Aston University

Birmingham, UK

Engineering and Applied Science

M.Sc Degree in Smart Telecoms and Sensing Networks

AI Enabled Market Research Platform
Literature Review

Graduating Student:
Dan Serbanoiu 200241287

Primary Supervisor:
Dr. Hussain Nazmul

Secondary Supervisor and Customer:
Ahmed Sohaib

Academic Year 2020/2021

Executive Summary

The aim of this project is to develop a general-purpose marketing research platform that employs big data analytics to provide affordable and high-quality data-driven marketing insights as-a-service to small and medium sized companies. Other such platform exist but they solve very specific business problems or require field specific knowledge. The proposed research platform aims at removing all the guesswork and provide businesses with a wide array of high quality marketing information to help them face better their challenges and improve competitiveness.

Contents

List of Figures

Chapter 1

Introduction

The purpose of the current chapter is to describe the general characteristics of an AI enabled market research platform, explain the scope of the project, detail a list of aims and objectives and define the structure of the dissertation.

1.1 Problem Statement

When doing market research, most Small and Medium-Size Enterprises (SMEs) don't know what to search for, struggle to find appropriate sources or lack the digital tools to properly analyse and make use of information. Nonetheless such data is vital for the success and functionality of businesses, as 60% of SMEs in UK fail within 3 years and lack of marketing research is a key reason [1]. For example, a company selling phones needs to see if the demand for mobile phones is growing or shrinking and know which accessories and phones are more popular and why. A physical restaurant should have a detailed knowledge of the dining trends of the food industry in their area. Usually, large companies have their own complex systems for recording data, store it upon well-structured models, analyse it and generate useful reports. Such systems, however, are too expensive to build for SMEs and require lots of knowledge and data for their functionality to begin with [2]. To overcome this limitation, this project aims to develop a general-purpose marketing research platform that can provide affordable and high-quality data-driven marketing insights as-a-service to small and medium sized companies. Other such platform exist but they solve very specific business problems. One such platform is QuantCloud, a trading service that uses predictive algorithms to find patterns within stock market data and offer stock market insights [7]. QuantCloud can calculate the seasonal volatility of log returns for grouped data, the Autoregressive-moving average (ARMA) etc. Another platform is Google Analytics which tracks websites usage and provides businesses with product trends, users behaviour information, custom reports etc. [3]. Even so, Google Analytics cannot track everything that happens on a web site so the information it provides is often an approximation. Moreover, in some cases there is lots of missing, unwanted or approximative information. Else-ways, the Market Research platform proposed by &Facts aims to actively help businesses in their search by taking away all the guesswork and provide them with a wide array of marketing tools and descriptive analytics. By gathering information from a variety of sources (documents, logs, APIs, websites, online surveys etc.) and even other marketing applications, the proposed platform can create marketing knowledge within different marketing fields and companies need only to make use of the final results, presented through graphs, charts etc. This enables smart business decisions and improves competitiveness of businesses.

1.2 Scope of the Project

Our goal is to study and develop AI-enabled market research tools and how to gather, integrate and apply big data to it in order to generate general purpose marketing insights. We are working closely with the industrial partner over a series of meetings in order to properly understand and define the business requirements and platform development opportunities and challenges. We will choose and learn all the appropriate technologies necessary to implement the project requirements. We will then write a state-of-the-art literature review in which we present any related work, existing limitations, development methodologies and project plan. Finally, we will work on the implementation of the marketing research platform, while documenting its modules and continuously deploy it.

1.3 Aims and Objectives

The main goal of this project is to develop a Web-based Market Research Platform by leveraging AI, ML and Big data technologies that will integrate data from different sources and process and analyse the information in real time for marketing insights. In order to implement such a system, the following objectives are set:

- Create an API for integrating all data sources on a common model;
- Use AI and ML to Develop a business intelligence tool that produces general purpose marketing insights;
- Implement a Search Module;
- Build a data visualization dashboard;
- Configure a central repository to store, process and query existing market research data;
- Develop a Minimum Viable Product (MVP) using suitable development tools and technologies;
- Test and validate the final product.

1.4 Use Case – &Facts

Figure 1.1: Logo of &Facts Limited

The industrial partner for this project is &Facts Limited (https://www.andfacts.com), a market research service company based in Birmingham that provides end-to-end online market research services for Small and Medium-Sized Enterprises (SMEs) on the internet. The &Facts business challenge is a well-known online market research problem in general for SMEs. The primary objective of the MSc research is to implement the &Facts use cases by investigating and developing an AI enabled research platform that analyses data from various sources and produces invaluable marketing information.

