Data Science – Deep Learning (TF) & AI Ethics

Task 1

Practical: Artificial Intelligence (AI)

In Data Science we process a lot data through AI. With the GDPR, it is becoming increasingly important to understand the ethics behind the data that is collected, stored, processed and evaluated.

Your task is to:

- Find out what Responsible AI is?
- Find instances where AI has failed? Or been used maliciously or incorrectly.
- Implications of when AI fails. There is a specific article in the GDPR Law that covers this, especially with automated decision making. (opt in and out options).
- What should organisations do to ensure that they are being responsible with AI and the wider use of data in general?
- Maximum 500 words.

Artificial Intelligence collects, stores, processes immense amounts of data. With so much data being processed General Data Protection Regulations (GDPR) laws are gradually becoming more important.

Responsible AI is a governance framework that documents how a specific organization is addressing the challenges around artificial intelligence (AI) from both an ethical and legal point of view. Resolving ambiguity for where responsibility lies if something goes wrong is an important driver for responsible AI initiatives. The development of fair, trustworthy AI standards is up to the discretion of the data scientists and software developers who write and deploy a specific organization's AI algorithmic models. This means that the steps required to prevent discrimination and ensure transparency vary from company to company.

An important goal of responsible AI is to reduce the risk that a minor change in an input's weight will drastically change the output of a machine learning model. Within the context of conforming to the four tenets of corporate governance, responsible AI should be: each step of the model development process should be recorded in a way that cannot be altered by humans or other programming. The data used to train machine models should not be biased. The analytic models that support an AI initiative can be adapted to changing environments without introducing bias. The organisation deploying AI programming is sensitive to its potential impact, both positive and negative.

A famous case where artificial intelligence failed miserably was in 2017 where Apple's Face ID Defeated by a 3D Mask. Apple released the iPhone X with generally positive reviews. The phone's shiniest new feature was Face ID a facial recognition system that replaced the fingerprint reader as your primary passcode. Apple said that the Face ID used on iPhone X's advanced front-facing camera and machine learning to create a 3-dimensional map of your face. The machine learning/AI component helped the system adapt to cosmetic changes (such as putting on make-up, donning a pair of glasses, or wrapping a scarf around your neck), without compromising on security. But a week after the iPhone X's launch, hackers were already claiming to beat Face ID using 3D printing masks. Vietnam-based security firm Bkav

found that they could successfully unlock a Face ID-equipped iPhone by gluing 2D “eyes” to a 3D mask. Researchers in Vietnam claim to have bypassed Apple's Face ID facial recognition technology with a mask that cost less than $150 to make, but many questions remain about just how they achieved their hack.

The GDPR imposes legal requirements on whoever uses the AI system for profiling and/or automated decision-making purposes, even if they acquired the system from a third party. Various implications of artificial intelligence are that systems based on supervised learning could be trained on past individual reasoning and consequently replicate the strengths and weaknesses of the individuals who made these decisions, as well as their inclinations of inaccuracy and unfairness.

There are three ways in which organisations can be more responsible with artificial intelligence. By establishing internal governance, for example by an objective review panel, that is diverse and that has the knowledge to understand the possible consequences of AI infused systems. A key success factor is leadership support and the power to hold leadership accountable. Ensuring the right technical guardrails, creating quality assurance and governance to create traceability and auditability for AI systems. This is an important part of every organisation’s toolkit to allow operational and responsible AI to scale. Investing more in their own AI education and training so that all stakeholders, both internal and external, are informed of AI capabilities as well as the pitfalls.