université PARIS-SACLAY

Hands-on NLP Project
Emotion Recognition

Md. Naim Hassan Saykat

Aloïs Vincent

Marija Brkic

Introduction

- Predicting the emotions in tweets

- Comparison between traditional models and Deep Learning

- Use cases :
  - Collecting feedback
  - Adjusting marketing strategies

Introduction : Dataset

- Texts labeled using hashtags

- We used the 6 emotions of the updated dataset :
  - Sadness
  - Joy
  - Love
  - Anger
  - Fear
  - Surprise

Emotions	Amount	Hashtags
sadness	214,454	#depressed, #grief
joy	167,027	#fun, #joy
fear	102,460	#fear, #worried
anger	102,289	#mad, #pissed
surprise	46,101	#strange, #surprise
trust	19,222	#hope, #secure
disgust	8,934	#awful, #eww
anticipation	3,975	#pumped, #ready

Benchmark models

Models	Features	anger	anticipation	disgust	fear	joy	sadness	surprise	trust	F1 Avg.
BoW	**word frequency**	0.53	0.08	0.17	0.53	0.71	0.60	0.36	0.33	0.57
BoW$_{\text{TF-IDF}}$	**TF-IDF**	0.55	0.09	0.18	0.57	0.73	0.62	0.39	0.35	0.60
n-gram	**word frequency**	0.56	0.09	0.17	0.57	0.73	0.64	0.42	0.39	0.61
n-gram$_{\text{TF-IDF}}$	**TF-IDF**	0.58	0.12	0.17	0.60	**0.75**	0.67	0.47	0.45	0.63
char_ngram	**character frequency**	0.49	0.06	0.12	0.46	0.67	0.55	0.30	0.28	0.52
char_ngram$_{\text{TF-IDF}}$	**TF-IDF**	0.53	0.07	0.15	0.53	0.71	0.59	0.35	0.31	0.57
LIWC	**affective words**	0.35	0.03	0.11	0.30	0.49	0.35	0.18	0.19	0.35
CNN$_{\text{w2v}}$	**word embeddings**	0.57	0.10	0.15	0.63	**0.75**	0.64	0.61	**0.70**	0.65
EmoNet	**word embeddings**	0.36	0.00	0.00	0.46	0.69	0.61	0.13	0.25	0.52
DeepMoji	**word embeddings**	0.60	0.00	0.03	0.49	**0.75**	0.67	0.20	0.27	0.59
CNN$_{\text{BASIC}}$	**basic patterns**	0.65	0.10	0.22	0.64	0.73	0.56	0.15	0.08	0.52
CARER$_{\beta}$	**enriched patterns**‡	0.61	0.31	0.34	0.67	**0.75**	0.68	0.60	0.55	**0.67**
CARER	**enriched patterns**	**0.74**	**0.41**	**0.43**	**0.79**	**0.83**	**0.82**	**0.76**	**0.75**	**0.79**

Benchmark models : Deep Learning models

Model	Input	Epochs	Accuracy
RNN_{w2v}	word2vec (Mikolov et al., 2013)	24	0.53
CNN_{char}	character embeddings (end-to-end)	50	0.63
CNN_{w2v}	word vectors (Deriu et al., 2017)	33	**0.69**
EmoNet	word embeddings (end-to-end)	23	0.58
DeepMoji	word embeddings (end-to-end)	100	0.63
BiGRNN	our enriched patterns[‡]	12	**0.68**
CARER_{β}	our enriched patterns[‡]	12	**0.72**
CARER_{EK}	our enriched patterns	12	**0.81**

Benchmark models : Transformers

Table 1: Benchmark models on CARER dataset

Paper	Model	Biggest Accuracy
Saravia et al.[2]	Semi-supervised, graph-based algorithm	0.81
Wang et al.[3]	Fine-tuned BERT(transformer based)	0.93
Fengkai[4]	Fine-tuned distilBERT-base-uncased(transformer based)	0.94

