

# Fake News Detection Using Machine Learning

---

Author: Simon Lorent

Supervisor: Aswhin Ittoo

September 10, 2019

University Of Liège

# Introduction

---

1. Expliquer l'impact des fake news sur les élections
2. La suite définit ce que sont les fake news

2016 US presidential elections (Allcott et al.)[1]

- 62% of US citizens get their news from social media[2]
- Fake news had more share on Facebook than mainstream news[3].

## Definition

### Definition

*Fake news is a news article that is intentionally and verifiable*

*false[4]*

1. Dire que l'on peut caracteriser les fake news de plusieurs faon:  
le contenu et le context

News content features:

- **Source:** Where does the news come from, who wrote it, is this source reliable or not.
- **Headline:** Short summary of the news content that try to attract the reader.
- **Body Text:** The actual text content of the news.
- **Image/Video:** Usualy, textual information is agremented with visual information such as images, videos or audio.

# Fake News Characterisation ii



Different kind of models:

- **Expert-oriented:** relies on experts, such as journalists or scientists, to assess the news content.
- **Crowdsourcing-oriented:** relies on the wisdom of crowd that says that if a sufficiently large number of persons say that something is false or true then it should be.
- **Computational-oriented:** relies on automatic fact checking, that could be based on external resources such as DBpedia.

## Methodology

---



Goal:

- Comparing the performances of "traditional" machine learning techniques and deep learning techniques
- Comparing the performances of these techniques on two different datasets.

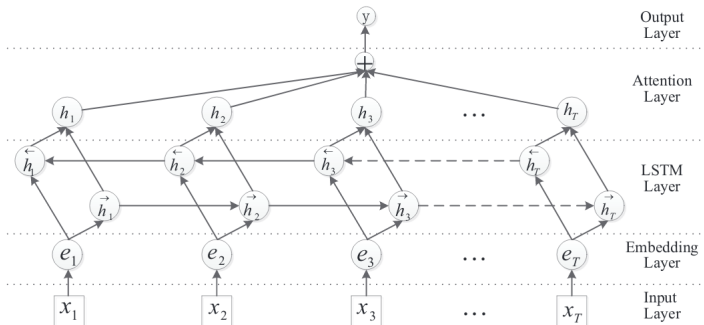
Models used:

- Naïve-Bayes
- SVM
- Decision Tree
- Ridge Classifier
- LSTM
- Attention Mechanism[5]

Two kind of text to vectors:

- TF-IDF
- word2vec[6]

# Attention Mechanism i





Hunt Allcott and Matthew Gentzkow.

**Social media and fake news in the 2016 election.**

*In Journal of Economic Perspective*, volume 31, 2017.



Jeffrey Gottfried and Elisa Shearer.

***News Use Across Social Medial Platforms 2016.***

Pew Research Center, 2016.



Craig Silverman and Lawrence Alexander.

**How teens in the balkans are duping trump supporters  
with fake news.**

*Buzzfeed News*, 3, 2016.



Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu.

**Fake news detection on social media: A data mining perspective.**

*ACM SIGKDD Explorations Newsletter*, 19(1):22–36, 2017.



Peng Zhou, Wei Shi, Jun Tian, Zhenyu Qi, Bingchen Li, Hongwei Hao, and Bo Xu.

**Attention-based bidirectional long short-term memory networks for relation classification.**

In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short*

*Papers*), pages 207–212, Berlin, Germany, August 2016.  
Association for Computational Linguistics.



Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean.

**Efficient estimation of word representations in vector space.**