Chapter 12 Machine Learning: Pattern Hunting

Notes 1

Introduction 1

AI Overview 2

Algorithms: Text Analysis 2

Narrative detection and analysis 2

Algorithms: Network Analysis 2

Image, Video and Audio Analysis Needs 3

Test datasets 3

Further reading 3

Notes

Data science nice, but big volume, patterns not obvious: use machines. AI good on other side, e.g. deepfakes; also for detection etcetc - Pablo's talk on AI/ML and disinfo.

Introduction

Disinformation analysis has changed a lot since 2016 when a search on \#qanon, and some simple checks would find you botnets and a disinformation campaign. There are people who are good at disinformation data science (Eliot Alderson, Conspirador Norteno etc), and there's been a lot of academic money in this area recently. This section covers useful tricks, processes and tools.

In the League, we also do many things by hand, but are working on ways to speed them up, and automating the parts that make sense (this will probably never be a fully-automated activity, but we can support and get a lot of load off disinformation analysts) as we validate process.

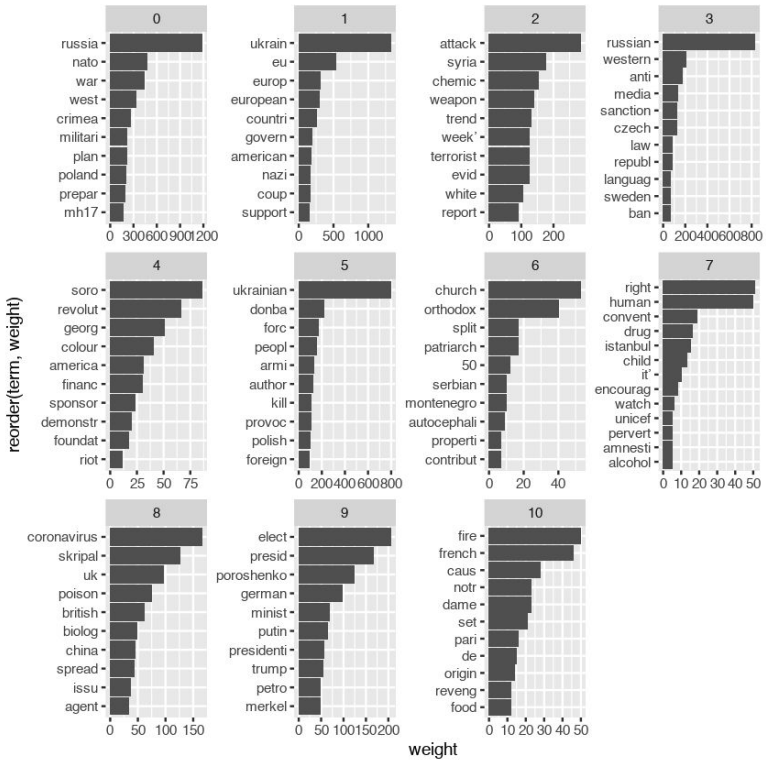
This chapter is about tackling the three Vs (volume, velocity, variety) with machine learning, automating the work that's too large, coming in too quickly or across too many channels for the team to concentrate on them all. There is no magic "plug in this algorithm and disinfo will go away" system: the big idea here is that the humans and algorithms can work together - e.g. this is augmented intelligence that lets the people focus on what they're good at, not a replacement for people.

One of the reasons we shift work from humans to algorithms is make sifting through the data and highlighting potential patterns in it more efficient; another is that disinformation artefacts (images, text etc) are often difficult to handle, and we want to reduce exposure to them as much as possible. If we can cluster artifacts so that instead of looking at 100 near-identical images, a human can view and classify one copy of that image, the reduction of stress on the humans is worth is.

