PROJECT: Skill Match Resume Matcher and Skill Recommender

MILESTONE 1 PRESENTATION

-By Omkar Yadav

ARTIFICIAL INTELLIGENCE:

Artificial Intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems. It enables computers to perform tasks that usually require human intelligence such as reasoning, learning, problem solving, perception, and understanding language.

Applications:

- Speech Recognition (e.g., Siri, Alexa)
- Self-Driving Cars
- Chatbots and Virtual Assistants
- Recommendation Systems (e.g., Netflix, Amazon)

MACHINE LEARNING:

Machine Learning is a subset of AI that enables computers to learn and improve from experience without being explicitly programmed. It uses algorithms to find patterns in data and make predictions or decisions based on them.

Types of Machine Learning:

1. **Supervised Learning:**
   The model learns from labelled data.
   Example: Predicting house prices using past data.
2. **Unsupervised Learning:**
   The model finds patterns in unlabelled data.
   Example: Customer segmentation or clustering.
3. **Reinforcement Learning:**
   The model learns by interacting with an environment and receiving feedback in the form of rewards.
   Example: Training robots or game AI.

NUMPY (NUMERICAL PYTHON):

NumPy is a powerful Python library used for numerical and scientific computing. It provides high-performance multidimensional arrays and tools to work efficiently with them.

Key Features:

- Fast mathematical computations on large arrays
- Statistical and linear algebra functions
- Random number generation
- Supports vectorized operations (no loops needed)

[14]
✓ 0s

```
a = np.arange(6).reshape(2,3); print(a)
```

```
[[0 1 2]
 [3 4 5]]
```

[15]
✓ 0s

```
print(np.mean([1,2,3]))
print(np.median([1,3,2]))
print(np.std([1,2,3]))
```

```
2.0
2.0
0.816496580927726
```

[16]
✓ 0s

```
A=np.array([[1,2],[3,4]])
B=np.array([[5,6],[7,8]])
print(np.dot(A,B))
```

```
[[19 22]
 [43 50]]
```

[17]
✓ 0s

```
A=np.array([[1,2,3],[4,5,6]])
print(A.T)
```

```
[[1 4]
 [2 5]
 [3 6]]
```

PANDAS:

Pandas is a Python library for data manipulation and analysis. It provides two key data structures:

- **Series** (1D labeled array)

```
import pandas as pd
s=pd.Series([10,20,30])
print(s)
```

```
0    10
1    20
2    30
dtype: int64
```

- **DataFrame** (2D table-like data structure)

```
df=pd.DataFrame({'A':[1,2],'B':[3,4]})
print(df)
```

```
   A  B
0  1  3
1  2  4
```

PANDAS

```
import pandas as pd
s=pd.Series([10,20,30])
print(s)
```

```
0    10
1    20
2    30
dtype: int64
```

```
df=pd.DataFrame({'A':[1,2],'B':[3,4]})
print(df)
```

```
   A  B
0  1  3
1  2  4
```

```
from sklearn.datasets import load_iris
iris = load_iris()
df = pd.DataFrame(data=iris.data, columns=iris.feature_names)
print(df.head())
```

```
   sepal length (cm)  sepal width (cm)  petal length (cm)  petal width (cm)
0                5.1               3.5                1.4               0.2
1                4.9               3.0                1.4               0.2
2                4.7               3.2                1.3               0.2
3                4.6               3.1                1.5               0.2
```

[22]
✓ 0s

```
print(df.describe())
```

```
       sepal length (cm)  sepal width (cm)  petal length (cm)  \
count         150.000000        150.000000         150.000000
mean            5.843333          3.057333           3.758000
std             0.828066          0.435866           1.765298
min             4.300000          2.000000           1.000000
25%             5.100000          2.800000           1.600000
50%             5.800000          3.000000           4.350000
75%             6.400000          3.300000           5.100000
max             7.900000          4.400000           6.900000

       petal width (cm)
count        150.000000
mean           1.199333
std            0.762238
min            0.100000
25%            0.300000
50%            1.300000
75%            1.800000
max            2.500000
```

[23]
✓ 0s

```
print(df.tail())
```

```
     sepal length (cm)  sepal width (cm)  petal length (cm)  petal width (cm)
145                6.7               3.0                5.2               2.3
146                6.3               2.5                5.0               1.9
147                6.5               3.0                5.2               2.0
148                6.2               3.4                5.4               2.3
149                5.9               3.0                5.1               1.8
```

It is widely used in data science for cleaning, transforming, and analysing datasets.

