

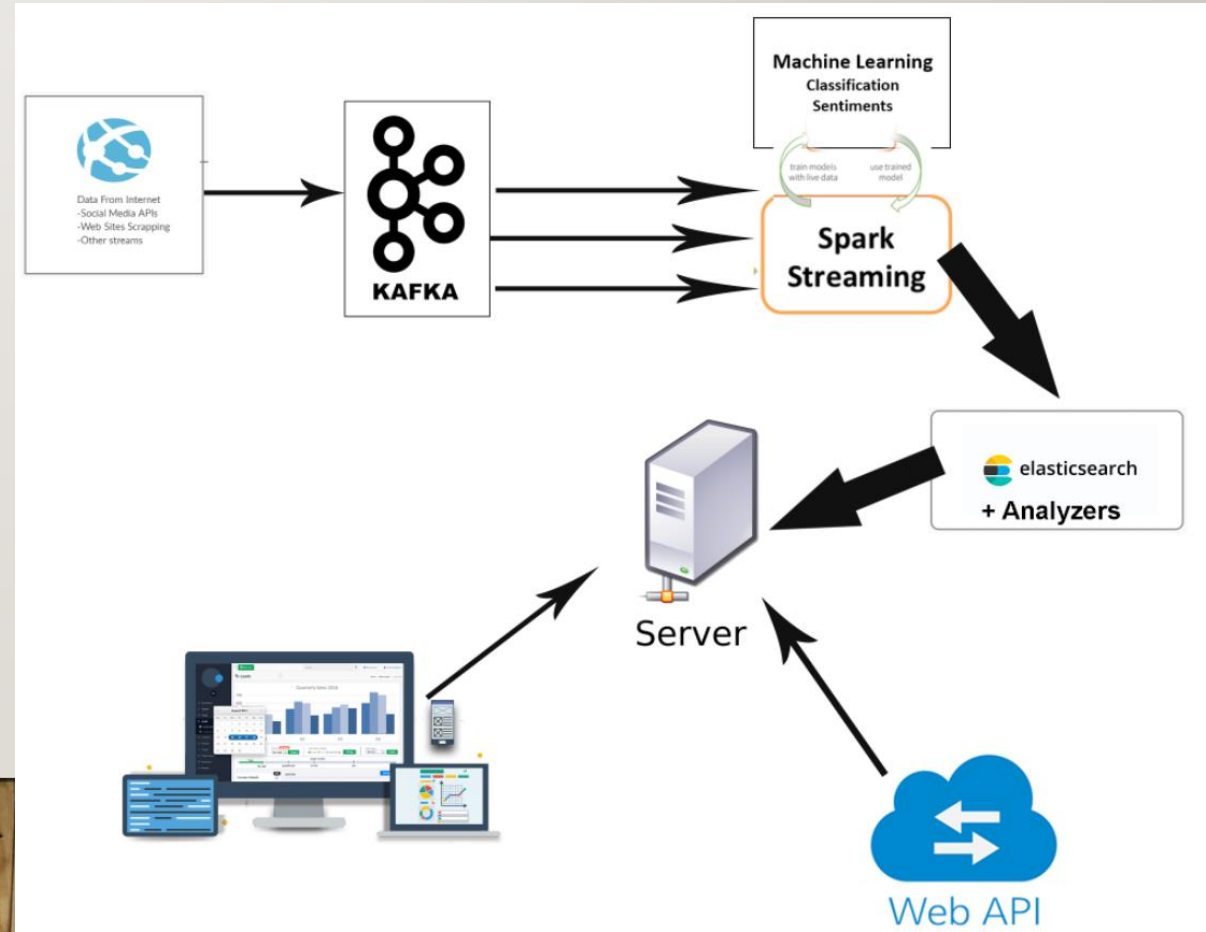
Anset Project Technical Overview



Project Architecture

Our project contains the following components and flow.

- Data Source Extractors
- Data Integration, Kafka
- Spark Streaming
- DB, Elasticsearch
- Web Application & API



Project Architecture – Data Sources Extractors

- The Social media represent the project data source, and each of platform has its own method to get its data, and each platform has deferent output.
- So we may need to build separate logic to extract data for each media, to deal with its output .
- Also for some data source we me need to use Web scrapping methods to extract data.



Data From Internet
-Social Media APIs
-Web Sites Scrapping
-Other streams

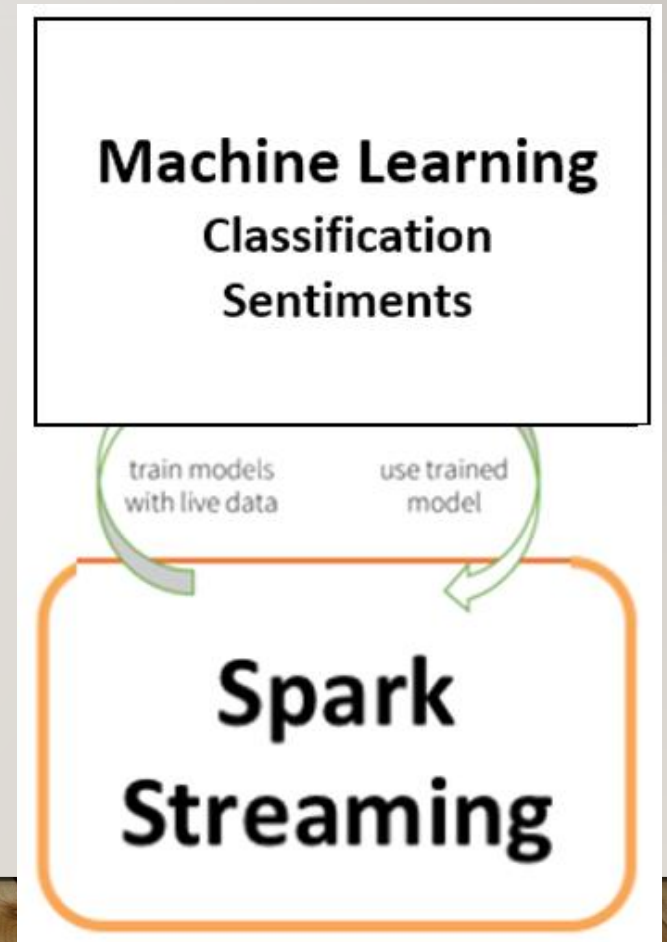
Project Architecture – Data Integration and Kafka

- Data extraction is continuous process, So we need to integrate new data and prepare it to the next step.
- Each data source has a different output, so need to be consumed by deferent process.
- Kafka is the best solution that receive the data from deferent sources and submit it to the appropriate consumer in real time



Project Architecture – Data Analysis & Spark Streaming

- For data processing and analyzing we are using Spark and its APIs and extensions.
- Spark Streaming will implement consumers for the data coming from Kafka server.
- Consumers will pass the data to trained model to predict its class and sentiment analysis.



Project Architecture – Elasticsearch

- After data analyzed, spark streaming will submit the data into Elasticsearch server.
- Elasticsearch maintain the data as a database
- It is the best solution to search through texts as it have some feature such as analyzer that will help to extract correct data based on text search



Project Architecture – Web Application & APIs

- We built a web application as Project UI based on Java spring boot framework
- Thought the application the clients can signup and access to system services
- The client can request his search statements and get the result as a static or interactive report
- We can implement APIs to request data or results



Web API



THANKS

