National University of Computer and Emerging Sciences

Department of Artificial intelligence & Data science

**

AL-2001

Programming for AI

Project

|  |  |
| --- | --- |
| **Instructor** | **Ms.Isma Waheed** |
| **Semester** | **BA/I-3A-Fall 2024** |

**Instruction:**

1. ***Anyone found in copying the project from any other group, both the groups will get zero marks.***
2. ***2-3 students are allowed per group.***
3. ***You must submit the whole project by one member of the group.***
4. ***There must be a block of comments at start of code, the block should contain brief description about functionality of code.***
5. ***Proper indentation of code is essential***
6. ***Write a python code using Jupyter on windows.***
7. ***First think about statement problems and then write/draw your logic on copy.***
8. ***Project deadline is 15-12-2024.***
9. ***At the end when you done your project, submit complete project.***
10. ***Submit your file in this format FirstGroupMemberRollNumber\_SecondGroupMemberNumber\_Project.zip.***
11. ***Do not copy code from any source otherwise you will be penalized with negative marks.***
12. ***Submit your complete project file with your code copy pasted on word file with all possible output screenshots. Both project file and word file must be included in .zip file. Make your submission on Google Classroom (Make sure your submission is completed and on time, Late Submission is not allowed).***

**Project Problem Statement:**

You are tasked with conducting an analysis of a popular product selling website. The goal is to scrape historical data, perform data manipulation using NumPy and Pandas, conduct various analyses, and present the results. Follow the steps below to accomplish this task. The dataset file attached with this doc.

**Objective:** Analyze sentiment trends for brands/products by scraping data.

**Step 1: Data Cleaning and Structuring:**

Organize the scraped data into a structured format. Create a class, CryptoData, to represent each brand data. The class should include appropriate attributes and methods for data cleaning and structuring.

**Step 2: Handling Missing Values**

Deal with any missing values sentiment scores or incomplete posts using default values in the scraped data.

**Step 3: Data Transformation**

Use NumPy to calculate rolling averages of sentiment/ product reviews. Store the transformed data.

**Step 4: Analysis**

Perform the following analyses on the transformed data:

* Calculate the mean, median, and standard deviation of the returns average sentiment and standard deviation for each brand.
* Identify posts with the most positive and negative sentiment
* Correlate sentiment trends between the brands.

**Step 5: Visualization:**

* Line and bar charts for sentiment trends over time.
* A pie chart showing sentiment distribution (positive, neutral, negative).

**Step 6: Save Results**

Save the cleaned and transformed data along with the analysis results in a CSV file for future reference.

**Step 7: Summarize Findings**

Highlight customer reviews patterns and key factors driving them.

**Step 8: OOP Concepts**

Ensure your implementation adheres to OOP principles. Utilize encapsulation, inheritance, and polymorphism where appropriate. Add methods for analyzing hashtags or keywords.

**Step 9: Documentation**

Lastly, document your code thoroughly, including comments explaining the purpose of each method and any assumptions made during the analysis. Clearly explain each step of sentiment analysis.