AI Overview

<Fixit: drop in Pablo's AI talk>

Algorithms: Text Analysis



co-occurrance graph for text clusters in EuVsDisinfo (Gabe)

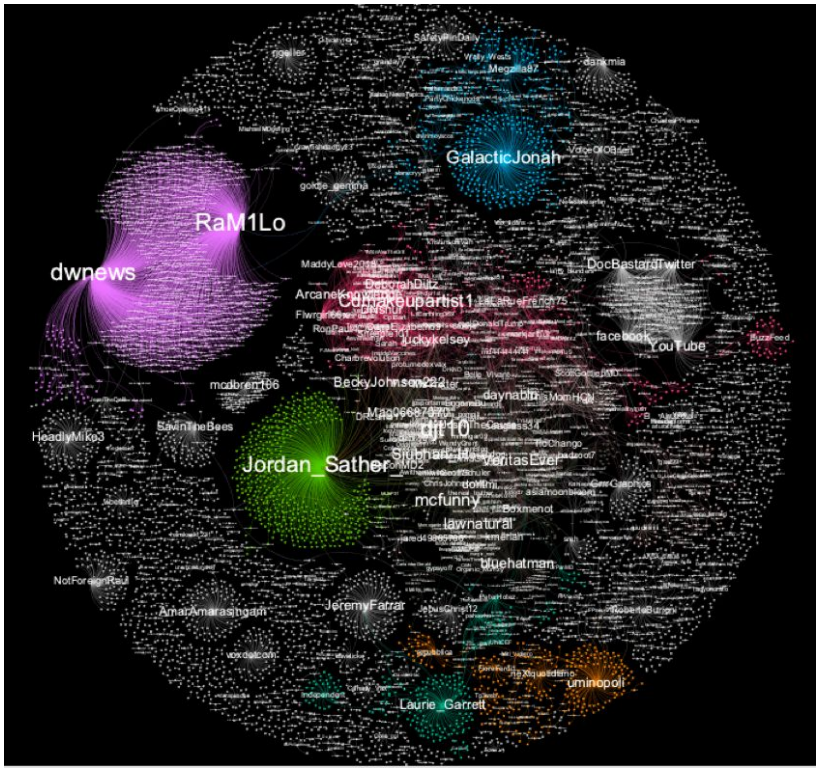
Text-based algorithm needs include:

- Find themes
- Classify to narratives
- Cluster text to narratives
- Search for similar text/narratives

Narrative detection and analysis

Topic modelling.

Algorithms: Network Analysis



Gephi visualisation around a Twitter hashtag

Network algorithm needs include:

- Finding super-spreaders
- Finding rumor origins
- Uncover new artefacts
- Track movement over time

Image, Video and Audio Analysis Needs

Image/Audio algorithm needs include:

- Cluster images

- Search for similar images
- Detect shallowfakes

Whilst we've seen deepfakes being used for things like fake profile pictures, most of our image/video etc needs have been more mundane: searching for reused and/or mistagged images, finding images that have been crudely doctored (shallowfakes), and clustering sets of near-identical images to make then faster to sift through with less exposure of potentially-harmful material to the people checking through them.

Test datasets

- Kaggle “getting real about fake news” Getting Real about Fake News - used a lot
- Twitter deleted 200,000 Russian troll tweets. Read them here. - NBC’s Russian twitter dataset
- fivethirtyeight/russian-troll-tweets - 538’s IRA dataset
- 538 dataset was from Salesforce’s Social Studio tool ($1000/month) Editions & Pricing: Social Media Marketing

PS a lot of the examples are in Python and Pandas - you don’t escape from learning these Python Data Science Handbook

Further reading

- https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624278/EPRS_STU%282019%29624278_EN.pdf
- Attention is All They Need: Combatting Social Media Information Operations With Neural Language Models - Fireeye on text generation and detection
- Looking for more: Google search “disinformation ‘data science’” - lot of posing. Found data science sites’ articles on disinformation. Github searches for “disinformation” and “misinformation” found hundreds of repos. Searching student sites like Towards Data Science for “fake news”, disinformation, misinformation can be useful because they’re scanning recent work.
- https://www.cnn.com/2019/05/23/politics/doctored-video-pelosi/index.html
- https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/