Experiment 1: Introduction to Python for Machine Learning

Title:

Introduction to Python Libraries for Machine Learning (NumPy, pandas, matplotlib)

Aim:

To become familiar with the Python environment and learn how to use essential libraries for data manipulation and visualization.

Objectives:

- Learn to use **NumPy** for numerical operations.
- Use **pandas** for handling tabular data.
- Use **matplotlib** for basic data visualization.

Theory:

Machine Learning requires powerful tools for numerical computation and data handling.

- **NumPy (Numerical Python):** Provides support for large, multidimensional arrays and mathematical functions.
- **pandas:** Provides data structures like *Series* and *DataFrame* to manipulate structured data.
- **matplotlib:** A plotting library for visualizing data (line plots, bar graphs, histograms, etc.).

Algorithm / Steps:

1. Import the required libraries.
2. Create and manipulate NumPy arrays.
3. Create pandas DataFrame and perform basic operations.
4. Visualize data using matplotlib.
5. Observe and interpret the outputs.

Sample Python Code:

```
# Experiment 1: Introduction to Python Libraries

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# 1. NumPy array operations
arr = np.array([10, 20, 30, 40, 50])
print("Array:", arr)
print("Mean:", np.mean(arr))
print("Standard Deviation:", np.std(arr))

# 2. pandas DataFrame
data = {'Name': ['Alice', 'Bob', 'Charlie', 'David'],
        'Marks': [85, 90, 78, 92]}
df = pd.DataFrame(data)
print("\nDataFrame:\n", df)
print("\nAverage Marks:", df['Marks'].mean())

# 3. matplotlib Visualization
plt.bar(df['Name'], df['Marks'], color='skyblue')
plt.title("Student Marks")
plt.xlabel("Name")
plt.ylabel("Marks")
plt.show()
```

Expected Output:

- Display NumPy array, mean, and standard deviation.
- Show pandas DataFrame with average marks.
- Display a bar chart of student marks.

Result:

The experiment successfully demonstrated the use of NumPy, pandas, and matplotlib for basic data manipulation and visualization.

Viva Questions:

1. What is NumPy used for in Python?
2. What are pandas DataFrames?
3. How do you install libraries in Python?
4. What function is used to plot a bar chart in matplotlib?

5. What are the advantages of using pandas over lists?