

Innovation Practices Final Review

Industrial Project

Team Members

Ms. J. Swathi
Academic Mentor

Mr. Subir Singh
Industry Mentor

Gondi Rajeev
19Z213

Koushik Balaji P.
19Z223

Lohith Sowmiyan P.S.
19Z224

Manoj Kumar
19Z226

Rakesh M.
19Z235

Sairam Vaidya M.
19Z238

Applying Predictive Analysis on Stock Markets

The problem statement deals with performing predictive analysis on an EIC (Economy - Industry - Company) basis, consolidating the various analysis performed and displaying them via a dashboard.



Motivation

Recently there has been an increasing interest in having two skills – being **legally aware** and **financially sound**.

By simply knowing the various types of investments, cash flow, etc., involved, a person can tune his **portfolio** ever so slightly to reap **benefits**.

The company we approached gave us a statement involving **Machine Learning** and **Finance** - we were very **interested**.

Literature Survey

As of May 24, 2022

Purpose

Prediction and analysis
of stock market data
using various
algorithms
(linear/non-linear)

References

Financial Markets
Prediction with
Deep Learning - Jia
Wang, et al.

Advantages

Probability of high
returns for short term,
ownership stake in the
company

Disadvantages

Volatility,
unsystematic risk is
unique to a specific
industry or a
company and can be
diversified

Literature Survey

As of May 24, 2022

Purpose

Data Science relies heavily on forecasting future outcomes and data modeling using Time-Series modelling

References

The prediction system for data analysis of stock market - Ching-Te Wang, et al.

Advantages

Stock market plays a pivotal role in the financial aspect of the nation's growth.

Disadvantages

It is highly volatile and complex in nature. It is affected by significant political issues and uncertainty in the future of a company.

Literature Survey

As of May 24, 2022

Purpose

It provides consistency in measures and measurement procedures across department and business units.

References

Pauwels, et al. (2009). Dashboards & Marketing: Why, What, How and Which Research is Needed?. *Journal of Service Research.* 12. 175-189.

Advantages

Dashboards can communicate what an organization values and can be used to disclose crucial marketing information to investors.

Disadvantages

The academic research is needed fully to exploit their potential in viewing the analysis for the users.



Hardware Requirements

Processor

Intel Xeon E2630 v4 - 10 core processor, 2.2 GHz with TurboBoost upto 3.1 GHz. 25 MB Cache

Motherboard

ASRock EPC612D8A

RAM

16 GB DDR4 2133 MHz

Hard Disk

Minimum 33 GB

GPU

GPU - NVidia TitanX Pascal (12 GB VRAM)

Heatsink

Intel Heatsink to keep temperature under control

Business and Data Understanding

Tableau, Power BI, Microsoft Excel

Data Collection

Yahoo Finance API, Beautiful Soup, Twitter API
(elevated access), Google News API