Key Features:

- Easy data cleaning and transformation
- Handle missing data effectively
- Import/export data from CSV, Excel, SQL, JSON, etc.
- Grouping and summarizing data

MATPLOTLIB:

Matplotlib is a Python library used for creating static, animated, and interactive data visualizations. It's one of the most popular tools for plotting graphs and charts.

Common Plot Types:

- Line plot

  A line plot is used to display data points connected by straight lines. It shows trends or changes in data over time.
- Bar chart

  A bar chart represents categorical data with rectangular bars where the height (or length) of each bar shows the value.
- Scatter plot

  A scatter plot shows the relationship between two numerical variables using dots on a graph.
- Histogram

  A histogram displays the distribution of numerical data by dividing it into bins (ranges) and showing how many values fall into each bin.

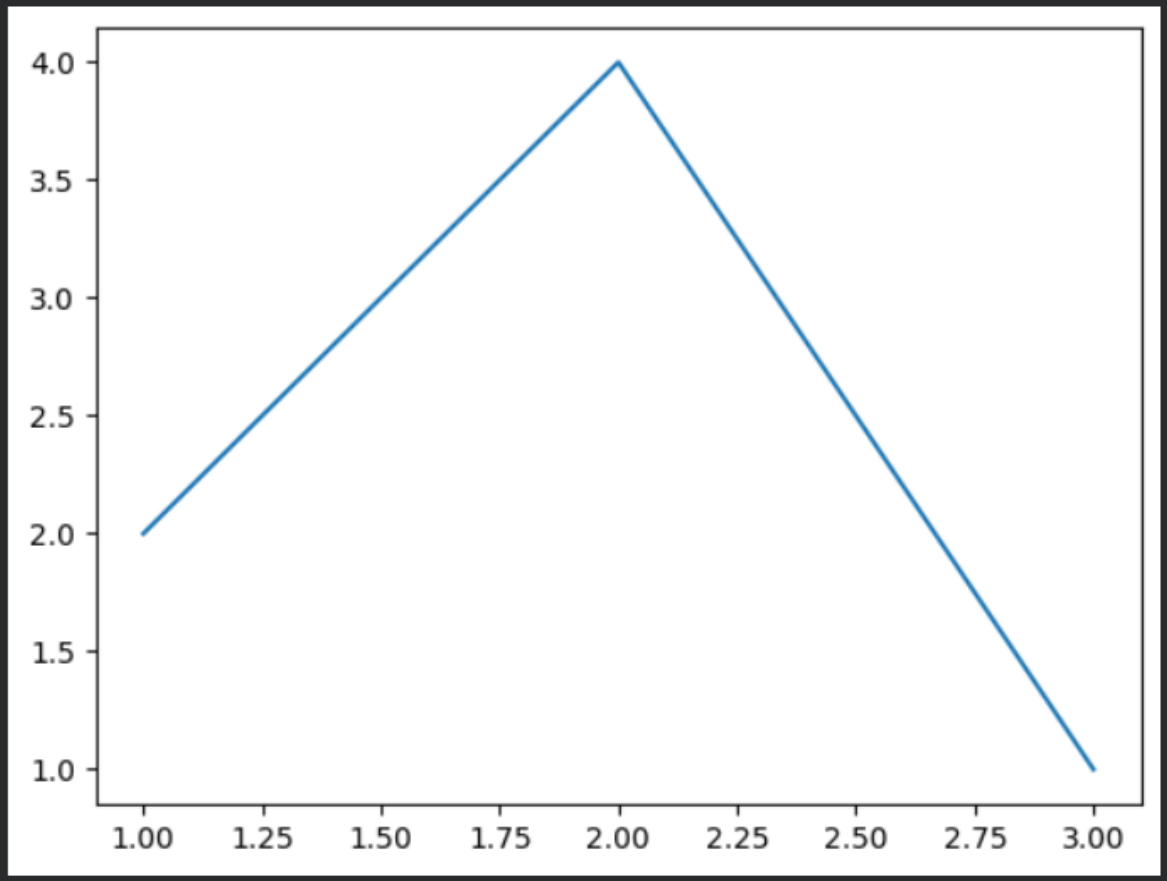
- Pie chart

  A pie chart represents data as slices of a circle, where each slice shows the proportion of a whole.

