Fake news

First Year Master's degree Advanced Software Engineering Techniques project

Theme:

Fake news identification. Users and their posts credibility on Twitter.

Description

The application will be presented as a Chrome plugin that will provide information related to the level of trust is assigned to a Twitter post and/or it's creator.

The app's algorithm will calculate, using a heuristic, the degree of a user's induced confidence based on the content he posts on the platform:

- the information will be collected from Twitter (posts, user information, reacts at certain posts (likes, comments, retweets, etc)
- user profiles will be analyzed and created(based on user details, it's activity, reacts of other users in connection with his activity, etc.)
- a metric will be provided to measure a user's credibility on Twitter
- a metric will be provided to measure a post's veracity
- methods for detecting „fake" users and „fake" news
- the information will be validated using external resources (Google, blogs, online newspapers, etc.)

Team members

- Dascalu Ana
- Glodeanu Irina
- Insuratelu Madalina
- Macariu Camelia
- Rusu Gabriel

Coordinators:

- Conf. dr. Iftene Adrian
- Prof.dr. Tiplea Ferucio Laurențiu

Github： https://github.com/IrinaGlodeanu/FakeNews

State of the art:

- What did the others, techniques and methodologies used, evaluation methods, results;
  1. B.S. Detector (http://bsdetector.tech/) is an Open Source browser extension for Chrome and Mozilla that alerts users to unreliable news sources through visual warnings. It searches all links on a given webpage for references to unreliable sources comparing them with a manually compiled list of domains;
  2. FiB by Devpost (https://devpost.com/software/fib) is a Chrome extension that goes through Facebook feed and verifies the authenticity of posts. The extension is built with JavaScript that uses advanced web scraping techniques to extract links, posts and images, which are sent to backend where artificial intelligence and a collection of APIs are used to determine if the received data represents fake news;
  3. TrustServista powered by Zetta Cloud (https://www.trustservista.com/) is a software that can determine the trustworthiness of online news, the origin of the information used in articles and understand semantic links between articles using artificial intelligence algorithms, processing natural language and semantic analysis.

- Important names in the field, research teams;
  1. AdVerifai (http://adverifai.com/) is a research team whose main purpose is to detect phony stories, nudity, malware and a host of other types of problematic content. Their research is focused on the intersection of Natural Language Processing and Machine Learning, covering topics such as recognizing textual entailment, question answering and argumentation mining;
  2. Facebook made fighting misinformation a priority. Not only they work with third-party fact-checkers, who are focused on articles, to reduce the spread of false news, but they built a machine learning model that uses various engagement signals, including feedback from people on Facebook, which is able to analyze photos and videos before sending those to their partners;
  3. Google News Initiative (https://newsinitiative.withgoogle.com/) represents Google's efforts to support the media industry by fighting misinformation, having three goals: highlight accurate journalism while fighting misinformation, help news sites continue to grow and create new tools to help journalists do their jobs.

- Related Articles and books;
  - Castillo, C., Mendoza, M., Poblete, B. (2011) Information Credibility on Twitter. WWW 2011 - http://chato.cl/papers/castillo_mendoza_poblete_2010_twitter_credibility.pdf - automatic methods for assessing the credibility of a given set of tweets
  - Automatic deception detection: methods for finding fake news (https://dl.acm.org/citation.cfm?id=2857152). The paper proposes guidelines for creating a fake news detection system, providing veracity assessment methods, using linguistic cue approaches with machine learning and network analysis approaches;
  - Fake News Detection on Social Media: A Data Mining Perspective (https://dl.acm.org/citation.cfm?id=3137600). Since fake news are written to mislead readers to believe false information, which makes it difficult to detect based

on news content, including auxiliary information, such as user social engagements on social media, would help make a determination;

- Fake news detection on social media (https://www.csustan.edu/sites/default/files/groups/University%20Honors%20Program/Journals/02_stahl.pdf). Using Naïve Bayes Classifier, Support Vector Machines and Semantic Analysis methods and a discussion on Linguistic Cue and Network Analysis approaches this paper proposes an accurate way to detect fake news on social media.

- Relevant links;
  - http://itransfer.space/identificarea-stirilor-similare-pe-twitter/
  - http://itransfer.space/identificarea-stirilor-false/
  - http://itransfer.space/identificarea-evenimentelor-pe-twitter/
  - http://itransfer.space/vizualizarea-informatiilor-pe-twitter/
  - http://itransfer.space/exploatarea-informatiilor-in-timp-real/
  - http://itransfer.space/exploatarea-informatiilor-de-pe-twitter/
  - http://aclweb.org/anthology/D17-2019
    https://www.researchgate.net/publication/319255985_Automatic_Detection_of_Fake_News
    http://sbp-brims.org/2018/proceedings/papers/challenge_papers/SBP-BRiMS_2018_paper_120.pdf
  - https://agency.reuters.com/en/insights/articles/articles-archive/reuters-news-tracer-filtering-through-the-noise-of-social-media.html
  - https://blogs.thomsonreuters.com/answerson/making-reuters-news-tracer/

- Resources and tools available.
  - **Botometer** is an online tool to classify Twitter accounts as human or bot. Formerly known as BotOrNot. - https://twitter.com/botometer
  - **Hoaxy -** search finds claims and related fact checking in a limited corpus of articles from low-credibility and fact-checking sources, dating back to 2016. - https://hoaxy.iuni.iu.edu/faq.html#faq-q1
  - **Fakey -** a web and mobile news literacy game that mixes news stories with false reports, clickbait headlines, conspiracy theories and "junk science." Players earn points by "fact-checking" false information and liking or sharing accurate stories. - https://fakey.iuni.iu.edu/

Conclusions regarding the relevant links:

- A strategy on classifying tweets :for a tweet that appears multiple times -> if a tweet will be resolved as being false, another with the same content/context will be also classified as false

- For the calculus of similarity: a strategy would be to use the string distance. The Jaro-Winkler algorithm is the most efficient when seen from time perspective

- Jaro-Winkler algorithm is described in http://itransfer.space/identificarea-stirilor-similare-pe-twitter resource

- "Stanford Named Entity Recognizer (NER)" – for detecting names in tweets : if a tweet contains the name of a person/entity is fake, we can relate other tweets about that person/entity

- we could take two approaches : social or linguistic

- Linguistic: the word composing the tweet will be processed, and, based on them, through different techniques(naive , syntactic analysis, machine learning) we can approximate, with a certain precision , if it's fake or not

- Social : using „knowledge network" -> concepts represented by nodes in a graph, connected by edges with certain costs, represented by the degree of relevance between them (Donald Trump and USA sunt are connected by a bigger score than McDonalds and Himalaya).

- The issue reduces to the computation of the shortest path between two concepts. The closer the concepts implied, the bigger are the odds that the news is true.

- We can inspire from the metrics of „quality" of tweets provided by Google (Google Similarity Distance) or Flickr (Flickr Distance)

- The filtering of data for the resolution of true/fake based on geographic area / language, etc. This helps the correlation between similar tweets about facts of interest of a certain area.

- Identification specific evaluation of events (duplicate events, precision, recall, F1). The precision of the evaluation depends on the approach of detection of those events.