BWT WEEK 1

TASK 1
SHERAZ BIN TAHIR

1st Sub-Topic

DATA

From this, I have learned that data is an integral part of our daily life. In terms of computing and the Internet, we can include text, numbers, images, and videos. We know computer operates data so it is a need to understand how computers store and process data.

DATA SCIENCE

Data Science is a field that focuses on extracting knowledge and insights from data using scientific methods. understanding data, uncovering hidden relationships, and applying insights to real-world problems are included in data science.

We have 3 types of data here in this documentation, Structured data, Unstructured data & semi-structured data. We can get data from different resources respectively. After getting data we will store data into relational databases, NoSQL, and Data lack. Then we will do data processing, data visualization, and at the end using it for training the model.

Data Science is also supported by related disciplines such as databases, big data, machine learning, artificial intelligence, and data visualization. In the context of digital transformation, businesses leverage data to improve operations and make informed decisions.

2nd Sub-Topic

DATA ETHICS

In it first I learned about data ethics which tells us about moral problems related to data, algorithms, and corresponding practices. Besides this Applied ethics (actively investigating ethical issues), and ethics Culture (adoption of ethical principles and practices) are also well explained in this sub-topic.

There are six core principles of data ethics which are transparency, accountability, fairness, reliability & safety, privacy & security, and inclusiveness.

Along with principles, there are also challenges that can arise when collecting data and designing algorithms which include Data ownership, Informed consent, data privacy, data Quality, algorithm fairness, dataset bias and some others.

There are a few ways to apply ethical principles in your project listed in this subtopic which includes professional code, ethics checklist, ethics regulation, and ethics culture.

3rd Sub-Topic

Defining Data

In the first subtopic, there was a definition of data and its 3 types, in this subtopic, there is more explanation about data, data types & sources of data.
let's have a look at how to describe data.
Raw Data - data that has come from its source in its initial state and has not been analyzed or organized.

Quantitative Data - numerical observations within a dataset and can typically be analyzed, measured, and used mathematically.
Qualitative Data - data that cannot be measured objectively like observations of quantitative data. Generally, it has various formats.

Structured Data - data that is organized into rows and columns, where each row will have the same set of columns.
Unstructured Data – No rows or columns and doesn't contain a format or set of rules to follow.

Semi Structure Data - a combination of structured and unstructured data.

Moreover, in this subtopic, there is more explanation of sources of data Common sources are databases, files, internet resources, APIs, and web scraping.
Understanding these concepts helps in identifying, classifying, and effectively working with various datasets, crucial for informed decision-making and data analysis.