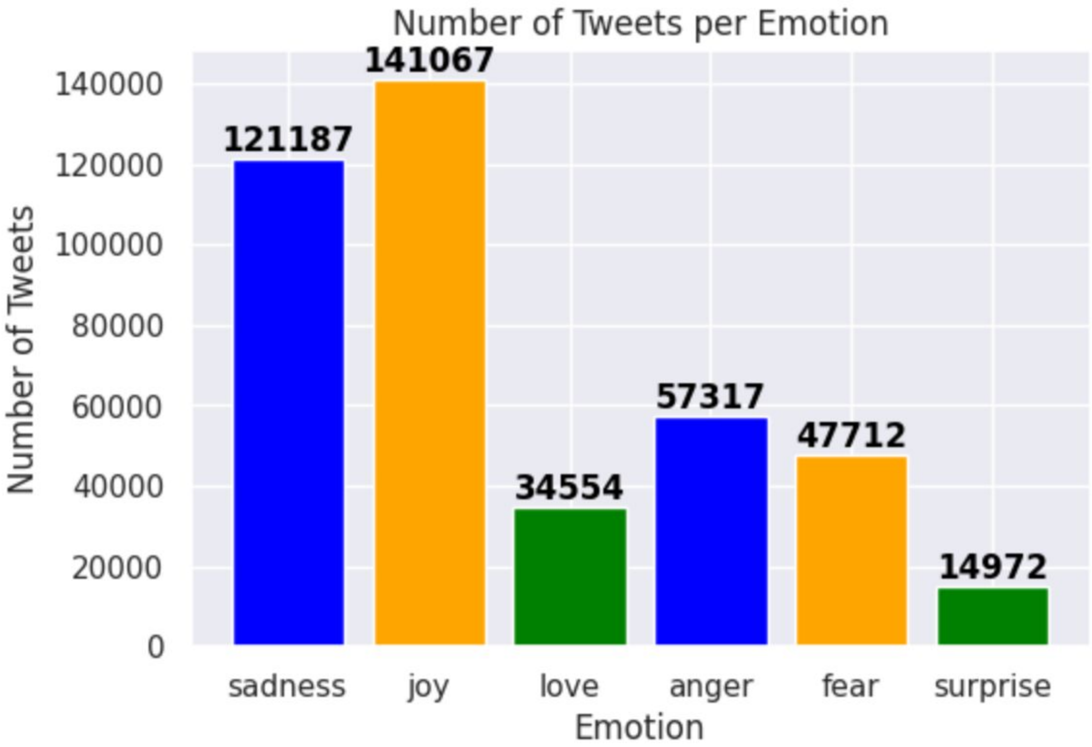
Evaluation Metrics

- Accuracy: *Measures the overall correctness of predictions.*
- Precision: *How many selected items are relevant?*
- Recall: *How many relevant items are selected?*
- **F1-Score**: *Harmonic mean of precision and recall.*

