# RESEARCH STAY WEEK 12, Deep learning in sentiment analysis

Daniel Cajas A01708637

# CONTEXT

Before deep learning techniques were implemented for sentiment analysis state of the art models used manual feature selection and user more traditional classifiers, such as SVMs, for the task. This placed a ceiling in the accuracy of these models as the features were not able to fully represent the sentiment of the corpus. With deep learning feature selection was automated which gave a boost to sentiment analysis model performance. As new and better deep learning architectures have been developed the mopeds have gotten better and better.

# SEARCH METHODOLOGY

[1] Does a comparison between different DNN arquitectures. Plain DNN, CNNs and RNNs. It concluded that RNNs using word embedding (word2vec) yielded the best results but where to costly. However CNNs had a similar score but were significantly les compute demanding. I wanted to see how this paradigm changed with the inclusion of the transformer and attention mechanisms developed the past few years.

# Comparison

| Title | Objective | Year | Datasets | Arquitectures compared | Results |
|---|---|---|---|---|---|
| Sentiment Analysis Based on Deep Learning: A Comparative Study | Compare different performance metrics and compute cost of different DNN arquitectures for sentiment analysis | 2020 | Sentiment140<br>Tweets Airline<br>Tweets SemEval<br>IMDB Movie Reviews (1)<br>IMDB Movie Reviews (2)<br>Cornell Movie Reviews<br>Book Reviews<br>Music Reviews | DNN, CNN, RNN | RNNs using word embedding (word2vec) yielded the best results but where to costly. However CNNs had a similar score but were significantly les compute demanding. |
| RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network | Propose a new sentiment analysis model that leverages BER as a word embedding fiune tuned for sentiment analysis and aids it with an LSTM layer for the sequence processing. | 2022 | IMDb<br>Twitter<br>US Airline<br>Sentiment140 | Naive Bayes<br>Logistic Regression<br>Decision Tree<br>KNN<br>AdaBoost<br>GRU<br>LSTM<br>BiLSTM<br>CNN-LSTM<br>CNN-BiLSTM<br>RoBERTa-LSTM | RoBERTa-LSTM with data augmentation gives the best acurracy, precision, recall and F!-score out of all the architectures in all the datasets tested |

# BIBLIOGRAFÍA

[1] Sahoo, C., Wankhade, M. & Singh, B.K. Sentiment analysis using deep learning techniques: a comprehensive review. Int J Multimed Info Retr 12, 41 (2023). https://doi.org/10.1007/s13735-023-00308-2

[2] Tan, K. L., Lee, C. P., Anbananthen, K. S. M., & Lim, K. M. (2022). RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network. IEEE Access, 10, 21517-21525. https://doi.org/10.1109/ACCESS.2022.3152828