Data Processing and Visualization

Matplotlib, Seaborn, Pandas, Streamlit, NLTK

Data Modelling

Scikit-Learn, Tensorflow (2.0), Tensorflow GPU, Keras

Software Requirements

System Analysis

Input

Requirements
Determination

Literature Survey,
Consultation,
Business Research

Evaluation

Requirements/
Data Gathered

Validation using
Screeners, Historic
Trends

Process

Software
Development

Model
Development for
Predictive Analysis
and Sentiment
Analysis

Evaluation

Analysis Result

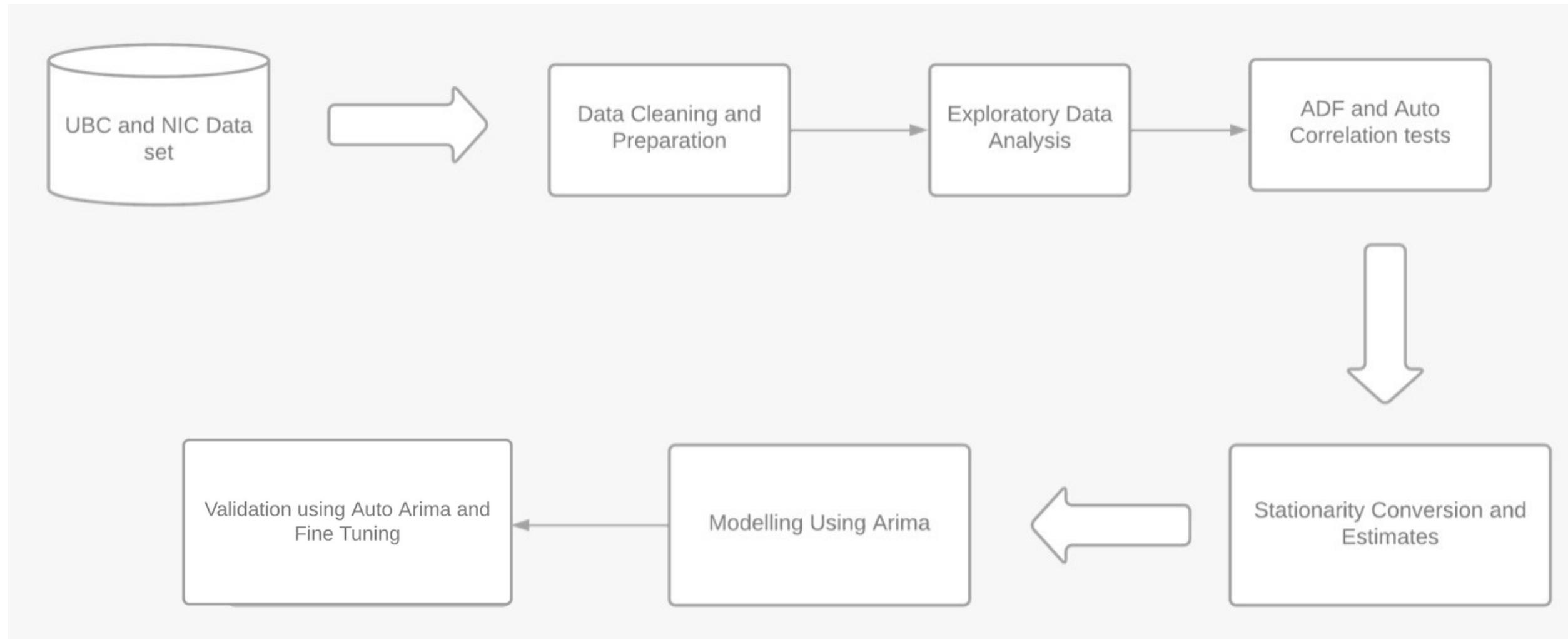
Accuracy, Growth

Output

Software
Implementation &
Maintenance

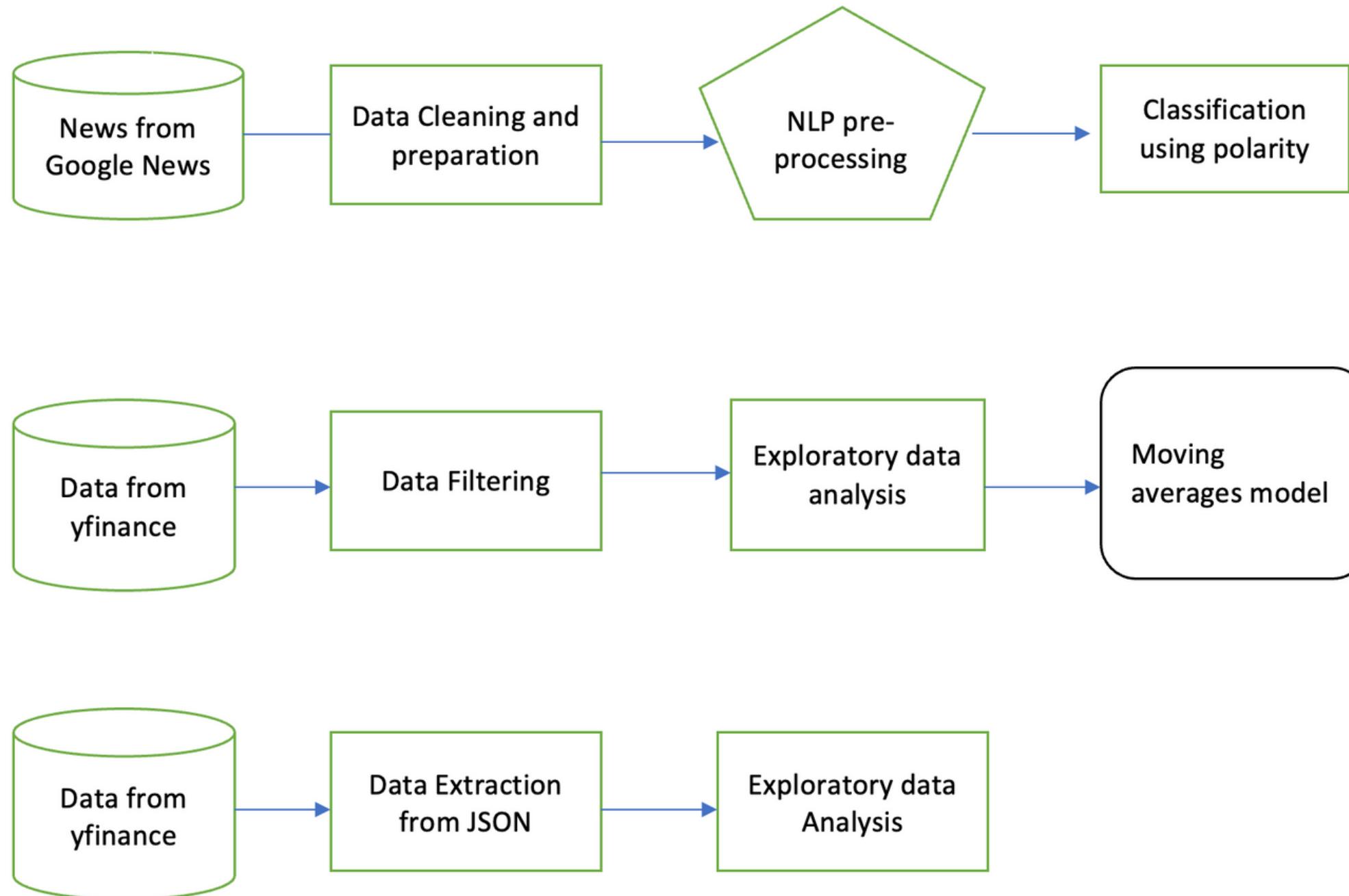
Portfolio,
Dashboard

System Design



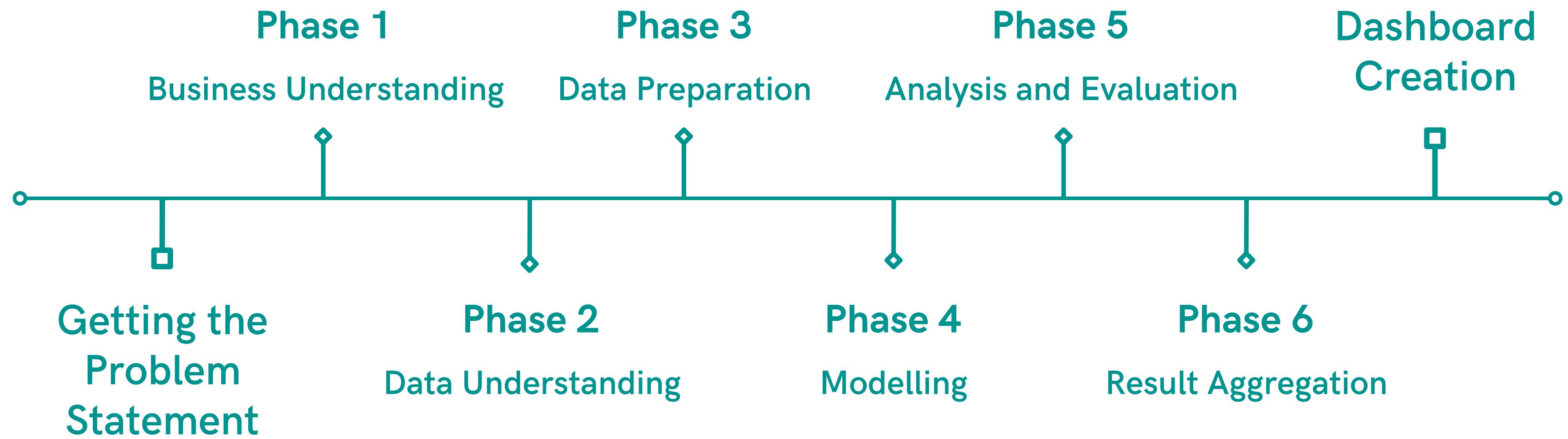