1.5 Internship

The project is a 600 hours long internship done under the Enabling Technologies & Innovation Competences Challenge program (ETICC) and as a part of the ERASMUS-funded Smart Telecom and Sensing Networks program (SMARTNET). I will write my thesis on the work done during this internship. The ETICC project at Aston University offered a feasibility study of an existing &Facts research system and &Facts agreed to pay for and provide the necessary API keys and infrastructure required for development. The customer agreed to remote working with frequent by-weekly meetings on streaming platforms to track progress and re-adjourn.

1.6 Dissertation Structure

The workload is split into the following activities:

- Chapter 1 gives an overview of the project;
- Chapter 2 presents the literature review;
- Chapter 3 offers an analysis of the technologies that have been used and how;
- Chapter 4 provides an analysis of the requirements;
- Chapter 5 explains the architecture of the research platform;
- Chapter 6 presents a retrospective analysis of the project and offers some conclusions.

The activities and their subtasks are scheduled through a Gantt chart, which is a useful tool for seeing the various dependencies between activities and sub-activities. In a Gantt chart is possible to represent, through black rhombuses, milestones that coincide with the end of a period or objectives to be achieved. It's important to respect the allocated time for the main activities, in order not to cause delays in all the other related tasks. The planning is described in chapter 3.

Chapter 2

Literature Review

The purpose of the current chapter is to provide knowledge of the subject areas relevant to the project including current developments, controversies, breakthrough, previous research and relevant background theory.

2.1 Market Research Platform

The purpose of the project is to develop a general purpose Market Research Platform that employs big data analytics to produce marketing information and useful insights that small and medium companies can use to understand their market and be more competitive. Marketing research is useful in the area of product planning and development such as when evaluating the need for a new product or its positioning in the market. It has applications in the area of advertising and copy testing or to identify existing and potential distribution channels. It can help to deduce the pricing expectations of consumers and their reactions and responses to different price levels of products. Marketing research can identify the forces operating in a market and assess the market trends, the size of present and potential markets of the company, the evolution of the impact of government legislation, policies, and schemes on the performances of marketing operations of the company. It can also study the sales potential of the company's products and the evolution of the company's sales performance. Finally, it can assess the environmental fitness of the firm. [4]. A study conducted by the Alibaba Group in China has also shown a way of calculating the impact of brands on society, by combining the impact brands have on the media, the government Impact and the personal impact. [5] The marketing platform can obtain all these analytics by integrating several data sources and where necessary apply AI and Machine learning to deduce complex results [6]. Currently, there is no one platform that provides such a varied and unified view of the markets of a business. Platforms like QuantCloud, a trading service [7], or Google Analytics, a websites activity tracking platform [3], all offer insights within specific areas and require specialized knowledge to be used. The goal is to build a product that does not require specialized knowledge and that is able to provide as much information as possible.

2.2 Data Extraction, Integration and Processing and Storage

Data usually comes from a very different a pool of sources such as social media, web server logs, sensors, the stock market, e-mails, etc. and it is no easy task to gather and integrate it all efficiently on a common model. The main challenges of data integration are how to integrate structured and unstructured data, how to manage it and sync it across many data sources. One of the best ways to performs this task is using the Extract, Transform and Load pattern (ETL) [1]. The data is collected from multiple types of sources, mapped to a mutual model and finally stored to a data repository. This process can be made very efficient using a framework such as Spark [8].

2.3 AI, ML and Big Data

Big data is a term used to describe the huge amount of data that is being generated by organizations in today's business environment. Such data usually comes from three primary sources: social data such as likes, tweets, comments, video uploads, queries in search engines etc., machine data generated by sensors installed in industrial equipment and transaction data. Machine learning is a branch of artificial intelligence that provides tools and techniques for working on all this data to produce useful results. It based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention. Machine learning algorithms are classified into supervised learning, unsupervised learning and reinforcement learning [6]:

2.3.1 Supervised learning

Supervised learning is the task of training a function with labelled training data in order for this function to be able to learn a general rule and map any input to its appropriate output. Machine learning models are usually very good at performing regression and classification:

