# Affect-Oriented Fake News Detection Using Machine Learning

***Mr. Anoop Kadan****

*Email: anoopk_dcs@uoc.ac.in*

Hoo! Cashews instead of Prozac! Treating depression with Prozac is fearful while thinking about a few things such as loss of weight and appetite. But if two handfuls of cashews are equivalent to one dose of Prozac, it can be safely substituted.


*"Two handfuls of cashews can alleviate depression just as much as a dose of Prozac"*

These are some of the health news articles in social media from the past, and still circulating to astonish us. But how many of us are aware that these are just fake news circulated to make people believe them?

*"Placing a raw, cut onion in contact with your foot overnight purifies your blood, removes toxins, and heals your body."*

Expeditious growth in technology, along with the advent of the Internet, has minimized the challenges to reach a news source. Among all other media platforms, online social media

---

* Mr. Anoop Kadan, PhD Scholar from the Department of Computer Science, University of Calicut, Kerala, is pursuing his research on "Emotion Detection and Application in Fake News Detection". His popular science story entitled "Affect-Oriented Fake News Detection Using Machine Learning" has been selected for AWSAR Award.

plays an important role in sharing news and information along with user opinion. This quick propagation and accumulation of information form a data deluge where it is very hard to believe all the pieces of information even though it appears to be very realistic. Hence, the success of social media networks marked through its assistance and situational awareness during disasters, crisis, and emergencies are harmed by the creation and propagation of fake information.

Most of the domains in online social news media are adversely affected by fake news. The spread of *fake news* through online social media during natural disasters such as Hurricane Sandy at Houston during 2012, the earthquake in Chile during 2010, and Tsunami in Japan during 2011, has caused panic and chaos among people. A tweet stating an explosion that injured Barack Obama, which wiped out 130 billion dollars in stock value within a few minutes, is an example of large-scale investments and stock market prices being affected by fake news. In political news, fake information is used to spread false beliefs among people. Besides these domains, fake news on health and well-being pose serious adverse effects, mainly by delaying necessary medical care and attention to a patient, making patients doubtful on the doctors' advice or going behind treatments that are not medically proven.

Among all other media platforms, online social media plays an important role in sharing news and information along with user opinion. This quick propagation and accumulation of information form a data deluge where it is very hard to believe all the pieces of information even though it appears to be very realistic.

Fake news creation is not new. We can even see stories of fake news dating back to the early 13th-century BC. One example is the story of an Egyptian pharaoh Ramasses II, who spread propaganda stating the victory of the Egyptian Empire over Hittite in a battle by depicting some scenes of himself striking the opponents on the walls of nearby temples. Then to now, fake news is one of the greatest issues, and hence fake news was the Macquarie Dictionary Word of the Year and *post-truth* was the Oxford Dictionaries Word of the Year in 2016.

In this context, we at the Computational Intelligence and Data Analytics Lab, University of Calicut, are working toward characterizing and recognizing fake news in online social news media. We started our work in collaboration with Dr. Deepak P of Queen's University, UK, who has strong backgrounds in the areas of *Natural Language Processing* and *Affective Computing* that could help our research. To understand the holistic picture of this research area, our first step was to study the conventional fake news detection systems. We observed that conventional systems employ techniques to detect fake news using the *content of the news, social network properties*, or *knowledge-based* methods. The *content of the news* includes both textual data and images associated with the text. Machine learning algorithms are generally used on the features (properties) extracted from these textual or image data to detect fake news. *Social network properties* used for fake news detection include the knowledge about a social media user like how he/she is propagating the news, structure, and the behavior of a social media network, time-dependent propagation features, etc. The *knowledge–based* approaches assess the genuineness of news by checking the source and history of each news with the vast amount of information available around us. We have published this initial study as a chapter *Leveraging Heterogeneous Data for Fake News Detection'* in the *book 'Linking*

*and Mining Heterogeneous and Multi-view* Data published by Springer International Publishing in December 2018.

Apart from the conventional systems, we tried to model a simple and efficient fake news detection system based on the *affective character* (emotion content) of the news article. The idea of using *affective character* was based on our observation that fake news articles are generally blended with strong emotion content and exaggerations intended to attract eyeballs and mislead people. The following are some examples of fake and real news headlines, which show the presence of strong emotion content in fake news than in real news.

**Fake news headlines:**

- Warning: This household plant can kill a child in less than a minute and an adult in 15 minutes!
- Revolutionary juice that can burn stomach fat while sleeping

**Real News headlines:**

- Chain-smoking children: Indonesia's ongoing tobacco epidemic
- Breastfeeding makes kids more likely to eat vegetables

Such news articles with strong emotion content are what is most significant about contemporary fake news. These kinds of emotionally targeted news produced by journalism are referred to as empathic media. The commercial and political phenomenon of automated empathic fake news creation is on the near horizon, which requires significant attention.

Our model *Affect-oriented Fake News Detection*, utilizing the emotion character embedded in fake news, was built over the popular *discrete emotion theory* of six basic emotions *anger, disgust, fear, happiness, sadness*, and *surprise* suggested by Paul Ekman. We set our first objective of fake news detection over a corpus of health and well-being news domain, considering it as a novel direction of inquiry toward an important domain where injection of fake information posed serious issues.

To carry out the experiments there was a lack of textual corpus/dataset in the health and well-being domain. We, for the first time, to the best of our knowledge, procured a *Health and Well-Being Real versus Fake (HWB-RvF)* news dataset from 7 real and 15 fake news web portals with the help of multiple fact-checking sites such as Snopes.com. The *HWB-RvF* dataset consisted of 500 real and 500 fake news, which shall be publically released for the future research community with our corresponding ongoing publication *On the Coherence of Fake News Articles.*

We developed an algorithm that utilized an emotion lexicon (word–emotion dictionary, for example, unlucky–sadness, joy–happiness) to amplify the emotion content in a document and fed this *emotion amplified document* to machine learning algorithms. Our empirical study illustrated that such an amplification helped significantly improve the accuracy of fake news detection. We have communicated this work of *Affect-oriented Fake News Detection* as a research paper entitled *Emotion Cognizance improves Health Fake News Identification.* As future work, we plan to extend our algorithm to other news domains such as science, politics, etc., by analyzing the emotion content in different domains.

Fake news creation is not new. We can even see stories of fake news dating back to the early 13th-century BC. One example is the story of an Egyptian pharaoh Ramasses II, who spread propaganda stating the victory of the Egyptian Empire over Hittite in a battle by depicting some scenes of himself striking the opponents on the walls of nearby temples.