

The EBU logo is displayed in a bold, white, sans-serif font. The background of the slide is a photograph of a modern glass-walled building with several large satellite dishes mounted on its roof. The sky is clear and blue, and there are some lens flare effects in the upper left corner. On the right side, there are three overlapping semi-transparent circular shapes in shades of blue and white.

EBU

OPERATING EUROVISION AND EURORADIO

LANGUAGE AS A KEY TO DETECT FAKE NEWS

ALEXANDRE ROUXEL - T&I

PIERRE FOUCHÉ - T&I - SWISS DATA SCIENCE CENTER

EBU

OPERATING EUROVISION AND EURORADIO

LANGUAGE AS A KEY TO DETECT FAKE NEWS





FAKE NEWS DEFINITION



- › We consider as Fake News , articles (news report, editorial, expose, etc.) intentionally deceptive.
- › Fake News articles are written and published with the intent to mislead the reader, the ultimate goal being to benefit from it in many different ways. This often results in damaging an agency, entity, or person.
- › Don't falsify the term Fake News it is not a question of opinion !



AUTOMATED DETECTION



- › Considering the definition as mentioned above. There are different strategies and algorithms to detect Fake News :
 - › Analyse the sources of information,
 - › Fact-checking,
 - › Network traces.
- › We are focusing on language analysis.



LINGUISTIC CLUES



- › The Fake News publishers often have malicious intent to spread misleading information and to influence large communities
- › It requires a particular writing style: this is our key assumption!
- › But human performance in detecting Fake News by analysing the content without knowing the context and the facts is very poor, close to a random choice!
- › That's why we use Natural Language Processing, a subfield of Machine Learning, to analyse the writing content



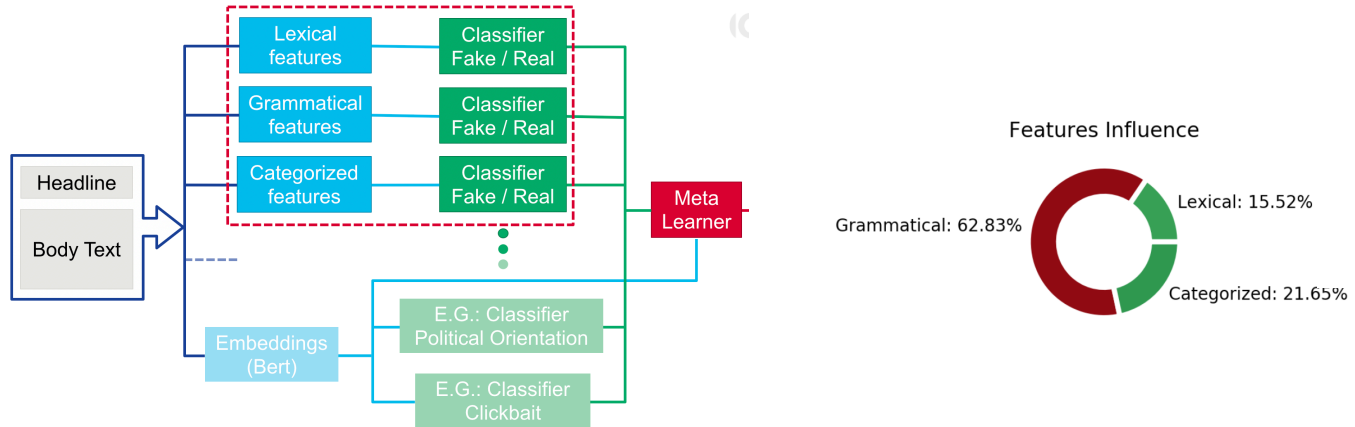
APPLICATION TO JOURNALISM



- › Fact-checking is a complex problem, which requires a thorough understanding of the context and, in this game, journalists are much better than Machine Learning techniques.
- › On the other hand, Machine Learning and NLP are more efficient than humans at extracting rules from the content of the writing.
- › That is why the tool we have developed helps journalists to extract linguistic clues from articles.
- › It can also be used to analyse massive amount of articles to assess the quality of the information source without any prior assumption.



EXPLAINABLE FAKE NEWS DETECTOR



- › A classifier generating the probability of being fake is applied for each category of features :
 - › Lexical, Grammatical, Categorized and Embeddings
- › We then pass these probabilities to a meta-learner that takes the final decision
- › It facilitates the tuning of hyperparameters per group of features
- › The explainability can have different granularity, it is a hierarchical structure



PERSPECTIVE AND FUTUR WORKS

- › Extend the tool to generate meta-data related to content like the level of formality, style, news categorization
- › Adapt the model to the detection of neural fake news

Gather data to train and test our models !

EBU

OPERATING EUROVISION AND EURORADIO

THANK YOU !