- Statistical modelling techniques are used to find a parametric function that best fits a set of observations and train it appropriately so that is can predict new observations. Statistical models (i.e. logistic models, polynomials, linear regression models) are very good at modelling and predicting trends;
- Neural networks are powerful tools used to perform regression and classification of non linearly separable data. For example, convolutional neural networks can be used to classify consumers according to their preferences so that they can be targeted by specific ads or marketing strategies [9]. Neural Networks, paired with Locality sensitive hashing [10] have been also used to improve de-duplication, that is to identify chunks that are already stored in a large database;
- Support Vector Machines are excellent regression and classification algorithms that work by separating data by a hyperplane and maximizing the margin between the closest points. They can be employed to classify social media data such as tweets, posts, comments etc. according to the emotions they emit (positive, negative and neutral). These results can aid in understanding if people have a good or bad opinion about something and take action accordingly [11].

2.3.2 Unsupervised learning

Unsupervised learning is a machine learning technique in which a model works on its own to discover hidden patterns and separate unlabelled information into appropriate clusters, if they exist. It is useful for finding associations between different parameters in the available data, for example people that buy X might also buy Y. It can also be used to find the anomalous elements in a data set [12]. There are several clustering methods:

- Hierarchical clustering is a naive clustering method that builds a hierarchy of clusters by iteratively joining together the closest elements;
- K-means is a set of unsupervised clustering algorithms that can divide data points into k clusters [12];
- K-Nearest Neighbors is an algorithm that finds the k nearest points to the point of interest.

2.3.3 Reinforcement learning

Reinforcement learning is the training of machine learning models to make a sequence of decisions. The agent learns to achieve a goal in an uncertain, potentially complex environment. [13] Some relevant reinforcement learning algorithms are:

- Monte Carlo methods are a way of obtaining estimates when working with uncertain phenomena. They use randomness to obtain meaningful information and are effective for calculating business risks and predicting costs or scheduling overruns;

- For any finite Markov decision process, Q-learning algorithms are able to find an optimal path by maximizing the expected value of the total reward over any and all successive steps, starting from a specific state. These algorithms can be used to make decisions in business environments, for example in order to determine whether to hold, buy, or sell at any point in time.

2.4 ML workflow

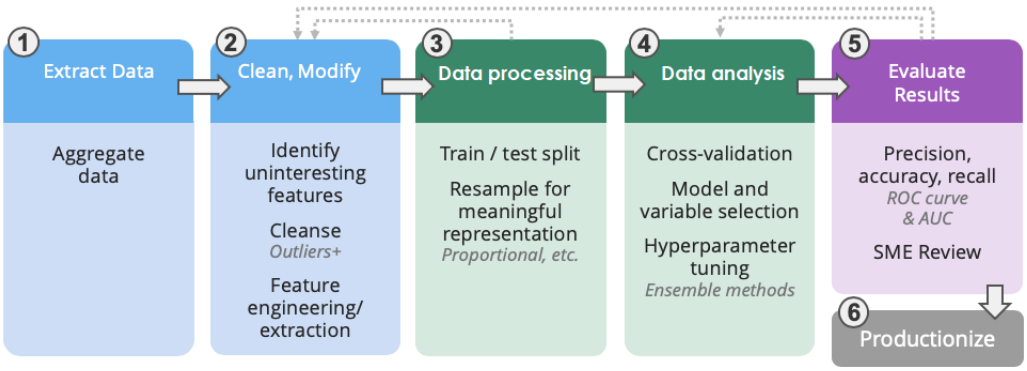

Figure 2.1: ML workflow

A machine learning task typically follows a cycle similar to the diagram above [14] and includes data gathering/extraction, data cleaning, data processing, data analysis, training and testing and finally the evaluation phase.

2.4.1 Data aggregation

Data aggregation is the process of gathering data and presenting it in a summarized format. The data may be gathered from multiple data sources with the intent of combining these data sources into a common format.

2.4.2 Data cleaning

Before data can be fed to a training algorithm or used for numerical analysis, it is important to prepare it by removing or changing rows that are incorrect, incomplete, irrelevant, duplicated, or improperly formatted. Skipping this step is not recommended because it can lead to runtime errors, inaccurate results and false conclusions [7].

2.4.3 Data processing

The data is divided into 3 three different sets, the training set, the validation set and the test set. During this phase some dimensions of the data might be removed, merged together or reshaped depending on the kind of task to be performed.

