ABDUL SAMAD & SYED TOUSEEF AHMED
01-134181-003 & 01-134181-061

Should I buy it? A Smart Guide for Consumers

Bachelor of Science in Computer Science

Supervisor: Dr. Momina Moetessum

Department of Computer Science
Bahria University, Islamabad

November, 2021

Certificate

We accept the work contained in the report titled “Should I buy it? A Smart Guide for Consumers”, written by Abdul Samad & Syed Touseef Ahmed as a confirmation to the required standard for the partial fulfillment of the degree of Bachelor of Science in Information Technology.

Approved by . . . :

Supervisor: Dr. Momina Moetessum (Associate Professor)

____

Internal Examiner: Name of the Internal Examiner (Title)

____

External Examiner: Name of the External Examiner (Title)

____

Project Coordinator: Dr.Mozzam Ali (Assistant Professor)

____

Head of the Department: Dr. Arif-ur-Rehman (Sr. Associate Professor)

____

Nov, 2021

Abstract

The objective of this project is to develop a mechanism that can automatically respond to a consumer's query regarding the purchase of a product based on the feedback provided by experts and other consumers available on the website. This report explores different techniques used for smart guidance for consumers. Different stages involving Classification, Relevance ranking, and opinion mining will be studied and discussed.

This project uses Natural Language Processing (NLP), Pattern Recognition (PR), and Machine Learning (ML) techniques to develop the model. Through these techniques, it is feasible to design an automated system that can compile and summarize relevant feedbacks to generate the most relevant response to a consumer's query. Such a system can facilitate Convenience and guidance. This model first classifies the query entered by the customer as either open-ended or binary, then summarizes feedbacks of reviewers into their relevant parts and ranks reviews by relevance to apply a voting-based model to generate a relevant response Opinion mining is applied to find keywords in reviews relevant to user queries to differentiate facts from subjective opinions. Recommendations for future development and conclusions are also included in the report.The outcome of the project can add value to existing e-commerce practices and improve customer experience.

Acknowledgments

In the name of Allah, the most beneficent and the most merciful. We are highly grateful to the One who created us and blessed us with a privileged life. We would like to thank our parents who have always been a pillar of strength and support. We are thankful to our family who taught us that how hard work, devotion and working towards your goal with pure intention can take you to places. Furthermore, we would like to thank and appreciate our supervisor Dr. Momina Moetessum who has given us a chance to work on a project that can help in creating value for the society. Her professional guidance, time and effort in this critical time of global pandemic will always be remembered. Last but not the least, we would like to appreciate our friends who make study and hardtimes less challenging. May ALLAH (SWT) guide us to the right path.

SYED TOUSEEF AHMED & ABDUL SAMAD ISLAMABAD, PAKISTAN
Islamabad, Pakistan

November, 2021

Contents

List of Figures

Chapter 1

Introduction

Big e-commerce websites like Amazon host a large collection of products and each product has an independent reviews or Q/A section for specific information about a product. In order to gather relevant information regarding a product, the user has to manually examine each and every response/review to decide, which is time consuming and inefficient. We need to develop a system which automatically responds to the customer queries. In this project we have developed a mechanism in which user selects the product and asks a query regarding that product.The system read's all the reviews and responds to the user query in binary format.

1.1 Background

In today's age e-commerce has become a widely used platform for purchasing goods and products[1]. Ecommerce has evolved to make products easier to discover and purchase through online retailers and marketplaces. Independent freelancers, small businesses, and large corporations have all benefited from e-commerce, which enables them to sell their goods and services at a scale that was not possible with traditional offline retail.Especially during Covid, due to restrictions and fear of infection, worldwide e-commerce retail sales reached their highest peaks shown in Figure 1.1.

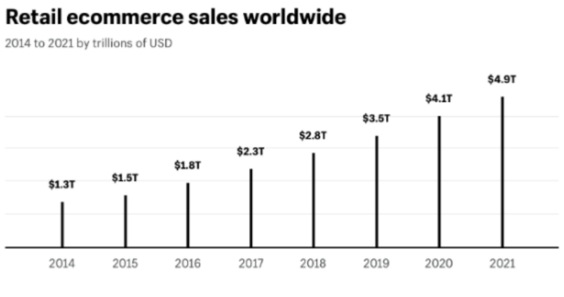
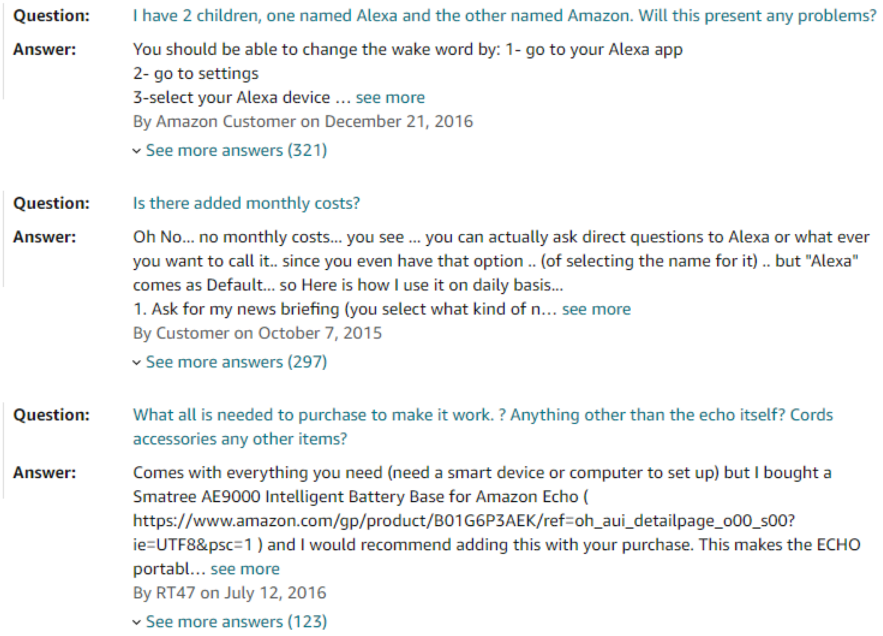


Figure 1.1: Retail Commerce Sale Worldwide[1]

When we buy an item physically, we can check each and everything by asking about it and gain insight directly from the store owner, and if it is satisfactory then we buy it. However, on the other hand in e-commerce this facility of real-time engagement in guidance is absent. E-commerce does provide the facility to ask questions regarding the product ,however, the users would either have to wait for a reply or search the Q/A section for a similar query with a useful and relevant response to it. This proves as an inconvenience for consumers who wish to gain more info or clear their doubts regarding a product however do not wish to spend an excess amount of time on research just to clear a few doubts. Consumer reviews are an invaluable source of data to help people form opinions on a wide range of products. Online reviews are our first port of call when considering products and online purchase . Reviews may contain a wide range of both objective and subjective product related information, including features of the project, evaluation of its positive and negative attributes and various personal experiences. Amazon hosts a large collection of products and each product has an independent review or Q/A section for specific information about a product[1]. In order to gather relevant information regarding a product, the user has to manually examine each and every response/review to decide. This is a time consuming and inefficient practice. As we see in Figure 1.2, Amazon provides the facility of asking questions but it is difficult to read all the reviews and get the exact answer. As we see in Figure 1.2, the questions are of two types: “Binary” and “Open ended”. In binary questions, a large fraction of them are real world opinion QA data, where the answers have a fairly straight forward reply as either ‘yes’ or ‘NO’. Mining opinions from relevant reviews provides us with that answer. In addition to binary questions, a significant number



Question: I have 2 children, one named Alexa and the other named Amazon. Will this present any problems?

Answer: You should be able to change the wake word by: 1- go to your Alexa app
2- go to settings
3-select your Alexa device ... see more
By Amazon Customer on December 21, 2016
⌄ See more answers (321)

Question: Is there added monthly costs?

Answer: Oh No... no monthly costs... you see ... you can actually ask direct questions to Alexa or what ever you want to call it.. since you even have that option .. (of selecting the name for it) .. but "Alexa" comes as Default... so Here is how I use it on daily basis...
1. Ask for my news briefing (you select what kind of n... see more
By Customer on October 7, 2015
⌄ See more answers (297)

Question: What all is needed to purchase to make it work. ? Anything other than the echo itself? Cords accessories any other items?

Answer: Comes with everything you need (need a smart device or computer to set up) but I bought a Smatree AE9000 Intelligent Battery Base for Amazon Echo (https://www.amazon.com/gp/product/B01G6P3AEK/ref=oh_aui_detailpage_o00_s00?ie=UTF8&psc=1) and I would recommend adding this with your purchase. This makes the ECHO portabl... see more
By RT47 on July 12, 2016
⌄ See more answers (123)

Figure 1.2: Amazon provide section of questioning[1]

of product related questions are open ended . Through an automated system, it is a complex and severely challenging process to answers such questions directly[7]. Due to this reason, we are more interested in creating a good relevance function for binary only.Our aim is to retrieve Queries and responses useful information from reviews, so that customer gain guidance to reach out toward the conclusion they are satisfied with. With the advancements in Artificial Intelligence (AI) and more specifically, Natural Language Processing (NLP), Pattern Recognition (PR) and Machine Learning (ML), it is now possible to design an automated system that can compile and summarize relevant feedback to generate the most relevant response to a consumer's query. Such a system can facilitate online customers Convenience and guidance.

1.2 Objectives

The main objective of this project is to develop a system that automatically responds to the customer query regarding the product and saves their time and provides the desired result in a short interval of time by providing the most relevant answer about the query. The following are the main objectives that are to be targeted for the completion of our final year project.

- To classify a user query whether it is 'binary' or 'open-ended'.
- To summarize feedback of reviewers into their relevant parts and rank reviews by assigning a relevance score.
- To apply opinion mining to find keywords in reviews relevant to user queries in order to differentiate facts from subjective opinions and find the intention of the review as negative or positive to the query
- Apply a voting-based model to generate a final yes/no output.

1.3 Problem Description

Huge e-commerce websites, such as Amazon, contain a large number of products, each with its independent reviews or Q/A area for more detailed information. To get meaningful information about a product, the consumer must carefully evaluate each response/review before making a decision, which is time-consuming and difficult. We need to create a system that answers customer queries automatically. In this project we have developed a mechanism to intelligently select the most relevant responses and then summarize them to provide the final response. Our system allows the user to select a particular product available on the website and ask a query. The system will select the most relevant reviews available and summarize them to derive the final inference. It's a very challenging task to get the answer of each query because sometimes we cannot get the exact answers, sometimes it is a binary response or sometimes open ended, so sometimes it is very difficult to respond. Keeping these issues in mind we addressed these things by first training our model to classify the query entered by the customer as open ended or binary.Only binary queries are selected for providing response. The model then summarizes feedback of reviewers into their relevant parts and rank reviews by relevance to apply a voting-based decision to generate a relevant response. To apply opinion mining to find keywords in reviews relevant to user queries in order to differentiate facts from subjective remarks.

1.4 Project Scope

The Amazon dataset contains several categories of products like beauty, food, hair products, however, we only considered the electronics category. Although this system can be applied to other categories as well, the magnitude of data and relevance of electronics in today's e-market makes it ideal to work on. This application can be used as a web extension for websites with Q/A feedback and reviews, however, currently we considered it as a stand-alone application. The system considers feedback provided on the same platform, it cannot collect relevant reviews online from various platforms.Also currently,we are only focusing on binary queries.Open-ended queries are beyond the scope of this project.

1.5 Document Organization:

This report is ordered as follows: Chapter 2 will discuss the previous work that has been done related to responding to customer queries. Chapter 3 will discuss requirement specifications of our project. Chapter 4 will discuss the design of the project.Chapter 5 details the system implementation while chapter 6 presents the system testing and evaluation.Finally,chapter 7 concludes the report.

Chapter 2

Literature Review

This chapter summarizes related works.Two main modules of our system are based on the techniques employed in text summarization where mining and summarization are performed based on opinion minings and perspectives from texts (particularly from review collections) and Q/A systems in general are the most closely linked lines of work to ours.

2.0.1 Summarization of documents

Summarization reduces the amount of data, and sentiment analysis assists the user in identifying positive, negative, and other information in the documents. Summarization techniques generate a condensed version of the text, and sentiment analysis determines the emotion expressed in the text[8]. The difficulty of multi-document summarization is perhaps the most closely related to our goal of picking meaningful reviews from a big corpus of reviews. This task intelligently combines relevant or "salient" elements of reviews. The most closely similar approaches are those that use document summarization techniques to generate an overview of opinions or product features through evaluation of reviews.

2.0.2 Ranking by relevance

Learning whether a document (or a phrase inside a text) is relevant to a given query is a critical component of the above line of work. The term "relevance" might refer to a variety of characteristics, including the text's "quality," its lexical significance, or its diversity as compared to previously selected materials[9].

2.0.3 Opinion research

Consumer opinion research, particularly through rating and review datasets, is a broad and varied subject. Review text has been used to supplement 'traditional' recommender systems by identifying elements or facets that are significant to people's opinions, and, more importantly, to locate ‘helpful’ reviews or experts on specific issues. There has also been work on generating product feature summaries, including the use of multi-document summary, as previously indicated.[6] Although the goal is not normally to respond generic inquiries as we do here, this approach is related in terms of the data used and the necessity to understand some sense of "relevance"[10].

2.0.4 Question-and-answer systems.

Many of the principles from multi-document summarization, relevance rating, and subject expert-finding mentioned above have been used to create cutting-edge automated Q/A systems. The first is summarization that is ‘query-focused’. It is similar to our work in that phrases must be chosen from a set of papers that match a query; however, the relevance function is often not learnt from training data as it is here.

