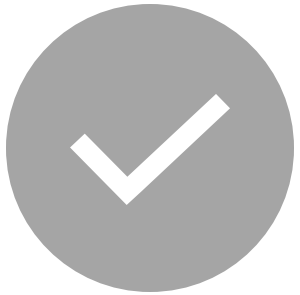
News Sentiment Analysis

Jakub Kozieł, Jakub Lis, Bartosz Sawicki

News sentiment analysis

Assess sentiment analysis of news articles - their parts and the whole text, as well as aspects of sentiment towards different mentioned issues



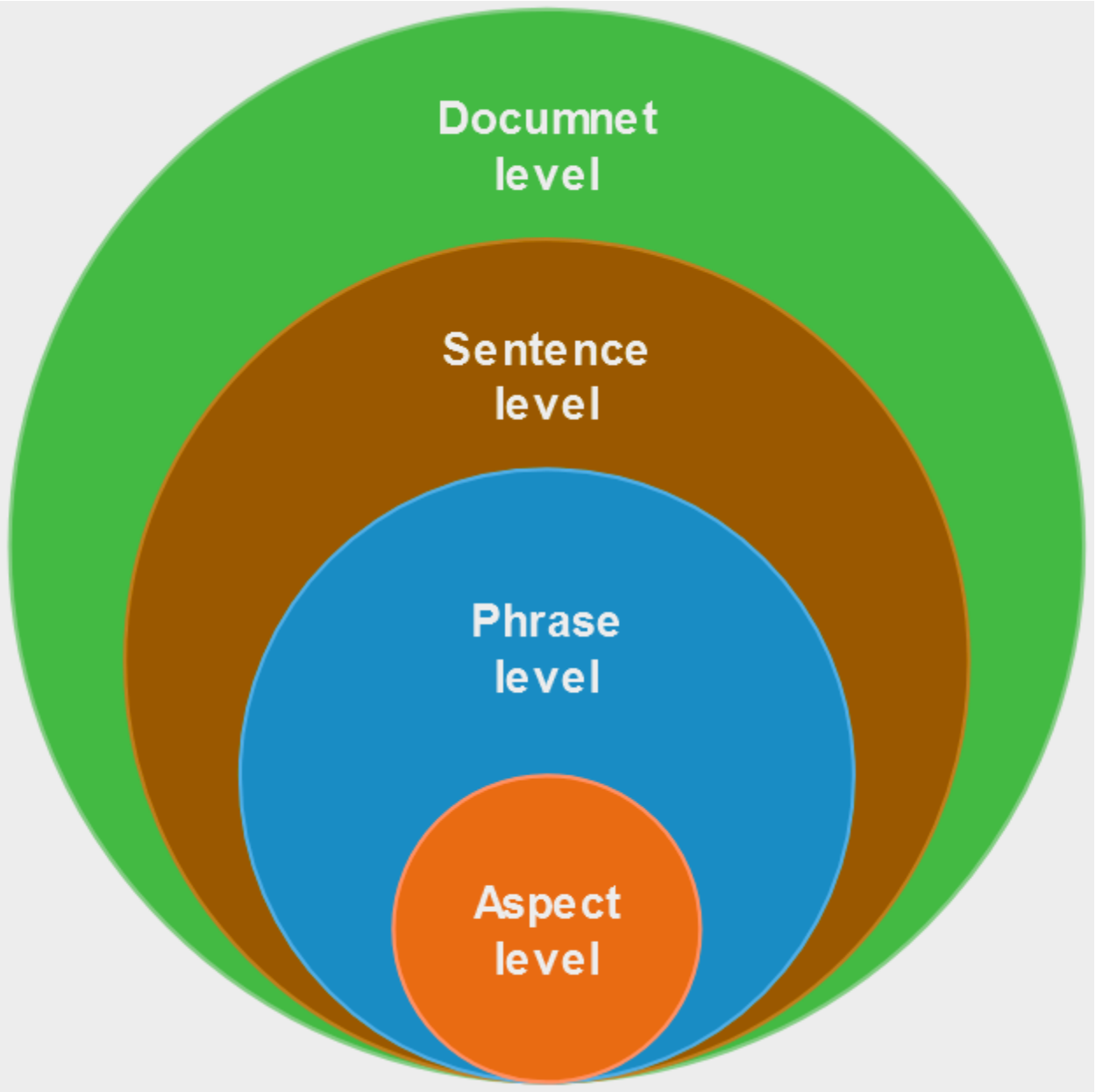
Exploration of current SOTA approaches

Make vizualizations using predicted sentiment

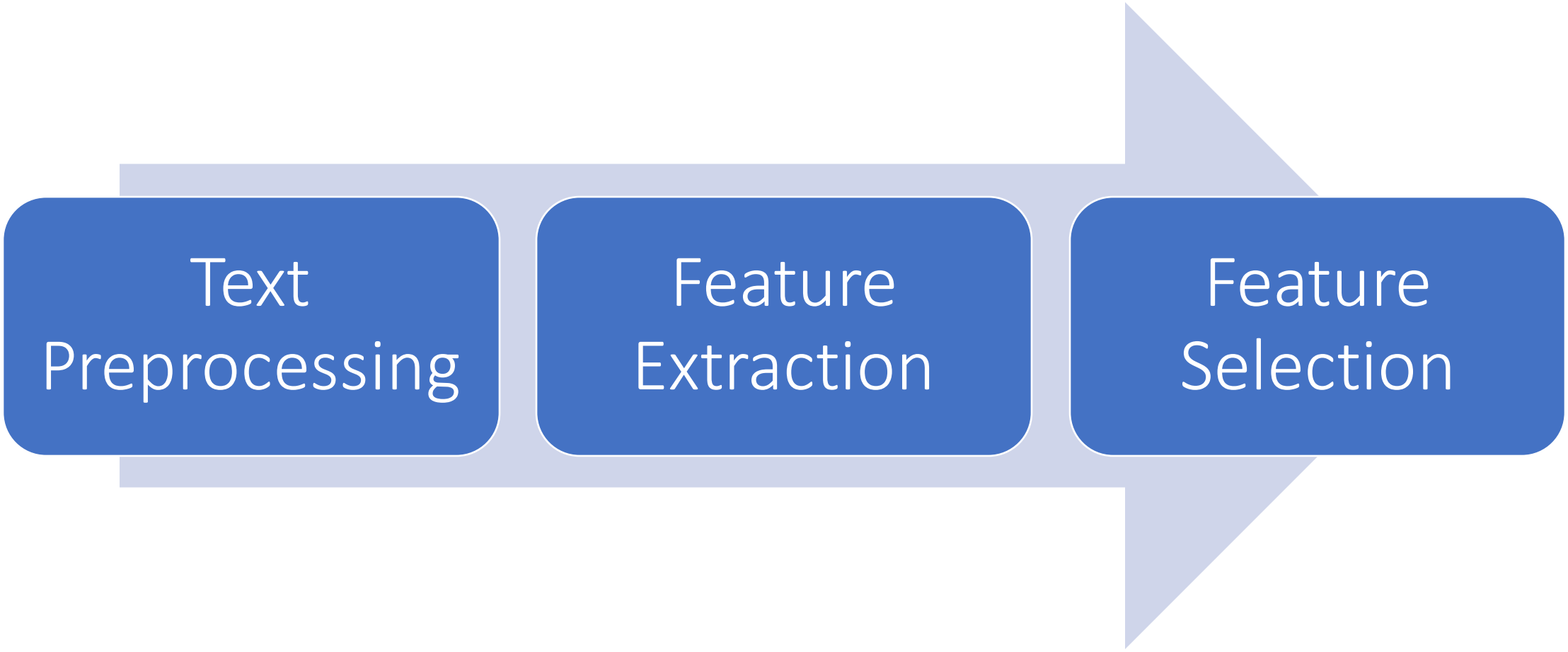
