Project Description

Department of Economics, University of Copenhagen

Summer 2022

1. **Group number?**

   Group 7

2. **Names of students in group?**

   Jørgen Baun Høst and Marius Heltberg Lassen

3. **What is your research question?**

   What 'fake news trends' from a news site such as the right-wing breitbart.com (or similar) can we detect in the lead-up to the 2020 election through a supervised ML model?

4. **What kind of data are you planning on using? How will you get access to these data?**

   We are goíng to be using mostly text-as-data. Our 'training' data will come from PolitiFact which we have webscraped (adhering to the terms of use).

5. **What will your data analysis be like? Will you use machine learning? How?**

   We have scraped PolitiFact for their human-annotated fact checks. As such, we will make a text-as-data analysis. This includes an overview of what 'tags'/subjects are fact checked, the distribution of true/false statements, topic modelling. We are going to use machine learning (logistic regression w/ TfidfVectorizer word preprocessing).

6. **Have you already identified other papers within this area that you can use in a literature review? If so, name a few and explain what they do in one sentence only.**

(a) Allcott and Gentzkow (2017): A theoretical framework for fake news (they mention PolitiFact). Good to refer to when describing the phenomenon of fake news.

(b) Reis et al. (2019): Fake news detection using supervised ML (similar approach to us)

(c) Shu et al. (2017): Fake news detection on social media - technical paper on methods.

(d) Kshetri and Voas (2017): Short paper theoretical paper on the 'economics' of fake news.

7. **How do you 'contribute' to the literature?**

We would contribute to the growing literature on fake news detection, focusing on the 2020 US election.

References

Allcott, Hunt, and Matthew Gentzkow. 2017. Social media and fake news in the 2016 election. *Journal of economic perspectives* 31 (2): 211–36.

Kshetri, Nir, and Jeffrey Voas. 2017. The economics of "fake news". *IT Professional* 19 (6): 8–12.

Reis, Julio CS, André Correia, Fabrıcio Murai, Adriano Veloso, and Fabrıcio Benevenuto. 2019. Supervised learning for fake news detection. *IEEE Intelligent Systems* 34 (2): 76–81.

Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: a data mining perspective. *ACM SIGKDD explorations newsletter* 19 (1): 22–36.