# The Antai Project

Athul Muralidhar

# The Advent of Machine Learning (1)

## Machine Learning

A subset of **Artificial Intelligence**, which deals with the study of algorithms based on Statistical modelling. **No explicit instructions** are given to the algorithms to performing tasks, rather **relying on inference** instead.

## Data usage of ML

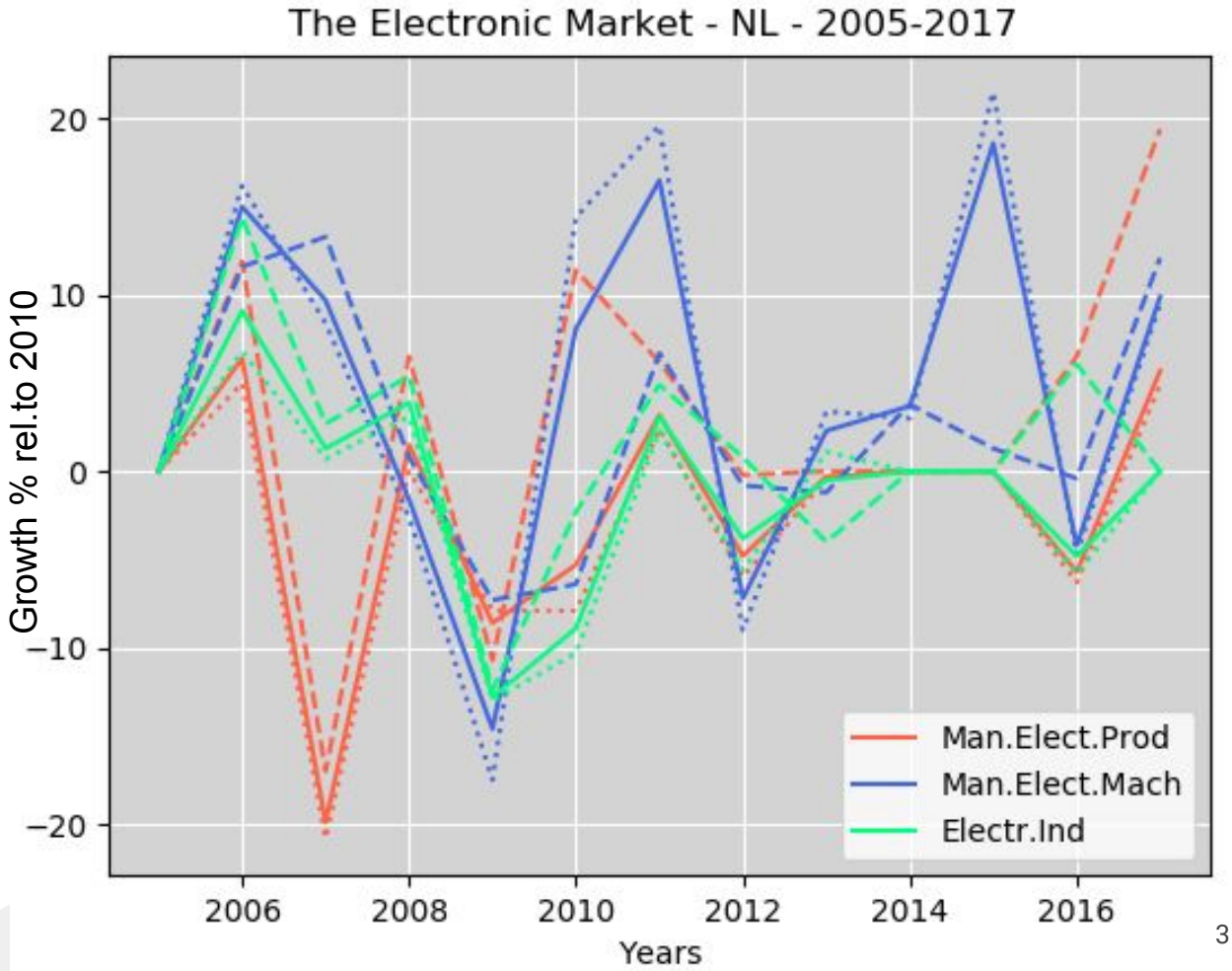
Due to the very design of ML algorithms, they tend to be "**data hungry**" i.e need huge amounts of data to be able to "**train**" the models and make predictions.