2.1 Existing Applications.

There have been several websites/models developed to help user queries such as Quora, Amazon feedback section, stack Overflow etc that focus on answering queries either complex or simple, these systems incorporate newer improvements to each previous work before them with our project also incorporating different aspects of each work into our models and hoping to improve them further focusing on consumers more. Previous works relating to binary classification , addressing complex and subjective product-related queries with customer reviews and further improvements to this model in modeling ambiguity, subjectivity, and diverging viewpoints in opinion question-answer systems . These Systems have used complex machine learning models to counter problems that occur when dealing with subjective and complex queries. Through trial and error have found methods that produce most accurate results during training and final output Some of the feature’s comparisons are shown in the table below

Q/A websites	Voting	Classification	Automated response
Quora	Yes	No	No
Amazon Feedback section	Yes	Yes	No
Stack Overflow	Yes	No	No
Our proposed system	Yes	Yes	Yes

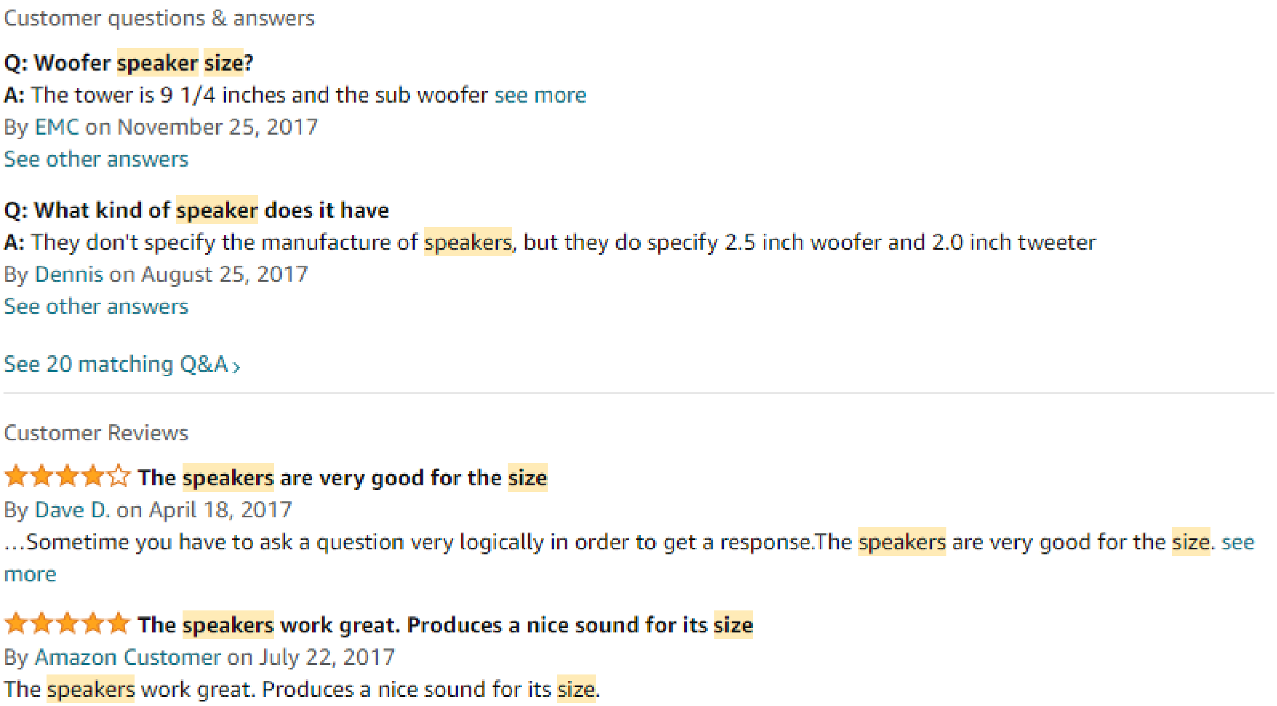
Table 2.1: Feature Comparison

2.1.1 Quora

In Quora, Students create a search engine that unseats Google?” Users can then follow questions and offer responses. The Quora community in a new update can vote answers, similar to the popular social platform Reddit[11]. The answers with the most upvotes are highlighted and viewed the most. The main disadvantage of this system is it works through popularity and less on relevance, although judgment of the masses can bring satisfactory answers however the popular answer may not even be relevant to what some people were asking.

2.1.2 Amazon feedback section

Amazon allows the customers to leave their feedback about the product as shown in figure [2.2]. At the Amazon feedback section, customers give their review about a product. It feeds into the buyer's judgement regarding whether or not to purchase the product from the seller by using your average feedback score ratings. But you have to read all the reviews manually which is impossible and time consuming



Customer questions & answers

Q: Woofer speaker size?
A: The tower is 9 1/4 inches and the sub woofer see more
By EMC on November 25, 2017
See other answers

Q: What kind of speaker does it have
A: They don't specify the manufacture of speakers, but they do specify 2.5 inch woofer and 2.0 inch tweeter
By Dennis on August 25, 2017
See other answers

See 20 matching Q&A ›

Customer Reviews

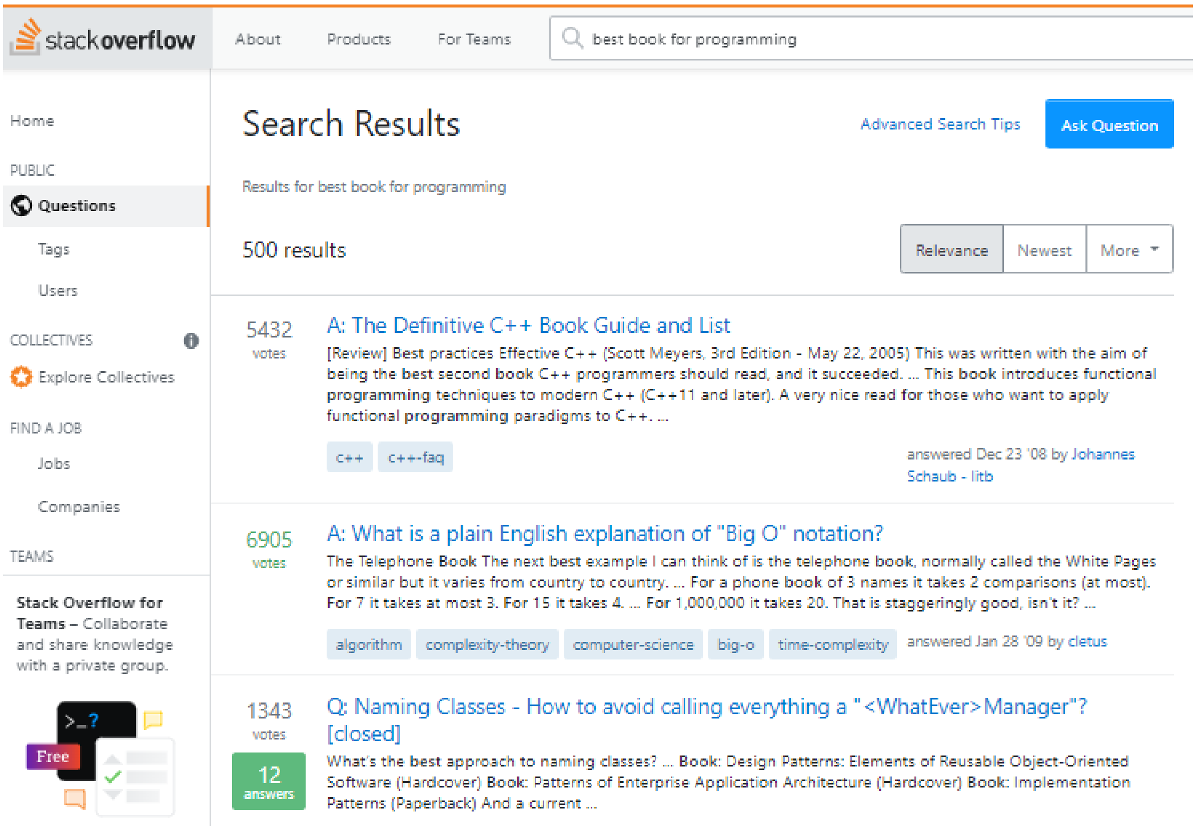
★★★★☆ **The speakers are very good for the size**
By Dave D. on April 18, 2017
…Sometime you have to ask a question very logically in order to get a response.The speakers are very good for the size. see more

★★★★★ **The speakers work great. Produces a nice sound for its size**
By Amazon Customer on July 22, 2017
The speakers work great. Produces a nice sound for its size.

Figure 2.1: Amazon feedback section [2]

2.1.3 Stack Overflow

Stack Overflow is a question-and-answer website that supports programmers in questioning. Provide questions and answers on different topics related to computer science as shown in figure [2.3]. It was built to be better than the earlier Q/A websites such as Exchange. This allowed users to ask a question and read the people-related answers as shown in the figure, but again it is very time-consuming because after asking a question users have to wait for a long time to get an answer.



Search Results

Results for best book for programming

500 results

5432 votes

A: The Definitive C++ Book Guide and List

[Review] Best practices Effective C++ (Scott Meyers, 3rd Edition - May 22, 2005) This was written with the aim of being the best second book C++ programmers should read, and it succeeded. ... This book introduces functional programming techniques to modern C++ (C++11 and later). A very nice read for those who want to apply functional programming paradigms to C++. ...

c++ c++-faq

answered Dec 23 '08 by Johannes Schaub - litb

6905 votes

A: What is a plain English explanation of "Big O" notation?

The Telephone Book The next best example I can think of is the telephone book, normally called the White Pages or similar but it varies from country to country. ... For a phone book of 3 names it takes 2 comparisons (at most). For 7 it takes at most 3. For 15 it takes 4. ... For 1,000,000 it takes 20. That is staggeringly good, isn't it? ...

algorithm complexity-theory computer-science big-o time-complexity

answered Jan 28 '09 by cletus

1343 votes

12 answers

Q: Naming Classes - How to avoid calling everything a "<WhatEver>Manager"? [closed]

What's the best approach to naming classes? ... Book: Design Patterns: Elements of Reusable Object-Oriented Software (Hardcover) Book: Patterns of Enterprise Application Architecture (Hardcover) Book: Implementation Patterns (Paperback) And a current ...

Stack Overflow for Teams – Collaborate and share knowledge with a private group.

Figure 2.2: Stack overflow questioning[3]

2.2 Summarization

As seen from the above comparisons, the existing systems don't have an ability to answer as an automated reply. However, our proposed solution satisfies the missing features remaining in previous examples.

Chapter 3

Requirement Specifications

This chapter focuses on gathering requirements for our project, details including existing systems already developed, our proposed system, limitations of our project, functional and non-functional requirements, ending with use cases for our proposed project.

3.1 Proposed system

Our project is a Query Based Answering system designed to give a swift answer to any binary query the user might have. Our system focuses on exploring Q/A in E-commerce websites to scrounge through different reviews and return with a satisfactory answer. It Concerns itself with providing relevant answers, scoring them by relevance and computing the answer for each answer as either yes or no and compiles it as a final answer in majority voting The proposed systems user flow diagram explains that the user will first enter a question,then the system will check whether the asked question is binary or not.If the question is not binary the system will prompt the user to rephrase your question.If the question is binary, then the system will match the question with the related reviews then with the help of machine learning algorithm the system will analyze whether the answer is yes or no, then based on the voting mechanism the system will display the final answer in yes or no.

3.2 Models and Integration

Our System Proceeds with a model-based approach where each model performs its specific job to reach the final output given the input. Our first model consists of a query identification function which given a specific query from user identifies it as either binary or open-ended which it finds using different regular expression which were made by google resulting in 91% accuracy and 80% recall. Our relevance ranking Model consists of several different weak classifiers acting as a single expert in out mixture of expert model which

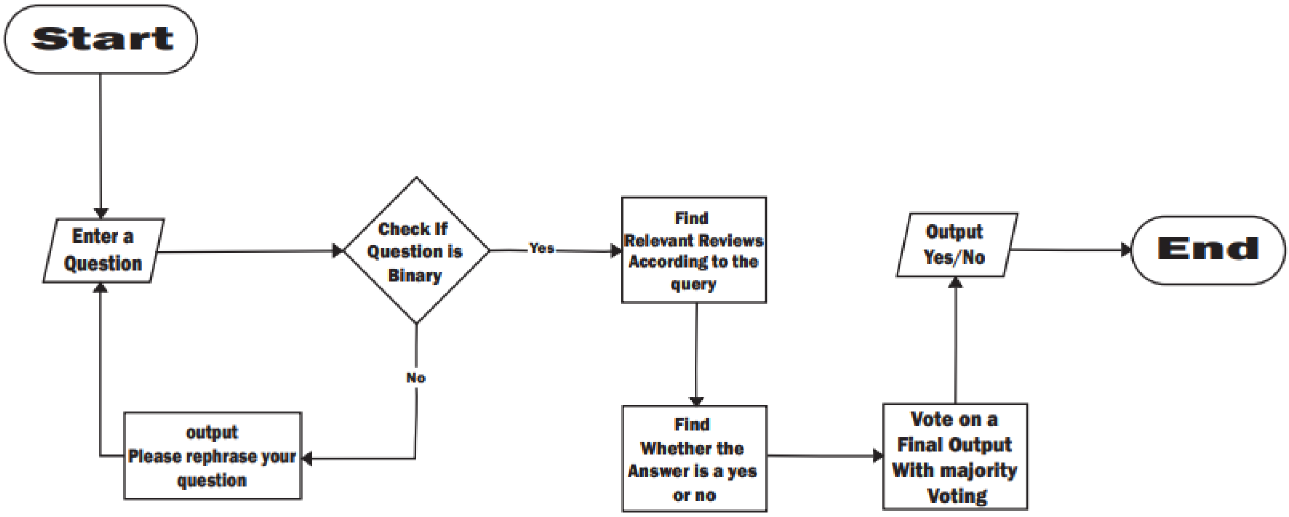

Start
Enter a Question
Check If Question is Binary
Yes
No
Find Relevant Reviews According to the query
output Please rephrase your question
Find Whether the Answer is a yes or no
Vote on a Final Output With majority Voting
Output Yes/No
End

Figure 3.1: Data flow diagram for project

takes each review as an opinion which opposes or is in favor of the query, with finding relevant reviews our opinion mining model finds the intention of the review as either a yes or no, with final voting in a majority pole to generate a final output.

3.3 Limitations

Our projects aim is to help people when purchasing goods by providing an automated system that simulates real time query answering, however our project lacks the ability to cater open ended or more subjective queries given by the user. These complex queries require further research and more far more resources expertise to address these issues. However, for the time being our project fails to address these issues.

3.4 Functional Requirements

A functional requirement specifies a system's or component's function where a function is defined as a definition of how inputs and outputs behave[12]. Functional requirements for our project include:

- UI for entering Question a user interface is required for interaction between the user and the system which we have created a web page for interface.
- Query classification is necessary to first distinguish the question as either a binary question or open ended one.
- Relevance function is required to compile relevant reviews from the Q/A or review section and assign a relevance score to them.

- Computing intention of relevant reviews as either yes or no.
- Voting through compiling a final answer and giving output.

3.5 Non-Functional Requirements

Non-Functional requirements include:

- Performance as our project aims to simulate real time query answering, fast performance is a main requirement for this system to provide.
- Accuracy our system focuses more on accuracy than recall where unless a query is a binary question it drops it and asks for rephrasing in order to get queries best for accuracy.
- Ease of use an interactive and easy to understand UI provides users with a more comfortable and good experience, where a fun and interactive system can provide as good leverage for.

