

Big Data Storage and Processing

REAL-TIME X(TWITTER) SENTIMENT ANALYSIS

Group 21

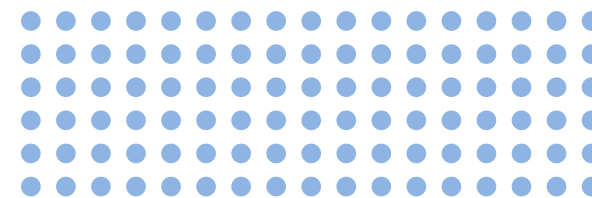
- Nguyễn Đức Anh- 20225432
- Trần Thị Hiền - 20214896
- Nguyễn Quốc Thái- 20225456
- Bùi Thị Thu Uyên-20225537
- Nguyễn Hải Dương - 20214887





Introduction

- X (Twitter) generates a huge volume of data every second.
- Real-time sentiment analysis helps businesses, governments, and organizations understand public opinion instantly.
- Requires scalable storage and fast, streaming data processing (Big Data Storage & Processing).

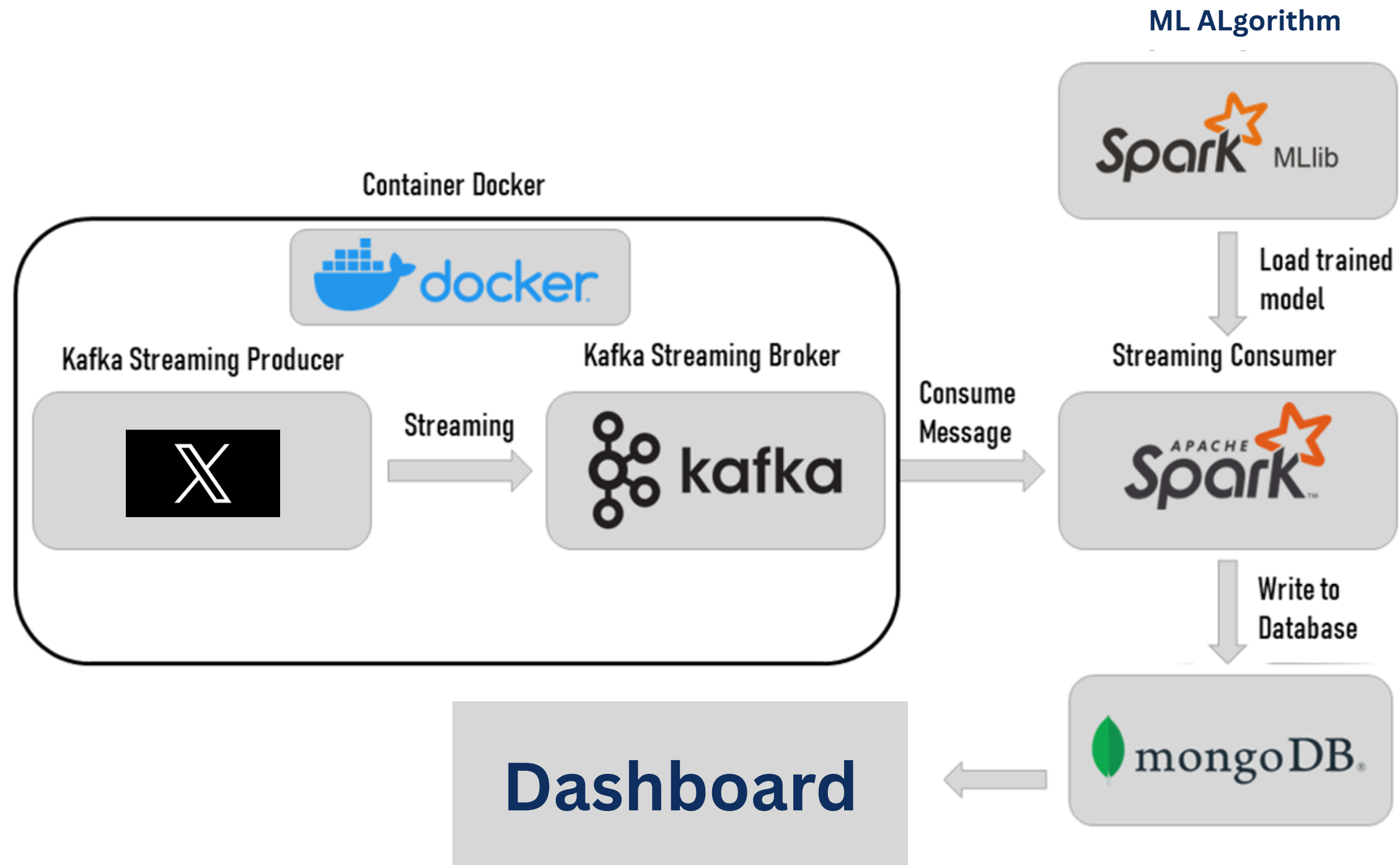


Project Objectives

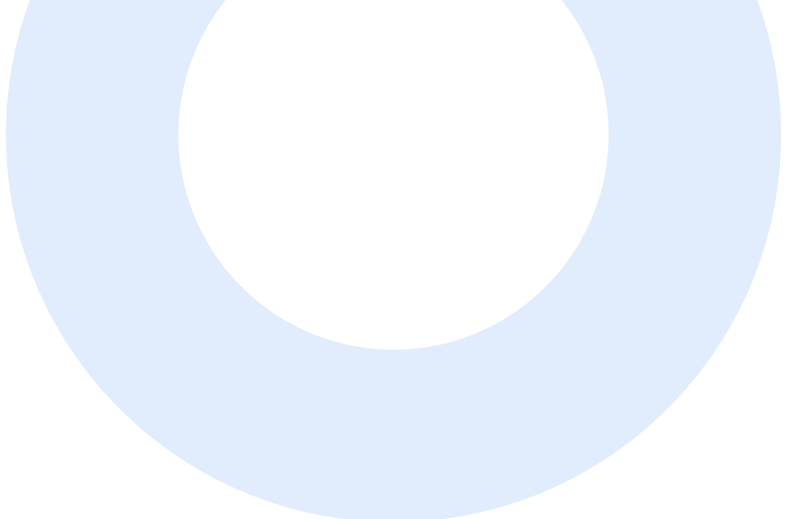
- Collect and store large-scale X(Twitter) data in real-time
- Process streaming data for sentiment analysis
- Store and visualize results using scalable big data technologies



General Architecture

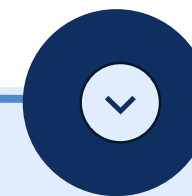


Features



Real-time Data Ingestion

Collects live tweets using Kafka from the Twitter DataSet.



Stream Processing

Utilizes Spark Streaming to process and analyze the data in real-time.



Sentiment Analysis

Classifies tweets into different sentiment categories using NLP techniques.



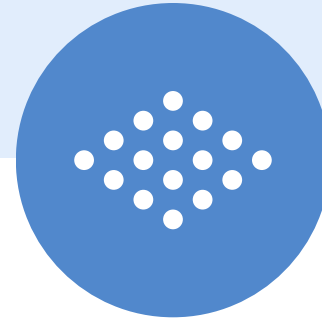
Data Storage

Stores the sentiment analysis results in MongoDB for persistence.



Visualization

Provides a real-time dashboard to visualize the sentiment trends and insights.



**THANK
YOU!**