Data Analysis

- Very imbalanced dataset

- Contains 416,000+ tweets



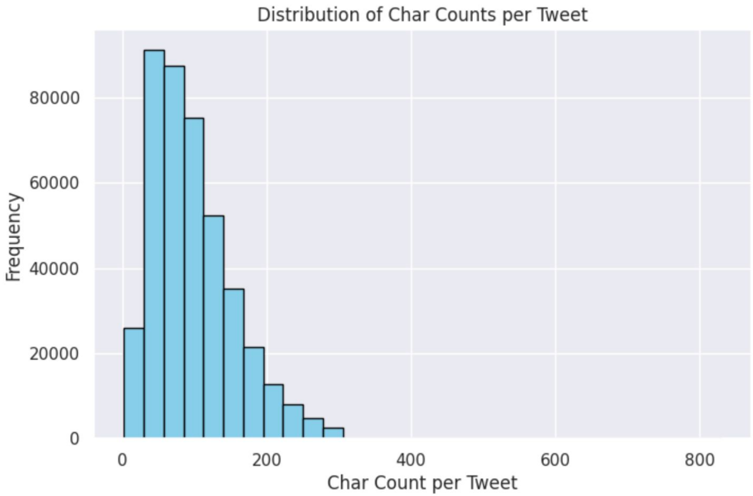
universite PARIS-SACLAY

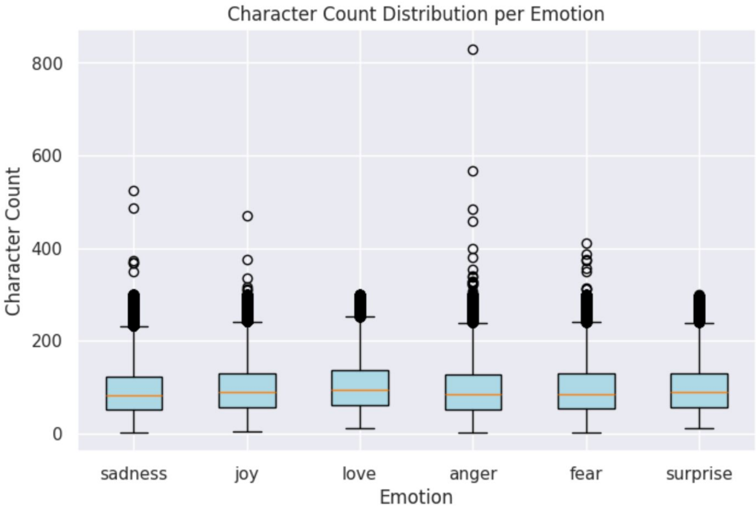
Data Analysis

- Same text format
- Familiar language
- Introspective
  - Starting with “I …”

	text	label
0	i feel awful about it too because it s my job ...	0
1	im alone i feel awful	0
2	ive probably mentioned this before but i reall...	1
3	i was feeling a little low few days back	0
4	i beleive that i am much more sensitive to oth...	2

