10 ACADEMY

**Kifiya AI Mastery Program 2**

Week 1, Task 1

Interim Report

by

Jerusalem Fetene Mekonnin

August 30, 2024

**Introduction**

Python is a simple, clear and intuitive programming language used for many scientific and mathematical, and statistical and financial applications. Owing to this many scientists and data analysts gain valuable insights in the least amount of tine and resource while their statistical data analysis.

Data analysis is simply working with numbers and extracting valuable insights from them using descent skills of mathematics, programming and domain area working on i.e., financial, statistical, descriptive etc.

Financial analysis is the process of evaluating businesses, projects, budgets, and other finance-related transactions to determine their performance and suitability. Typically, financial analysis is

used to analyze whether an entity is stable, solvent, liquid, or profitable enough to warrant a monetary investment. Stock market prediction involves trying to determine the future value of a company stock or other financial instrument traded on an exchange.

**Objective**

The express aim of this report is to analyze the correlations between news sentiment and stock market movements of the financial news data. Here below are some of the objectives of this report:

- To perform descriptive statistics
- To perform sentiment analysis on the 'headline' text to quantify the tone
- To perform time series analysis
- To establish statistical correlations between the sentiment derived from news articles and the corresponding stock price movements
- To perform publisher analysis

**About Dataset**

Financial News and Stock Price Integration Dataset

FNSPID (Financial News and Stock Price Integration Dataset), is a comprehensive financial dataset designed to enhance stock market predictions by combining quantitative and qualitative data.

**Python Libraries**

1.Pandas

Pandas is the open-source python library that is widely used for data analysis and data. Its main purpose is to perform data analysis on the structured data and focuses on the fundamental data processing.

2.NumPy

NumPy package comes with a wide collection of numerical functions that makes it an important library in academia and finance industry.

3. SciPy

Scipy is used for financial computation and other numerical integrations in the finance industry4.
4. Polars

Polars is a blazingly fast DataFrame library for manipulating structured data.

5. Pyfolio

It is an open-source library that provides risk analysis reports and performance results of financial portfolios based on the returns.

6. QuantPy

A framework for quantitative finance in python.

7. Statsmodel

Statsmodel is used for statistical tests and statistical data exploration. Some best models of statsmodel includes linear regression model, discrete model, time series analysis, bayesian analysis.

8. Statistics

This module provides functions for calculating mathematical statistics of numeric

9. Pynance

Pynance will work wonders for a stock market trader. It is an open-source python package that retrieves, analyses and visualizes the data from stock market derivatives.

10. Quandl

This is the biggest and powerful marketplace where lives the financial, economic, and alternative data in modern formats for financial analysts. The Quandl Python library will get your financial data directly into the Python.

### Dataset Features

1) **headline**: Article release headline, the title of the news article, which often includes key financial actions like stocks hitting highs, price target changes, or company earnings.
2) **url**: The direct link to the full news article.
3) **publisher**: Author/creator of article.
4) **date**: The publication date and time, including time zone information (UTC-4 time zone).
5) **stock**: Stock ticker symbol (unique series of letters assigned to a publicly traded company). For example (AAPL: Apple)

**<!these below shall be in detail discussed after the implantation is completed coming week on Tuesday Sept 3, 2024!>**

### Data Preprocessing

### Data Analysis

### Data Visualization

### Discussion

### Conclusion