



LANGUAGE AS A KEY TO DETECT FAKE NEWS





FAKE NEWS DEFINITION



- We consider as Fake Fews, 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



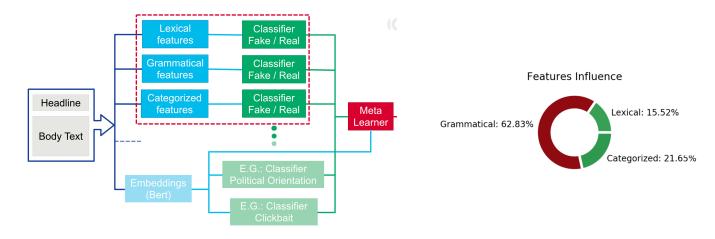
APPLICATION TO JOURNALISM



- > 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



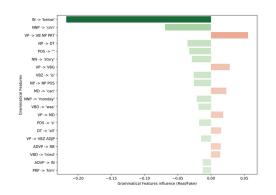
- A classifier generating the probability of being fake is applied for each category of features :
 - Lexical, Grammatical, Categorized and Embeddings
- We then pass these probabilities to a meta-learner that takes the final decision
- > It facilitates the tuning of hyperparameters per group of features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

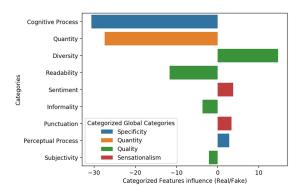


DEEPER EXPLAINABILITY

We provide humanly understandable inputs showing how the features categories influence the classification decision.

This is hot off the presses folks – something dangerous is going on in the large York (it) metro area. At around 12:40 AM Monday morning police were alerted to a suspicious backpack near a trashcan in an Elizabeth, large Jersey trashcan. The men who reported it said they could see a pipe and wires sticking out of the bag. When police arrived they discovered FIVE different devices but when a bomb squad robot approached the bag one of the devices exploded, fortunately no one was injured. UPDATE: Elizabeth, NJ mayor says five devices found in bag near train station. One exploded while bomb squad robot tried to disarm it – Fox News (@FoxNews) explembed 19, 2016 The latest bomb scare comes large one day after two other bombs exploded (one in Seaside Park, large Jersey and one in Manhattan) and three more were found by authorities before they could be detonated. Late Sunday evening authorities announced that they believed all five bombs from Saturday were the work of the same people, then later if lannounced later than a manufactor of a "vehicle of interest" in the Chelsea bombing Sunday night in Briooklyn. A law enforcement told Fox News that a "number of individuals" who are possibly connected to the explosion were taken into custody. Law enforcement officials told the Associated Press that at least 5 men





Lexical features

Grammatical features

Categorial features



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!