Business Understanding, Data Understanding and Analysis

System Design



Sentiment Analysis using Google News

Timeline



Percentage of Work Completed

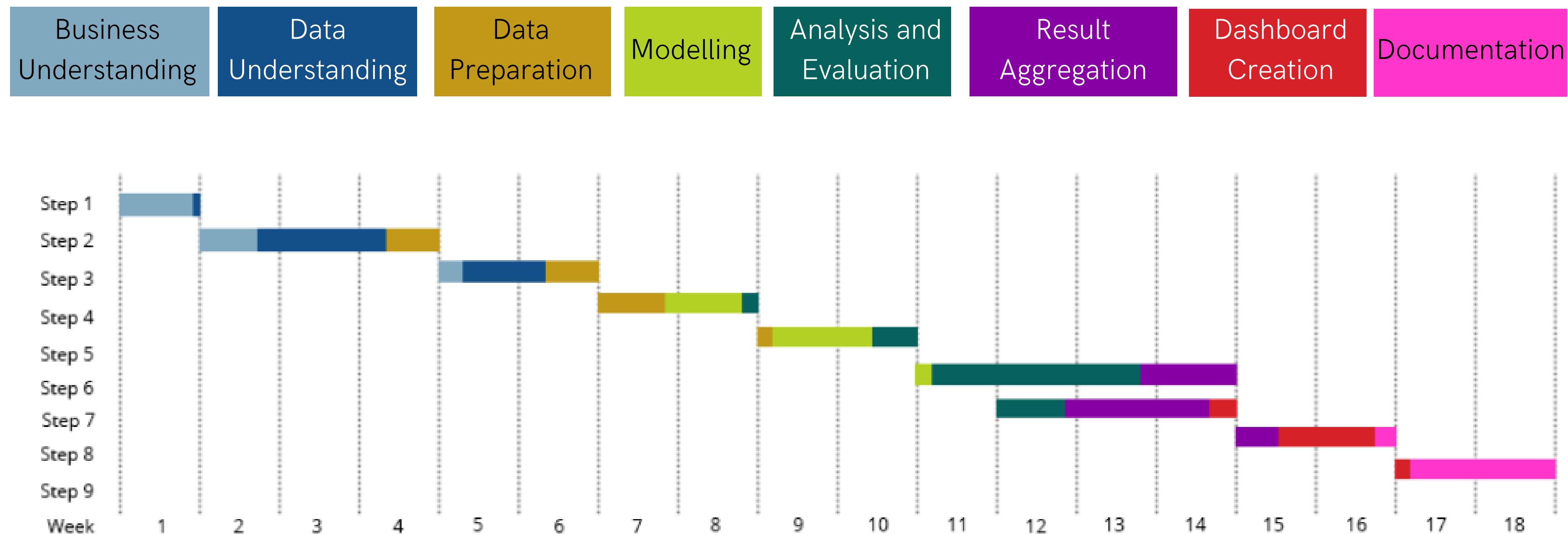
Getting The Problem Statement



The problem statement was obtained from PredictRAM, a fintech startup based in New Delhi, and it was based on data analysis and blockchain.



Planned Steps



Percentage of Work Completed

Business Understanding



This was completed by reading Varsity - Modules 1, 2 and referring economics concepts from Investopedia and Khan Academy.



Percentage of Work Completed

Data Understanding

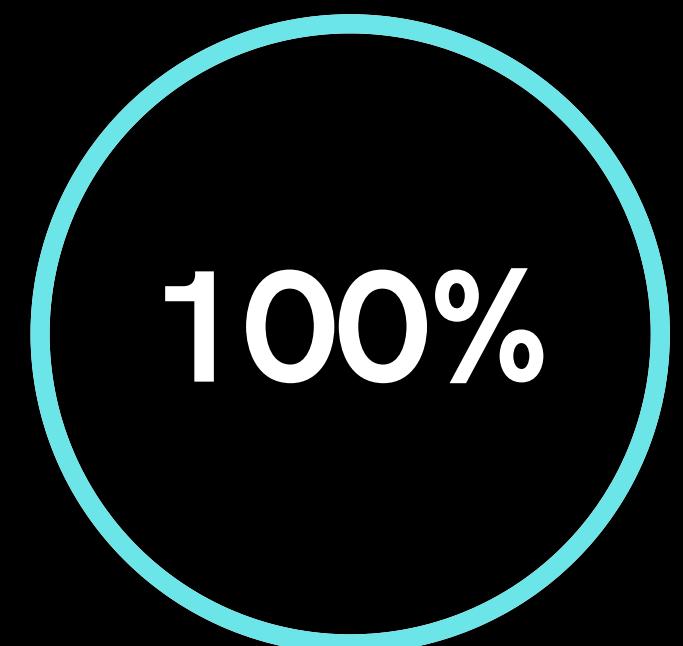


Performing economic analysis on IIP and GDP.



