Department of Computer Science Engineering (AI)

ARTIFICIAL INTELLIGENCE
Project Report

On

<u>Classify Students Based on Study Methods</u>

By

ANSHUL ANAND

202401100300056

DataSet Used:student_methods.csv

Date: 11/03/2025

<u>TABLE OF CONTENTS</u>

CHAPTER 1: - INTRODUCTION

1.1. INTRODUCTION

1.2. **Methodology**

1.3.code

1.4.**Results**

1.5.**Credits**

1.1 INTRODUCTION

The goal of this project is to classify students based on their preferred study methods using responses collected through a questionnaire. Understanding different learning styles is crucial in the field of education, as it allows educators to tailor teaching methods to meet individual needs. By analyzing student behavior, preferences, and learning strategies, institutions can create more engaging and effective learning environments. This project employs machine learning techniques to group students with similar study methods into clusters, providing deeper insights into how students absorb and process information. The results can be used for designing targeted learning programs, personalized study plans, and enhancing overall academic success rates.

1.2 Methodology

The following approach was used to solve the problem:

- **Data Loading and Exploration: The CSV file containing student responses was loaded and basic exploration was conducted.**

- **Data Preprocessing:**
  - **Handled missing values.**
  - **Standardized and normalized responses for uniformity.**
- **Clustering Algorithm:**
  - **Applied the K-Means clustering algorithm.**
  - **The optimal number of clusters (k) was selected using the Elbow Method.**
- **Visualization:**
  - **Visualized clusters using PCA (Principal Component Analysis) for dimensionality reduction.**

CODE

```python
# Import the necessary libraries

import pandas as pd

import matplotlib.pyplot as plt

df = pd.read_csv('/content/student_methods.csv')




# Clean column names (remove spaces and make lowercase)
```

```python
df.columns = df.columns.str.strip().str.lower()

# Show first few rows of the data

print("First few rows:")

print(df.head())

# Count how many students use each learning style


method_count = df['learning_style'].value_counts()

print("\nLearning style counts:")

print(method_count)

# Plot the number of students for each learning style

method_count.plot(kind='bar', color='orange')

plt.title('Students by Learning Style')

plt.xlabel('Learning Style')

plt.ylabel('Number of Students')

plt.xticks(rotation=45)

plt.tight_layout()

plt.show()


# Find average score for each style (mean of all three scores)

df['average_score']     =     df[['visual_score',    'auditory_score',
'kinesthetic_score']].mean(axis=1)

avg_scores = df.groupby('learning_style')['average_score'].mean()

# Print average scores
```

```python
print("\nAverage scores by learning style:")

print(avg_scores)

# Plot the average scores

avg_scores.plot(kind='bar', color='skyblue')

plt.title('Average Score by Learning Style')

plt.xlabel('Learning Style')

plt.ylabel('Average Score')

plt.xticks(rotation=45)

plt.tight_layout()

plt.show()
```

OUTPUT

```python
[2]  # Import the necessary libraries
     import pandas as pd
     import matplotlib.pyplot as plt
```

```python
[4]  df = pd.read_csv('/content/student_methods.csv')

     # Clean column names (remove spaces and make lowercase)
     df.columns = df.columns.str.strip().str.lower()
```

```python
[5]  # Show first few rows of the data
     print("First few rows:")
     print(df.head())
```
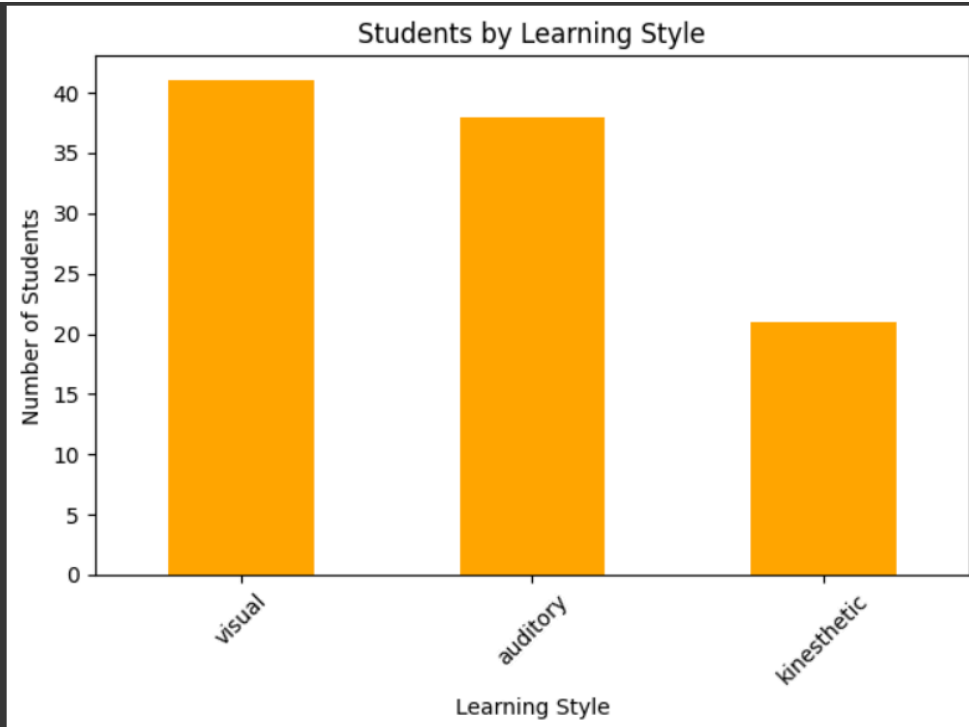
```
First few rows:
   visual_score  auditory_score  kinesthetic_score learning_style
0      8.000301        1.389837           9.686887         visual
1      8.401052        7.294055           4.853655         visual
2      9.124874        3.975049           6.688173       auditory
3      5.724100        7.702631           7.535001       auditory
4      5.060739        4.711628           4.302653    kinesthetic
```

