AI-Powered Stock Market Prediction & Analytics

Project Report

1. Project Overview

This project focuses on ingesting, processing, analyzing, and visualizing real-time stock market data to identify trends and predict stock prices. It integrates **ETL workflows, machine learning models, and interactive dashboards** to provide actionable insights into stock movements.

Key Features:

- Real-time stock data ingestion using APIs
- Data processing and storage in cloud databases (Azure)
- Stock price forecasting using LSTM, ARIMA, and XGBoost models
- Interactive Power BI dashboards with advanced visualizations

---

2. Technology Stack

ETL & Data Integration

- **Tools:** Azure Data Factory
- **Functionality:** Data extraction, transformation, and loading

Database & Cloud Storage

- **Tools:** Microsoft SQL Server, Azure Blob Storage
- **Functionality:** Storing structured and unstructured stock data

Machine Learning & Analytics

- **Models:** LSTM, ARIMA, XGBoost
- **Libraries:** TensorFlow, Scikit-learn, PyTorch
- **Functionality:** Predicting stock prices and trend analysis

Visualization & Reporting

- **Tools:** Power BI, Matplotlib, Seaborn
- **Functionality:** Interactive dashboards and statistical insights

Development & Automation

- **Languages:** Python, SQL
- **Version Control & Automation:** GitHub Actions

---

3. Project Workflow

Step 1: Data Collection & Storage

- Real-time stock data is fetched from APIs (**Alpha Vantage, Quandl**)
- Data includes **Open, High, Low, Close, Volume, Adjusted Close**
- Stored in **Azure Blob Storage and Microsoft SQL Server**

Step 2: Data Processing & ETL Pipelines

- **Tools Used:** Azure Data Factory
- **Processing Tasks:**
  ✔ Fill missing values (Forward-fill stock prices)
  ✔ Normalize stock prices for ML models
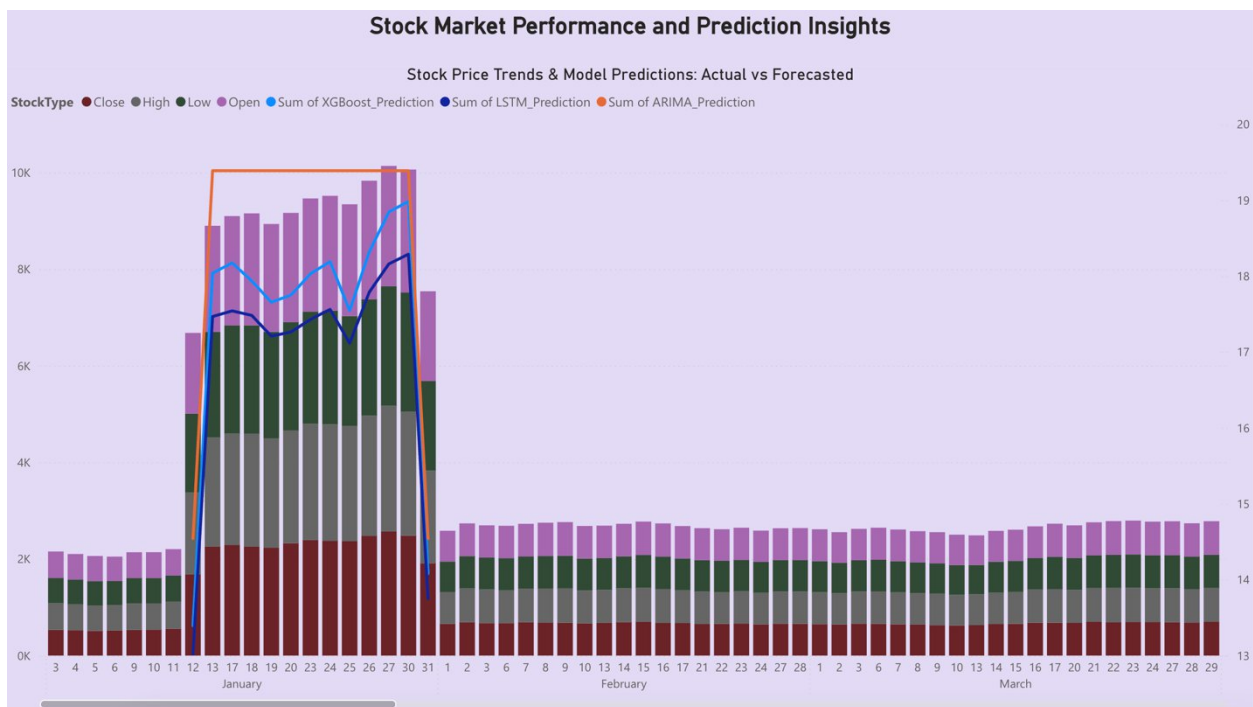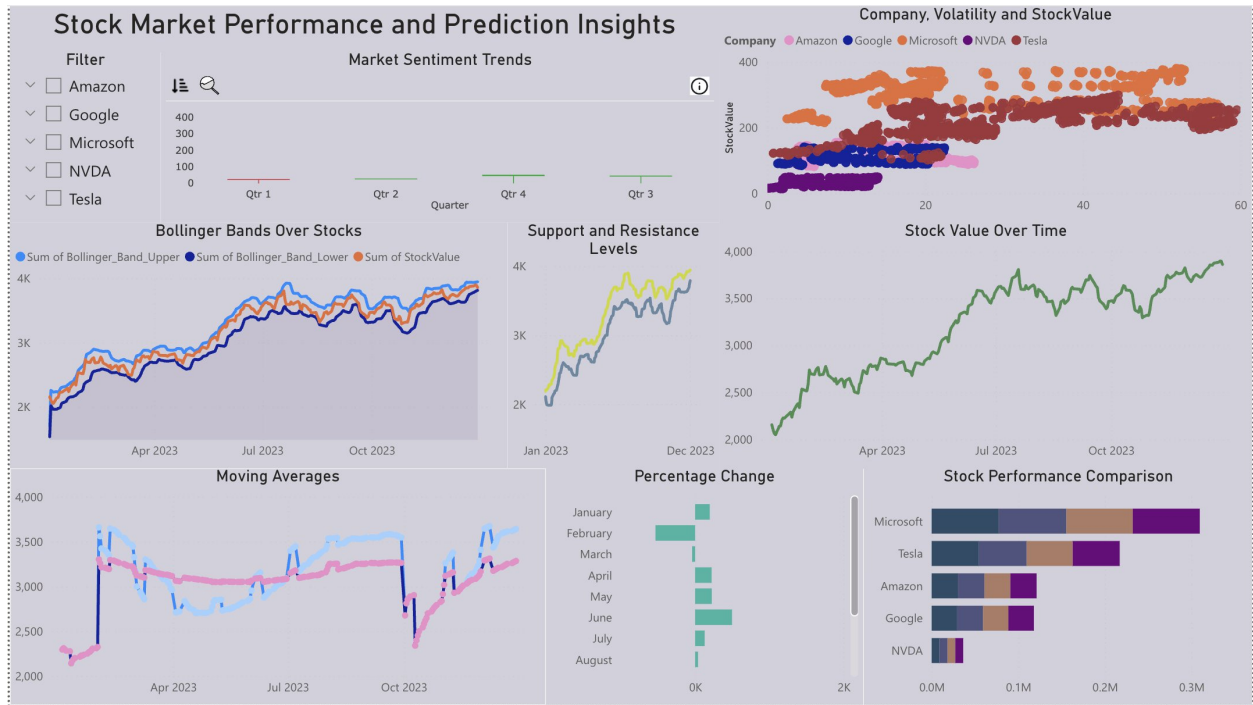 ✔ Convert timestamps for time series modeling

