Fake News: Desarrollar y entrenar un modelo para detectar Fake News.

Data sets:

Kaggle

1. **Manual and crowdsourced annotation approaches (FakeNewsAMT)**

* Temas: technology, education, business, sports, politics, and entertainment.
* Sources of legitimate news: ABCNews, CNN, USAToday, NewYorkTimes, FoxNews, Bloomberg, and CNET among others.
* The fake and legitimate news are provided in two separate folders. Each folder contains 40 news from the six different domains.
* The file names indicate the news domain: business (biz), education (edu), entertainment (entmt), politics (polit), sports (sports) and technology (tech).
* The fake and legitimate labels are also provided as part of the filename.

1. **Collected directly from the web (Celebrity).**

* Temas: entertainment, fashion and style
* Legitimate: news websites and from entertainment magazines websites.
* The fake news were obtained from gossip websites such as Entertainment Weekly, People Magazine, RadarOnline, and other tabloid and entertainment-oriented publications.
* The articles were manually verified using gossip-checking sites such as "GossipCop.com", and also cross-referenced with information from other entertainment news sources on the web.
* The fake and legitimate news are provided in two separate folders. The fake and legitimate labels are also provided as part of the filename.