Paper : SpotFake: A Multi-modal Framework for Fake News Detection

Summary

The paper proposes SpotFake – a multimodal framework for fake news detection. It considers features from two different modalities and classifies the sample into real or fake without taking into account any other sub-task.

A survey was conducted to do an empirical analysis on human performance, difficulties of fake news detection, and the importance of multiple modalities, especially the combination of text and image for fake news detection. A total of 88 candidates participated in the survey. 63% of them were male. The majority of candidates (about 64%) were 15-50 years old.

Based on the survey and based on previous work, the paper claims that it is evident that a multimodal system is necessary for fake news detection.

SpotFake considers the imagery and the text in a sample to determine whether it is fake or not. It has three submodules. One is a Textual Feature Extractor that is responsible for extracting the contextual text features from the posts. The next is a Visual Feature Extractor that is responsible for pulling the visual elements from the posts. The third is a Multimodal Fusion submodule that fuses the two features such that they can be passed through a fully connected neural network for fake news classification.

Valuable Contributions of the Paper

- The survey conducted by the team to help verify the hypothesis that a multimodal system is necessary for fake news detection.
- SpotFake.

Critique and Future work/improvements

SpotFake only takes into account the content, i.e. the text and the image. It doesn't consider the account that posted the content, the hashtags along with it, and other such details. Looking at the account that posted the content, things like whether the account is verified, how often the account posts things, whether the account was reported/blocked for posting false information before, and other such information might help identify fake news more accurately.