---

Step 3: Machine Learning for Stock Prediction

- **Models Used:**
  - LSTM (Long Short-Term Memory) - Best for time-series data
  - ARIMA (AutoRegressive Integrated Moving Average) - Detects trends & seasonality
  - XGBoost - Uses gradient boosting for stock price regression

- **Preprocessing Steps:**
  a. Convert timestamps into numerical values
  b. Normalize stock prices (Min-Max Scaling)
  c. Split data into **80% Training, 20% Testing**

Step 4: Data Visualization & Insights (Power BI Dashboard)

The following visualizations were implemented in Power BI for **real-time stock analysis**:

Chart Name	Columns Used	Visualization Type	Purpose
Stock Price Over Time	Date, StockValue	Line Chart	Track stock price movements
Moving Averages Comparison	Date, MovingAverage_50_Day, MovingAverage_200_Day	Dual-Line Chart	Identify buy/sell signals
Bollinger Bands Analysis	Date, Bollinger_Band_Upper, Bollinger_Band_Lower, StockValue	Line Chart	Detect volatility periods
Candlestick Chart (OHLC)	Date, Open, High, Low, Close	Custom Candlestick Chart	Visualize stock price trends
Support & Resistance Levels	Date, Support_Level, Resistance_Level	Line Chart	Identify key price zones
Stock Performance Comparison	Company, StockValue	Bar/Column Chart	Compare multiple stocks
Volatility vs. Stock Price	Date, Volatility, StockValue	Scatter Chart	Understand risk levels
Stock Price Percentage Change	Date, PercentageChange	Bar Chart	Track daily market movements
Model Predictions vs. Actual	Date, LSTM_Prediction, ARIMA_Prediction, XGBoost_Prediction, StockValue	Bar & Line Chart	Compare ML model accuracy

Stock Market Performance and Prediction Insights

Market Sentiment Trends

Company, Volatility and StockValue

Bollinger Bands Over Stocks

Support and Resistance Levels

Stock Value Over Time

Moving Averages

Percentage Change

Stock Performance Comparison

Stock Market Performance and Prediction Insights

Stock Price Trends & Model Predictions: Actual vs Forecasted

Interactive Features:
a. **Dynamic Slicers** - Filter data by company, date range, or stock type
b. **Conditional Formatting** - Highlights volatility and trend changes
c. **Tooltips** - Provides instant insights on stock movements

4. Key Takeaways & Future Enhancements

- **Key Insights**
  a. The Power BI dashboard provides **real-time stock insights** with AI-driven predictions
  b. Machine learning models (LSTM, ARIMA, XGBoost) **accurately track stock movements**
  c. Bollinger Bands and RSI **help traders identify buying/selling opportunities**
  d. Automated ETL ensures **data pipelines run seamlessly without manual intervention**

Future Enhancements

a. Improve ML model performance with **Transformer-based Time Series Models** (e.g., Temporal Fusion Transformers)
b. Add **sentiment analysis** using news and social media to influence predictions
c. Implement **high-frequency trading signals** for intraday traders

---

5. Conclusion

This project successfully integrates **real-time stock data, machine learning predictions, and Power BI analytics** to provide a **comprehensive stock market analysis system**. The insights generated can help traders and analysts **make data-driven investment decisions** with confidence.