Perform XAI on top of our predictions

Sentiment analysis

Level of sentiment analysis



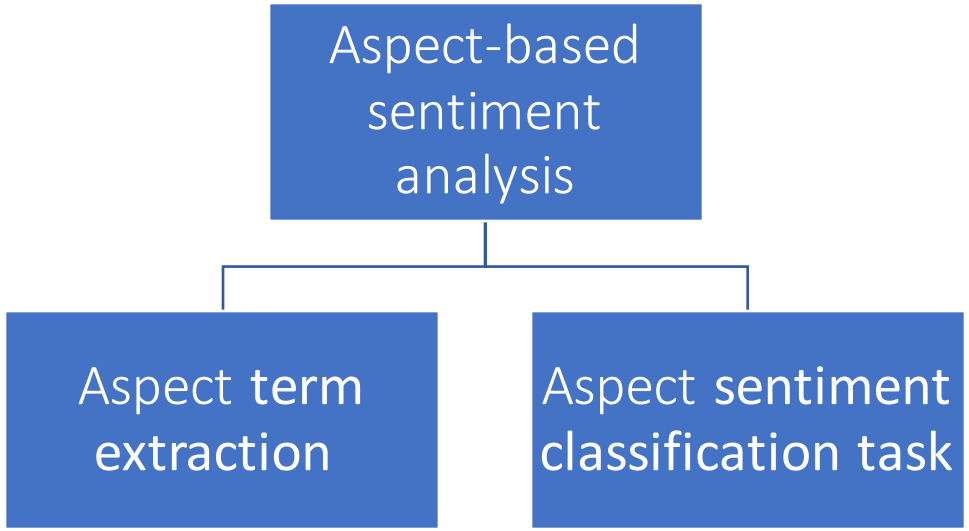
Data preprocessing

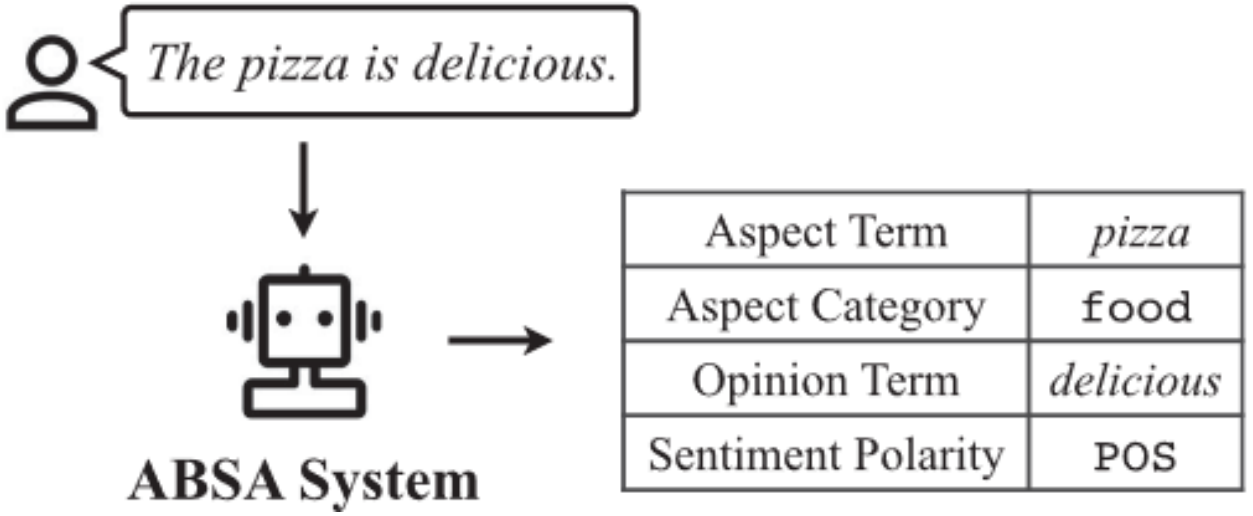


Challenges in sentiment analysis

- Sarcasm detection
- Negation handling
- Spam detection
- Anaphora and coreference resolution
- Word sense disambiguation
- Low-resource languages
- Sentiment analysis of code-mixed data