2.4.4 Data analysis

Depending on the task at hand (regression or classification) an array of ML models and sets of optimal hyperparameters are chosen for training.

2.4.5 Evaluation

The chosen models are trained with the training data and their precision, accuracy and recall are compared in order to find the best model and understand how well the chosen model can work in the future.

2.5 Tools and Techniques

There are many tools and technologies that have been developed for data extraction, integration, processing and analysis. The following tools and techniques have several features that are useful for Big Data integration and Machine learning tasks.

2.5.1 Aurelia

Aurelia is a client-side JavaScript framework that implements the Model-View-View-Model (MVVM) pattern and supports ES6 and TypeScript. It has a very good performance, an extensive ecosystem, it has a very solid and intuitive routing system and employs reactive binding. In Aurelia the presentation layer is completely separated from the application logic and as a result Aurelia applications are very maintainable, testable and extensible.

2.5.2 Django

Figure 2.2: Django

Django is a Python framework [15] and is used to write web applications. It comes with a lot of useful APIs, such as an Object relational mapping API, an User authentication module, a HTTP Session handling module etc. The framework employs the Model Template View pattern (MTV), which is a variation of the Model View Controller (MVC). In Django the Model components represent the data and the model classes are mapped automatically to database tables. The View components handle the HTTP requests and return responses. The templates are used to serve the HTML or data for the websites.

2.5.3 Spark

Figure 2.3: Spark

Spark is a framework developed in 2009 at UC Berkeley [8] that provides primitives for data mining and offers a very good implementation of MapReduce, a programming model for processing and generating big data sets with a parallel, distributed algorithm on a cluster [16]. In Spark every object is a Resilient Distributed Dataset (RDD). The elements of an RDD are not stored in memory but rather every RDDs maintains lineage information for fetching data sets from their source when needed. RDDs can be computed from a HDFS Hadoop Distributed File System (HDFS) or from an existing RDD. In Spark data is split among the different workers by a main driver. Every worker performs its assigned task and then sends the result back to the origin. Spark is up to

100 times faster than Hadoop, because the latter only processes data on disk, while Spark can also load data in memory. As a result, Spark can read 1 TB of data in 4 or 7 seconds. Spark can run on clusters created using many different technologies, such as Apache Mesos, Kubernetes, Amazon EMR etc. and a Spark cluster can scale very easily.

2.5.4 Docker

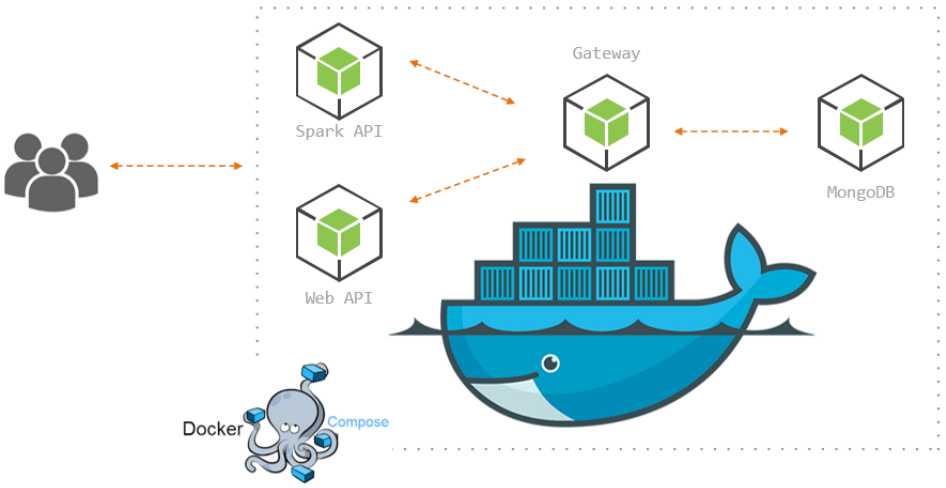

Figure 2.4: Docker

Docker is a platform for packaging and delivering independent software units in portable and self-sufficient containers that have all the dependencies required for the functionality of their applications. Every container is derived from an image, which essentially is a read-only template. An image provides its container with environmental variables, a command line and a filesystem that can store operating systems, packages, libraries and every other dependency a container needs. The main advantage of encapsulating all the dependencies of a service inside a self-sufficient entity it's ensuring that the development platform mimics perfectly the target platform. This makes it very easy to continuously integrate and deploy, increases productivity and encourages a frequent feedback loop. With Docker it's possible to package the different types of nodes of an application. [17].