3.6 Use Cases

The following is each use case along with their respected table, where the first covers the overall project and the other explore each model individually:

Feature	**Should I but it?(Main system)**
Actors	User
Description	The overall project for questioning and receiving an answer
Pre-condition	Select a product
Post-condition	Final answer is attained

Table 3.1: Use Case descriptive table of Main system

Feature	**Query Classification**
Actors	User
Description	Classifies Queries as Binary or Open-ended
Pre-condition	Query has been entered
Post-condition	Classified Binary query is reached as output.

Table 3.2: Use Case descriptive table of Query Classifier

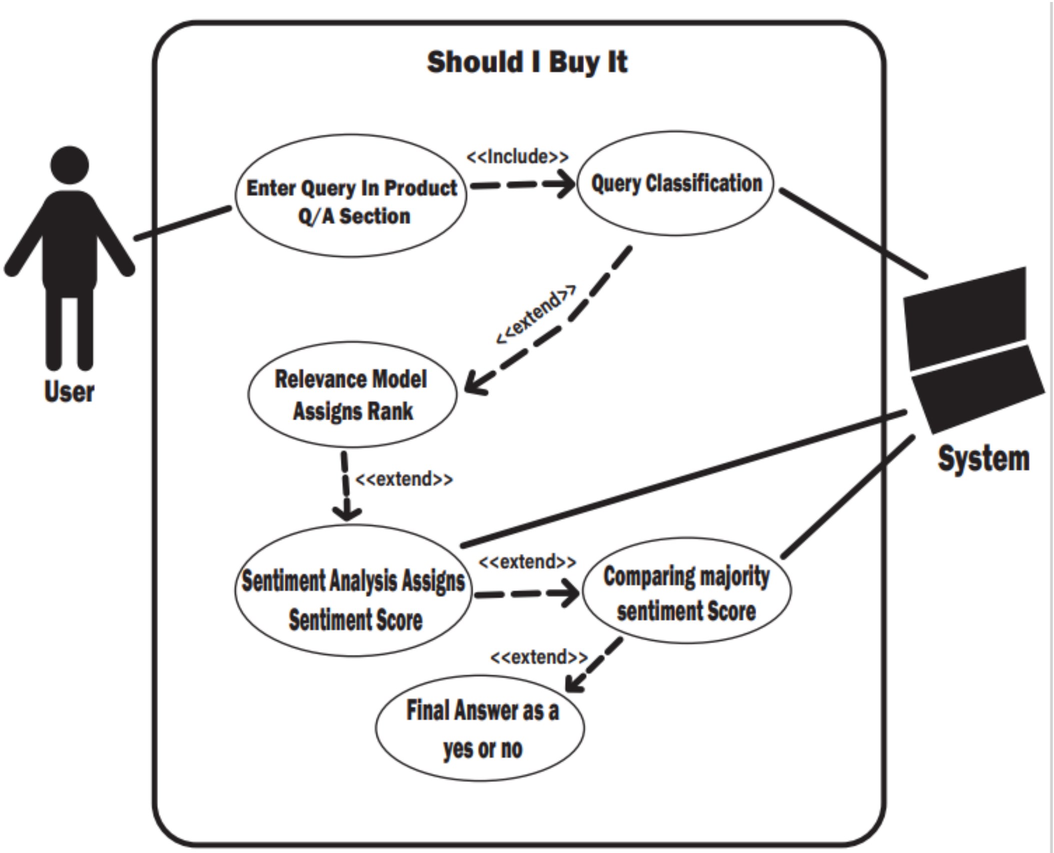

Should I Buy It
User
Enter Query In Product Q/A Section
<<Include>>
Query Classification
<<extend>>
Relevance Model Assigns Rank
<<extend>>
Sentiment Analysis Assigns Sentiment Score
<<extend>>
Comparing majority sentiment Score
<<extend>>
Final Answer as a yes or no
System

Figure 3.2: Use case for system

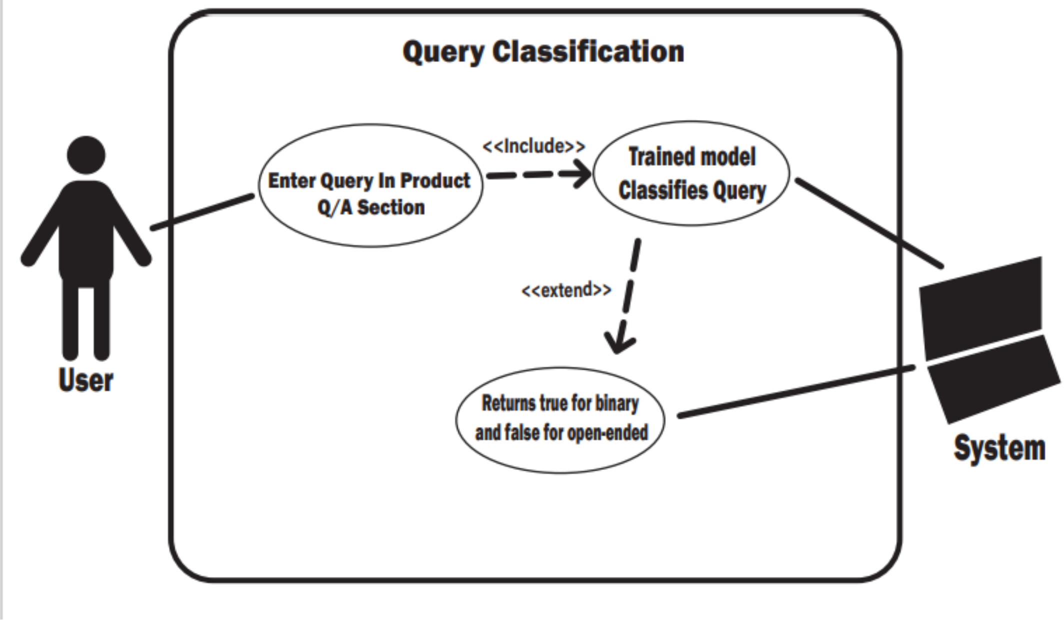

Query Classification
User
Enter Query In Product Q/A Section
<<Include>>
Trained model Classifies Query
<<extend>>
Returns true for binary and false for open-ended
System

Figure 3.3: use case diagram for query classification

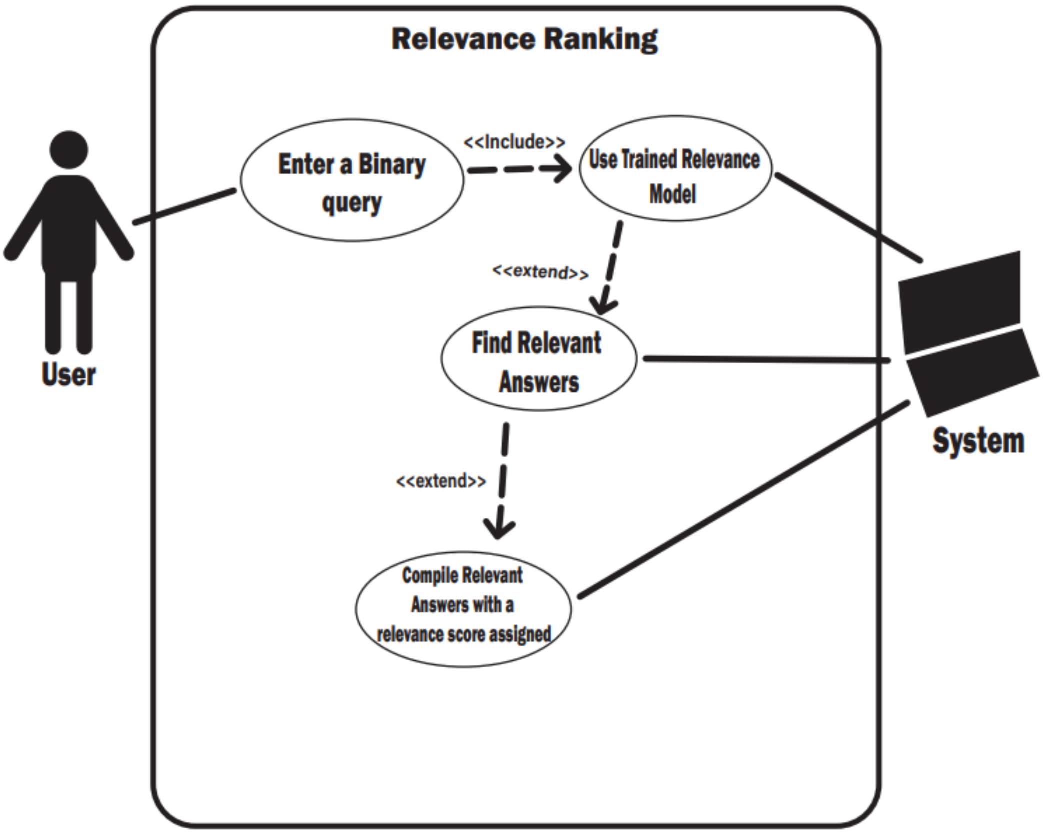


Figure 3.4: Use Case diagram for relevance model

Feature	Relevance Model
Actors	User
Description	Find relevant reviews according to the query.
Pre-condition	Binary query is entered.
Post-condition	List of relevant reviews is formed.

Table 3.3: Use Case descriptive table of Relevance model

Feature	Prediction and Voting
Actors	User
Description	Predicted answers are given to each review and rounded up in voting.
Pre-condition	Relevant reviews are found.
Post-condition	Final answer is attained.

Table 3.4: Use Case Diagram for Prediction and Voting

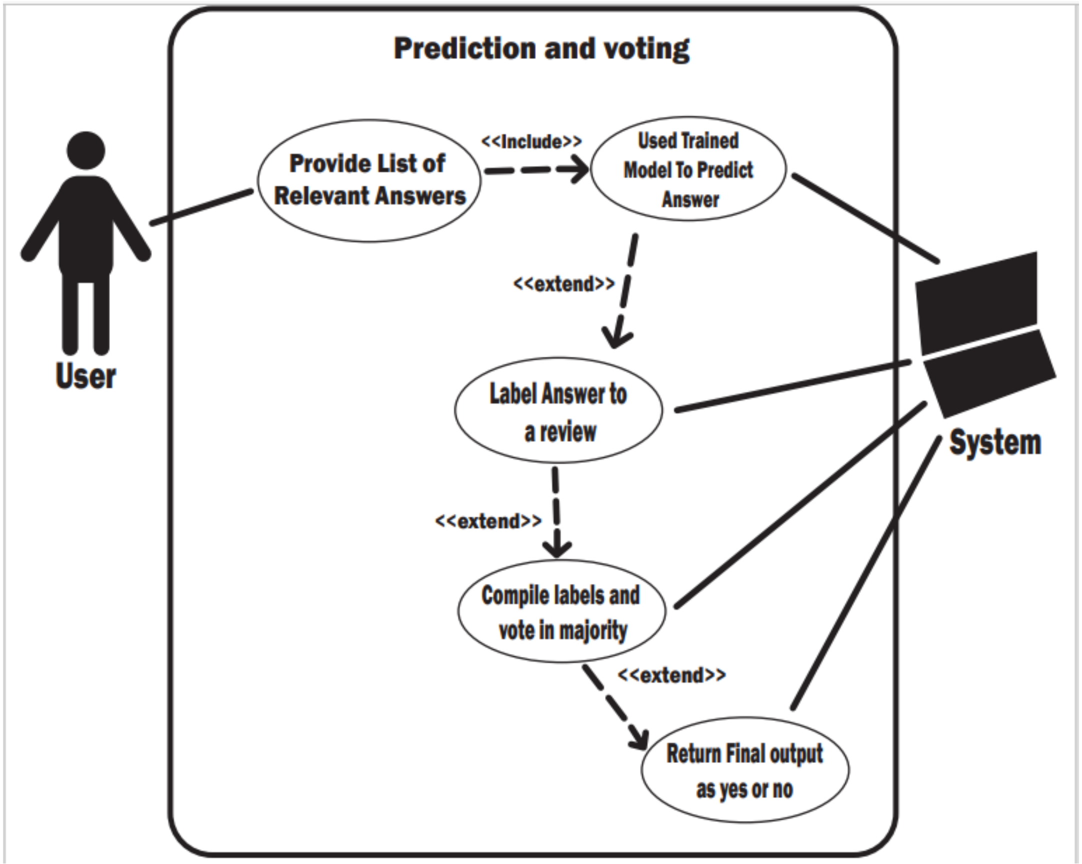

Prediction and voting
User
Provide List of Relevant Answers
<<Include>>
Used Trained Model To Predict Answer
<<extend>>
Label Answer to a review
<<extend>>
Compile labels and vote in majority
<<extend>>
Return Final output as yes or no
System

Figure 3.5: Use Case Diagram for Prediction and Voting

Chapter 4

System Design

The process of establishing the architecture, components, modules, interfaces, and data for a system in order to meet specific criteria is known as systems design..

4.1 Architecture

Our project aims to provide users with quick real time answers using several different machine learning models to answer them. Our system consists of 4 models where each model is dependent on the output of the previous one to reach the next stage. Each model has a specific architecture to follow, query classification uses simple regular expressions to classify queries, relevance model uses a mixture of expert's model to find relevant reviews ordering them in terms of relevance, with opinion minions using a similar model to compute intention and final model uses simple majority voting. Below Shows the overall model in figure 4.1 with each model further detailed in figure 4.2, 4.3 and 4.4

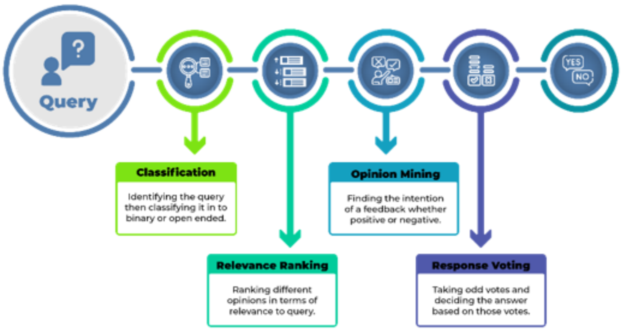

Query
?
YES
NO
Classification
Identifying the query then classifying it in to binary or open ended.
Opinion Mining
Finding the intention of a feedback whether positive or negative.
Relevance Ranking
Ranking different opinions in terms of relevance to query.
Response Voting
Taking odd votes and deciding the answer based on those votes.

