

StockWatch

ATLS 5214: Big Data
Architecture



Team Members:



Table of contents

01 Project
Overview

02 Tools and
Technologies

03 Architecture
Diagram

04 Demo

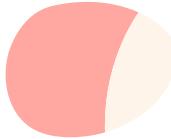
05 Learning
Curve

06 Challenges

07 Future Scope



Project Overview



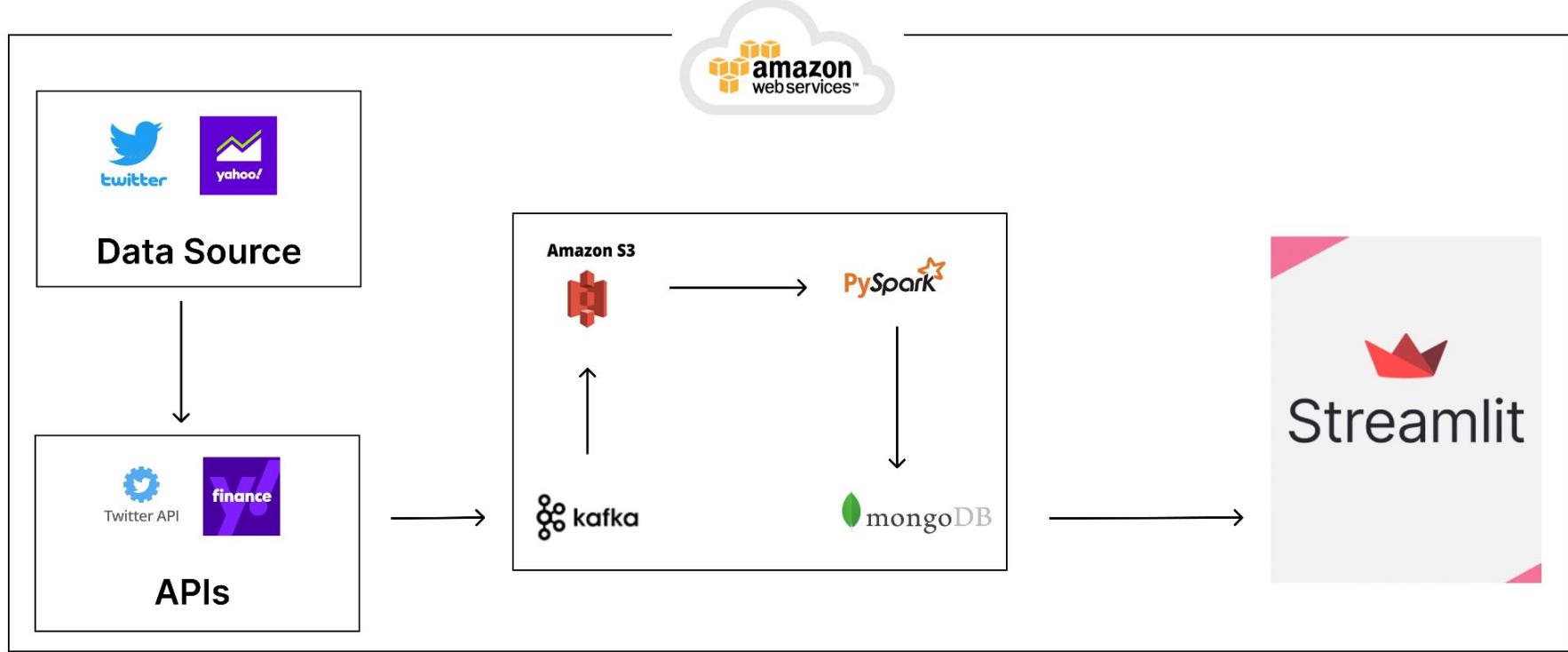
- Big data pipeline that collects twitter data pertinent to various stocks.
- Analytic framework for performing user sentiment analysis on various trending stocks.
- Correlation between user sentiments and stock price



Tools & Technologies

- Yahoo Finance API and Twitter API for Data
- Kafka for Data Ingestion
- PySpark for Data Processing
- MongoDB for Data Store
- Streamlit for Frontend
- Python for Backend
- AWS for Deployment
- Github for Version Control & Collaboration
- Notion for Project Management
- VSCode for IDE

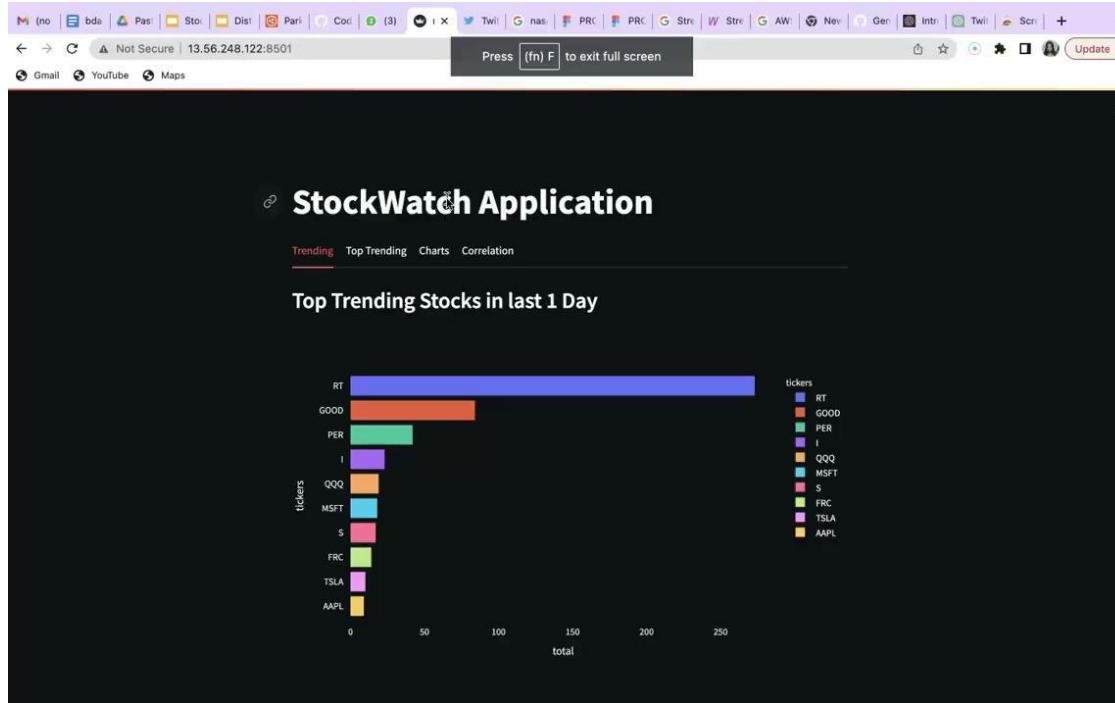
Architecture Diagram



Demo

<http://13.56.248.122:8501/>

Demo Link



Learning Curve

- Efficient data pipeline to handle real time big data.
- Learning new technologies like Kafka, PySpark, Streamlit.
- Understanding communication between various pipeline components and integrating them together.
- Deciding what data fields to use and how to visualize them in a user friendly way .
- Deployment on Amazon Web Services.
- Working in the AGILE methodology.

Challenges



- New Technologies
- Integrating Kafka
- Hosting and deployment
- Understanding Stock Market Dynamics

Future Scope

- Predicting Stock Prices Using LSTM
- Predicting a Bear or bull market
- Recommendation
- Containerization
- Exploring other API's





Thank you!