

Task 3. Matplotlib

1. What is Matplotlib and why is it used?

Matplotlib is a Python data visualization library used to create static, animated, and interactive plots. It is mainly used to represent data graphically so that patterns, trends, and relationships in data can be understood easily. Matplotlib helps in analyzing data, comparing values, and presenting results clearly through graphs like line plots, bar charts, and scatter plots.

2. Difference between line plot, bar plot, and scatter plot

A **line plot** is used to show trends or changes over time by connecting data points with lines. It is commonly used for time-series data.

A **bar plot** is used to compare values across different categories using rectangular bars. It is useful for showing differences between discrete groups.

A **scatter plot** is used to show the relationship between two numerical variables using individual points. It helps in identifying correlations, clusters, or outliers in data.

3. What are figure and axes in Matplotlib?

A **figure** is the overall window or canvas that holds the entire plot. **Axes** are the actual plotting areas inside the figure where data is drawn. A figure can contain one or multiple axes, allowing multiple plots in the same figure.

4. How do labels, titles, and legends improve data visualization?

- **Labels** describe what each axis represents
- **Titles** explain what the graph is about
- **Legends** identify different data sets in the same plot

They help viewers interpret the data correctly without confusion.

5. Why is data visualization important before building AI models?

Data visualization is important before building AI models because it helps in understanding the data. It allows us to identify patterns, trends, missing values, outliers, and relationships between variables. Visual analysis ensures better data preprocessing, feature selection, and model performance, leading to more accurate and reliable AI models.