**Week 8 homework**

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?**

Responsible AI is about developing , deploying and using AI in an ethical and legal manner.

Therefore, AI needs to be :

* Fair thereby having no intended or unintended bias.
* Trustworthy with users confident of its use.
* Reliable and safe thereby not causing intended or unintended harm.
* And be transparent ,thereby being explainable so that employees can question the ability and suitability of the AI.
* Furthermore, it needs to have privacy and security .

**Find instances where AI has failed? Or been used maliciously or incorrectly.**

There have been many failures of AI with some outlined below :

* A Microsoft chatbot operated using AI had declared that “Hitler was correct to hate the Jews”.
* At Jupiter Hospital in Florida an AI system advised to give cancer patients harmful remedies.
* When Amazon relied on bias modelling for an AI recruitment tool it recommended the hiring of only white males.
* AI facial recognition systems have failed at recognising black people possibly due to the training data being skewed.
* AI used as an automated investment tool resulted in losses of up to $20 million a day.
* A voice recognition system was maliciously used to impersonate a person and request monies from another.

**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).**

AI will fail as the quality of the models reduce overtime and besides AI essentially makes predictions which will never be 100% accurate always (e.g. errors by false positives and true negatives).

**What should organisations do to ensure that they are being responsible with AI and the wider use of data in general?**

In order to be responsible with AI and data organisations should ;

* have policies and strategies and governance framework in place.
* set fairness goals which are monitored and controlled . This would require having processes for validation, monitoring and verification.
* in addition organisations should test for bias and fairness and be prepared to remedy unintended consequences.
* ensure that AI is transparent and explainable so that it can be questioned and challenged.
* ensure that AI models are ethical and legal and that data and its use follow all laws.

Maximum 500 words.