

EAI 6010:

APPLICATIONS OF ARTIFICIAL INTELLIGENCE

Module 3: Text Classification with Transfer Learning

Submitted To:

Prof RJ Munthali, Faculty Lecturer

Submitted By: Abhilash Dikshit

Academic Term: Fall 2023

Northeastern University, Vancouver, BC, Canada

Master of Professional Studies in Analytics

November 29, 2023

**Title: A Comparative Analysis of Deep Learning and Traditional NLP Models for Sarcasm Detection in News Headlines**

1. **Introduction**

The goal of this study is to explore the effectiveness of deep learning and traditional NLP models for sarcasm detection in news headlines. The selected dataset, the "News Headlines Dataset for Sarcasm Detection," is sourced from [Kaggle](https://www.kaggle.com/datasets/rmisra/news-headlines-dataset-for-sarcasm-detection/data) which was taken from [Wikipedia](https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research#News_articles) and offers a diverse collection of sarcastic and non-sarcastic headlines. The motivation behind this choice is to address the practical business problem of automating the identification of sarcastic content in news, aiding in content curation and sentiment analysis.

1. **Dataset and Parameter Selection:**

The dataset was chosen for its relevance to the business problem and its availability for public use. The AWD\_LSTM pre-trained language model, implemented using the **fastai** library, was selected due to its strong performance in natural language processing tasks and the ULMFiT method, which allows for effective fine-tuning on domain-specific datasets. Parameters such as the learning rate and dropout multiplier were tuned to strike a balance between model expressiveness and generalization to the specific task.

A close-up of a white background

Description automatically generatedA screen shot of a computer code

Description automatically generated

1. **Constraints:**

To facilitate completion within the assignment's scope, constraints were added to limit the dataset size and training time. This decision was made to manage the computational resources required for fine-tuning the deep learning model and training the traditional NLP model.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. **Model Development:**

The AWD\_LSTM model was fine-tuned using the fastai library, and a traditional NLP model was implemented using TF-IDF vectorization and a Support Vector Machine (SVM). The accuracy of the deep learning model, as observed on the validation set, was compared to the traditional NLP model.

A screenshot of a computer

Description automatically generated

**Visualizations:**

Confusion Matrix - AWD\_LSTM:

A screenshot of a computer

Description automatically generated

A graph of a number of features

Description automatically generated

1. **Model Comparison:**

The AWD\_LSTM model achieved an accuracy of 90%, outperforming the traditional NLP model with an accuracy of 84%. While the deep learning model exhibited superior performance, its development effort was notably higher, requiring more computational resources and time for fine-tuning. The traditional NLP model, though less accurate, was computationally less intensive, suggesting a trade-off between accuracy and resource efficiency.

A screenshot of a computer program

Description automatically generated

**Additional Findings:**

**AWD\_LSTM:**

* The model excelled in capturing nuanced patterns and context in sarcastic language.
* Fine-tuning required careful parameter adjustments, with a notable impact on model performance.

**Traditional NLP Model:**

* Simplicity allowed for quick implementation and lower computational requirements.
* Limitations observed in handling subtle semantic nuances present in sarcastic headlines.

1. **Lessons Learned:**

During the experimentation, it became evident that the AWD\_LSTM model's architecture excelled in capturing nuanced patterns in language but required careful hyperparameter tuning. On the other hand, the traditional NLP model, being simpler, was more straightforward to implement but lacked the capacity to capture intricate semantic relationships.

1. **Recommendations:**

For productization, considering the trade-off between accuracy and development effort, the choice depends on the specific application requirements. If computational resources are limited and a balance between accuracy and efficiency is desired, the traditional NLP model may be a pragmatic choice. However, if higher accuracy is paramount, and computational resources are not a significant constraint, the AWD\_LSTM model proves to be a powerful solution.

1. **Conclusion:**

The study provides insights into the strengths and limitations of deep learning and traditional NLP models for sarcasm detection in news headlines. The selection between models should align with the specific requirements and constraints of the application.

1. **References:**
2. Misra, Rishabh and Prahal Arora. "Sarcasm Detection using News Headlines Dataset." AI Open (2023).
3. Misra, Rishabh and Jigyasa Grover. "Sculpting Data for ML: The first act of Machine Learning." ISBN 9798585463570 (2021).