Aspect-based sentiment analysis

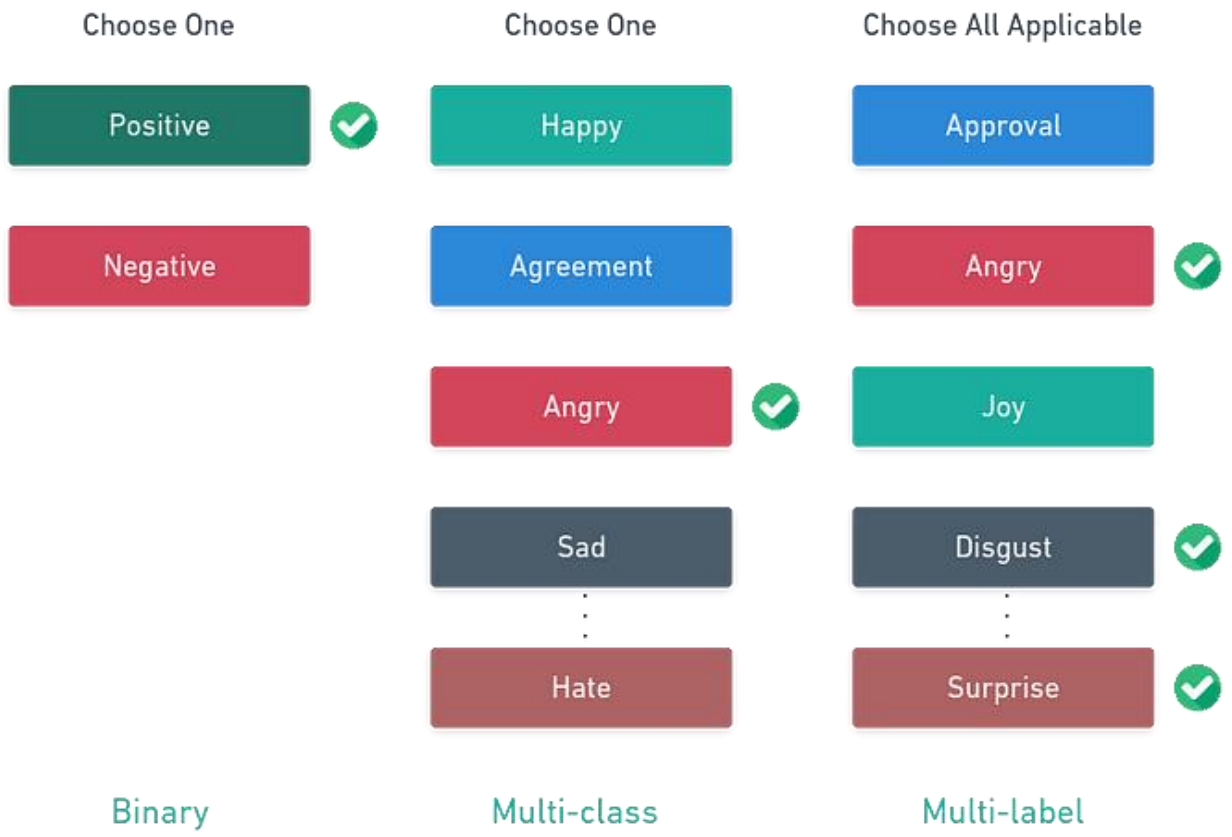




Zhang, W., Li, X., Deng, Y., Bing, L. and Lam, W. (2022). A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges. *arXiv (Cornell University)*.

Extensions of sentiment analysis

- Fine-grained scale of sentiment
- Multilabel classification : go emotions dataset



https://towardsdatascience.com/multi-label-emotion-classification-with-pytorch-huggingfaces-transformers-and-w-b-for-tracking-a060d817923

Slovenian language in NLP – available datasets

Annotated news corpora and a lexicon for sentiment analysis in Slovene (Bučar, J., Žnidaršič, M. & Povh, J.)

- 250 000 documents
- 10 000 manually annotated

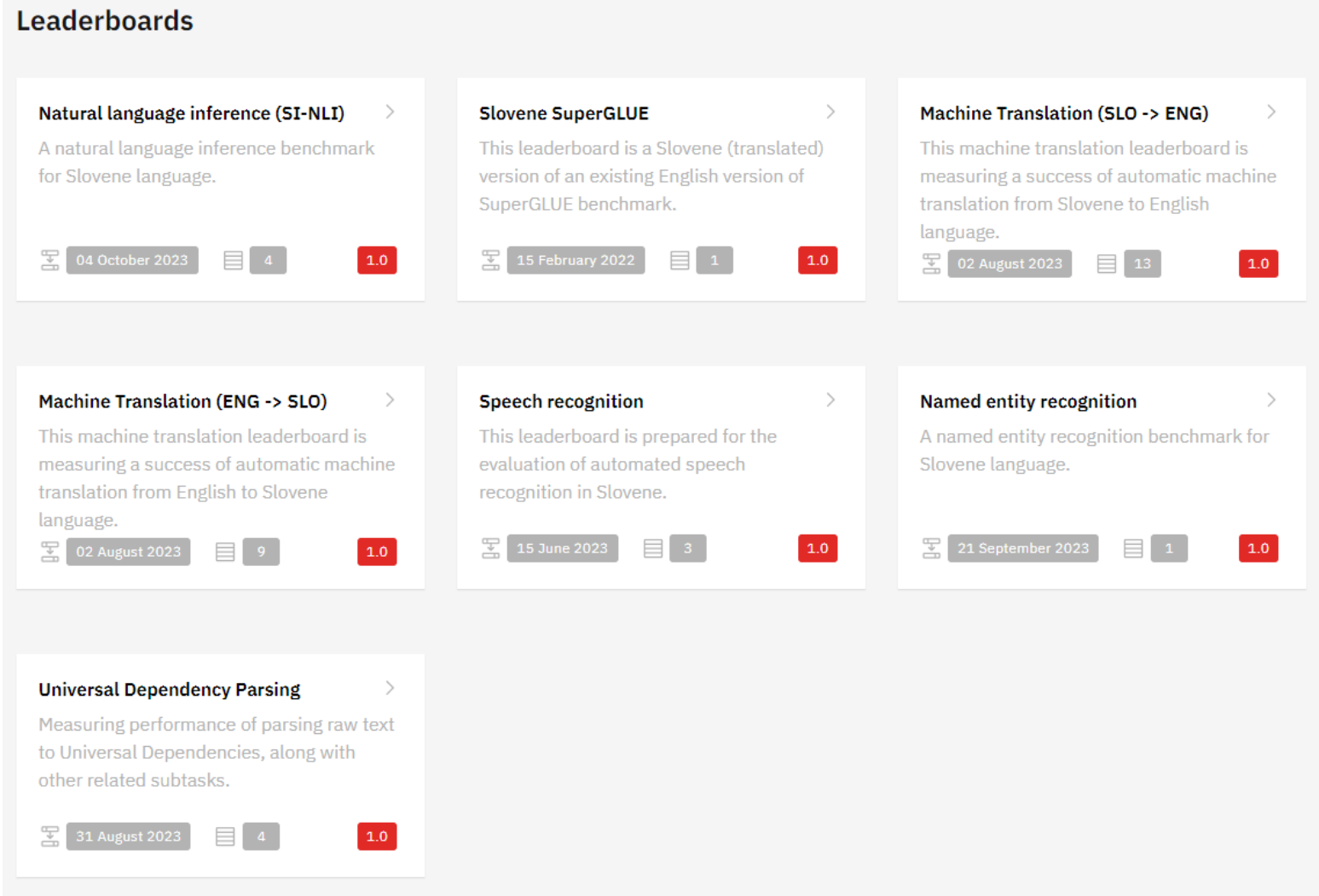
Slovene corpus for aspect-based sentiment analysis - SentiCoref 1.0 (Žitnik, S.)