2.5.5 Query optimization through hashing

A very good algorithm for optimizing comparison between items is locality sensitive hashing (LSH). LSH uses the properties of hashing in order to find similarities between big sized entities [10]. LSH can be used to improve query efficiency through an approximate Membership Query (AMQ) scheme. In such a system, the items are hashed with a similar process as LSH, with the only difference being that instead of hashing the signature into buckets, a Bloom filter is used in the last step to map every item into L bits. All the items that have the same L bits set to 1 are determined to be approximate members of the same data set S. AMQ can be used to produce results within a search engine in O(1) time and has demonstrated better accuracy than LSB-trees and other algorithms than run in at least logarithmic time. However, a good configuration is needed in order to minimize false positives and false negatives [18].

Chapter 3

Development methodology

This chapter describes the system architecture and explains the steps that will be taken for implementing the marketing platform.

3.1 System Architecture

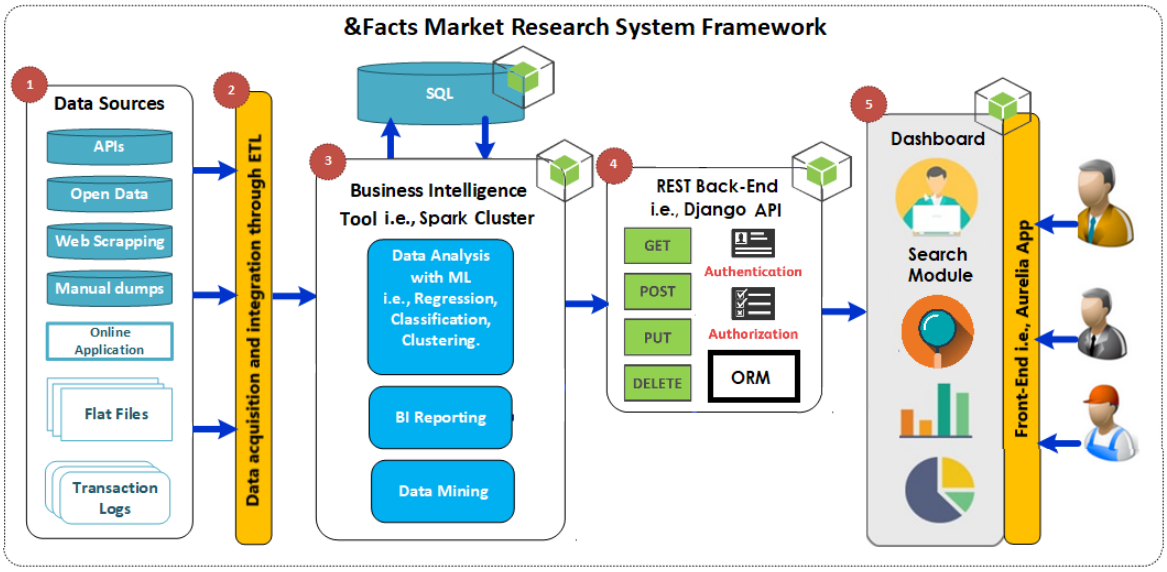

Figure 3.1: System Architecture

The conceptual view of the reference architecture shows the components of the &Facts Market Research platform solution. The marketing research platform is made of:

- A Front-End Application made of a Dashboard component and a Search module;
- A Back-End REST Application orchestrating communication between modules and handling the session;
- A Business Intelligence Tool that performs data analysis and data mining and generates useful reports;
- A SQL database that holds the data of the platform;
- A Data Acquisition module that fetches data from the various sources;
- A Data Integration component that performs data integration through the ETL paradigm.

