Multilingual Characterization and Extraction of Narratives from Online News
Subtask 2: Narrative Classification

Team 16 Apfelkuchen

Aleksandr Pavlenko, Vadim Chaikin, Nikita Morev, Fedor Gerasimov, Farkhat Almukhamedov

Problem Formulation:

Given a news article and a two-level taxonomy of narrative labels (where each narrative is subdivided into subnarratives) from a particular domain, assign to the article all the appropriate subnarrative labels. This is a multi-label multi-class document classification task.

Word Cloud of Training Data (excluding punctuation, articles, conjunctions)

Data Overview:

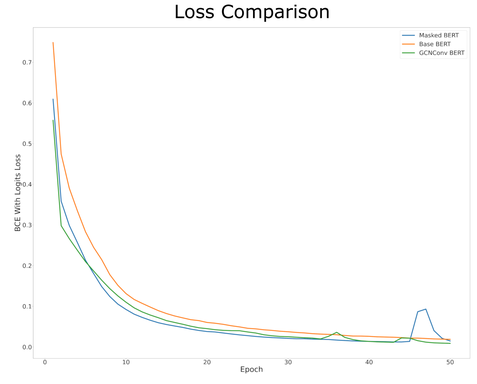
The dataset consists of text on different topics including Climate Change and Russian-Ukraine War, written in 5 languages. The preprocessing includes removal of punctuation, numbers and stopwords for the chosen languages, as well as lemmatizing and stemming.

The following specificities were acknowledged:

1) Duplication. Different IDs but same text.

2) Ambiguity. Narratives and subnarratives with labeled as "Other" contain varying content.

Approach	Model	Loss	Benefits
Base	bert-base-multilingual-uncased	Sum of narrative and subnarrative loss	Simple implementation; surprisingly good results
Masked	bert-base-multilingual-uncased	Masked loss for narrative and subnarrative, applying constraints based on allowed labels	Ensures predictions respect hierarchical label dependencies (e.g., only valid subnarratives for a given narrative are considered)
Tree	bert-base-multilingual-uncased with GCNConv layers	Sum of narrative and subnarrative loss	Takes into account the relation between different levels of labels



Quantitative Results

Language	F1 macro coarse	F1 st. dev. coarse	F1 samples	F1 st. dev. samples
English	0.392	0.384	**0.336**	0.420
Portuguese	0.670	0.295	**0.434**	0.380
Hindi	0.345	0.400	**0.143**	0.322
Bulgarian	0.526	0.429	**0.300**	0.412
Russian	0.516	0.319	**0.248**	0.370

Conclusion:

1. Taking into account and incorporating hierarchy of labels into model significantly increase performance.
2. Properly preprocessed data slightly increase performance
3. Graph convolutions could be outlined for a future research as labels in this task can be represented as a graph

Reference:

1. Devlin J. Bert: Pre-training of deep bidirectional transformers for language understanding //arXiv preprint arXiv:1810.04805. – 2018.
2. Wang Z. et al. Incorporating hierarchy into text encoder: a contrastive learning approach for hierarchical text classification // arXiv preprint arXiv:2203.03825. – 2022.
3. Kipf T. N., Welling M. Semi-supervised classification with graph convolutional networks //arXiv preprint arXiv:1609.02907. – 2016.