- 837 documents
- 31 000 named entities
- 5 level sentiment annotation for each entity

https://slobench.cjvt.si/

Slovenian language in NLP

- No sentiment analysis task in the benchmark
- 1 model available at huggingface.com referring to Slovenian and Croatian



https://slobench.cjvt.si/

eXplainable Artificial Intelligence

Integrated Gradients

$$IntegratedGrads_i(x) ::= (x_i - x'_i) \times \int_{\alpha=0}^{1} \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x_i} d\alpha$$

- https://medium.com/@kemalpiro/xai-methods-integrated-gradients-6ee1fe4120d8

Integrated Gradients

$$IntegratedGrads_i(x) ::= (x_i - x_i') \times \int_{\alpha=0}^{1} \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x_i} d\alpha$$

- https://medium.com/@kemalpiro/xai-methods-integrated-gradients-6ee1fe4120d8

Application in NLP

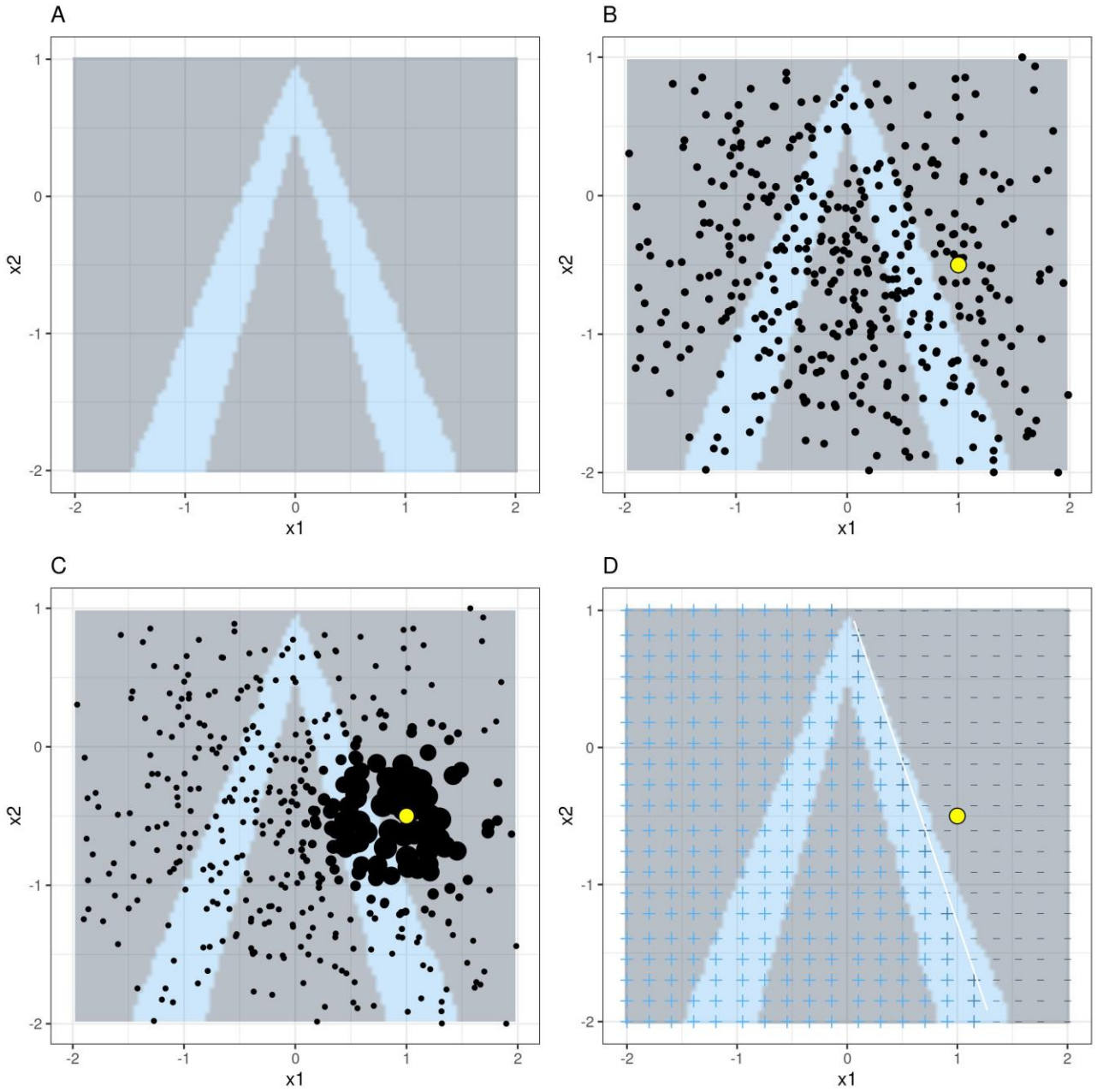
- Baseline = [#pad] (embedding dependent)
- Attributions calculated with respect to embedding tensor
- Attributions summed across all embedding dimensions to get word/token level attribution

Legend: Negative Neutral Positive

True Label	Predicted Label	Attribution Label	Attribution Score	Word Importance
pos	pos (0.96)	pos	1.29	it was a fantastic performance ! #pad
pos	pos (0.87)	pos	1.56	best film ever #pad #pad #pad #pad
pos	pos (0.92)	pos	1.14	such a great show ! #pad #pad
neg	neg (0.29)	pos	-1.11	it was a horrible movie #pad #pad
neg	neg (0.22)	pos	-1.03	i 've never watched something as bad
neg	neg (0.07)	pos	-0.84	that is a terrible movie . #pad

https://github.com/pytorch/captum/blob/master/tutorials/IMDB_TorchText_Interpret.ipynb

LIME

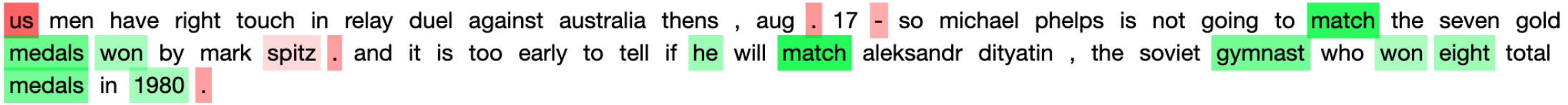


https://christophm.github.io/interpretable-ml-book/lime.html

LIME in NLP

- Generation of perturbed instances using masks
- Similarity computed in embedding space

Output:



https://github.com/pytorch/captum/blob/master/tutorials/Image_and_Text_Classification_LIME.ipynb