## Issues

- **How to protect human generated data from being learnt?**
- **Bringing accountability to the field of ML/DL**
- **Exploring the vulnerabilities in ML, digitally and ethically**

# The Advent of Machine Learning (3a)

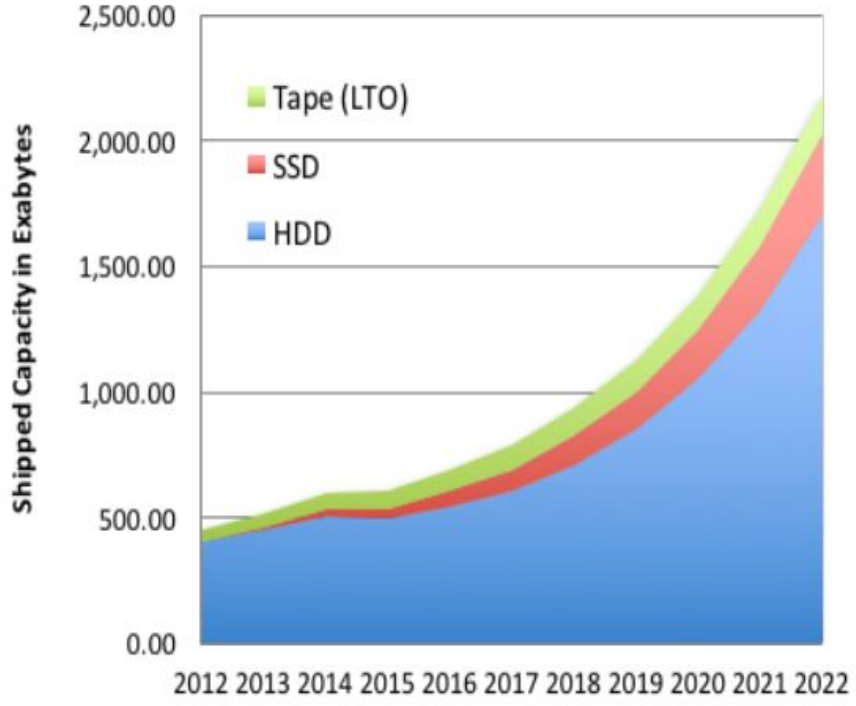
Data from :
https://www.cbs.nl/



# The Advent of Machine Learning (3b)

Electronic memory Projections from IEEE-17

Figure 12. History and projections for annual shipped capacity of magnetic tape, SSDs (with NAND flash), and HDDs. (Coughlin Associates, 2017.)



# Humans are generating more data than ever before!

Given this and the advent of ML, it is inevitable that this data is being used to “train” ML algorithms

# Solution

An initiative to protect human generated data from usage in ML

- Non-Profit, transparent and open source
- Cutting edge research
- Improve upon already established vulnerabilities
- Create awareness and conscious usage
- Pioneer the next revolution in ML/DL

# Implementation

# The Antai Project (1)

Stage 1

**The Antai App**

Make an app/Web service that monitors individual data usage and generation of clients.

Make suggestions regarding vulnerabilities

Stage 2

**Encryption services**

Start encrypting client data against use in ML based systems.

Clients have direct control of their data encryption and usage

Stage 3

**Integration**

Try to bring all the data that the user generates over various digital devices into a central monitoring system, for ease of use

# The Economics

- Build affordable subscription based data protection services
- Generate revenue by offering consultation services in the field of ML
- Grants and tie-ups from Industry, Academia and Sponsors
- Crowdsourcing and campaigning

# The Numbers - Initial projections

- 93.7% of the population of NL use the internet.
- Initial Target group: 0.000589% (~100 clients)
- With an individual subscription of ~10 EUR per month
- An average server cost (AWS*) ~ 0.004 EUR per GB/month
- Initial investment of ~3000 EUR until beta deployment (time frame: 3 months)

http://www.internetlivestats.com/internet-users/netherlands/

*Amazon Web Services

# The Team

- Initial core group ~ 3 people majoring in technical fields (Physics, Mathematics, AI)
- Expansion after beta using regular Hackathon events, with tie ups to the university of Amsterdam and other educational institutions.

# What does the project lack?

- Initial investment(s)
- Economics based consultations
- Law and legal advice
- Expansion and scale up strategies
- Industry based advice from experts

# Thank you

Questions? comments?

Credits to : Rico van Midde for his valuable contributions

# Implementation strategies

- Established:
  - “Adversarial attack” (https://arxiv.org/abs/1712.09665)
  - Data dilution
  - Recursive memory attacks
- Novel:
  - Architecture infiltration
  - Forced biasing - Machine unLearning (MuL) techniques
  - Many more...

# Implementation strategies

- On the app side, every (android) app has a content provider class, which talks to a RDBMS server for the particular app instance.  This can be monitored (source: https://developer.android.com/guide/components/fundamentals)
- On the web side of the project, all POST type requests can be monitored and checked.
- There are also strategies involving standalone software extensions that could be downloaded from the TAP site, through which image(s) and or text(s) could be machined out to be ML resistant.

# Implementation strategies

- Two distinct problems exist:
  - How to make a file/image restant to ML? (throwing off predictions on pre-trained ML setups)
    - Widgets attached to browsers
    - Adversarial attack patches
  - How to make a huge folder resistant ? (making it unfit for training ML)
    - Stand alone (offline) software solutions
    - Forced biasing, Architecture infiltration
- Both are addressed by TAP