Percentage of Work Completed

Data Preparation



Raw data was extracted from various sources such as Yahoo Finance, CSO, World Bank, Google News, etc.



Percentage of Work Completed

Modelling



Different types of models were created for predictive analysis and sentiment analysis.



Percentage of Work Completed

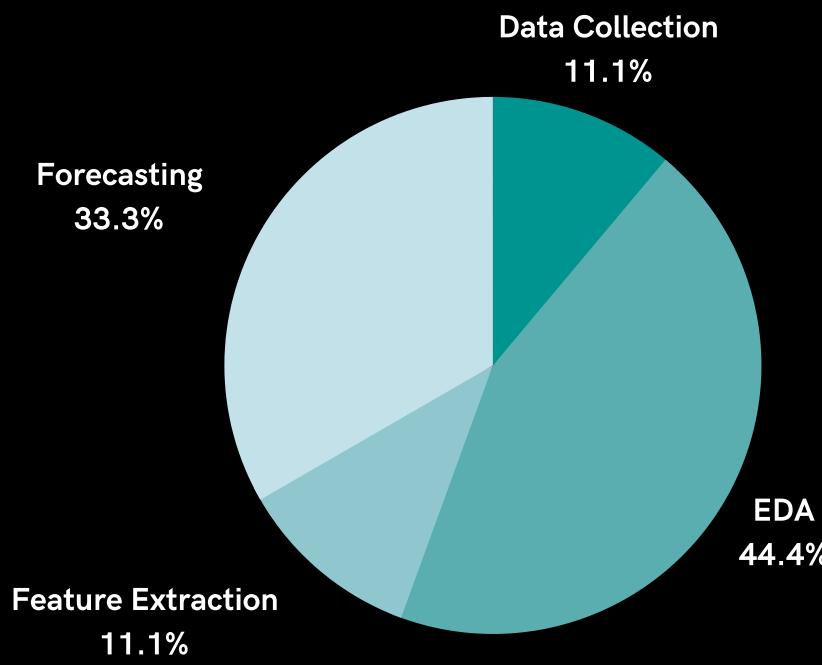
Analysis and Evaluation

100%

Different models were compared on the basis of accuracy, time taken for querying, quality of inferences, etc.



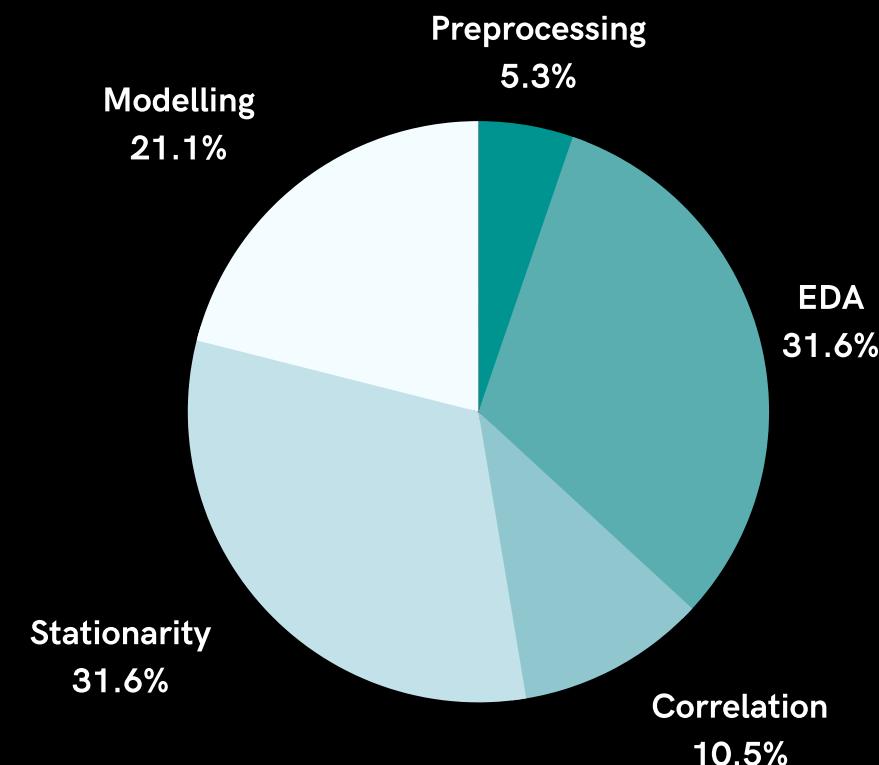
Economy and Industry Analysis



Industry

Through IIP (from CSO)
and GDP (from WB)