3.2 Functional and Technical User Stories

The project goal is to develop a web-based Minimum Viable Product (MVP) general purpose Market Research Platform integrating AI and Big Data technologies that will be running on the cloud. In order to build the &Facts platform, the following steps are followed:

- The Front-end application will be written using Aurelia JS and Bootstrap. The application will run in a Docker container and will include a secure access module, updating and configuration capabilities, a search system and an intelligent marketing information visualization dashboard (using tables, graphs, and charts);
- The Back-end REST application will be developed using Python over Django and Spark. This component will reply to queries and perform disk-efficient and computationally optimized information gathering, data aggregation and analysis;
- The User Search module will take a list of keywords provided by the user and fetch them through appropriate APIs iteratively in order to generate more related keywords. The keywords saved into a keystore as they are being generated. The main APIs that will be used for the implementation of this module are Facebook Graph API and DataforSEO;
- The Data Acquisition module will fetch relevant data from APIs given by the customer and other sources using the keywords generated by the User Search module. Different data sources should be used accordingly depending on the use case;
- The Data Integration module will map the retrieved data to a common structure, sort it, filter it and verify it using the ETL paradigm (Extract, Transform, Load). Duplicated rows may be removed from the data. Spark framework will be used for preprocessing huge data sets from multiple sources and as it can run on a cluster;
- The Business Intelligence tool-set will use AI and ML on the data to produce results such as forecasting, causal factors and economics of trends, classification of features and clustering, NLP analysis on Facebook ads, social listening and more;
- The data will be stored in a relational database that will be connected with the Data acquisition and Data integration modules on one end and the Business Intelligence tool-set on the other.
- The whole platform will be developed locally and continuously deployed inside Docker containers running on Amazon Web Services clusters;

Chapter 4

Planning

This chapter describes the project development model and a plan of the activities and show their dependencies.

4.1 Life-cycle model

The chosen life-cycle model is the incremental model and it has the following characteristics:

1. Architectural Analysis and Design activities are not repeated: the goal is to spend sufficient time studying the problem and gathering all the requirements so that the architecture of the system is can be fully identified during the initial stages of the project. This allows for a stronger planning;
2. The implementation of the different parts of the system is incremental and is also planned in the initial stages of the project;
3. The activities of Design, Coding and Verification are repeated in order to improve existing parts of the system, correct mistakes or add new features;
4. Maintenance is a continuous activity and aims at making the product solid and bug free.

The advantages of using this model are that the requirements are treated according to their strategic importance and the implementation starts from the primary ones. Furthermore, each increment adds new functionality to the baseline and allows the costumer to frequently test the product prototype and give valuable feedback. By using the evaluations of the customer, every successive iteration of the application should become more complete, converging to an appropriate final version of the product.

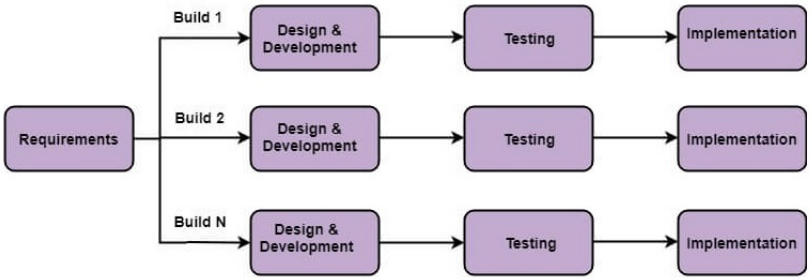

Figure 4.1: Incremental Model

4.2 Gantt chart

The following chart shows the planification of all the activities involved in the project:

