Evidence Detection

An Overview of Two Approaches: LR + TF-IDF & Bi-LSTM

By Bilal Yaqub and Abdelrahman Abdeldaim

The University of Manchester

Introduction

Evidence detection (ED) is a vital NLU task, involving analyzing text to determine if provided evidence supports or contradicts a claim. It's crucial in domains like law, media and academia where information accuracy is key. Manual assessment of evidence relevance is labor-intensive and prone to subjective interpretations.

Objective

Develop two distinct computational methods for Evidence Detection (ED). The first method employs TF-idf vectorization, a traditional machine learning technique, while the second utilizes a sophisticated deep learning model, Bi-directional LSTM.

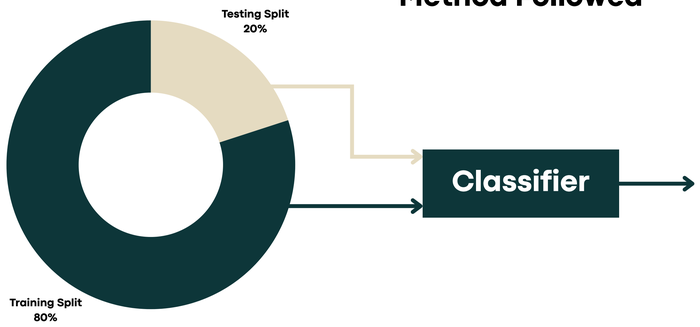
Methodology

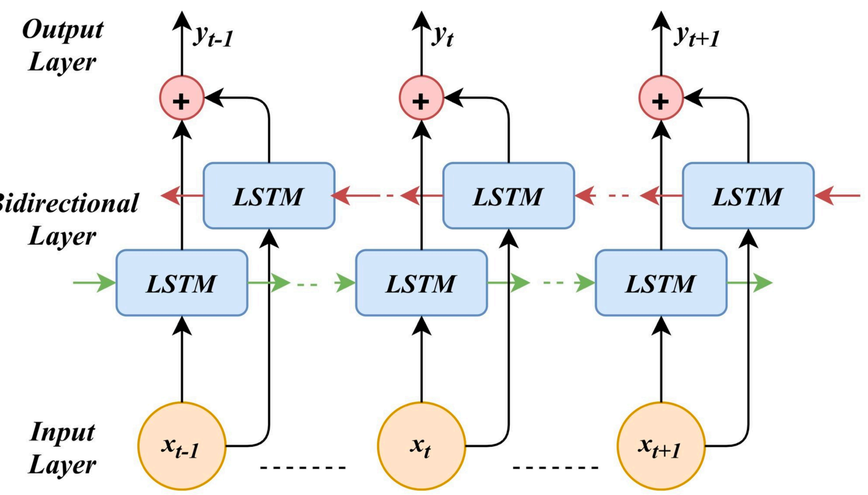
The project employs a dataset with over 23,000 training pairs and nearly 6,000 validation pairs, with rigorous preprocessing to improve data quality for modeling.

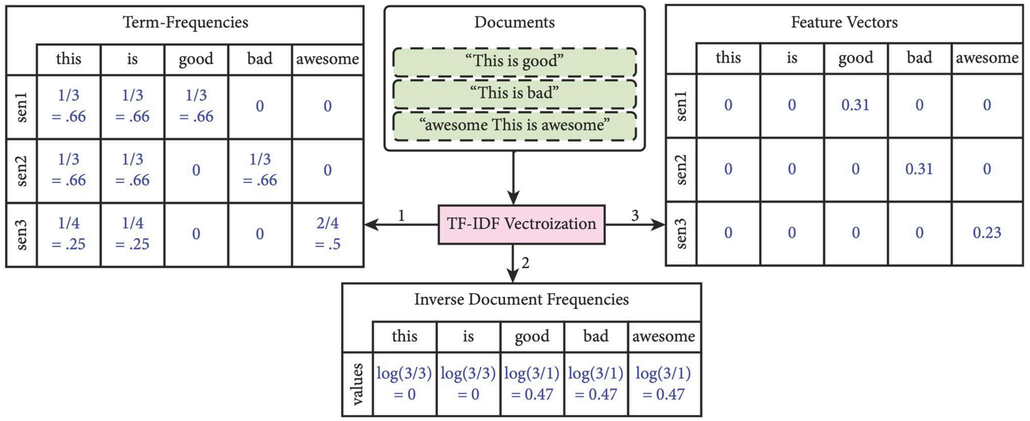
Model Development

- TF-idf Vectorization Model: Utilizes statistical analysis to assess word relevance, establishing a baseline by highlighting term frequency and significance.
- Bi-directional LSTM Model: Employs a neural network that processes text bidirectionally, capturing complex contextual relationships more effectively than traditional models.

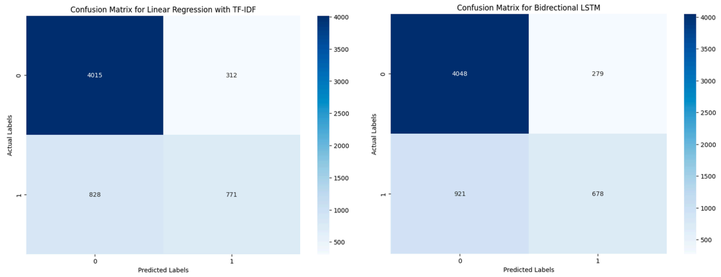
Method Followed







Results



	LINEAR REGRESSION + TFIDF	BI-LSTM
PRECISION	0.80	0.79
RECALL	0.81	0.80
F1-SCORE	0.79	0.78
ACCURACY	0.808	0.798

Conclusion

The model utilising Logistic Regression with a TF-IDF vectorizer achieved an accuracy of >80% with an F1 score of 0.79, recall score of 0.81 and precision score of 0.80. Showing that even a simple linear relation was able to achieve promising results in this experiment. On the other hand, using a deep learning technique like a Bidirectional-LSTM achieved very similar results with an accuracy of nearly 80%, F1 score of 0.78, recall of 0.80 and precision of 0.79, however, training it took more than 2 hours of run time.

References

- Khan, M. Y., Qayoom, A., Nizami, M., Siddiqui, M. S., Wasi, S., & Syed, K.-U.-R. R., 2021. Automated Prediction of Good Dictionary Examples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning, and Word Embedding-Based Deep Learning Techniques. Complexity, DOI: 10.1155/2021/2553199.