Models: ARIMA, Auto ARIMA

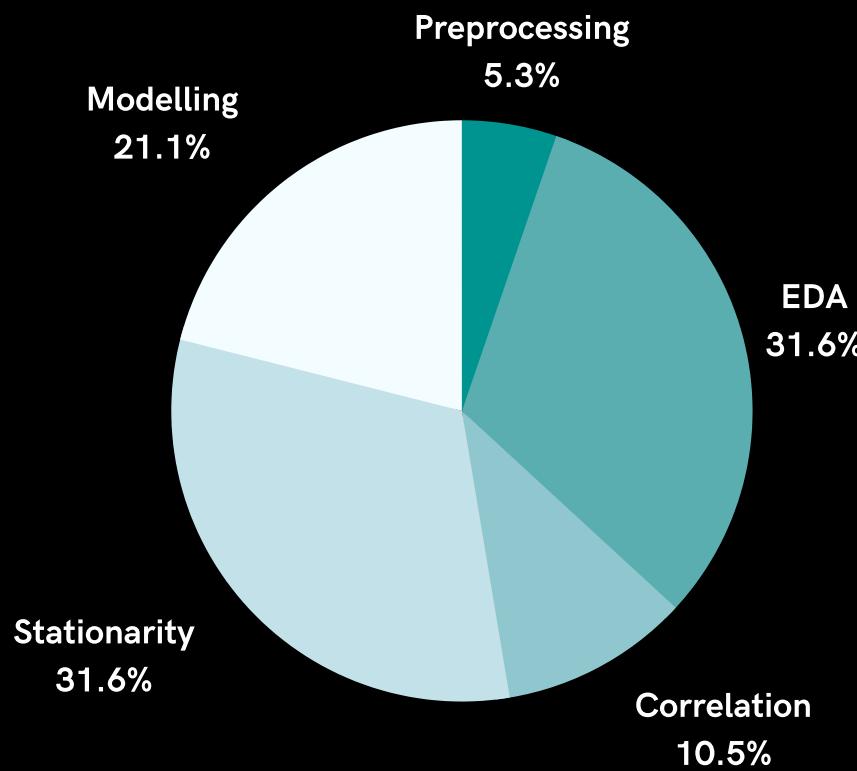


Economy

UBC and NIC datasets

Models: ARIMA, Auto ARIMA

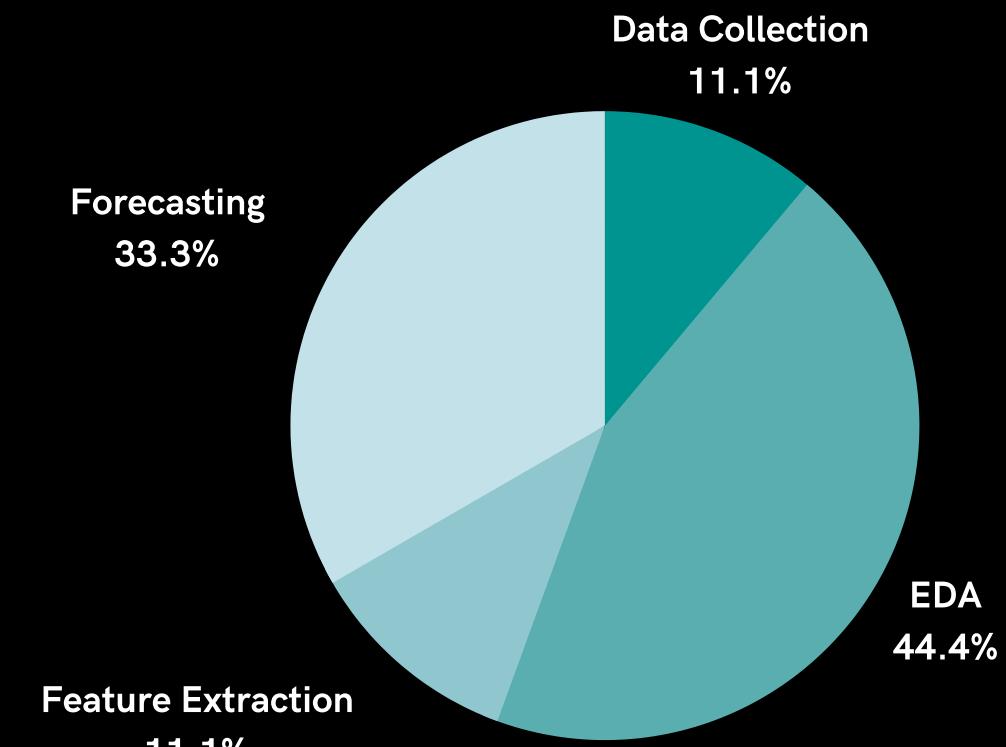
Economy and Industry Analysis



Economy

UBC and NIC datasets

Models: ARIMA, Auto ARIMA

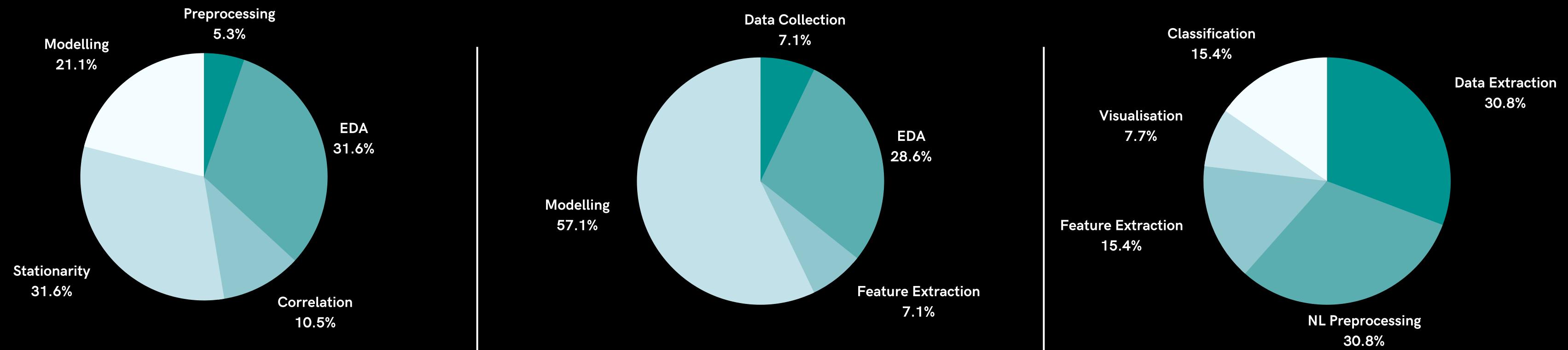


Industry

Through IIP (from CSO)
and GDP (from WB)