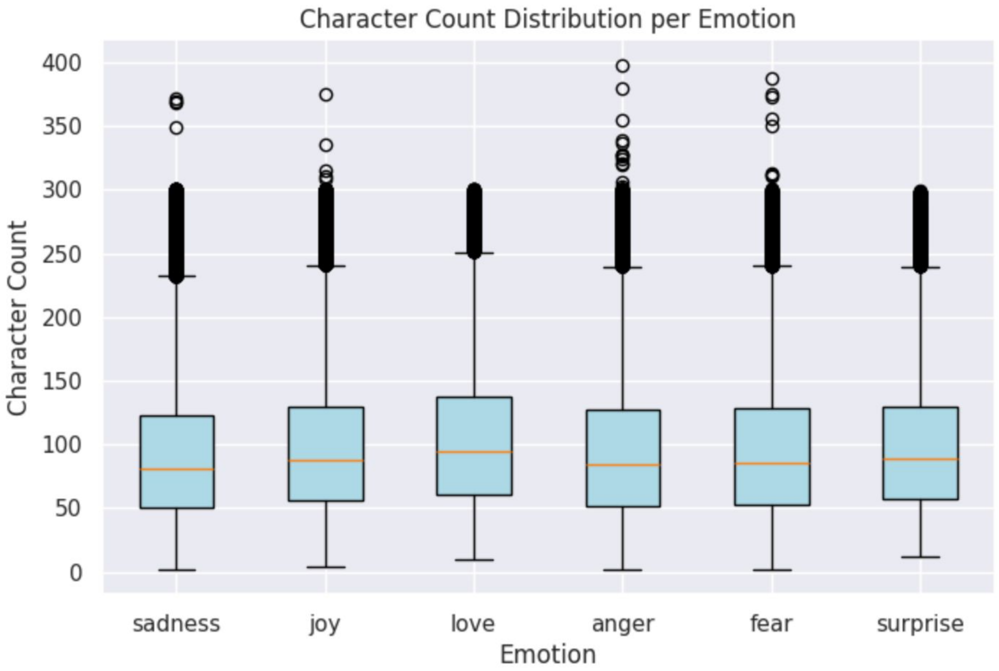
Data Analysis





Preprocessing

- Removed outlier

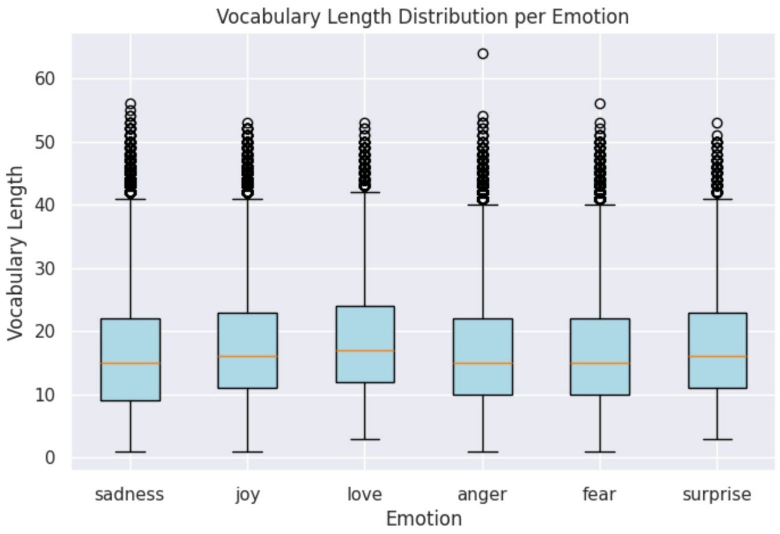


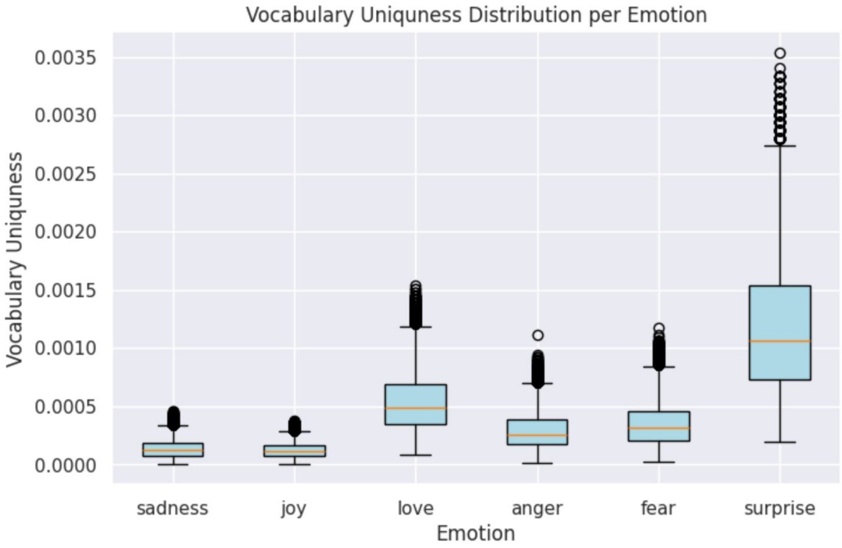
Preprocessing

- There was no punctuation in the original dataset
- As expected : I, feel, etc…

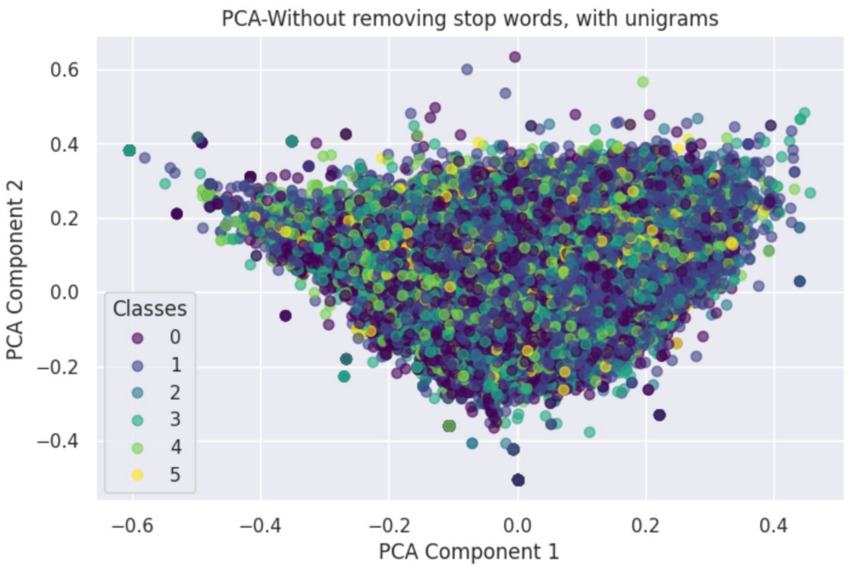
	token	frequency
0	i	676149
1	feel	289936
2	and	250251
3	to	233087
4	the	216591
...	...	...
75289	galleryimageborder	1
75290	danbo	1
75291	truc	1
75292	entrails	1
75293	usaully	1