Figure 4.1: Showcasing the main architecture of the system

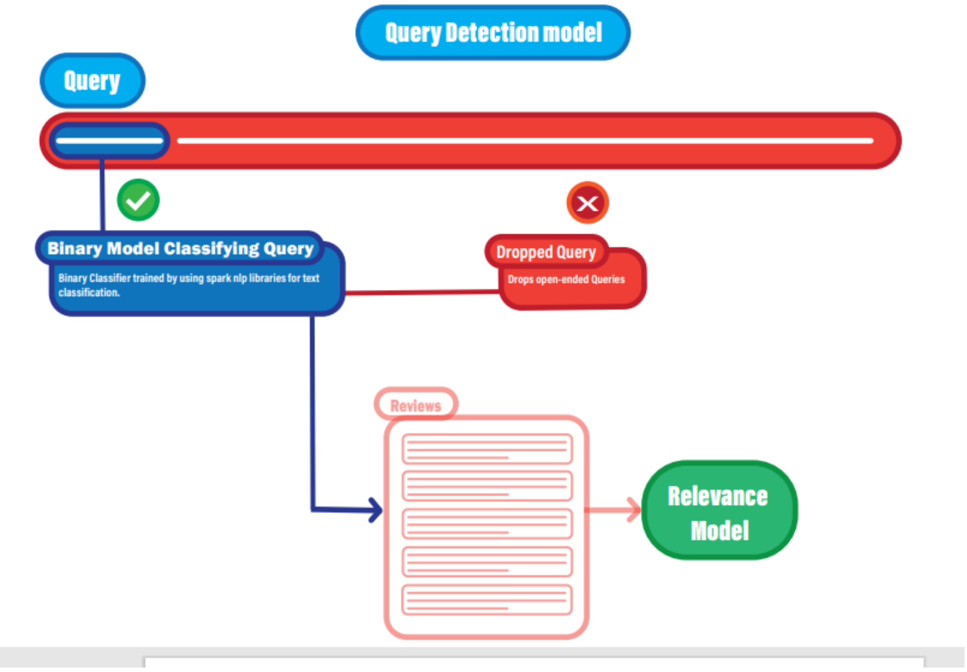

Query Detection model
Query
Binary Model Classifying Query
Binary Classifier trained by using spark nlp libraries for text classification.
Dropped Query
Drops open-ended Queries
Reviews
Relevance Model

Figure 4.2: Query Classification model with labeling initial answers which can provide a yes, no answer.

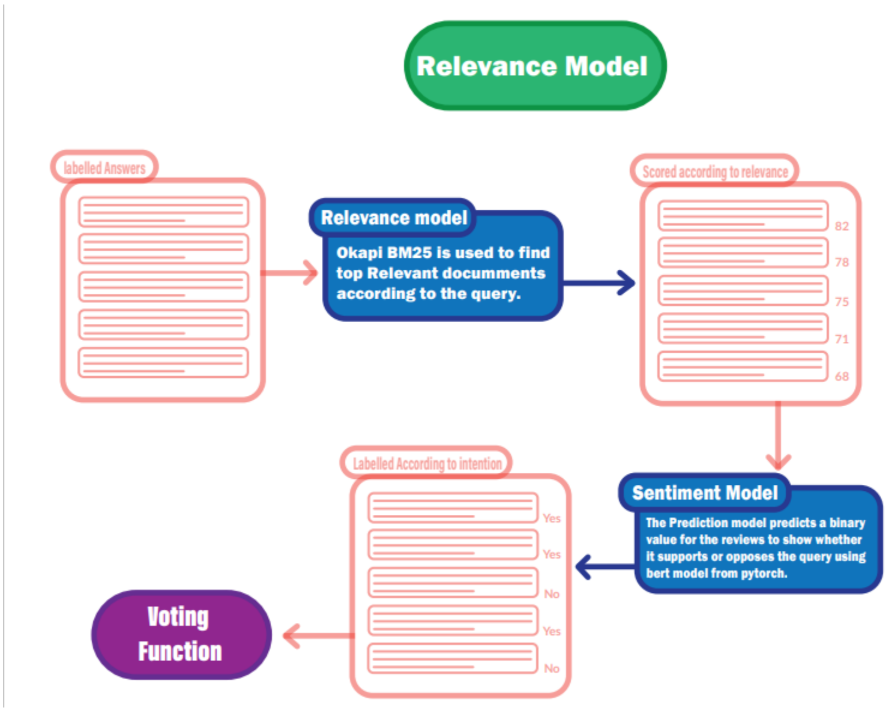

Relevance Model
labelled Answers
Relevance model
Okapi BM25 is used to find top Relevant docummments according to the query.
Scored according to relevance
82
78
75
71
68
Sentiment Model
The Prediction model predicts a binary value for the reviews to show whether it supports or opposes the query using bert model from pytorch.
Labelled According to intention
Yes
Yes
No
Yes
No
Voting Function

Figure 4.3: Relevance model along with prediction model.

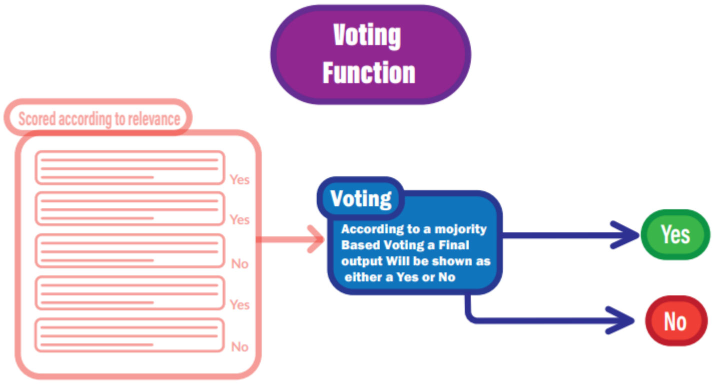

Voting Function
Scored according to relevance
Yes
Yes
No
Yes
No
Voting
According to a mojority Based Voting a Final output Will be shown as either a Yes or No
Yes
No

Figure 4.4: Voting function giving final output.

4.2 Design Constraints

There are some design constraints in our project.

- User interfaces must be operable.
- All the features that the system offers to its users must be user-friendly
- **Mobile or computer to run a browser:** To run a web page we have a need of mobile or computer to run a browser.Web browsers provide a user interface, often known as the front-end, where you may click, scroll, and navigate across web pages and websites to display information graphically on your device.
- **Satisfactory number of reviews to answer from:** There must be some reviews on the basis of which the system will answer the user's queries.
- **Binary query:** The query must be binary,so that the system answers in a binary format.

4.2.1 Data Requirements

- Amazon Electronics Reviews Dataset.
- Amazon QA Electronics Dataset

4.2.2 Programming Languages

- **Python** :We used python to make our backend.NLTK is one of the most useful Python modules for any natural language processing work.The NLTK library offers a number of useful tools for manipulating and analysing linguistic data. Text classifiers are among the sophisticated features, which can be used for a variety of classifications, including sentiment analysis.The method of utilizing algorithms to classify numerous samples of linked text into overall positive and negative categories is known as sentiment analysis. With NLTK, we used these methods to extract insights from linguistic data using strong built-in machine learning procedures.

- HTML: We made our front end with the help of HTML.We used to make the webpage and all of the information on it, including the graphics and text.
- CSS: CSS is used to design a webpage and tell it where to put the words and images, such as an image at the top of the page, another in the middle, and so on.so,we the help of CSS we made our website more attractive and made easy for user where they see the product more better and ask a question in a easy way.

4.3 Design Methodology

Our systems take a query as input and assigns it a label of either binary or open-ended query, taking only binary queries into consideration we further compile answers that give a yes or no answer and continue with finding reviews relevant to the queries and finding the answer to those reviews as yes or no and providing a final output. Further description of each model is defined below:

4.3.1 Binary Classification

The model works to classify queries as either binary or open ended by incorporating regular expressions to evaluate queries. It works by first checking the first word in the query and evaluates it against either be, model or auxiliary verbs, this method is developed by google to find binary questions with 90 percent precision.This method though effective has 2 instances where it fails, the first occurs by alternative questions example “Does this come in black or blue”, the second error starts with words like “Do you know…” or “Does anyone know….” these types of statements pass through the first regular expression of verbs however are not taken as binary questions. To Avoid these errors regular expressions can be used to drop questions like the above 2 examples.

4.3.2 Relevance Ranking

After receiving a binary query, the next stage of our system begins, finding relevant reviews against the query. There are several techniques developed relating to relevance ranking, examples such as cosine similarity or a far more advanced method Okami BM25 [6] these methods are good and prove satisfactory results however a far more effective approach found is the mixture of expert’s architecture, which uses each review as an individual expert that either opposes to votes in favor of it with their final output aggregated. It Works with 2 main parameters in effect, 1 finds the relevance of each review against the query using pairwise similarity and Bilinear Models, the second predicts the answer detailed in the prediction model section.

4.3.3 Prediction Model

The system matches with the most relevant answers to the user query and predict the relevant reviews later where the system performs voting. Using bilinear scoring function to label the relevant reviews in either "yes" or "No".

4.3.4 Voting

After Attaining all relevant reviews with each review labeled as either yes or no, final voting is performed to count in the majority where the number of positive and negative reviews are matched and a final output is shown to the user.

4.4 High Level Design

The high-level Design covers the general layout of the system covering all the functional requirements of the system[9]. Input is taken as text which falls to the classification layer then the relevance ranking layer scrounging relevant reviews to feed into the prediction model labeling each answer and finally voting on the final output shown to the user.

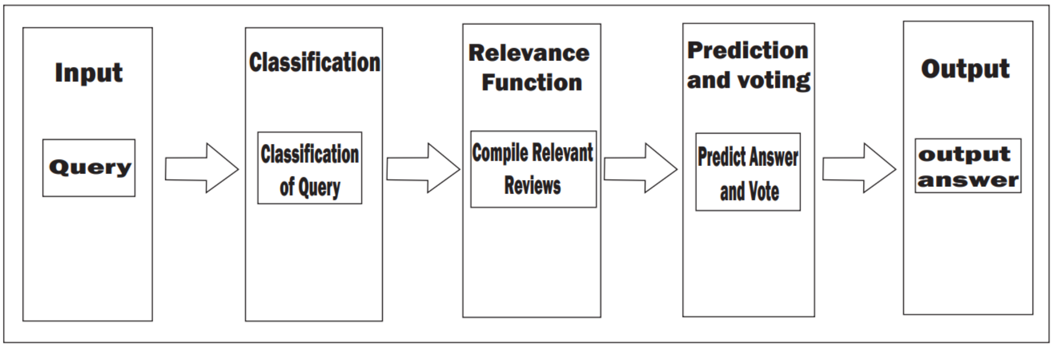


Figure 4.5: High Level Design diagram for the system

4.5 Low Level Design

Low level also known as class diagram Displays each model as an individual class with attributes and functions. There are 4 classes with their own respective attributes and functions with relations to other classes shown through connected lines; each class has an aggregated relationship with the indicated connected lines.

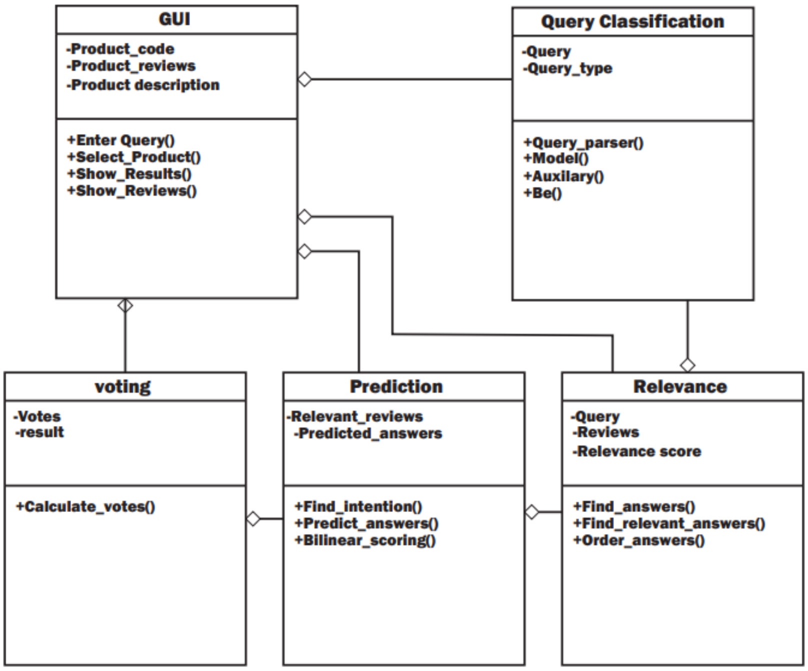


Figure 4.6: Low Level Design diagram for the system

4.5.1 Sequence diagram

Sequence diagrams are useful for showcasing the different timeline of events each model takes during execution. Here the user first interacts with the system through the GUI enters a query and the rest follows with query classification with a request of rephrasing the query if it failed to be binary query, continued with relevance ranking and prediction and final voting where the output is returned to the GUI to show to the user.