Concept-based interpretability : TCAV

- Concepts
- Predefined, high-level features
- In image processing: "stripes"
- In NLP : "positive sentiment"
- Represented as dataset

```
. . so well . . .
. . so good . . .
. . . love it . pad
. . . like it . pad
. . even greater . . .
. antic . . . . .
. . . fantastical . . .
. grotesque . . . . .
fantastic . . . . . pad
grand . . . . . pad
```

- Concept Activation Vectors (CAVs)
- CAVs are obtained by perturbing input data along the concept's direction and observing the model's responses.

https://github.com/pytorch/captum/blob/master/tutorials/TCAV_NLP.ipynb

Captum library

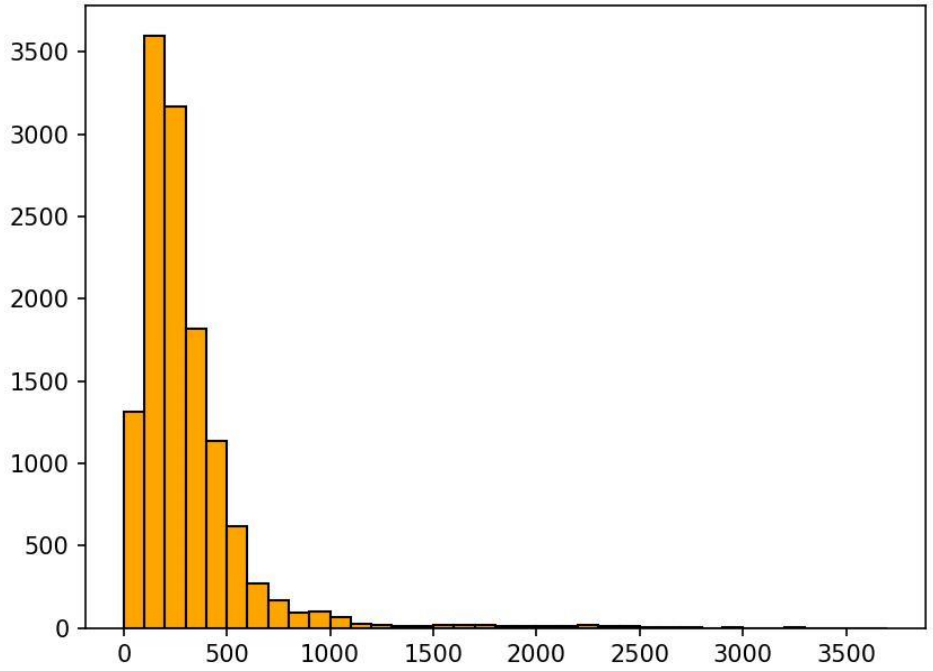
- Integrated Gradients, LIME and TCAV are implemented
- Well-maintened and well-documented

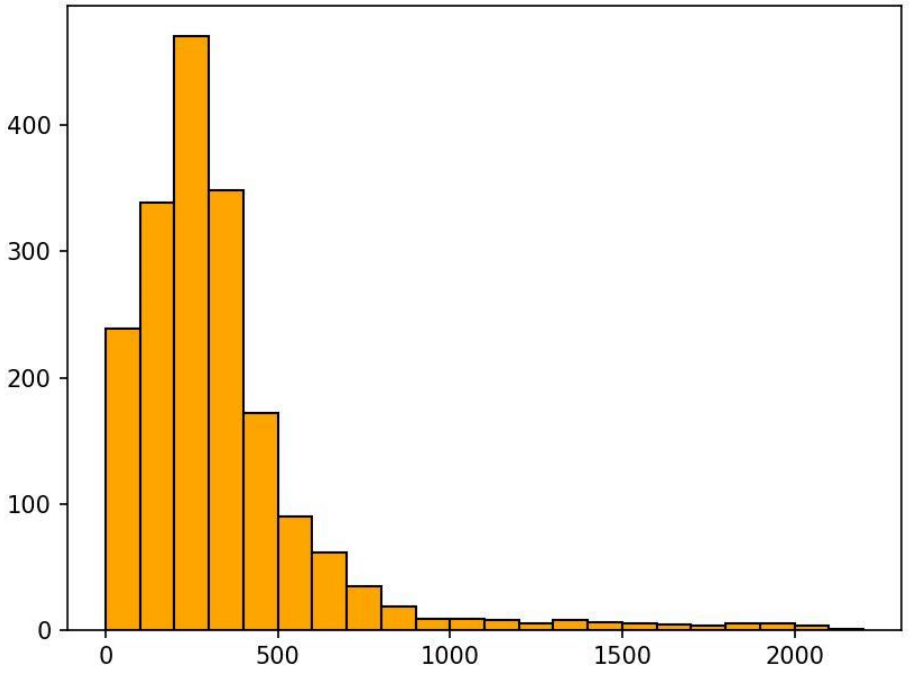
https://captum.ai/

STA database and its API

- News text;
- Authors (their initials) of the news;
- Headline;
- List of categories;
- List of keywords;
- Places;
- Timestamp of the creation of the news.

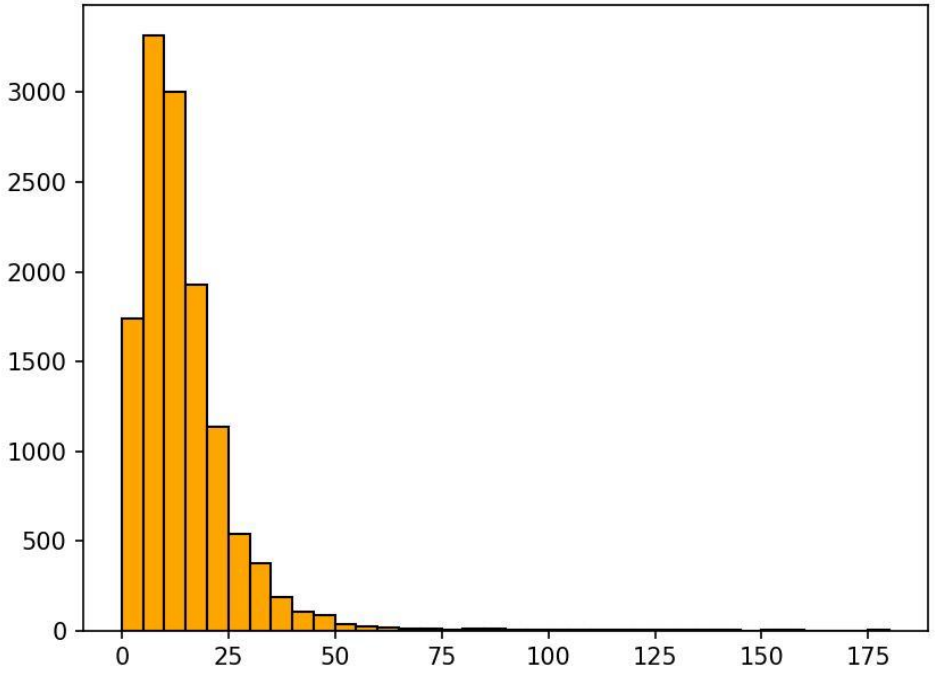
Preliminary exploratory data analysis (EDA)