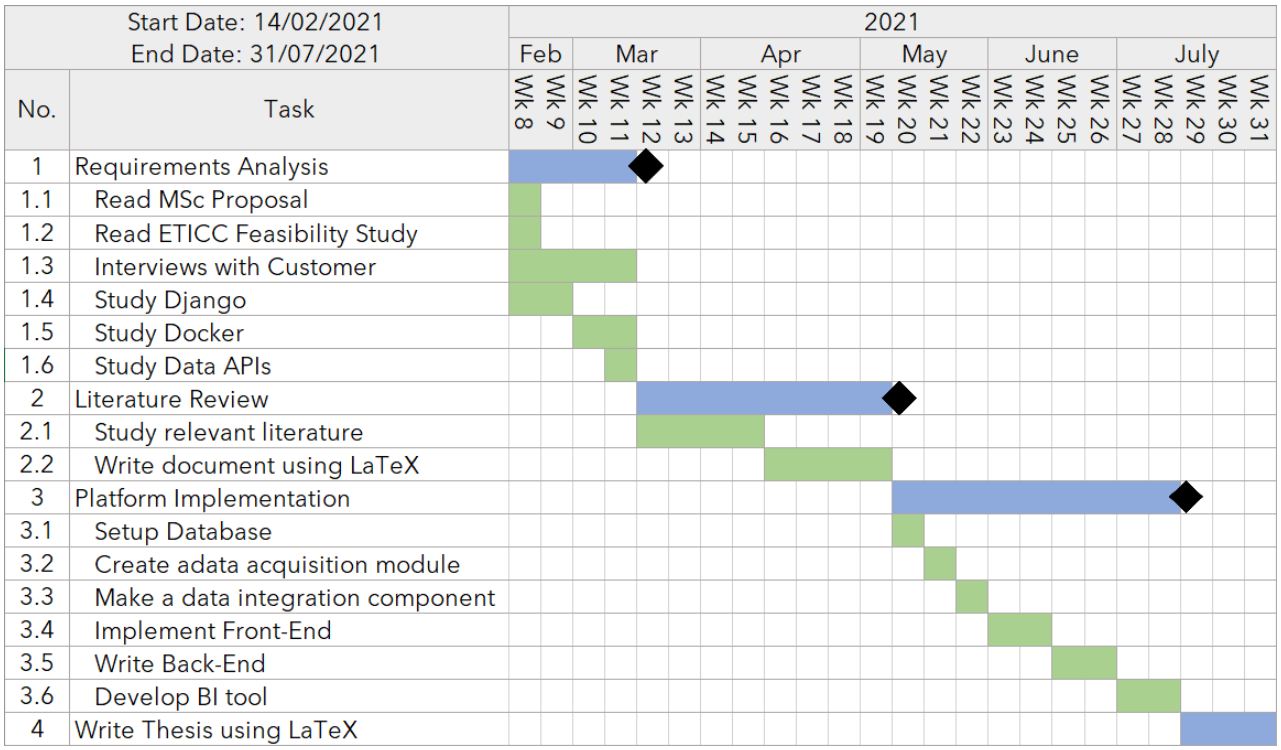

Figure 4.2: Gantt Chart

Bibliography

[1] D. N. Hussain, "Feasibility study and software requirements of &facts market research platform," 09 2020.

[2] D. Ulas, "Digital transformation process and smes," *Procedia Computer Science*, vol. 158, pp. 662–671, 2019. 3rd WORLD CONFERENCE ON TECHNOLOGY, INNOVATION AND ENTREPRENEURSHIP"INDUSTRY 4.0 FOCUSED INNOVATION, TECHNOLOGY, ENTREPRENEURSHIP AND MANUFACTURE" June 21-23, 2019.

[3] S. Galbraith, "Google analytics," *Journal of the Canadian Health Libraries Association*, vol. 34, pp. 119–122, 08 2013.
https://www.researchgate.net/publication/272894748_Google_Analytics.

[4] B. Ayuba, O. Aina, and Kazeem, "The role of marketing research on the performance of business organizations," vol. 7, 01 2015.
https://www.researchgate.net/publication/344177049_The_Role_of_Marketing_Research_on_the_Performance_of_Business_Organizations.

[5] L. Huang and X. Zhang, "A new marketing effectiveness metric based on web data mining," in *2009 1st IEEE Symposium on Web Society*, pp. 5–9, 2009.
https://ieeexplore.ieee.org/document/5271733.

[6] P. Vats and K. Samdani, "Study on machine learning techniques in financial markets," in *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, pp. 1–5, 2019.
https://ieeexplore.ieee.org/document/8878741.

[7] P. Zhang, X. Shi, and S. U. Khan, "Quantcloud: Enabling big data complex event processing for quantitative finance through a data-driven execution," *IEEE Transactions on Big Data*, vol. 5, no. 4, pp. 564–575, 2019.
https://ieeexplore.ieee.org/document/8386682.

[8] M. Zaharia, M. Chowdhury, M. Franklin, S. Shenker, and I. Stoica, "Spark: Cluster computing with working sets," *Proceedings of the 2nd USENIX conference on Hot topics in cloud computing*, vol. 10, pp. 10–10, 07 2010.
https://www.researchgate.net/publication/234790155_Spark_Cluster_Computing_with_Working_Sets.

[9] H. N. Bhor, T. Koul, R. Malviya, and K. Mundra, "Digital media marketing using trend analysis on social media," in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pp. 1398–1400, 2018.
https://ieeexplore.ieee.org/document/8399038.

