin{bmatrix} 346.6 & 97.4 \\ 97 & 329 \end{bmatrix}$
LSTM	0.7548	**0.7711**	$\begin{bmatrix} 296.8 & 147.2 \\ 66.6 & 361.2 \end{bmatrix}$
Vanilla RNN	0.7103	0.7552	$\begin{bmatrix} 229.2 & 214.8 \\ 37.8 & 390.2 \end{bmatrix}$

Test Results

- **GRU:**
  - **Standard Deviation of Test Accuracy:** 0.0093
  - **Standard Deviation of F1 Score:** 0.0195
- **LSTM:**
  - **Standard Deviation of Test Accuracy:** 0.0157
  - **Standard Deviation of F1 Score:** 0.0078
- **Vanilla RNN:**
  - **Standard Deviation of Test Accuracy:** 0.0112
  - **Standard Deviation of F1 Score:** 0.0089

1.2.1. Task 4

Default Parameters: lr: 1×10^{-4}, batch size: 32, epochs: 5, clip: 0.25

Table 1.2: Comparative Analysis of Different RNN Cells and Hyperparameters

Model	Parameter	Value	Test Accuracy
Vanilla RNN	Hidden Size	40	0.5092
		150	**0.6984**
		400	0.6560
	Num Layers	2	0.6984
		4	0.6778
		7	**0.7099**
	Dropout	0.0	**0.6984**
		0.25	0.6571
		0.5	0.6732
	Bidirectional	True	**0.7523**
		False	0.6984
GRU	Hidden Size	40	0.5149
		150	**0.7466**
		400	0.7076
	Num Layers	2	0.7466
		4	**0.7592**
		7	0.7534
	Dropout	0.0	**0.7466**
		0.25	0.7294
		0.5	0.7202
	Bidirectional	True	**0.7695**
		False	0.7466
LSTM	Hidden Size	40	0.5115
		150	**0.7328**
		400	0.7213
	Num Layers	2	0.7328
		4	**0.7787**
		7	0.7683
	Dropout	0.0	0.7328
		0.25	**0.7534**
		0.5	0.7454
	Bidirectional	True	**0.7706**
		False	0.7328

I have opted for LSTM cell because it yields best results. I have combined the best parameters from each section.

LSTM Model Results With Best Hypeparameters

Parameters:

- Learning rate: 1×10^{-4}
- Batch size: 32
- Epochs: 5
- Clip: 0.25
- Hidden size: 150
- Number of layers: 4
- Dropout: 0.25
- Bidirectional: True

Test Results:

- **Average Test Accuracy:**

$$\frac{0.7833 + 0.7856 + 0.7924 + 0.7787 + 0.7764}{5} = 0.7833$$

- **Average F1 Score:**

$$\frac{0.8108 + 0.7929 + 0.8026 + 0.8031 + 0.7937}{5} = 0.8006$$

- **Standard Deviation of Test Accuracy:** 0.0056
- **Standard Deviation of F1 Score:** 0.0042
- **Average Confusion Matrix:**

$$\frac{1}{5}\left(\begin{bmatrix}303 & 141\\ 48 & 380\end{bmatrix} + \begin{bmatrix}327 & 117\\ 70 & 358\end{bmatrix} + \begin{bmatrix}323 & 121\\ 60 & 368\end{bmatrix} + \begin{bmatrix}309 & 135\\ 58 & 370\end{bmatrix} + \begin{bmatrix}302 & 142\\ 53 & 375\end{bmatrix}\right) =$$

$$\frac{1}{5}\begin{bmatrix}1564 & 656\\ 289 & 1851\end{bmatrix} = \begin{bmatrix}312.8 & 131.2\\ 57.8 & 370.2\end{bmatrix}$$

Comparative Analysis of Baseline and LSTM Models Without Embeddings

Parameters:

- Learning rate: 1×10^{-4}
- Batch size: 32
- Epochs: 5

Baseline Model Results

- suffers by 7%

Test Results:

- Average Test Accuracy: 0.67156
- Average F1 Score: 0.63388
- Average Confusion Matrix:

$$\begin{bmatrix} 337.4 & 106.6 \\ 179.8 & 248.2 \end{bmatrix}$$

LSTM Model Results

- suffers by 10%

Test Results:

- Average Test Accuracy: 0.6961
- Average F1 Score: 0.6925
- Average Confusion Matrix:

$$\begin{bmatrix} 307.2 & 136.8 \\ 128.4 & 299.6 \end{bmatrix}$$

1.3. Additional Search of hyperparameters

Parameters:

– Batch size: 32
– Epochs: 5
– Clip: 0.25
– Hidden size: 150
– Number of layers: 4
– Dropout: 0.25
– Bidirectional: True
– Min Frequency: 1
– Acitvation: ReLu
– Learning rate: 1×10^{-4}
– Optimizer: Adam
– Freeze: True

Table 1.3: Comparative Analysis of LSTM Model with Additional Parameters

Parameter	Value	Test Accuracy	F1 Score
Min Frequency	10	0.7397	0.7458
	100	0.6319	0.7058
	500	0.6112	0.5830
Activation	LeakyReLU	0.7041	0.7041
	Tanh	0.6984	0.7169
	ELU	0.7099	0.7452
Learning Rate	0.005	0.5092	0.0
	0.0002	0.7225	0.7056
	0.0005	0.7420	0.7239
Optimizer	SGD	0.5092	0.0
	RMSprop	0.6812	0.6160
	Adagrad	0.5092	0.0
Freeze	True	**0.7833**	**0.8006**
	False	0.7213	0.7356

Since noting made the improvement no hyperparameter optimizing has been done on baseline.

Attention

Test Results

- **Average Test Accuracy:**

$$\frac{0.7443 + 0.7328 + 0.7236 + 0.7397 + 0.7271}{5} = 0.7335$$

- **Average F1 Score:**

$$\frac{0.7392 + 0.7288 + 0.7343 + 0.7486 + 0.7430}{5} = 0.7388$$

- **Standard Deviation of Test Accuracy:** 0.0074
- **Standard Deviation of F1 Score:** 0.0072
- **Average Confusion Matrix:**

$$\frac{1}{5}\left(\begin{bmatrix}333 & 111\\ 112 & 316\end{bmatrix} + \begin{bmatrix}326 & 118\\ 115 & 313\end{bmatrix} + \begin{bmatrix}298 & 146\\ 95 & 333\end{bmatrix} + \begin{bmatrix}307 & 137\\ 90 & 338\end{bmatrix} + \begin{bmatrix}290 & 154\\ 84 & 344\end{bmatrix}\right)$$

$$= \frac{1}{5}\begin{bmatrix}1554 & 666\\ 496 & 1644\end{bmatrix} = \begin{bmatrix}310.8 & 133.2\\ 99.2 & 328.8\end{bmatrix}$$