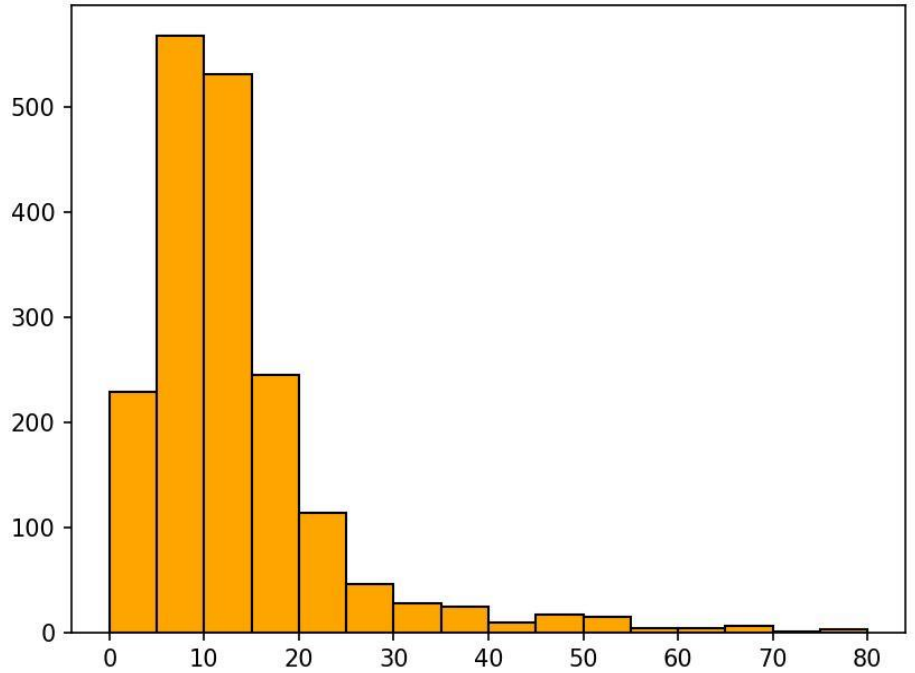




- Number of used words in news for Slovenian (on the left) and English (on the right).

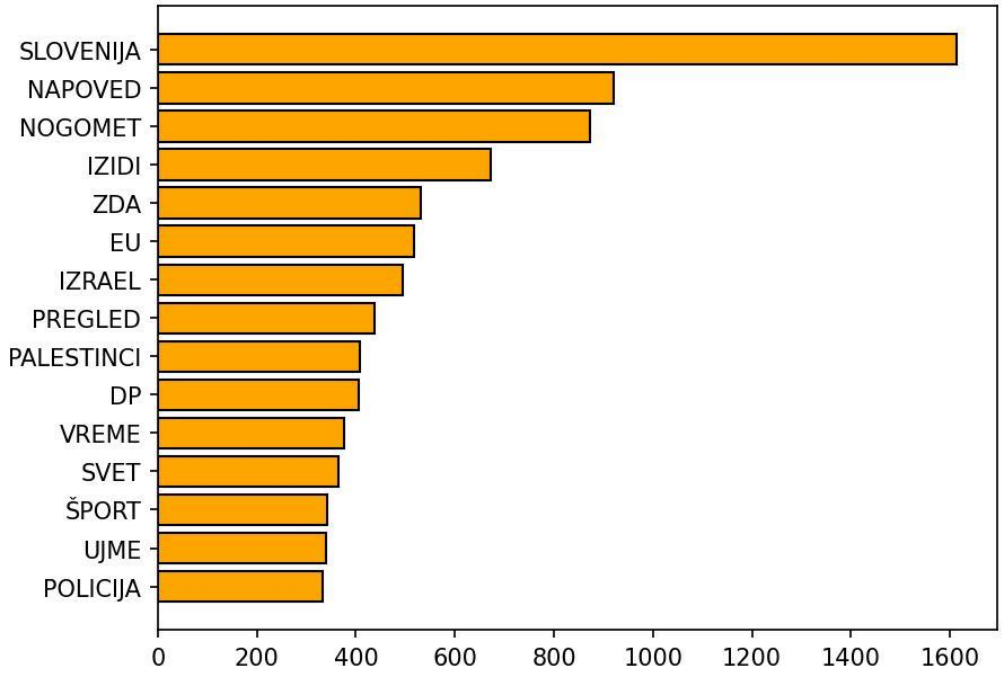
Preliminary exploratory data analysis (EDA)

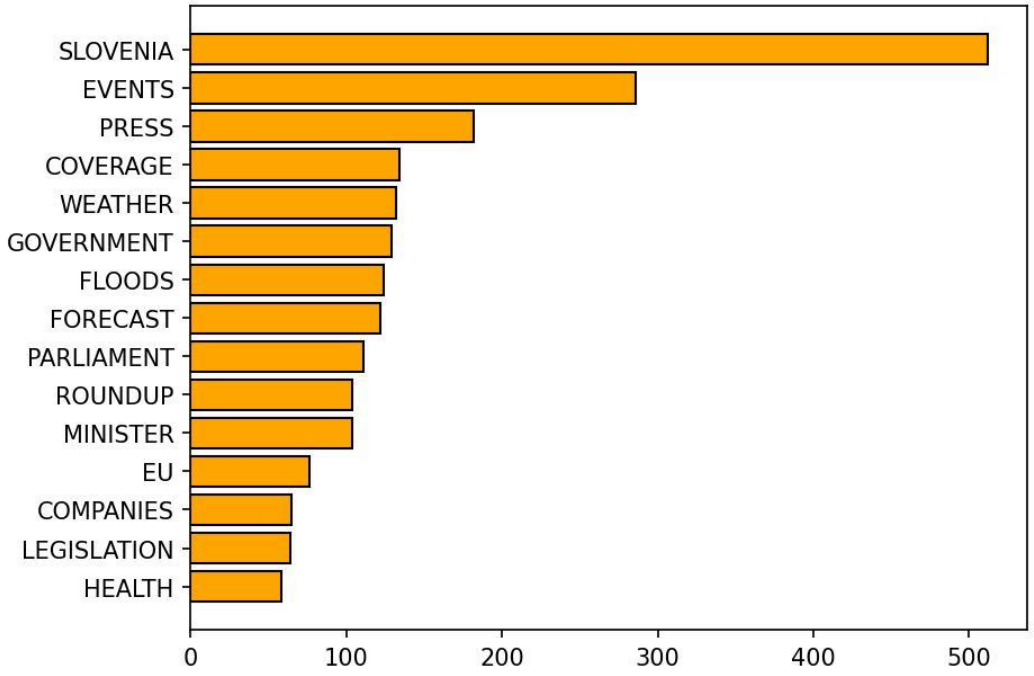




- Number of sentences in news for Slovenian (on the left) and English (on the right).

