**HOME LEARNING TASK WEEK 7**

1. **Find out what Responsible AI is?**

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.

Now that software programs with artificial intelligence (AI) features are becoming more common, it is increasingly apparent that there is a need for standards in AI.

AI and the machine learning model that support it should be comprehensive, explainable, ethical, and efficient.

* **Comprehensiveness** – comprehensive AI has clearly defined testing and governance criteria to prevent machine learning from being hacked easily.
* **Explainable** AI is programmed to describe its purpose, rationale and decision-making process in a way that can be understood by the average end user.
* **Ethical** AI initiatives have processes in place to seek out and eliminate bias in machine learning models.
* **Efficient** AI can run continually and respond quickly to changes in the operational environment.

The data used to train machine models should not be biased if not the result will be too.

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

An incorrect use of current intelligence is that of Deepfakes. Deepfakes are short videos that use an AI algorithm to make one person pass for another, using images. There have been some famous cases like that of Michelle Obama, where a private video of her was virilized so well done through this intelligence that it was difficult to know if it was really her or not.

1. **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).**

The entry into force of the GDPR coincides with the more widespread adoption of artificial intelligence as the technology is integrated into more and more business applications. The GDPR is designed to protect the privacy of EU citizens and give them more control over their personal data. Its objective is to establish a new relationship between the user and the system, one where transparency and a privacy standard are not negotiable. Artificial Intelligence (AI) is a set of technologies or systems that allow computers to perform tasks that involve a simulation of human intelligence, including decision-making or learning. To do so, the technology or the system collects large amounts of data (called Big Data) and, namely, personal data as we indicated above, the GDPR regulates AI to the same extent as any other tool / technology used to process personal data. However, there are provisions of the GDPR that are of relevance to AI systems. These include:

* Requirement that processing be fair.
* Principle of data minimization
* Data protection impact assessments

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

Ensuring that corporate values ​​can more easily be applied to decisions in the AI ​​development process is critical to developing AI responsibly, but it is not enough. In many cases, well-meaning data scientists and talented teams in the field of data science unwittingly ventured into swampy ground and their organizations were caught in a whirlwind of negative press. Added to this, advances in techniques and the increasing use of AI complicate this situation, since they continually modify the framework of action of scientists. So, CEOs need to dig deeper, urging their analytics teams to evaluate their actions in five key areas in the face of public opinion.

* Appropriate data acquisition
* Relevance of the data
* Fairness of artificial intelligence results
* Regulatory compliance and commitment
* Explanation of the model