Capstone Two: Project Proposal

Fake News Detection

By: Angela Dunston

The rapid spread of fake news has become a major issue worldwide. The spread of false and misleading news has led to significant social and economic consequences, impacting industries from finance to healthcare. In this context, this project will explore opportunities that exist the identify fake news as soon as it is available on social media to limit its influence on people, communication, and to prevent confusion.

The criteria for success would be a model that accurately identifies patterns in news that a considered to be fake. The scope of the solution space is using sample data for detecting fake articles; however, the process can be used to detect fake tweets or anything similar. A major constraint would be the amount of data that is available on the internet. Privacy concerns would be another constraint as it may limit access to data. Stakeholders include any user of the internet. Key data sources are as follows. Kaggle.com and FavioVazquez/fake news.

The fake news detection model will be solved by the doing the following steps:

1) Data reading and concatenation
2) Data cleansing
3) Data exploration
4) Modeling

Deliverables will be a GitHub repo, a slide deck, and project report.