MATPLOTLIB

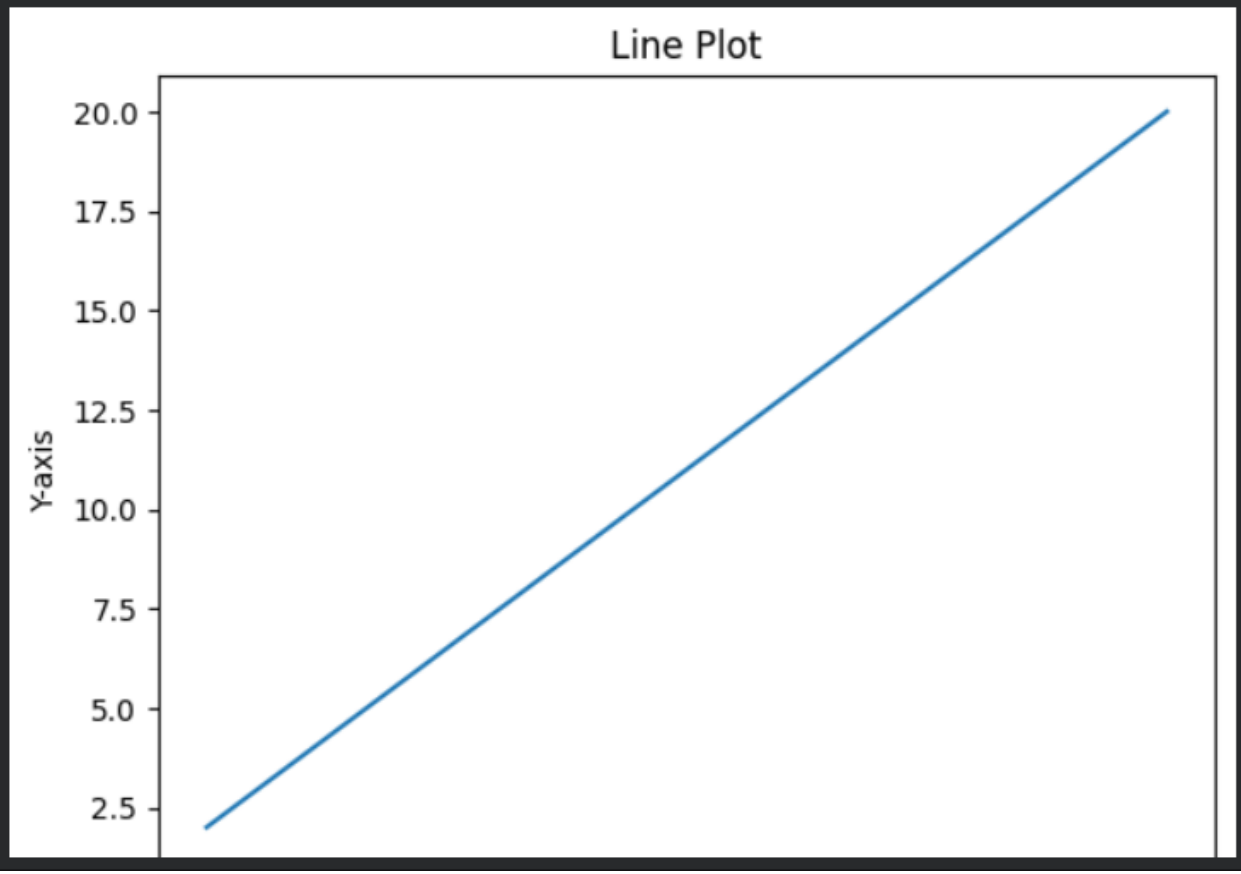
[26]
✓ 0s

```
import matplotlib.pyplot as plt
plt.plot([1,2,3],[2,4,1])
plt.show()
```



[27]
✓ 0s

```
x=np.arange(1,11)
y=2*x
plt.plot(x,y)
plt.title("Line Plot")
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.show()
```



SEABORN:

Seaborn is a statistical data visualization library built on top of Matplotlib. It provides a high-level interface to create attractive and informative graphs with minimal code.

Key Features:

- Built-in themes and color palettes
- Works seamlessly with Pandas DataFrames
- Visualizes complex statistical relationships • Supports heatmaps, violin plots, pair plots, etc.