Models: ARIMA, Auto ARIMA

Company Analysis



Revenue

Collected from Tinker
function in yfinance API

Stock Price

Datasets from Yahoo fin
API

Sentiment

Sentiments collected
through top recent news
articles from Google news

Models: Vader Sentiment Intensity
Analyzer

Percentage of Work Completed

Result Aggregation

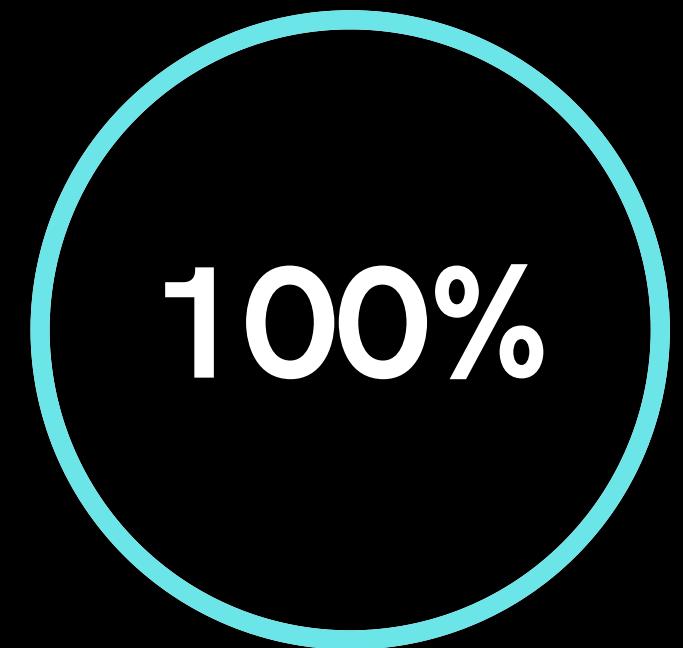


The results were aggregated into multiple CSV, XLSX, TWB, PDF, etc. and consolidated together.



Percentage of Work Completed

Dashboard Creation



The dashboard was created using Streamlit and consisted of the principal inferences obtained from the analysis.



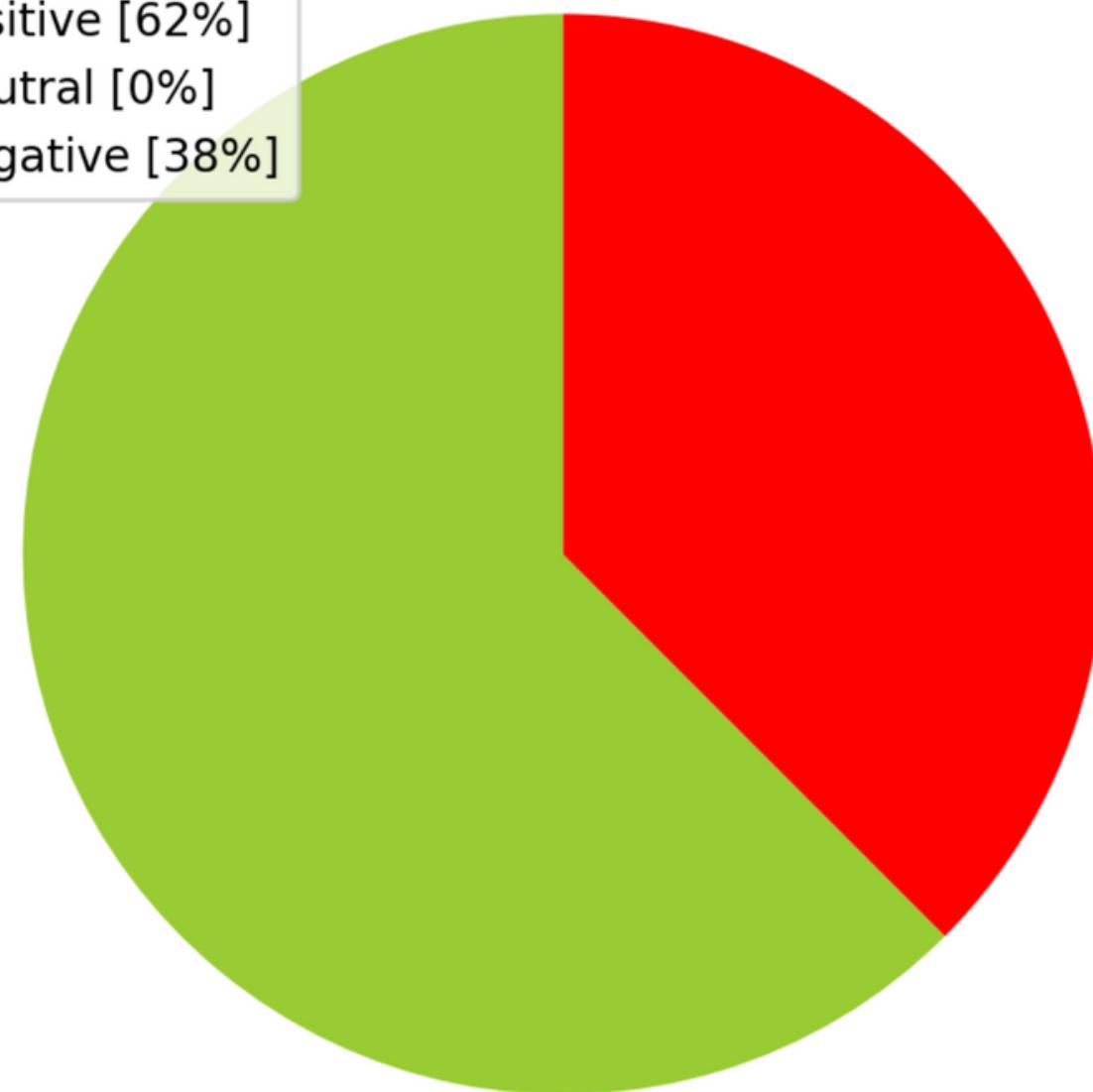
Dashboard

Enter the company name

apple

Customer Sentiments

- Positive [62%]
- Neutral [0%]
- Negative [38%]



