Day 9 – 3 July 2024

Training Day 9 Report

Date: 3 July 2024
Topic: Importing and Analyzing Data using Pandas

Summary:
On Day 9, we focused on **importing and exploring real datasets** using Pandas. The trainer demonstrated importing data from **CSV, Excel, and JSON files** using `read_csv()`, `read_excel()`, and `read_json()`.

Once the data was loaded, we learned to **explore the dataset** using functions such as `info()` to check column data types, `describe()` to get statistical summaries, `value_counts()` to count occurrences of categorical variables, and `unique()` to list unique values. We also practiced **filtering rows** using conditional statements and **sorting data** based on column values.

The session emphasized the importance of understanding the dataset's structure, identifying potential issues, and preparing data for analysis. We practiced exploring student performance datasets, identifying patterns, and calculating basic statistics for insights.

Key Learnings:

- Imported datasets from CSV, Excel, and JSON formats.

- Explored datasets using `info()`, `describe()`, `value_counts()`, and `unique()`.

- Filtered and sorted data for analysis.

- Identified patterns and anomalies in datasets.

- Prepared data for visualization and Machine Learning tasks.

Conclusion:
Day 9 enhanced my skills in data exploration. I now feel confident importing, analyzing, and preparing real datasets for further Machine Learning processes.