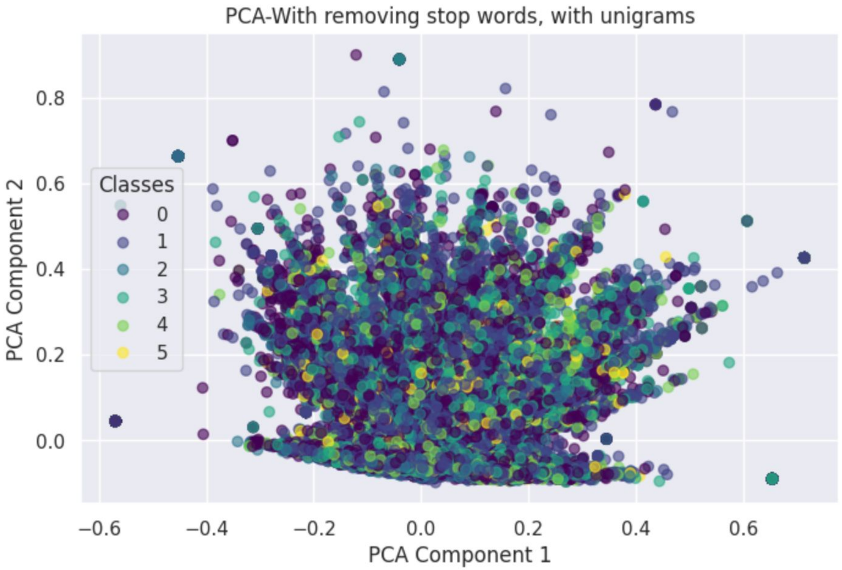
Preprocessing





Vectorizing and Visualization





Traditional models : accuracies

- SVM: 88.60%

- Random Forest: 87.00%

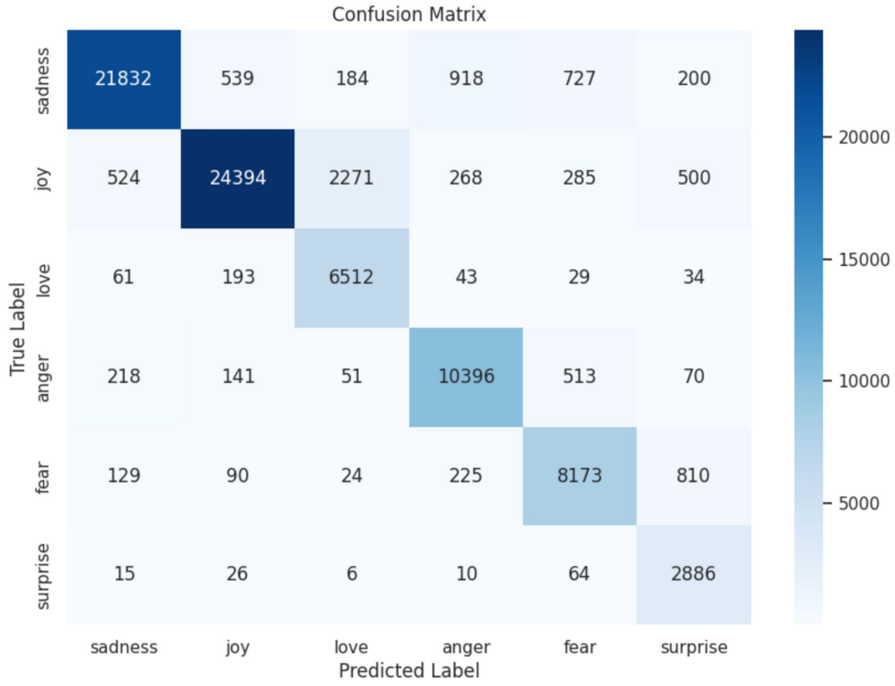
- Logistic Regression: 86.00%

Convolutional Neural Network

- 4 convolutional layers, 3 max-pool layers, fully-connected layer
- ReLU, Softmax
- Dropout, EarlyStopping
- Sparse Categorical Crossentropy, Adam optimizer
- 50 epochs, 32 batch-size

CNN : Results

	precision	recall	f1-score	support
0	0.96	0.89	0.93	24400
1	0.96	0.86	0.91	28242
2	0.72	0.95	0.82	6872
3	0.88	0.91	0.89	11389
4	0.83	0.86	0.85	9451
5	0.64	0.96	0.77	3007
accuracy			0.89	83361
macro avg	0.83	0.91	0.86	83361
weighted avg	0.90	0.89	0.89	83361