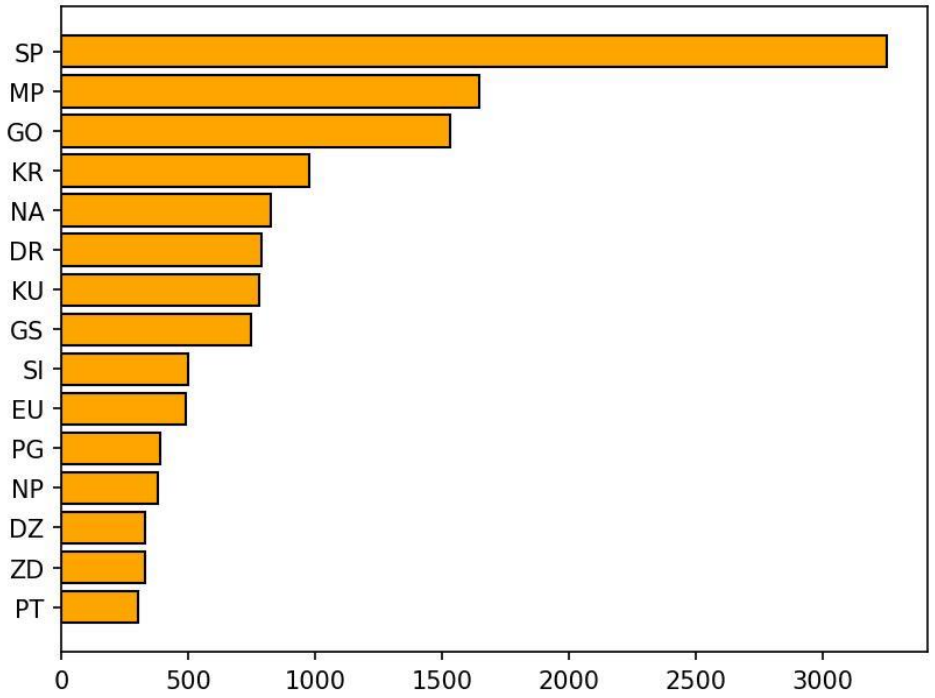
Preliminary exploratory data analysis (EDA)

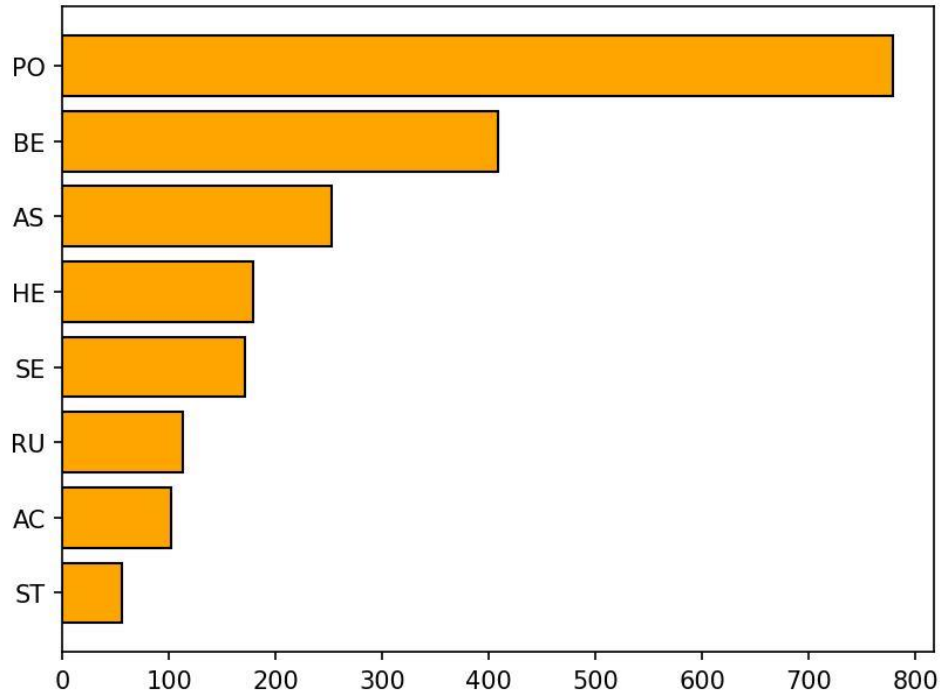




- Most common keywords in Slovenian (on the left) and English (on the right) news.

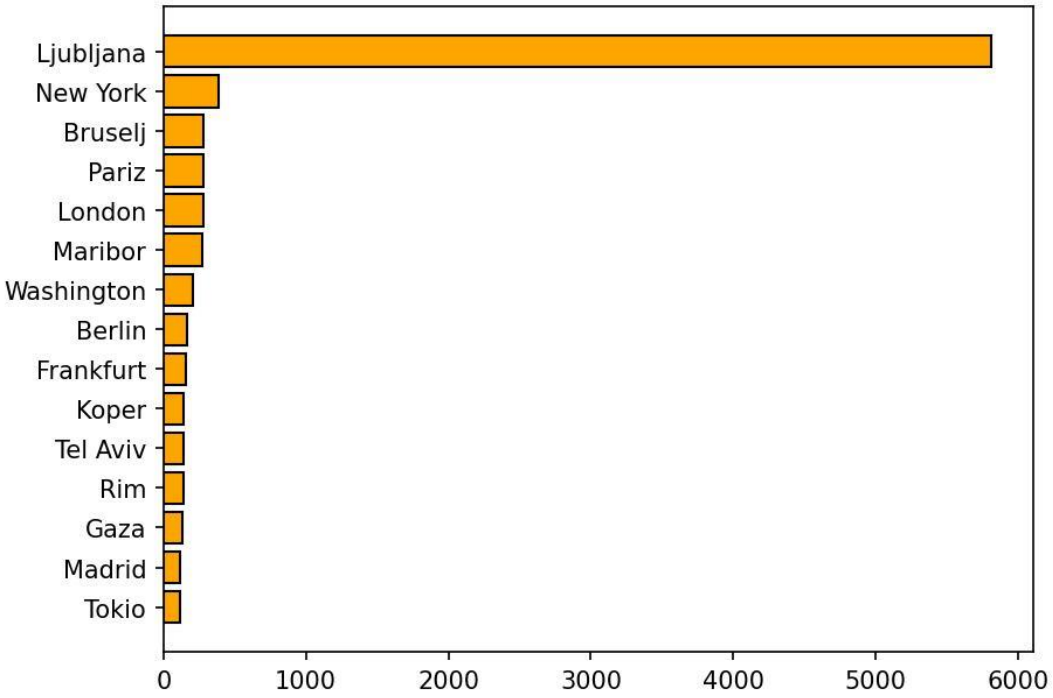
Preliminary exploratory data analysis (EDA)

