

Day 7 – 1 July 2024

Training Day 7 Report

Date: 1 July 2024

Topic: Introduction to Pandas and DataFrames

Summary:

On the seventh day, we started learning **Pandas**, a powerful Python library for **data analysis and manipulation**. The trainer explained that Pandas is widely used in Machine Learning projects to handle tabular data efficiently.

We were introduced to **Pandas Series**, which is a one-dimensional labeled array, and **DataFrames**, which are two-dimensional labeled structures resembling spreadsheets. We learned how to create Series and DataFrames from Python lists, dictionaries, and NumPy arrays.

Next, we explored **creating DataFrames from CSV files**, which is a common way to import real-world datasets. The trainer demonstrated how to use `read_csv()` and how to inspect data using `head()`, `tail()`, and `info()` functions. We practiced selecting columns, filtering rows using conditions, and accessing specific data using `loc` and `iloc`.

The session also covered **basic operations** like adding new columns, performing arithmetic operations on columns, and summarizing data using `describe()`. We practiced working on small datasets like student scores and sales records to make learning practical and relatable.

Key Learnings:

- Understood Pandas Series and DataFrames.
- Created DataFrames from lists, dictionaries, and CSV files.
- Learned to inspect, filter, and select data in DataFrames.
- Applied arithmetic operations and summarized data.
- Practiced real-world dataset handling with Pandas.

Conclusion:

Day 7 gave me a strong foundation in Pandas. I now understand how to work with structured data efficiently, which is essential for any Machine Learning project. I feel confident creating, viewing, and manipulating DataFrames.