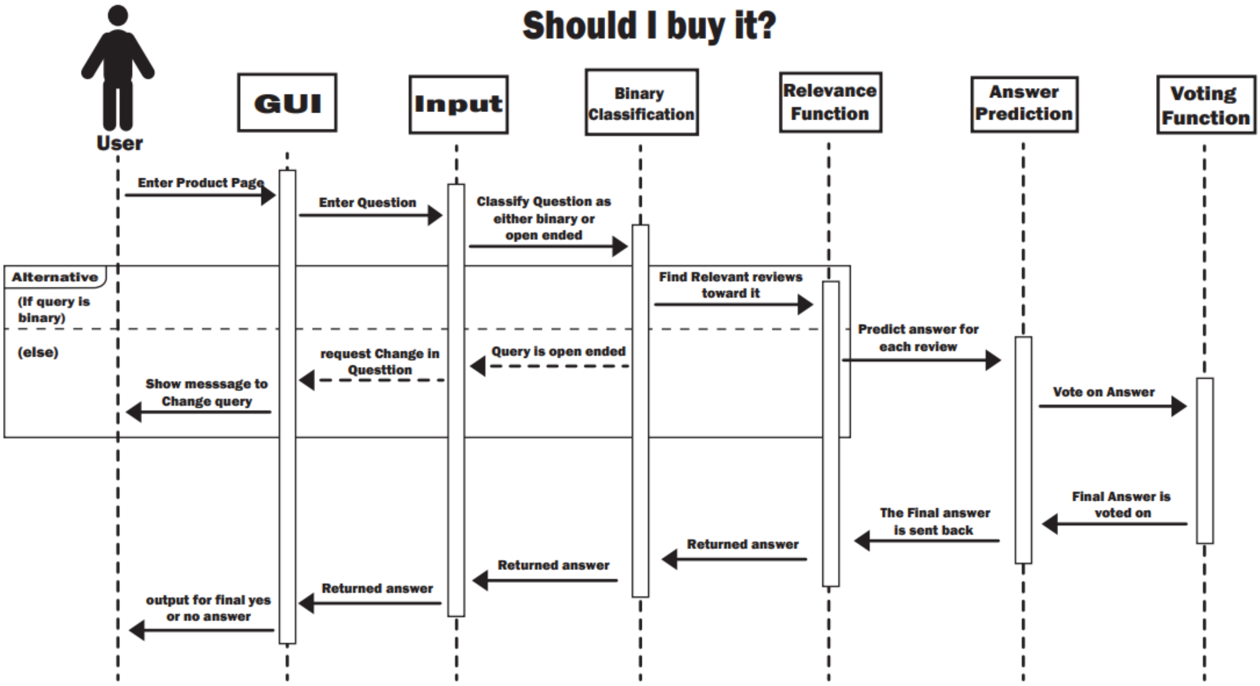


Figure 4.7: Sequence Diagram for the System

4.6 Level-0 DFD

Showcases the input and Output of the system with external parties in our case the user, where input is a binary query and output is a yes or no answer with a list of found relevant reviews

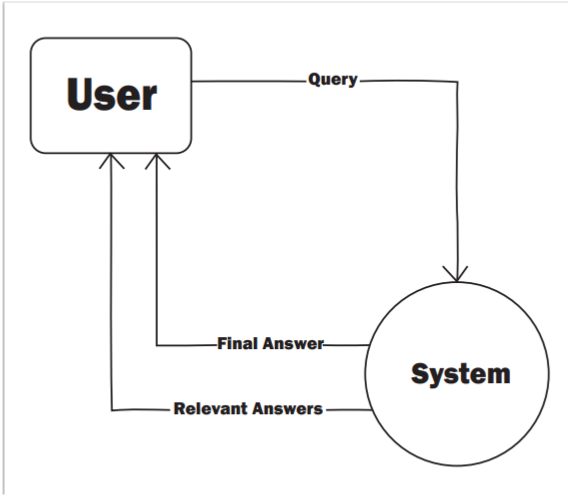


Figure 4.8: Level 0 DFD of the system

4.7 Level-1 DFD

This Diagram Showcases inputs and outputs of each model with each other and external entities. The Input and output of each model can be seen clearly in the model shown below with only the prediction and voting model returning output towards the user.

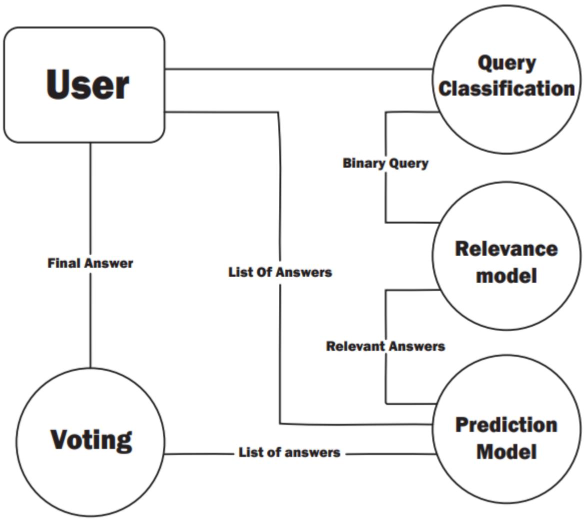


Figure 4.9: Level 1 DFD of the system

4.8 Package Diagram:

Package diagrams group different classes as groups called packages with their dependencies shown by the arrows connecting them. There are 3 packages for input, models and output with input containing GUI and Query Classification as both take input from the user. Since all of the models are dependent on the output of the previous one except GUI all of them are connected to the other in a descending order.

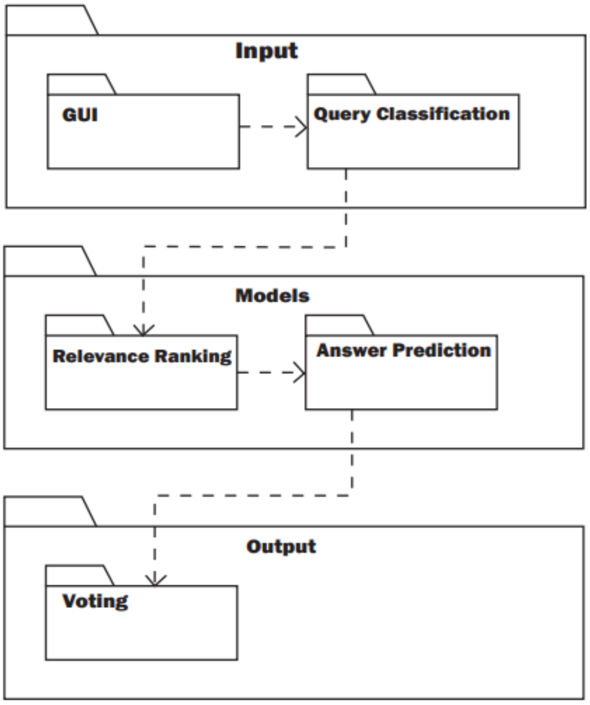


Figure 4.10: Package Diagram for the system

4.9 Data Set

Out System Uses 2 Datasets for training the relevance ranking and prediction models. The first one is a simple reviews data set containing an individual review with an assigned product code, rating and the reviews itself with proper Feature listing as follows:

- asin - ID of the product, e.g. 0000031852
- title - name of the product
- feature - bullet-point format features of the product
- description - description of the product
- price - price in US dollars (at time of crawl)

- imageURL - url of the product image
- imageURL - url of the high resolution product image
- related - related products (also bought, also viewed, bought together, buy after viewing)
- salesRank - sales rank information
- brand - brand name
- categories - list of categories the product belongs to
- tech1 - the first technical detail table of the product
- tech2 - the second technical detail table of the product
- similar - similar product table

4.9.1 QA dataset

The Second One is a Q/A Dataset with a product id accompanied by a number of questions where each question is given a different number of answers, with question type and question user and answer type and answer user is also assigned. A Basic Diagram to explain the structure of the dataset:

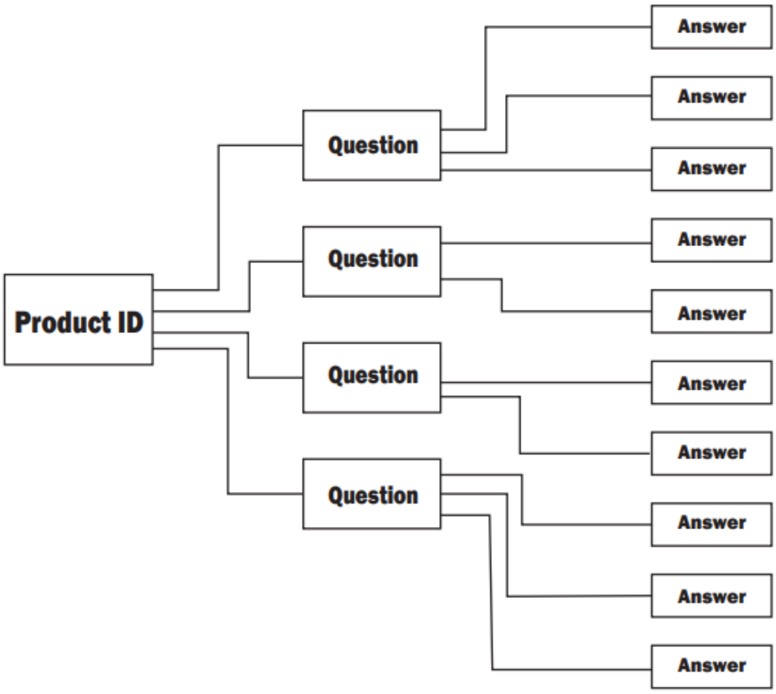
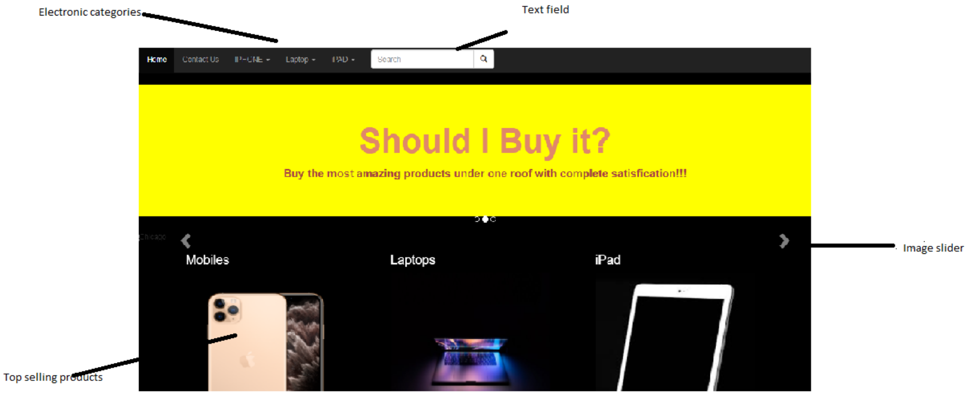


Figure 4.11: Dataset Diagram

4.10 GUI Design

Our System GUI is designed as a web page where different drop boxes containing products are available and when a product is selected the item is shown below it with an image and description of the item along with ratings and reviews. Below the item description is a text box to enter a query where a binary question can result in a yes and no answer below the box and a list of relevant reviews along with their answer at the side of it. Front-end page programming is an essential part of the project because its goal is to make a web application.

- **Responsive:**This becomes a good feature because all components in a page are responsible for the size of windows. This makes it simple and convenient for users to perform the test case when the application is running on several platforms. The CSS framework Bootstrap provides excellent support for responsive pages[9].
- **Modularisation:**The core of this concept is based on Krug's famous book . According to him, a good online application should be as simple and direct as feasible for users . The project is initially modularized based on this theory.

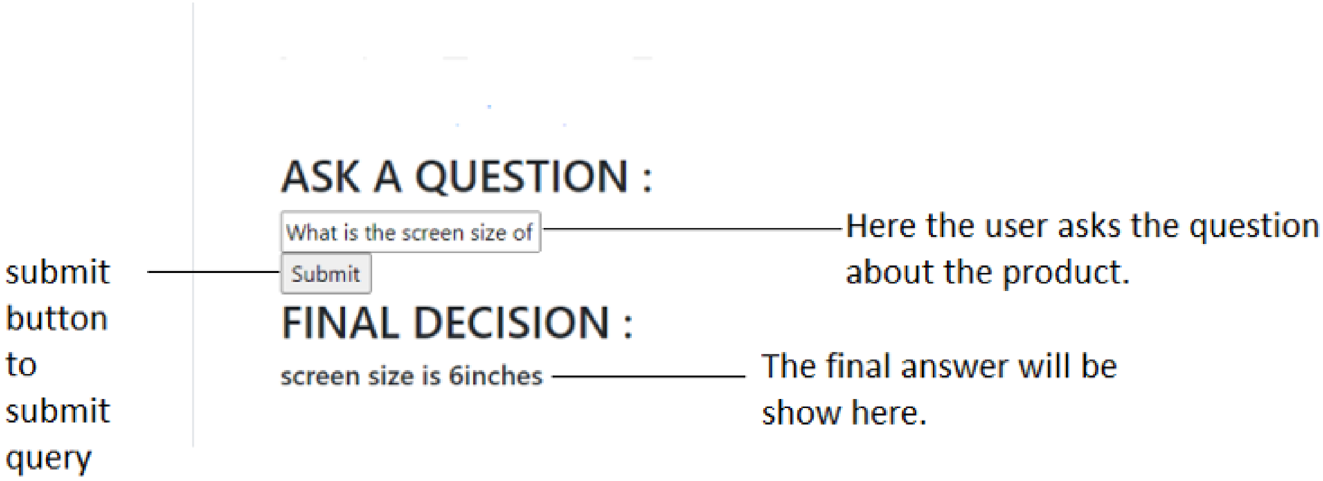


Buy the most amazing products under one roof with complete satisfication!!!

Figure 4.12: GUI

4.11 External Interfaces

- The system that is being used to build this project must have 8GB + RAM and 256GB+ SSD HardDrive for smooth working experience
- The internet connection must be good for research purposes.



ASK A QUESTION :

What is the screen size of

FINAL DECISION :

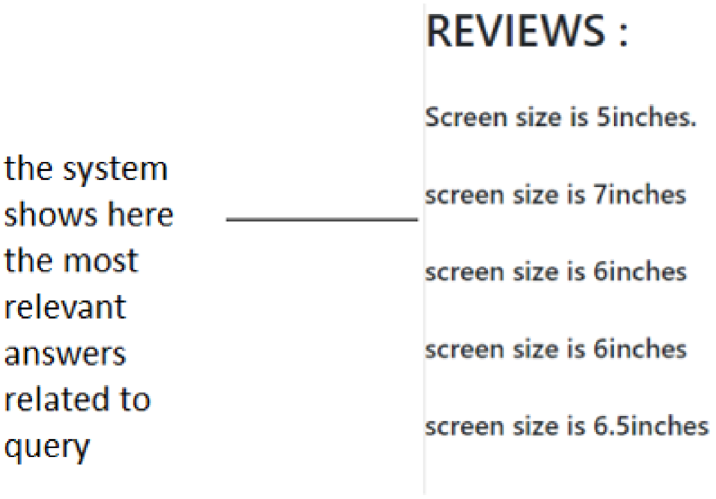
screen size is 6inches

Here the user asks the question about the product.

submit button to submit query

The final answer will be show here.

Figure 4.13: GUI for query entering and final output



REVIEWS :

Screen size is 5inches.

screen size is 7inches

screen size is 6inches

screen size is 6inches

screen size is 6.5inches

the system shows here the most relevant answers related to query

Figure 4.14: GUI for query entering and final output

Chapter 5

System Implementation

Implementation is the process of moving an idea from concept to reality. The System implementation is a realization of a technical specification or algorithm as a program, software component, or other computer system through programming and deployment. Our Project used several machine learning algorithms and techniques to achieve our ultimate end goal, using advances libraries like spark nlp and experimenting on several others to achieve our results[13].