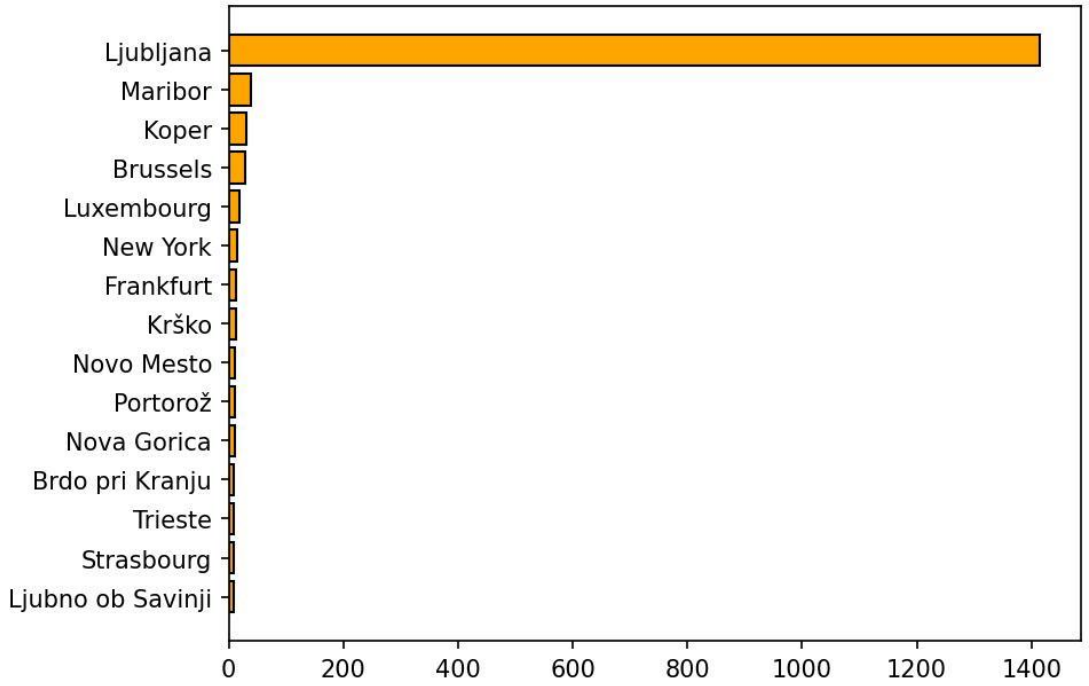




- Most common categories in Slovenian (on the left) and English (on the right) news.

Preliminary exploratory data analysis (EDA)





- Cities connected with news for Slovenian (on the left) and English (on the right).

Solution concept

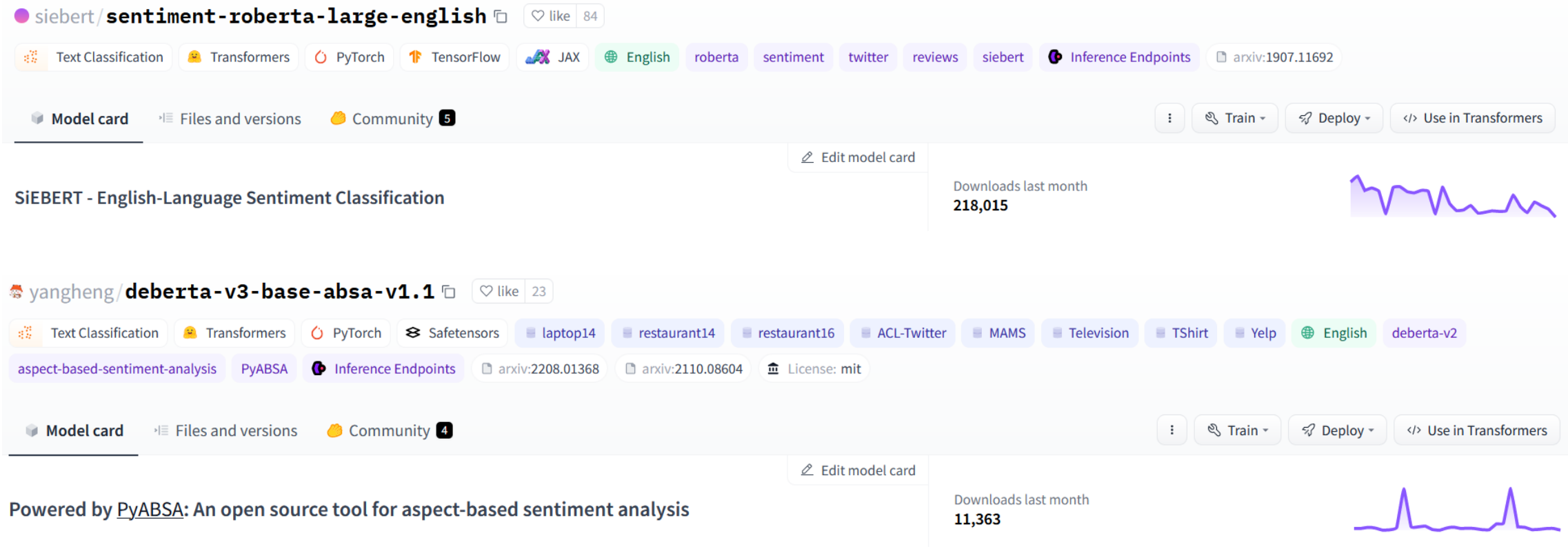
- No annotated data – we plan to annotate English data manually at least for the test set. For the first project, we only focus on the English news.

- We'll test existing pre-trained models and see how good they perform. In case on poor performance, we'll annotate also train data and fine-tune the models, or we could use existing annotated domain-specific data, in this case data with news.

Solution concept

- We'll prepare visualizations using outputs of our final model.
- We'll also perform XAI to explain performance of the model.

Starting models



Questions?