[10] A. Gionis, P. Indyk, and R. Motwani, "Similarity search in high dimensions via hashing," *Proceeding VLDB '99 Proceedings of the 25th International Conference on Very Large Data Bases*, vol. 99, 05 2000.
https://www.researchgate.net/publication/2634460_Similarity_Search_in_High_Dimensions_via_Hashing.

[11] V. S. Rajput and S. M. Dubey, "Stock market sentiment analysis based on machine learning," in *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, pp. 506–510, 2016.
https://ieeexplore.ieee.org/document/7877468.

[12] B. Bahmani, B. Moseley, A. Vattani, R. Kumar, and S. Vassilvitskii, "Scalable k-means++," *Proc. VLDB Endow.*, vol. 5, 03 2012.
https://www.researchgate.net/publication/221966107_Scalable_K-Means.

[13] L. Ma and Y. Liu, "Application of a deep reinforcement learning method in financial market trading," in *2019 11th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pp. 421–425, 2019.
https://ieeexplore.ieee.org/document/8858758.

[14] K. Johnson, J. Soto, B. Glicksberg, S. Khader, R. Miotto, M. Ali, and E. Ashley, "Artificial intelligence in cardiology," *Journal of the American College of Cardiology*, vol. 71, pp. 2668–2679, 06 2018.
https://www.researchgate.net/publication/325585229_Artificial_Intelligence_in_Cardiology.

[15] R. Jan, "Django fundamentals,"
https://app.pluralsight.com/.

[16] J. Dean and S. Ghemawat, "Mapreduce: Simplified data processing on large clusters," vol. 51, pp. 137–150, 01 2004.
https://www.researchgate.net/publication/220851866_MapReduce_Simplified_Data_Processing_on_Large_Clusters.

[17] N. Brown, "Containerizing a software application with docker,"
https://app.pluralsight.com/.

[18] Y. Hua, B. Xiao, B. Veeravalli, and D. Feng, "Locality-sensitive bloom filter for approximate membership query," *IEEE Transactions on Computers*, vol. 61, no. 6, pp. 817–830, 2012.
https://ieeexplore.ieee.org/document/5928322.

[19] A. Berman, Y. Birk, and A. Mendelson, "Accelerating duplicate data chunk recognition using nn trained by locality-sensitive hash," in *2014 IEEE 28th Convention of Electrical Electronics Engineers in Israel (IEEEI)*, pp. 1–5, 2014.
https://ieeexplore.ieee.org/document/7005887.

Appendices

Appendix A

Glossary

C

Continuous Integration and Continuous Delivery

Continuous Integration and Continuous Delivery are the processes in which a development team integrates changes frequently into the main branch while ensuring that it does not impact any changes made previously or by other developers.

Customer

It is the person with the highest stakes in the project and understands both the business needs and operational constraints of the project. He guides the development on and knows the priorities that should be emphasized.

D

Data Analysis

It is the process of collecting, transforming, cleaning, and modelling data with the goal of discovering information.

Data Warehouse

It is a large centralized repository of data that transform all multiple data formats into a single format and consolidate them in one place.

Data Mart

This is a a data warehouse oriented to specific business lines. It contains repositories of summarized data collected for analysis of a specific section or unit within an organization.

Descriptive Analytics

It is the interpretation of historical data to better understand changes that have occurred in a business. Descriptive analytics describes the use of a range of historic data to draw comparisons. Most commonly reported financial metrics are a product of descriptive analytics, for example, year-over-year pricing changes, month-over-month sales growth, the number of users, or the total revenue per subscriber. These measures all describe what has occurred in a business during a set period.

Development Methodology

This is the approach chosen to manage the different phases of the life cycle of an engineering project. The most common models are Waterfall, Agile, Lean, Iterative, Prototyping, DevOps, Spiral or V-model. It includes phases for tracing the progress of a product from a planned idea to its final release into operation and maintenance.

M

Marketing Research

It is the systematic design, collection, analysis, and reporting of data of the market of a company. Market-related data such as sales data, advertising expenditure, customer information, etc.is thus recorded over time.

P

Minimum Viable Product

A minimum viable product, or MVP, is a product with enough features to attract early-adopter customers and validate a product idea early in the product development cycle. In industries such as software, the MVP can help the product team receive user feedback as quickly as possible to iterate and improve the product.

S

Stakeholder

A stakeholder is a party that has an interest in a company and can either affect or be affected by the business. The primary stakeholders in a typical corporation are its investors, employees, customers, and suppliers.