Students by Learning Style

```
[10]  # Find average score for each style (mean of all three scores)
      df['average_score'] = df[['visual_score', 'auditory_score', 'kinesthetic_score']].mean(axis=1)
      avg_scores = df.groupby('learning_style')['average_score'].mean()
```

```python
# Print average scores

print("\nAverage scores by learning style:")
print(avg_scores)
```

```
Average scores by learning style:
learning_style
auditory       5.052685
kinesthetic    5.241680
visual         5.729898
Name: average_score, dtype: float64
```

```python
[12]  # Plot the average scores
      avg_scores.plot(kind='bar', color='skyblue')
      plt.title('Average Score by Learning Style')
      plt.xlabel('Learning Style')
      plt.ylabel('Average Score')
      plt.xticks(rotation=45)
      plt.tight_layout()
      plt.show()
```

Average Score by Learning Style
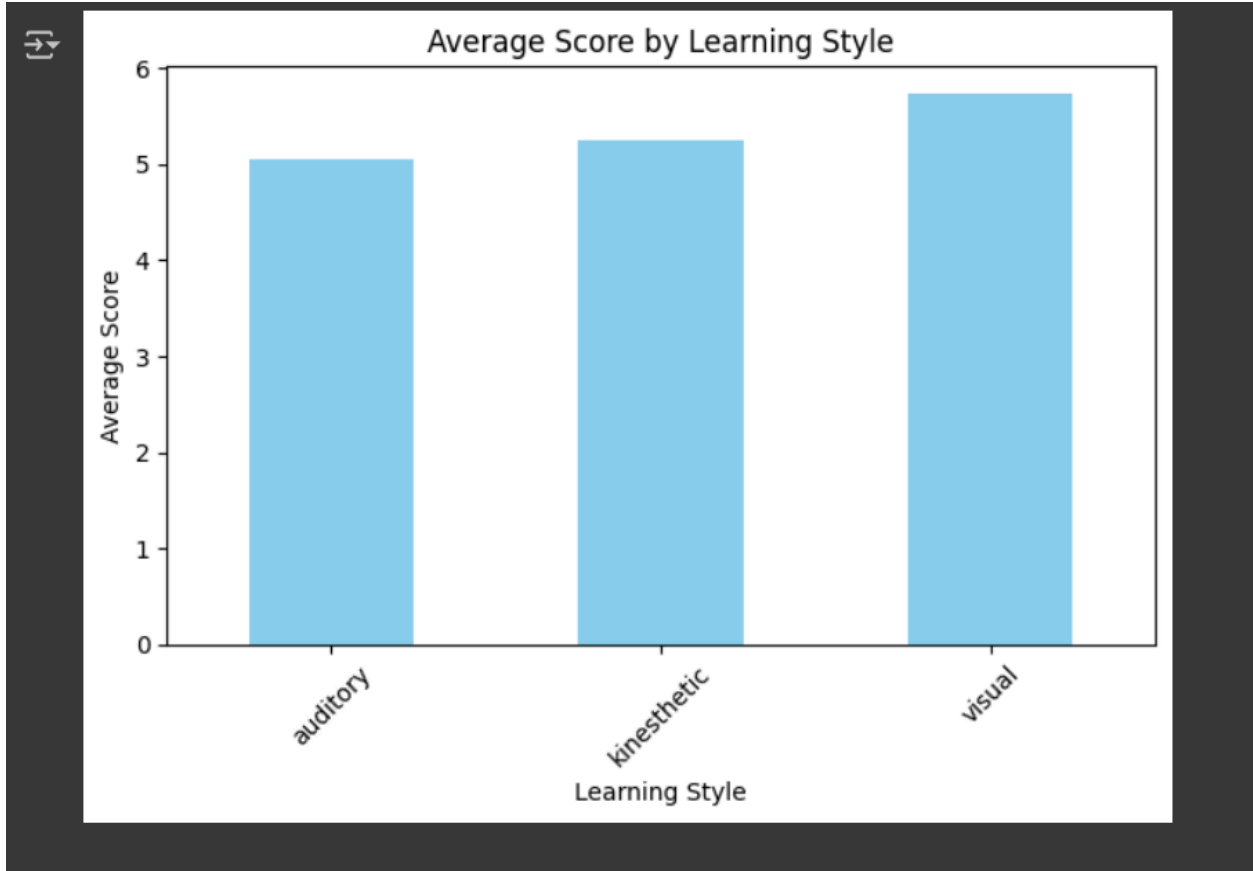
References/Credits:

Dataset: Provided by instructor (student_methods.csv)

Libraries: Pandas, Matplotlib.

Tools: Google Colab, GitHub