5.1 System Architecture

'Should I Buy It " interface is developed in HTML, CSS. The internal System is developed with Python, starting with text classification using ClassifierDL and USE in spark NLP. The relevance model works using a ranked document retrieval process using Word2Vec based space Model (VSM) and for sentiment analysis we are using a Bert model for scoring. BERT' tried to analyze the masked word from context, using 15–20 percent of words as masked words, causing the model to converge slower than left-to-right approaches at first.

5.1.1 Web Application

Users will simply use a web application. Users can select the specific product and ask any query about the product and the system will automatically respond to the query.

5.1.2 Web browser

A web browser (usually known as a browser) is software that allows you to access the World Wide Web. The web browser receives the necessary content from the website's web server and then displays the page on the user's device when a user follows the URL of a web page from a certain website.

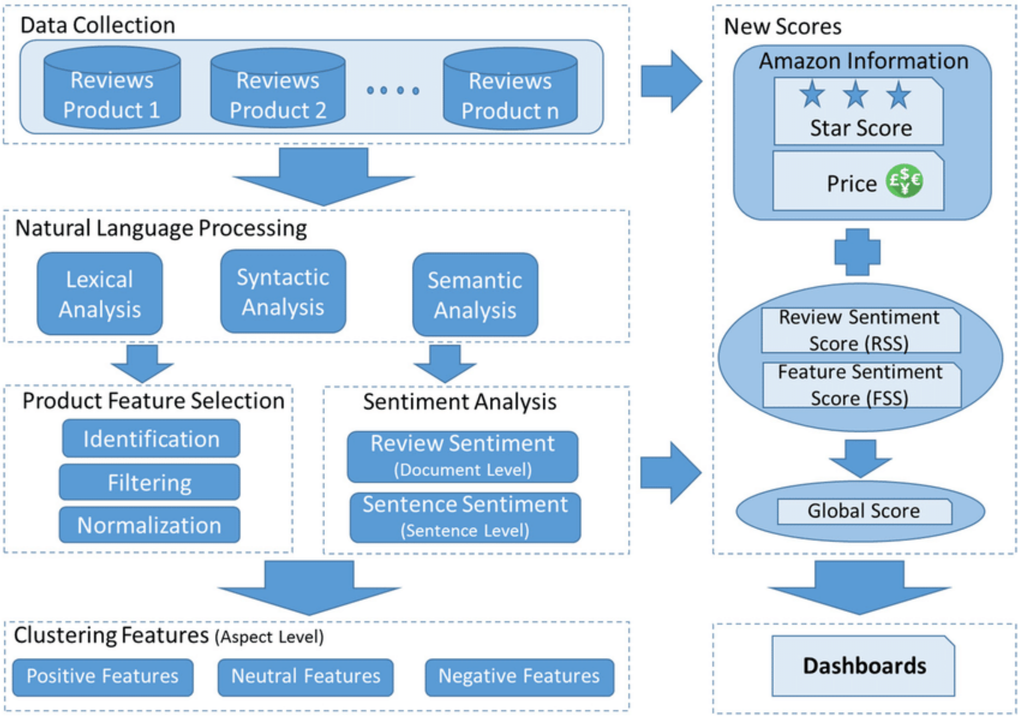


Figure 5.1: system Architecture[4]

5.1.3 Database server

A database server is used in web-based database applications to deliver the data for the application. Business logic is sometimes provided by the database server in the form of stored procedures. In a multi-tiered design, stored procedures can provide considerable performance benefits[14]. Limitations: The Amazon dataset contains several categories of products like beauty, food, hair products, however, we only considered the electronics category. Although this system can be applied to other categories as well, the magnitude of data and relevance of electronics in today's e-market makes it ideal to work on. This application can be used as a web extension for websites with Q/A feedback and reviews, however, currently we considered it as a stand-alone application.

5.2 Structural Components

Two structural components of our application are Client and Server A user interacts with a client, which is a user-friendly representation of a web app's capabilities. It's written in HTML, CSS, and JavaScript. It doesn't require any OS/device-specific changes. On server side we use Python Development skills

5.3 Tools and Technology

Several different tools and technologies have been used to develop this type of system, such as Apache spark,Okapi BM25,Bert, Anaconda etc.

5.3.1 Anaconda

Anaconda is a Python and R programming language distribution aimed for simplifying package management and deployment in scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, and so on). In Python sentiment analysis is a way for examining a piece of text and determining the hidden sentiment. This is achieved through the use of a combination of machine learning and natural language processing (NLP). Sentiment analysis is a technique for analyzing the comments made in a piece of writing.

5.3.2 Apache spark

Apache Spark is a distributed processing framework for big data workloads that is open-source[15]. For quick analytic queries against any size of data, it uses in-memory caching and efficient query execution. Classification techniques are capable of analyzing large amounts of data and can predict the class of newly provided data. Spark is a large-scale data processing engine with a unified analytics engine. It includes high-level APIs in Scala, Java, Python, and R, as well as an optimized engine for data analysis that supports broad processing graphs. Spark SQL for SQL and Data Frames, pandas API on Spark for pandas workloads, MLlib for machine learning, GraphX for graph processing, and Structured Streaming for stream processing are among the higher-level technologies it offers.

5.3.3 Sci-Kit Learn

Scikit-learn (formerly scikits.learn and also known as sklearn) is a free software machine learning library for the Python programming language[16]. This library contains multiple ML algorithms to train and test on these include classification, regression, clustering, random trees and several more. It uses numpy libraries for the complex algebra equations when implementing these ML models.

5.3.4 TFIDF Values

A numerical metric called term frequency–inverse document frequency is meant to show how essential a word is to a document in a collection or corpus. It's frequently used for information retrieval, text mining, and user modelling searches as a weighting factor. The tf–idf value rises in proportion to the number of times a word appears in a document and is offset by the number of documents in the corpus that contain the term, which helps to compensate for the fact that some words appear more frequently than others.

5.3.5 Random Forest Algorithm

As the name implies, a random forest is made up of a huge number of individual decision trees that work together as an ensemble. Each tree in the random forest produces a class prediction, and the class with the most votes becomes the prediction of our model.

5.3.6 BERT

We use the Bert model for ranking the reviews.It is a Google-developed transformer-based machine learning technique for natural language processing (NLP) pre-training.BERT was pretrained on two tasks:language modelling and next sentence prediction. Before being entered into BERT, text inputs must be converted to numeric token ids and grouped in multiple Tensors. For each of the BERT models outlined above, TensorFlow Hub provides a matching preprocessing model that implements this transformation using TF ops from the TF.text package. To preprocess text, you don't need to execute pure Python code outside of your TensorFlow model.

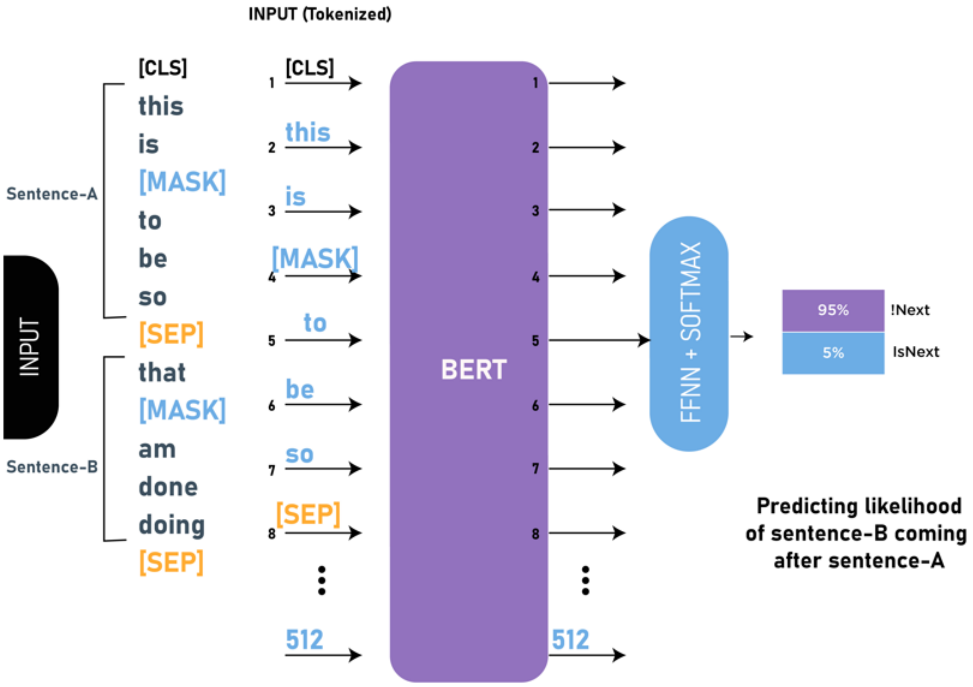


Figure 5.2: Bert Working.[5]

5.3.7 Okapi BM25:

BM25 is a bag-of-words retrieval function that prioritizes a set of documents based on the query keywords that appear in each document, regardless of where they appear in the text. It's a group of scoring functions with slightly different parameters and components.

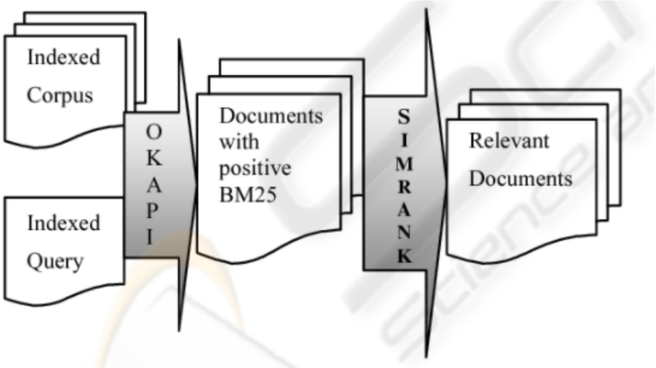


Figure 5.3: Okapi BM25[6]

5.4 Methodology

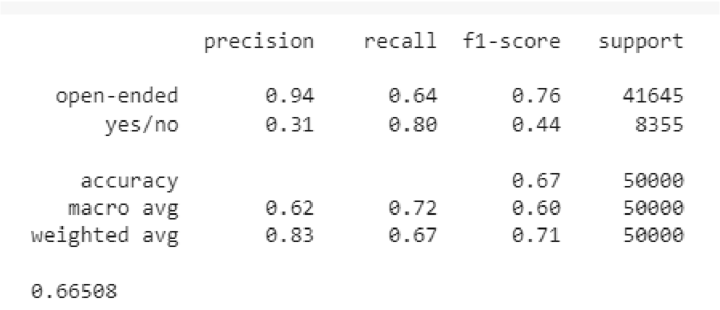
Our system is developed using a WaterFall Model, keeping in mind the enhancements that might be made in the future. The system was developed in phases.

5.4.1 Phase 1

In the first phase of system, we were gathering the required essential information regarding the development of the system. By made research how this system will be going to work how other Query Automatically responds systems are working and what are the basic requirements. And suggested a name for the system which is Fellow "Should I Buy It?"

5.4.2 Phase 2

In the second phase, we were making a plan by which tools we were going to initiate the development of the system. Initially, we used Apache spark for classification as it had an extensive list of NLP centered tools to used for our Query Classification. However, the results were not desirable with strikingly high precision score for open-ended queries on 94 percent and only 31 percent precision regarding yes/no queries with an overall accuracy of 66 percent with the low accuracy and high demand of resources for the model we decided to pursue another model for our projects performance and working. Our next attempt was made using TFIDF data where we converted our text data to more meaningful values to train our model on, these values were used with a random tree algorithm to train our model. The final results were more favorable then the previous model with 91 precision on open ended Queries and 71 percent precision on yes/no queries, the final accuracy was 77 percent 11 percent more than the previous model. For relevance ranking we tried the common and famous techniques the most used was okapi BM25 for relevance ranking and also word2Vector retrieval. For classifying the answers as positive of negative sentiment analysis techniques were used using grammar libraries like nltk.



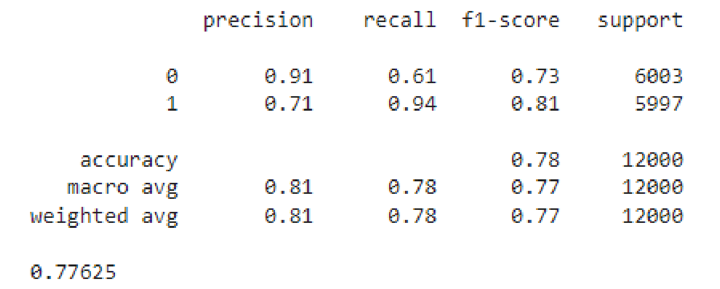
```
              precision    recall  f1-score   support

  open-ended       0.94      0.64      0.76     41645
      yes/no       0.31      0.80      0.44      8355

    accuracy                           0.67     50000
   macro avg       0.62      0.72      0.60     50000
weighted avg       0.83      0.67      0.71     50000

0.66508
```

Figure 5.4: Appache Spark Classifier Validation Results



```
              precision    recall  f1-score   support

           0       0.91      0.61      0.73      6003
           1       0.71      0.94      0.81      5997

    accuracy                           0.78     12000
   macro avg       0.81      0.78      0.77     12000
weighted avg       0.81      0.78      0.77     12000

0.77625
```

Figure 5.5: Random Forest Algorithm Results

5.4.3 Phase 3

In the third phase, we designed an interface for a system where customers select the desired product and ask a question about the product. The Design was standard Html With Bootstrap Css Styling for visual aesthetic. The interface is a standard e-commerce product page with product name description and reviews regarding that product there is a text box where users can write their queries and return with with their answers and a list of 5 relevant reviews in regards to the query.

5.4.4 Phase 4

In the final phase of our system, integrate the backend developed in python with the front end for ease in a proper web application. The Flask library was used for this phase as it was the easiest to use and implement and our results were met with this library.

Chapter 6

System Testing and Evaluation

After the Implementation phase of software, the next phase is the testing phase. System testing is an essential step for the development of a reliable and error free system. After Implementation phase of software next phase is testing phase. System testing is an essential step for the development of a reliable and error free system. Once the source code has been generated, software must be tested to uncover and correct as many errors as possible before releasing the final product. Our goal is to design a series of test cases that have a high likelihood finding errors but how, there are various methods that provides a systematic guidance for designing tests.

- Exercise the internal logic software component, and
- Exercise the input and output domains of the program to uncover errors in the program function, behavior, and performance. Software testing is a crucial element of software quality assurance and represents the ultimate review of specification, design, and code generation. The worked product is a set of test cases designed to exercise both internal logic and external requirements are designed and documented, expected results are defined, and actual results are recorded.[16] The primary objectives of testing software are to execute a program with the intent of finding an error, a good test case will find an as-yet –undercover error, and a successful that uncover an as-yet-undercover error.