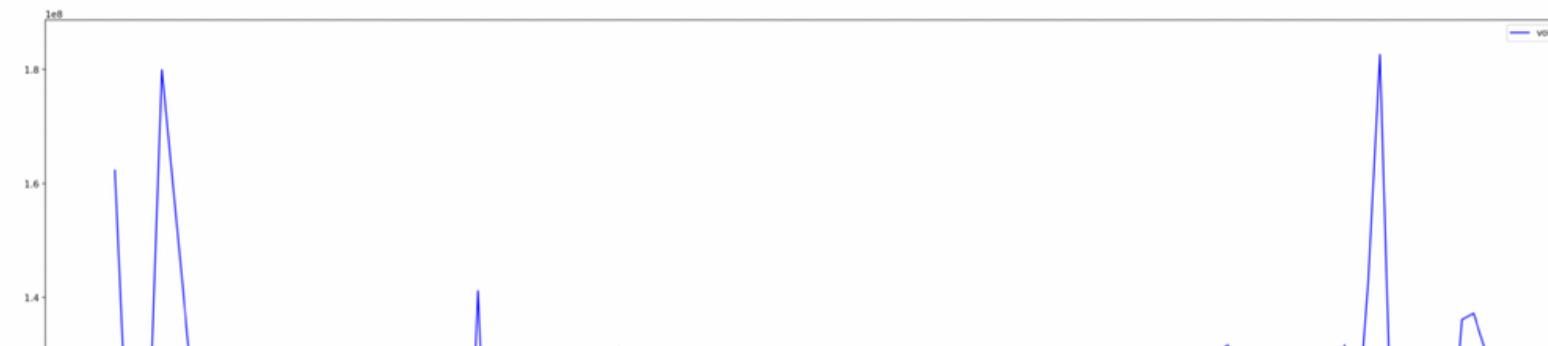
Dashboard

Stock Performance

Moving averages for apple

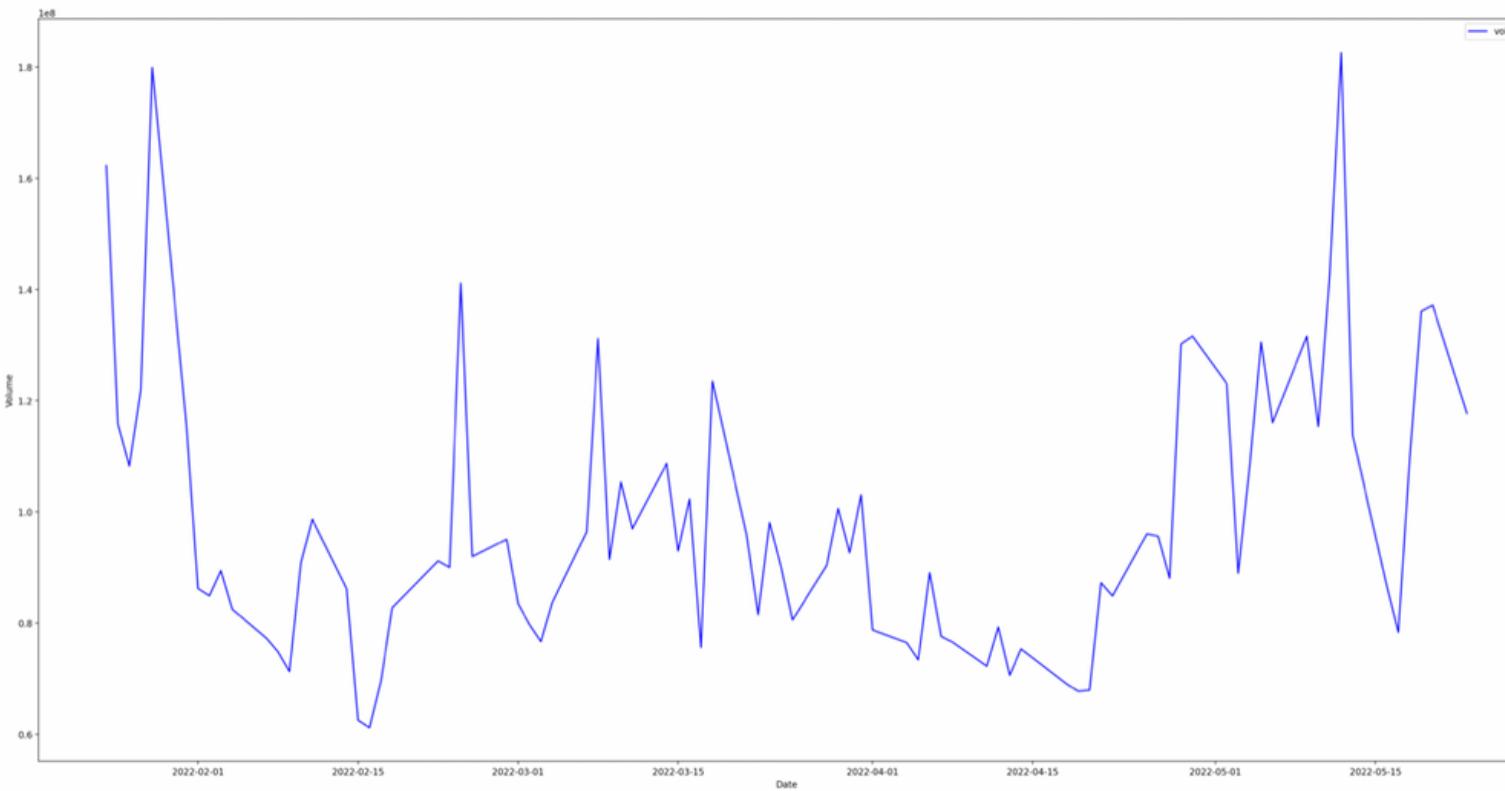


Volume for apple



Dashboard

Volume for apple



Moving averages for apple



Dashboard

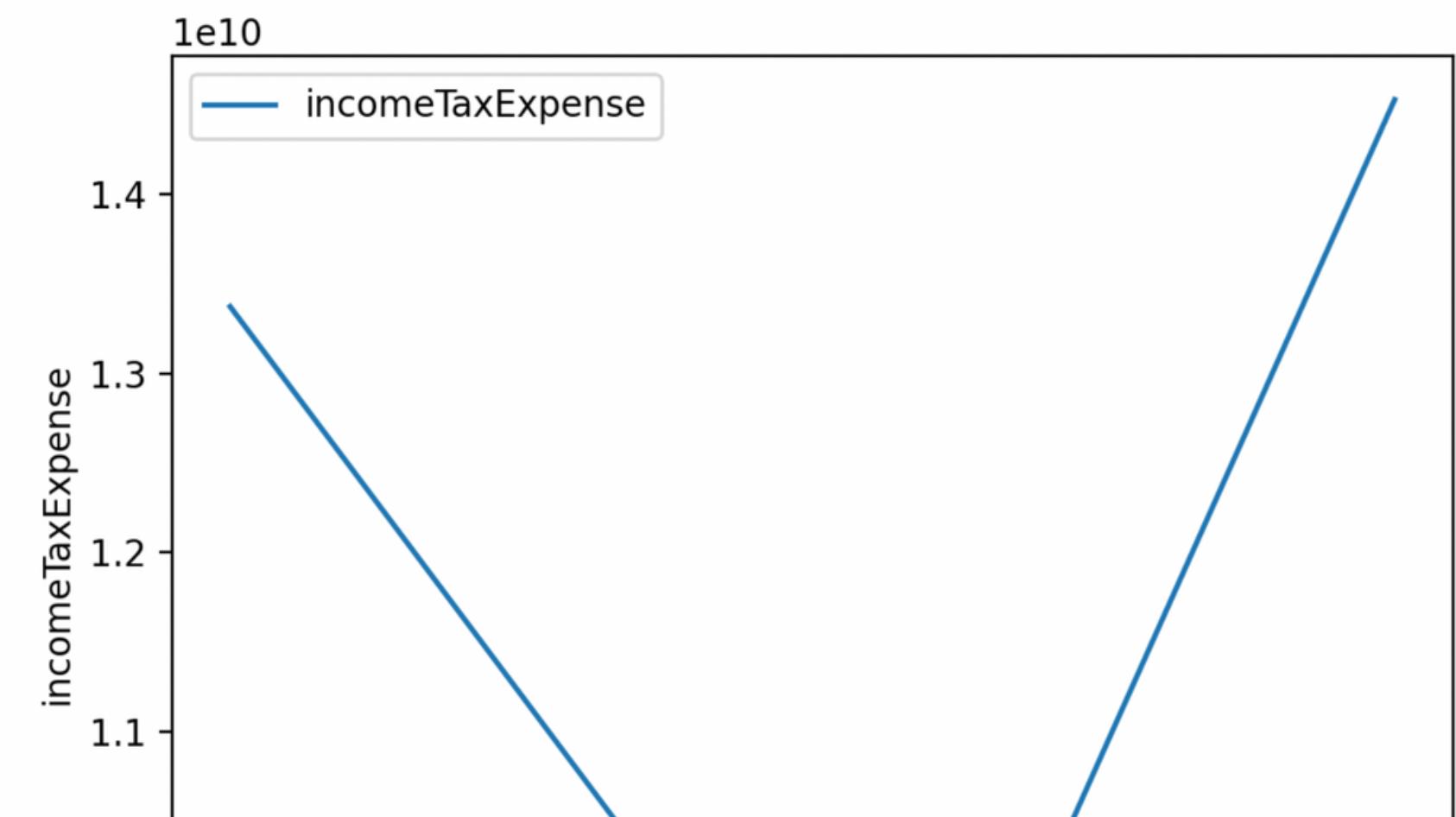
Select an attribute

sector

Technology

Select an attribute

incomeTaxExpense



References

- Strader, Troy J.; Rozycki, John J.; ROOT, THOMAS H.; and Huang, Yu-Hsiang (John) (2020) "Machine Learning Stock Market Prediction Studies: Review and Research Directions," Journal of International Technology and Information Management: Vol. 28 : Iss. 4 , Article 3. Tensor Flow resources and documentation.
- https://www.tensorflow.org/resources/mlgclid=Cj0KCQjw1N2TBhCOARIsAGVHQc6jl9iLh8aTXGjrmu0k6Tj6fVnYBJLIBCXmwmAL4WRJ2XFGWJOY03gaAsleEALw_wcB
- Naman Adlakha; Ridhima; Avita Katal, IEEE, 2021, "Real Time Stock Market Analysis", <https://ieeexplore.ieee.org/document/9526506>

Thank You

Company Specific Analysis

<https://bit.ly/3PDaONJ>

Dashboard

<https://bit.ly/3NvBh2H>

EIC Analysis

<https://bit.ly/38FPmvV>

Screen Scraping

<https://bit.ly/3wHvSis>

Sector Analysis

<https://bit.ly/3MJWp5h>

Sentiment Analysis

<https://bit.ly/3sTiBSI>

Done as an Industrial Project for,

PredictRAM (Params Data Provider Pvt Ltd.), New Delhi - 10.