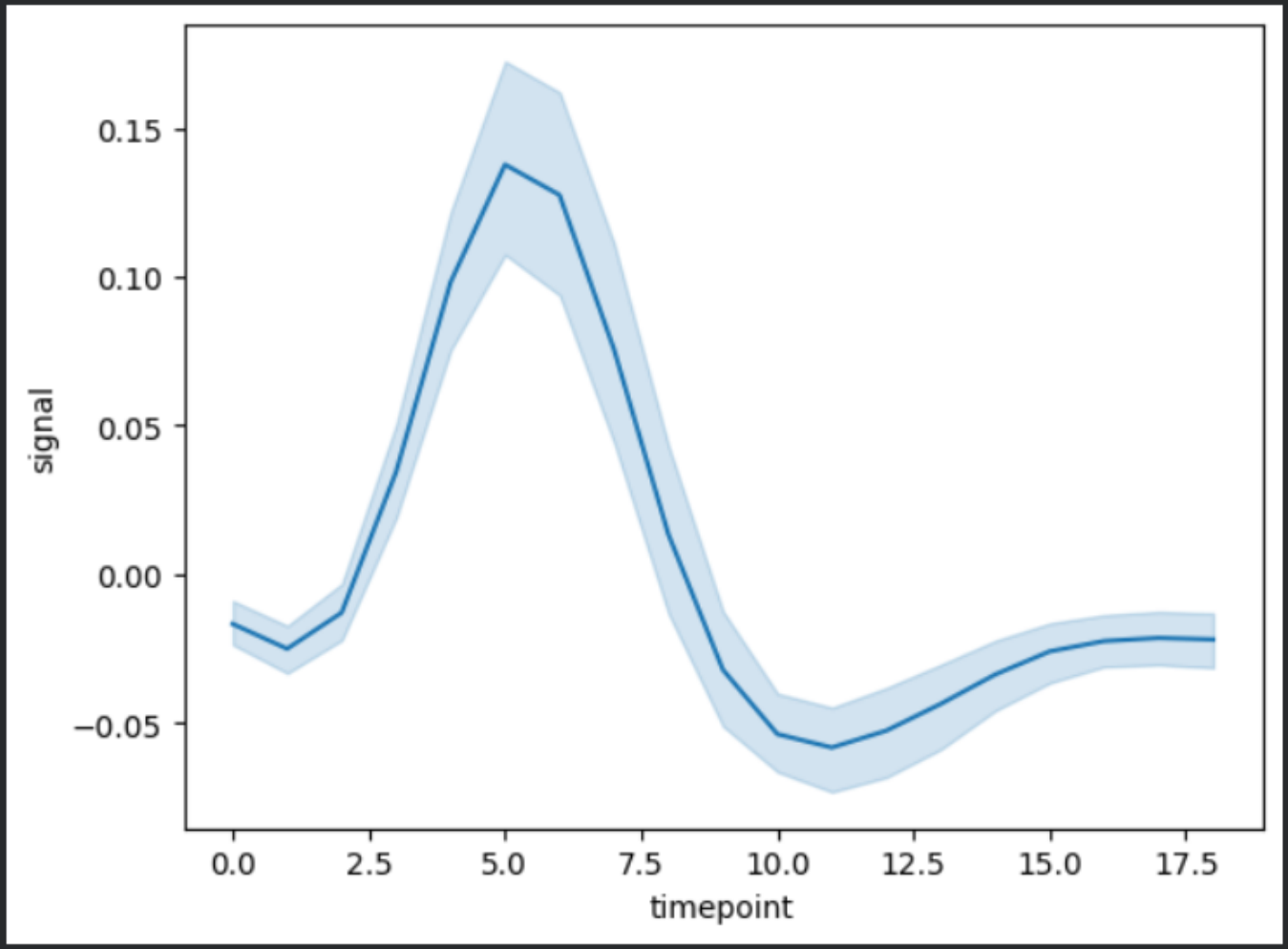
SEABORN

```
import seaborn as sns
from matplotlib import pyplot as plt
fmri = sns.load_dataset('fmri')
fmri.head()
```

	subject	timepoint	event	region	signal
0	s13	18	stim	parietal	-0.017552
1	s5	14	stim	parietal	-0.080883
2	s12	18	stim	parietal	-0.081033
3	s11	18	stim	parietal	-0.046134
4	s10	18	stim	parietal	-0.037970

[40]
✓ 0s

```
sns.lineplot(x='timepoint',y='signal',data=fmri)
plt.show()
```



CONCLUSION:

Artificial Intelligence and Machine Learning empower computers to think and learn like humans.
Libraries like **NumPy**, **Pandas**, **Matplotlib**, and **Seaborn** provide the backbone for data analysis, transformation, and visualization making AI and ML development faster, more accurate, and efficient.