6.1 Testing strategies

The basic strategies that were used for testing were following:

- Specification testing
- Black Box Testing
- White Box testing

- Regression testing
- Acceptance testing
- Assertion testing
- Unit testing
- System testing

Each of the strategy is discussed as following:

6.1.1 Specification testing

The word "specification testing" refers to the process of conducting tests in compliance with the requirements. Requirements Specifications are a document that serves as a link between the developers and the intended consumers. They are a precise explanation of a system's capability and limits. Even if the code testing is performed exclusively, it doesn't ensure against program failure. Code testing doesn't answer whether the code meets the agreed specifications documents. It doesn't also determine whether all aspects of the design are implemented. Therefore, examining specifications stating what a program should do and how it should behave under various conditions performs specification testing. Test cases are developed to test the range of values expected including both valid and invalid data. It helps in finding discrepancies between the system and its original objective. During this testing phase, all efforts were made to improve accuracy and results.

6.1.2 Black Box Testing

Black Box Testing is a software testing method that involves testing the functions of software applications without knowing the internal code structure, implementation details, or internal routes. Black Box Testing is a type of software testing that focuses on the input and output of software applications and is totally driven by software requirements and specifications.

6.1.3 White Box Testing

White-box testing, sometimes called glass-box testing is a test case design method that uses the control structure of the procedural design to drive test cases. In white Box testing internal code written in every component was tested and it was checked that the code written is efficient in utilizing various resources of the system like memory or the utilizing of input output[17].For our binary classifier using our questions database which were labeled as open-ended and binary class, cleaned from buggy sentences which could not be processed. A total of 60000 questions of equal open-ended and binary queries. Among these 50000 were used for training and 10 used for testing the ClassifierDL and USE pipeline made using spark NLP the test showed 61 percent classification of binary queries and an 83 percent accurate classification of open-ended queries.

6.1.4 Regression Testing

Regression testing is a method of testing that is used to ensure that a software change does not affect the product's existing functioning. In order to verify the impact of the modification, previously executed test cases are re-executed. In regression testing the software was tested against the boundary conditions. Various input fields were tested against abnormal values and it was tested that the software does not behave abnormally at any time.

6.1.5 Acceptance Testing

Acceptance testing is the most crucial stage in the testing process because it determines whether or not the client approves the application/software. Ready. Normally the quality assurance department performs the acceptance testing that the software is ready and can be exported.

6.1.6 Assertion Testing

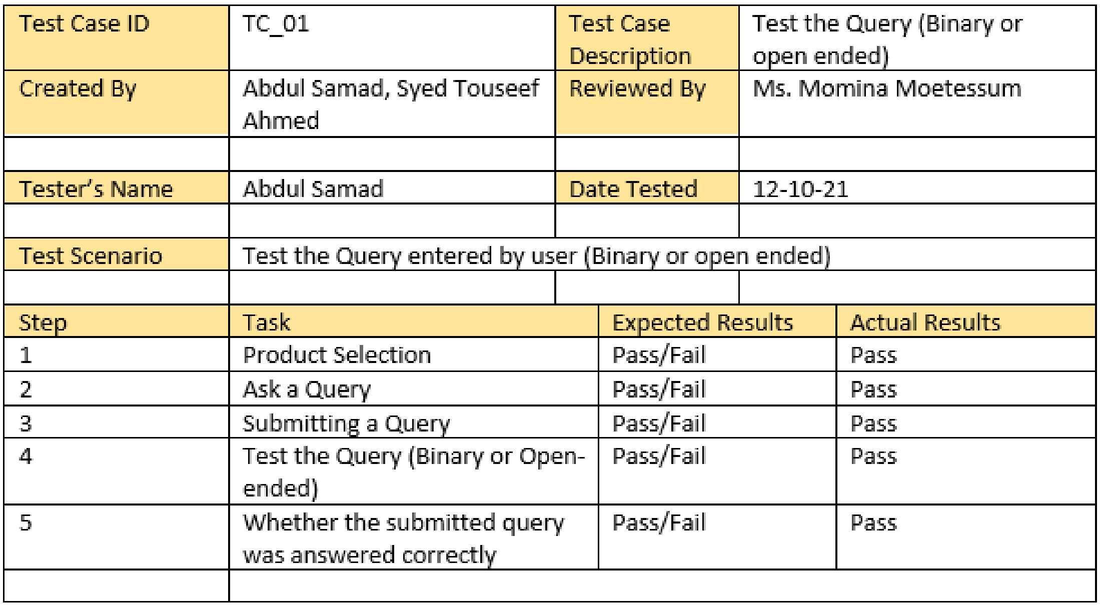
Acceptance testing is the most crucial stage in the testing process because it determines whether or not the client approves the application/software. Ready. Normally the quality assurance department performs the acceptance testing that the software is ready and can be exported.

6.1.7 System Testing

When all the units were working properly and unit testing was performed then comes the time for system where we checked all the integrated components as a whole and looked for possible discrepancies, which code have arisen after an integration[18]. Testing of "Should I Buy It?" test cases are built on to verify the main success scenario of the system. This section highlights the test cases build at design time and will be verified after the implementation is finished. It will give inside of testing the system to QA team and tester.

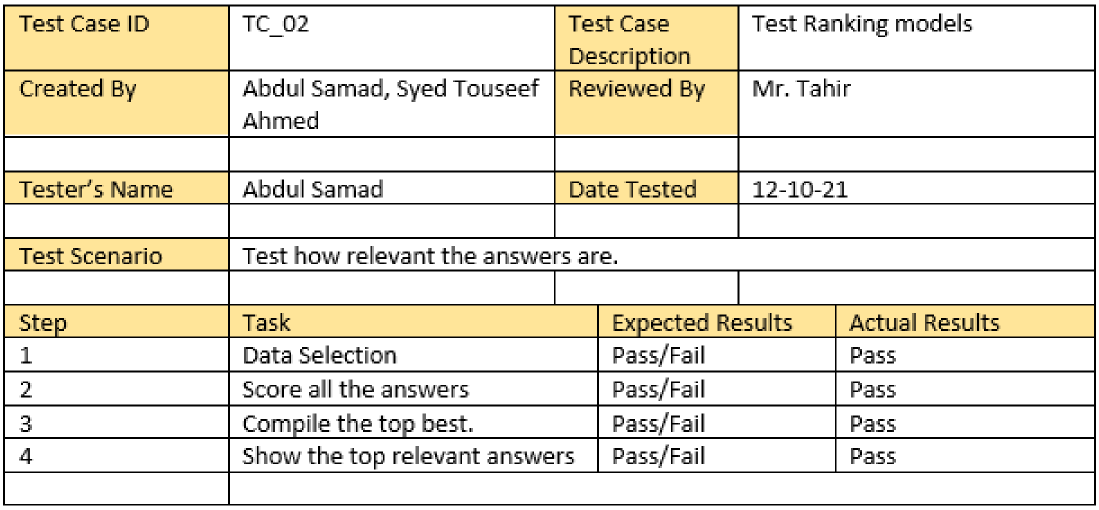
6.2 Test case

- In the first test case,the user should must select a product in order to ask a question and the system will check whether the query is Binary or open ended.



Test Case ID	TC_01	Test Case Description	Test the Query (Binary or open ended)
Created By	Abdul Samad, Syed Touseef Ahmed	Reviewed By	Ms. Momina Moetessum
Tester's Name	Abdul Samad	Date Tested	12-10-21
Test Scenario	Test the Query entered by user (Binary or open ended)		
Step	Task	Expected Results	Actual Results
1	Product Selection	Pass/Fail	Pass
2	Ask a Query	Pass/Fail	Pass
3	Submitting a Query	Pass/Fail	Pass
4	Test the Query (Binary or Open-ended)	Pass/Fail	Pass
5	Whether the submitted query was answered correctly	Pass/Fail	Pass

Table 6.1: Query Testing



Test Case ID	TC_02	Test Case Description	Test Ranking models
Created By	Abdul Samad, Syed Touseef Ahmed	Reviewed By	Mr. Tahir
Tester's Name	Abdul Samad	Date Tested	12-10-21
Test Scenario	Test how relevant the answers are.		
Step	Task	Expected Results	Actual Results
1	Data Selection	Pass/Fail	Pass
2	Score all the answers	Pass/Fail	Pass
3	Compile the top best.	Pass/Fail	Pass
4	Show the top relevant answers	Pass/Fail	Pass

Table 6.2: Test Relevancy of answers

- In the second test case,the system will test how the relevant answers are.The system will select the data and score all the reviews and compile the top best and finally show the best relevant answers.

- In the third test case,the system will test the relevant answers score.

- In the final test case,the system will do voting on all the relevant answers and will give the final answer based on that voting.

Test Case ID	TC_03	Test Case Description	Check the scoring of the Relevant answers
Created By	Abdul Samad, Syed Touseef Ahmed	Reviewed By	Mr. Jawad
Tester’s Name	Abdul Samad	Date Tested	11-11-21
Test Scenario	Test the score of relevant answers.		
Step	Task	Expected Results	Actual Results
1	Select relevant answers	Pass/Fail	Pass
2	Score the answers	Pass/Fail	Pass

Table 6.3: Test the relevant answers score

Test Case ID	TC_04	Test Case Description	Test the final output
Created By	Abdul Samad, Syed Touseef Ahmed	Reviewed By	Ms. Momina Moetessum
Tester’s Name	Abdul Samad	Date Tested	12-10-21
Test Scenario	Test the query entered by the user to get the relevant result.		
Step	Task	Expected Results	Actual Results
1	Product Selection	Pass/Fail	Pass
2	Ask a Query	Pass/Fail	Pass
3	Submitting a Query	Pass/Fail	Pass
4	Show list of relevant answers	Pass/Fail	Pass

Table 6.4: Test the final answer

Chapter 7

Conclusions and future work

This chapter concludes the project and future work.The system “Should I Buy It?” has been successfully designed and developed.The system, which aims to develop such a system where users can buy products with complete satisfaction. We have undertaken the task of creating such a system where the user can ask a question and the system automatically responds to the user by reading all the reviews and we are overwhelmed that it has been successfully done by us.

7.1 Conclusion

In this project,a general examination of the performance of sentimental analysis of Amazon reviews is conducted in this research.There are many different models proposed to meet the requirement of sentimental analysis.For scoring reviews, we used the Bert model.BERT and other Transformer encoder architectures have proven to be quite effective in a range of NLP tasks (natural language processing). They create natural language vector-space representations that can be used in deep learning models. The Transformer encoder architecture is used by the BERT family of models to process each token of input text in the context of all tokens before and after it. BERT models are often trained on a huge collection of text before being fine-tuned for specific tasks.For relevance ranking,okapi bm25 is used.It is a retrieval function that scores a set of documents based on the query phrases that appear in each text, regardless of where they appear in the document. It’s a group of scoring functions with slightly different parameters and components. This report goes from a literature review to comparative performance analysis of sentimental analysis. First, the work that has been done in the literature and related to the project is reviewed. One goal of the background study is to improve the model’s implementation. Next, the report presents the implementation of all the classifiers and models used in query classification, ranking to the reviews, and relevance ranking to get the answer related to the user query. After presenting the technical details of relevant algorithm implementations in the system, the study moves on to the web application development process. The system overview and architecture, front-end and back-end implementation, and testing are the main activities in this section. The overview provides a general understanding of how different algorithms are compared and evaluated.The architecture presents major components of the system. In terms of the project's outcome, there are some important factors considered in the stage of system architecture, such as the classification and ranking, etc. The front-end implementation section gives good features of the web application on the browser side. The back-end implementation section introduces some model that provides support for the NLP which automatically responds to user queries.Combining the front-end and back-end implementation, the web application has achieved a system where the user can easily ask a question and get the relevant answer. The final chapter concludes with a section on testing techniques, which briefly describes some of the main testing scenarios used to test the application.

Throughout the project we have learned about new developments in the text classification. Text categorization can be done in a variety of ways and with a variety of classifiers. It helped us gain experience and skills for our future ventures in our practical lives. We have learned a lot of things with the help of our supervisor. Natural Language processing is used in this system to read the reviews. We also hope to inspire young enthusiasts into the field. Side by side, we intend to focus on blooming an industry that has not been as accredited and endorsed in our country as of yet.

7.2 Future work

In the future, we plan to enhance our project by adding more categories and improving our system.Now this time we only have a category of electronics.we also have a plan to make the android application which user easily download from the playstore and buy the products with complete satisfication . Also we are trying to work on accuracy to get the better results.

Appendix A

User manual

A.1 Use Case Diagram of Should I Buy It?