CNN : Error Analysis

Text: i was feeling a little low few days back
True label: fear
Predicted label: surprise

Text: i don t feel comfortable around you
True label: fear
Predicted label: surprise

Text: i am only having day a week where i am feeling depressed or seriously anxious
True label: sadness
Predicted label: fear

BERT and ensemble method : Results

Ensemble Model Classification Report:

	precision	recall	f1-score	support
sadness	0.93	0.96	0.94	581
joy	0.88	0.97	0.92	695
love	0.84	0.64	0.73	159
anger	0.93	0.87	0.90	275
fear	0.87	0.86	0.87	224
surprise	0.87	0.50	0.63	66
accuracy			0.90	2000
macro avg	0.89	0.80	0.83	2000
weighted avg	0.90	0.90	0.89	2000

BERT and ensemble method : Misclassifications

```
Misclassified Examples (Ensemble Model - BERT + SVM + RF + Logistic Regression):
                                                     text  label  \
693   i can say is that as long as you enjoy the sto...      0
1533  i actually was in a meeting last week where so...      3
1714           i also do feel passionate about teaching      2
206   i wish to know whether i should feel sympathet...      2
476   i feel quite helpless in all of this so prayer...      0
254   i feel blessed beyond blessed to share my life...      2
1467  i seek out pain to feel tortured just to feel ...      4
433   i know that i have it nowhere near as worse as...      4
828      i feel unprotected even while travelling alone      4
861   i feel assaulted by this shit storm of confusi...      4

      ensemble_predictions
693                      3
1533                     0
1714                     1
206                      1
476                      4
254                      1
1467                     3
433                      0
828                      0
861                      0
```

université PARIS-SACLAY

BERT and ensemble methods : own sentences

```
Custom Sentence Predictions (Ensemble Model - BERT + SVM + RF + Logistic Regression):

Sentence: I am feeling very happy today!
Ensemble Model Prediction: joy

Sentence: This is the worst day of my life.
Ensemble Model Prediction: joy

Sentence: I can't stop smiling, this is the best surprise ever!
Ensemble Model Prediction: joy

Sentence: I am so scared to go outside alone.
Ensemble Model Prediction: fear

Sentence: I feel so loved and appreciated today.
Ensemble Model Prediction: love

Sentence: Why do you always make me so angry?
Ensemble Model Prediction: anger

Sentence: I feel like crying all day long.
Ensemble Model Prediction: sadness
```

universite PARIS-SACLAY

Results

Table 2: Model accuracies

Model	Accuracy
Logistic Regression	0.86
Random Forest	0.87
Support Vector Machine	0.89
Convolutional Neural Network	0.89
BERT	0.92
Ensemble (BERT + SVM + RF + LR)	0.90
Fine-tuned distilBERT-base-uncased(transformer based)[4]	0.94

Conclusion

- We achieved very high accuracies

- We managed to get pretty close to state of the art

universite PARIS-SACLAY

References

[1] **dair-ai/emotion dataset**. Available at: https://huggingface.co/datasets/dair-ai/emotion

[2] Elvis Saravia, Hsien-Chi Toby Liu, Yen-Hao Huang, Junlin Wu, and Yi-Shin Chen. **CARER: Contextualized Affect Representations for Emotion Recognition**. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), 3687–3697, 2018. Available at: https://aclanthology.org/D18-1404.pdf.

[3] Wang, Y., and others. **Large Language Models on Fine-grained Emotion Detection Dataset with Data Augmentation and Transfer Learning**. arXiv preprint arXiv:2403.06108v1, 2024. Available at: https://arxiv.org/html/2403.06108v1.

[4] Fengkai Yu Hugging Face Sollution Available at: https://huggingface.co/Fengkai/distilbert-base-uncased-Finetuned-emotion.

[5] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. BERT**: Pre-training of Deep Bidirectional Transformers for Language Understanding**. Proceedings of NAACL, 2019. Available at: https://arxiv.org/abs/1810.04805.

[6] Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. **Distributed Representations of Words and Phrases and Their Compositionality. Advances in Neural Information Processing Systems** (NeurIPS), 2013. Available at: https://arxiv.org/abs/1310.4546.

[7] Hochreiter, S., & Schmidhuber, J. **Long short-term memory**. Neural Computation, 1997. Available at: https://doi.org/10.1162/neco.1997.9.8.1735.