MANAGEMENT SUMMARY

Course: Natural Language Processing and Information Extraction
Topic: Extraction of Narratives from Online News ('propaganda detection')
Subtask 2: Narrative Classification
Instructor: Varvara Arzt
Group: 8, "Word Wizards"
Members: Denis Bugaenco, Iulia Mihaela Enache, Raquel Panadero Palenzuela, Tatiana Beliaeva

Overview of the task

The goal of the project is to address a major problem found in the current digital media environment: the widespread diffusion of manipulative content and propaganda. The Internet has enabled direct interactions between information producers and consumers, making audiences increasingly vulnerable to false narratives, particularly in times of crucial events such as the Ukraine-Russia conflict and discussions related to climate change. Identification and understanding of such narratives become essential to address the spread of disinformation.

The initiative is part of the SemEval 2025 campaign and is oriented toward the Multilingual Characterization and Extraction of Narratives from Online News. More precisely, it addresses the Narrative Classification task, which consists of automatically identifying and classifying narratives and sub-narratives in multilingual news articles. Such narratives are related to two relevant topics, the Ukraine-Russia War and Climate Change, and are considered in five languages: Bulgarian, English, Hindi, Portuguese, and Russian. The dataset is obtained from news sources and articles flagged by fact-checkers as potentially carrying misinformation.

The task is a multi-label, multi-class classification problem, where a system is required to assign one or more applicable narrative and subnarrative labels to a news article. For instance, an article that questions the validity of temperature measurements in climate change data might be assigned the narrative "Questioning the measurements and science" and the subnarrative "Methodologies/metrics used are unreliable/faulty."

External resources used

The project involved several key resources to achieve its objectives. Translation of non-English articles was performed using DeepL API (for Bulgarian and Portuguese) and Google Cloud Translation API (for Hindi), enabling the integration of multilingual data.

Libraries and frameworks included Stanza for tokenization and boilerplate removal, TfidfVectorizer from Scikit-learn for SVM feature extraction, and PyTorch for implementing LSTM and Transformer-based models. The Hugging Face Transformers library was used for fine-tuning the pre-trained BERT model. These tools, used with the annotated datasets provided and preprocessing techniques, were a requirement in developing the narrative classification systems.

Solution implemented

To classify narratives from multilingual news articles, the project used a combination of traditional and deep learning models. The raw text data was cleaned to remove irrelevant content such as advertisements and links in order to ease processing in the traditional model used. Deep learning models, on the other hand retain the original text since these types of models work better witch richer and unprocessed data.

We implemented four models in total:

- SVM, the traditional approach.

- LSTM and Transformers, advanced models analyzing text sequentially to capture its structure and context.
- BERT, a state-of-art model pre-trained on a large-scale text that needed very little additional processing and showed good performance.

The models were trained and tested on separate sets of articles relating to the Ukraine War and Climate Change, thus ensuring that each topic was examined in isolation.
Performance evaluation was done based on precision, recall, and F-1 score to evaluate how good models are in detecting narratives. SVM showed more simplicity and speed, but BERT was most reliable with the best balance between accuracy and speed.

Challenges & Limitations

The project faced several challenges in achieving its goals. A key issue was data imbalance, with significant differences in dataset size and quality across languages. Hindi, for instance, had far fewer articles than Bulgarian or Portuguese, complicating model generalization, particularly when using translated datasets.

The complexity of the multi-label, multi-class classification task was added by the fact that each article could belong to several overlapping narratives and sub-narratives. Capturing these relationships required more advanced models and precise tuning.

Another concern was over- and under-prediction: models, especially deep learning models, often over-predicted irrelevant labels (e.g., "renewable energy") and missed critical sub-narratives.

On top of that, language translation added another layer of complexity: while articles in Bulgarian, Portuguese, and Hindi were translated to English, it impacted model performance across languages.

Lastly, each of these types of models presented their challenges: the SVM traditional model, did not scale well and relied on manual feature engineering, while the deep learning models required hight computational resources. Pre-trained model (BERT) while effective, sometimes failed to fine-tune well for this particular task.

Possible future implementations

To overcome limitations and further improve the effectiveness of models, possible new strategies could be implemented.

One of the main priorities is data imbalance correction, where an effective approach will be data augmentation; in other words, generating examples for under-represented languages like Hindi. The technique, which include back-translation, consists in translating a text into another language and back into the original, creating variations of the existing articles and enriching the dataset to balance the languages. That way, the models will generalize better and their performance on low-resource languages will improve.

Another interesting method is synthetic data generation, where large language models could be used to create artificial articles that represent rare narratives or subnarratives. For exemple, generating examples for underrepresented labels such as “Criticism of institutions” with specific subnarratives would provide the model with additional training data, increasing its ability to handle nuanced cases.

On the modeling side, experimentation with ensemble methods that can combine the strengths of traditional and deep learning models could improve classification accuracy. Additionally, instead of relying on translations, am alternative approach for multilingual data could involve handling all languages together using pre-trained multilingual models such as mBERT or XLM-RoBERTa.