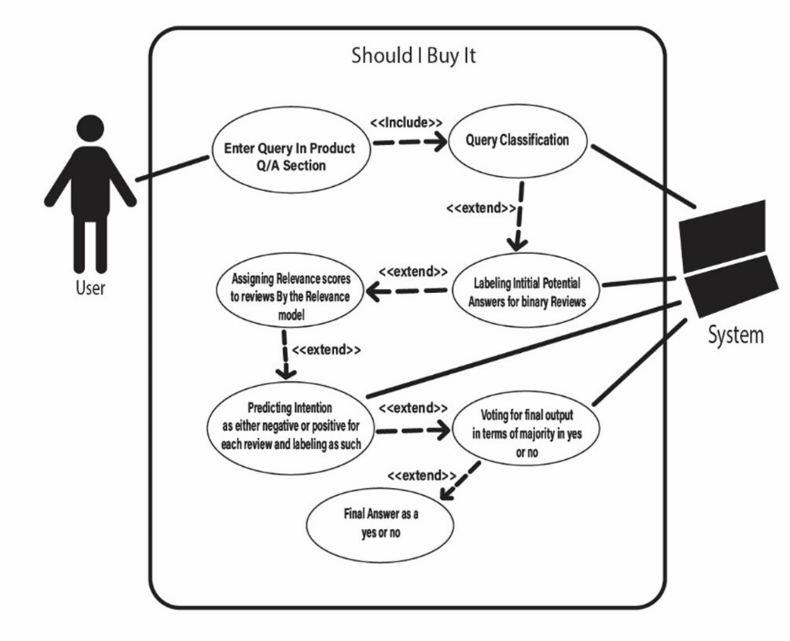

Should I Buy It
User
Enter Query In Product Q/A Section
<<Include>>
Query Classification
<<extend>>
Labeling Intitial Potential Answers for binary Reviews
<<extend>>
Assigning Relevance scores to reviews By the Relevance model
<<extend>>
Predicting Intention as either negative or positive for each review and labeling as such
<<extend>>
Voting for final output in terms of majority in yes or no
<<extend>>
Final Answer as a yes or no
System

Figure A.1: Use Case Diagram

A.2 Component Diagram

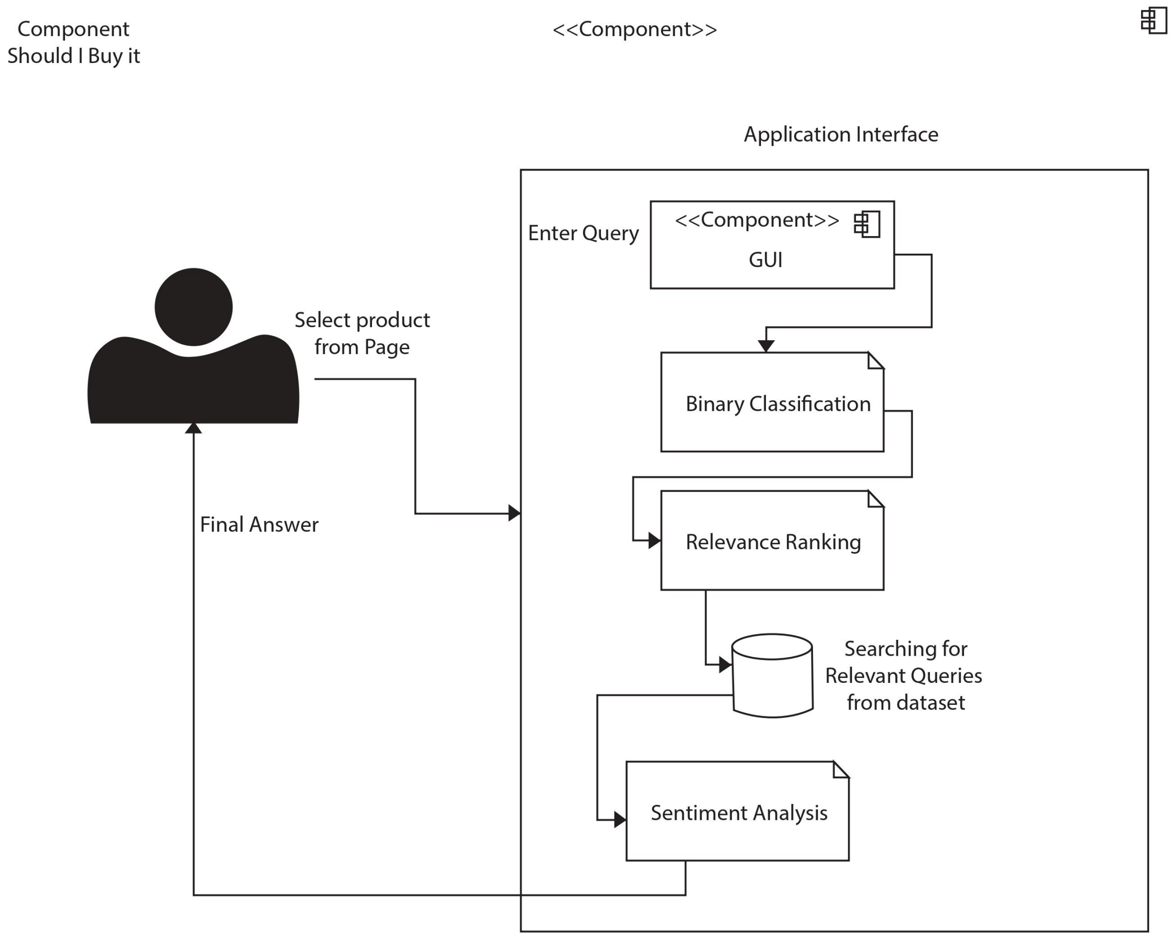


Figure A.2: Component Diagram

A.3 Sequence Diagram

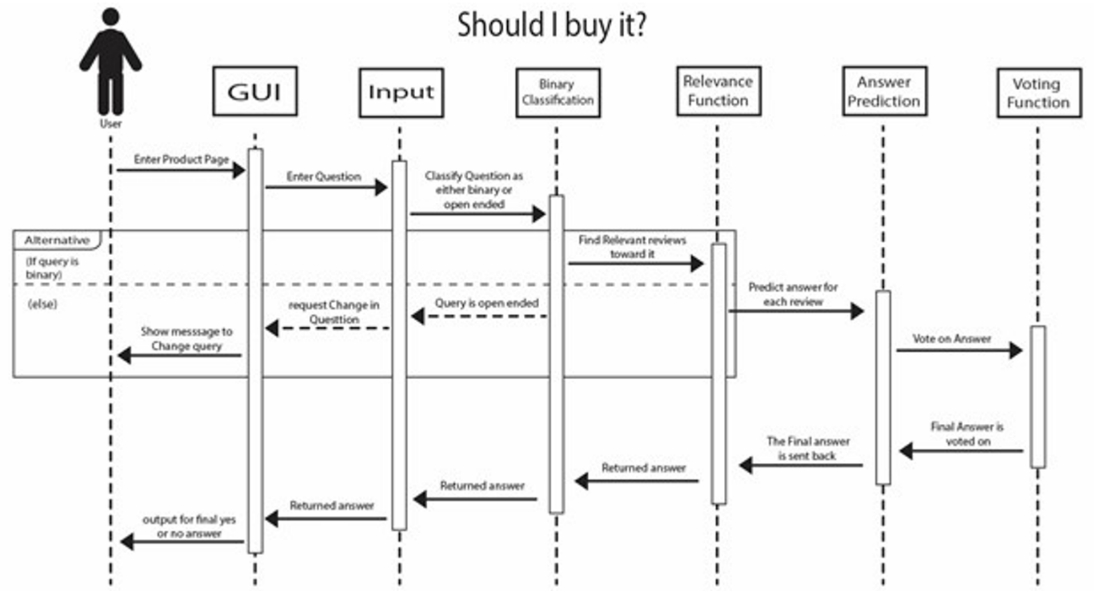
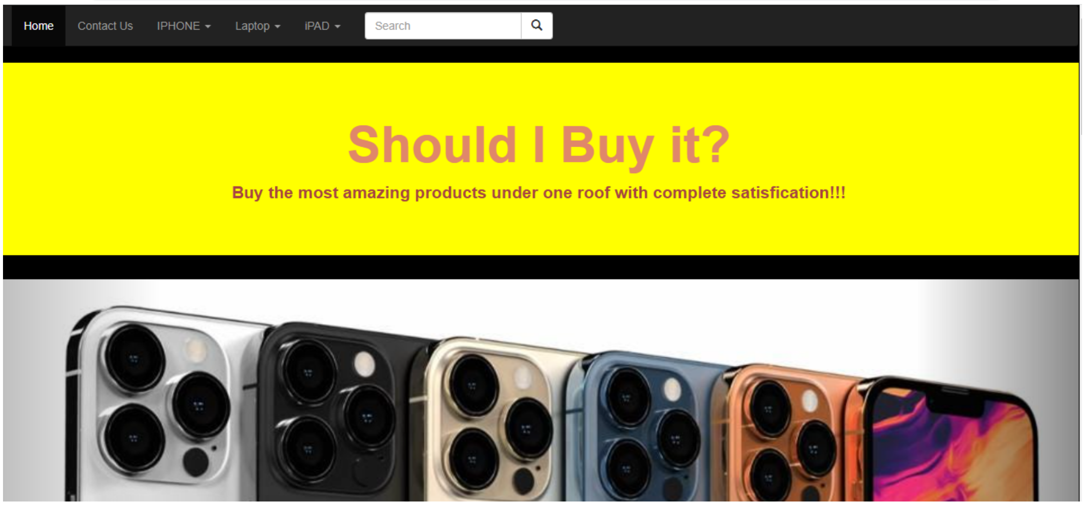


Figure A.3: Sequence Diagram

Appendix B

User Interface

B.1 Main page

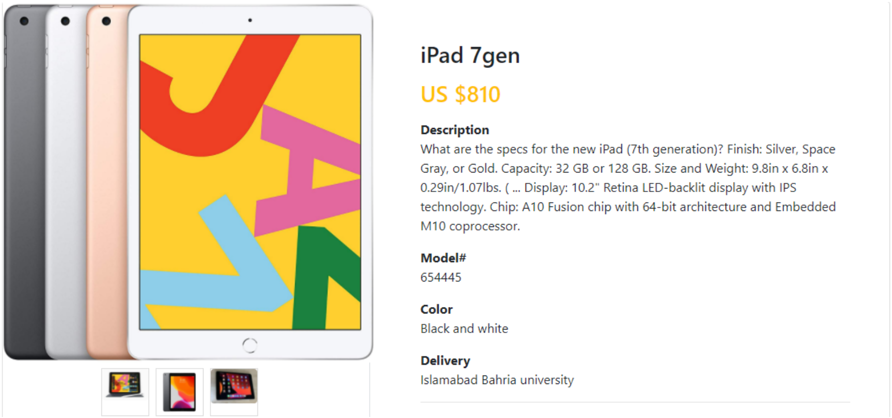


Should I Buy it?

Buy the most amazing products under one roof with complete satisfication!!!

Figure B.1: Main Page

B.2 product description page



iPad 7gen

US $810

Description

What are the specs for the new iPad (7th generation)? Finish: Silver, Space Gray, or Gold. Capacity: 32 GB or 128 GB. Size and Weight: 9.8in x 6.8in x 0.29in/1.07lbs. (... Display: 10.2" Retina LED-backlit display with IPS technology. Chip: A10 Fusion chip with 64-bit architecture and Embedded M10 coprocessor.

Model#

654445

Color

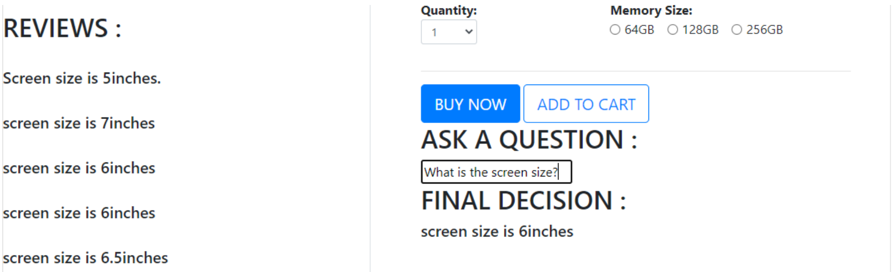
Black and white

Delivery

Islamabad Bahria university

Figure B.2: product description page

B.3 Demo



REVIEWS :

Screen size is 5inches.

screen size is 7inches

screen size is 6inches

screen size is 6inches

screen size is 6.5inches

Quantity: 1

Memory Size: ○ 64GB ○ 128GB ○ 256GB

ASK A QUESTION :

What is the screen size?

FINAL DECISION :

screen size is 6inches

Figure B.3: Demo

References

[1] Julian McAuley and Alex Yang. Addressing complex and subjective product-related queries with customer reviews. 12 2015. Cited on pp. 5, 1, and 2.

[2] Amazon. How much does amazon love dogs? just ask one of the 7,000 pups that "work" here. 07 2019. Cited on pp. 5 and 7.

[3] Stackoverflow. Human verification. 07 2021. Cited on pp. 5 and 8.

[4] Jesús Gil Ferrández Antonio Sellers-Rubio Ricardo Mora Kauffmann, Erick Peral. Architecture using in sentimental anlaysis. 09 2019. Cited on pp. 5 and 28.

[5] GeeksforGeeks. Explanation of bert model – nlp. 07 2021. Cited on pp. 5 and 30.

[6] Yaël Champclaux, Taoufiq Dkaki, and Josiane Mothe. Enhancing high precision by combining okapi bm25 with structural similarity in an information retrieval system. 2009. Cited on pp. 5, 18, and 31.

[7] Eugene Agichtein, Carlos Castillo, Debora Donato, Aristides Gionis, and Gilad Mishne. Finding high-quality content in social media. pages 183–194, 02 2008. Cited on p. 3.

[8] Pankaj Gupta, Ritu Tiwari, and Nirmal Robert. Sentiment analysis and text summarization of online reviews: A survey. pages 0241–0245, 2016. Cited on p. 5.

[9] Lixin Su, Ruqing Zhang, Jiafeng Guo, Yixing Fan, Jiangui Chen, Yanyan Lan, and Xueqi Cheng. Beyond relevance: Trustworthy answer selection via consensus verification. 2021. Cited on pp. 5, 19, and 25.

[10] Jing He and Decheng Dai. Summarization of yes/no questions using a feature function model. 20:351–366, 14–15 Nov 2011. Cited on p. 6.

[11] Quora. Qoura voting. 09 2017. Cited on p. 7.

[12] Julian McAuley and Alex Yang. Addressing complex and subjective product-related queries with customer reviews. *Proceedings of the 25th International Conference on World Wide Web*, 2016. Cited on p. 10.

[13] Guangyao Pang, Keda Lu, Xiaoying Zhu, Jie He, Zhiyi Mo, Zizhen Peng, Baoxing Pu, and Zhuojun Duan. Aspect-level sentiment analysis approach via bert and aspect feature location model. 2021, 2021. Cited on p. 27.

[14] Ruolin Wang, Zixuan Chen, Mingrui Ray Zhang, Zhaoheng Li, Zhixiu Liu, Zihan Dang, Chun Yu, and Xiang 'Anthony' Chen. Revamp: Enhancing accessible information seeking experience of online shopping for blind or low vision users. 2021. Cited on p. 28.

[15] Abdul Shoro. Big data analysis: Apache spark perspective. 01 2015. Cited on p. 29.

[16] Fabian Pedregosa; Gaël Varoquaux; Alexandre Gramfort; Vincent Michel; Bertrand Thirion; Olivier Grisel; Mathieu Blondel; Peter Prettenhofer; Ron Weiss; Vincent Dubourg; Jake Vanderplas; Alexandre Passos; David Cournapeau; Matthieu Perrot; Édouard Duchesnay. "scikit-learn: Machine learning in python". journal of machine learning research. 2011. Cited on p. 29.

[17] Mohammad Mowlaei, Mohammad Saniee Abadeh, and Hamidreza Keshavarz. Aspect-based sentiment analysis using adaptive aspect-based lexicons. *Expert Systems with Applications*, 148:113234, 06 2020. Cited on p. 34.

[18] Sadaf Janjua, Ghazanfar Siddiqui, Muddassar Sindhu, and Umer Rashid. Multi-level aspect based sentiment classification of twitter data: using hybrid approach in deep learning. *PeerJ Computer Science*, 7, 04 2021. Cited on p. 35.