ARTIFICIAL INTELLIGENCE BASED PERSONAL COMPUTER PARTS AND LAPTOPS RECOMMENDING ASSISTANT

Sewwandi E.D.D
(IT15146816)

B.Sc (Hons) in Information Technology
Specializing in Software Engineering and Information System Engineering

Information System Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

August 2019

ARTIFICIAL INTELLIGENCE BASED PERSONAL COMPUTER PARTS AND LAPTOPS RECOMMENDING ASSISTANT

B.Sc (Hons) in Information Technology

Specializing in Software Engineering and Information System Engineering

Information System Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

August 2019

Declaration

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Student ID Number	Name	Signature
IT15146816	Sewwandi E.D.D.	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

..... ……/……/……..

Prof. Koliya Pulasinghe Date

(Signature of the supervisor)

..... ……/……/……..

Ms. Hansi De Silva Date

(Signature of the Co-supervisor)

Abstract

Most of the computer and laptop users find it difficult to search for the ideal hardware component for themselves. Because there is a number of manufacturers in this sector. Thereby there are a variety of versions and brands available. This has created a need of having an online assistant to help the laptop and PC users to find the ideal hardware component matching their requirement. Another aspect that has created the need for such an assistant is the practice of assembling PCs' by people according to their requirements. This is a common practice, especially in the gaming industry. Because almost all the games demand PC specification if it needs to be played. Considering the above aspects an online assistant to assist you in finding the hardware components is designed in this research. This research was initiated with the deployment of an online questionnaire and interviews with the local vendors (PC and Laptops). With the information gathered an online assistant is designed to suggest compatible parts matching the requirements. In order to find the ideal component customer feedback analysis along with price optimization is used. This platform will provide assistance in finding the compatible PC parts, generating PC plans matching customer's budget and analysis of customer feedback and display the results in a rating format. Laptop users are given the service to compare laptops and based on the computational power the BEST will be recommended. In order to generate compatible PC part plan and laptop recommendation, Expert Systems will be used. For customer feedback analysis Natural Language Toolkit is used along with Python Libraries. Price optimization algorithms will be utilized to generate PC part plans according to the budget. The final output of this research is a web-based application built using Java, JavaScript and Python with Mongo DB as its database. Users will encounter this platform by the name "TechRing"- we make the right choices for you.

Keywords: Web scraping, assembling, expert systems, neural network, price optimization, sentiment analysis, lexicon, opinion mining, comparison, power.

Acknowledgement

I would like to express my sincere gratitude to our project supervisor Prof. Koliya Pulasinghe for his guidance, encouragement and useful critique at every milestone in our research. His willingness to give his time so generously is highly appreciated. His mentoring and continuous support helped our team immensely to be who we are now and complete our work successfully and professionally.

I would also like to thank our co-supervisor Ms.Hansi De Silva who assisted us in our research work. Her supervision on our performance greatly helped us to improve the research activities. She also generously shared her time for us to provide the proper guidance and advice to make this research a successful project.

Also, I would like to thank our former lecturer in charge, Mr. Jayanth Amarachchi for constantly providing the guidance and advice to carry out the research according to required standards. Also, special gratitude for our team for being there for one another and investing their time and effort on this project to achieve our research goals.

Finally a special thank you, for my parents and friends, who were always beside me helping and encouraging on this journey. If not for them this would not have to be successful as it is today.

Thank You.

Table of Contents

List of Figures

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	**Explanation**
PC	Personal Computer
DOTA	Defense of the Ancients
AMD	Advanced Micro Devices
GHz	Gigahertz
GB	Giga Byte
RAM	Random Access Memory
CPU	Central Processing Unit
IT	Information Technology
POS	Part of Speech
MVC	Model View Controller
AWS	Amazon Web Service
API	Application Programming Interface
JSON	JavaScript Object Notation
SDLC	Software Development Life Cycle

1 INTRODUCTION

Initially, computers were used as a tool for calculations, but now computers help people finish many aspects in life [1]. Therefore, computers do play vital role in almost every industry. Among all the industries one of the upcoming industries is the gaming industry. Initially computer games were designed as an entertainment kit. But this industry has achieved a greater success and it is considered as one of the most profitable industries today. Computers are the main backbone of this industry.

It is not a difficult task to find a computer today. But it is a challenging task to find a computer that would satisfy the specific requirements. Therefore, the computers available in the market fail to cater to those requirements. Thus, people tend to assemble their own PC's. But it takes various components to build a PC and with the specifications, components required differ. For example, computers used for editing photos and videos is different from the computers used by a cashier.

As mentioned in the beginning majority of PC's are assembled for gaming purposes. Because most of the games available nowadays have different specifications. For example, the game DOTA requires Processor: Dual core from Intel or AMD at 2.8 GHz, Memory: 4 GB RAM,

Storage: 15 GB available space. Therefore, when they assemble a PC according to a game system requirement they need to look into those aspects and decide on assembling procedure.

People find it difficult to match the PC parts with each other when assembling because the PC parts have compatibility issues with one another. When it comes to compatibility, motherboard plays a major role since it has to work with components like RAM, CPU and more. Mainly it is necessary to check the socket compatibility with the processor.

Assembling a PC is a task with high complexity [2]. Information Technology experts might be capable of identifying the compatibilities and the required remedies. But it is not the same with an average person. We cannot ignore the fact, not only IT experts use computers average people starting from school level also use computers. Thus, they also prefer assembling their own PC's. Through a background analysis we found out that this Non – IT PC assemblers find it difficult to match the compatibility and proceed the assembling. They use the help of internet, YouTube tutorials, gaming web sites to find the parts.

When purchasing a product, it is better to have options to select with. There are sites from vendors and e-commerce sites displaying the available PC parts. Some site provides both the ecommerce and vendor site together in one as well. Most of the local vendors do not reach to the online market. They believe that having the physical existence is enough for them to reach the market. This might be true couple of decades back but now it is must to have online access to the customers. Thus the need of having a common platform for products from local vendors, ecommerce sites and non-website holders is important. Because this type of platform will help them to choose the best product for the best price.

Laptops are also widely used in the industry. Today there are variety of laptops available. Different brands, models, versions and more. Depending on the requirements the laptop that is needed differ [3]. Due to the variety of options to select with cause confusions to a customer. Because they might have doubt as to which laptop is better. There are sites that provide the function where they can compare two laptops but none of the sites recommends a laptop. This makes the users to do a background checking on the laptops before purchasing. Sometimes people who does not have a reliable source of information will go to a shop and purchase a laptop that is recommended by the shop owners.

Before purchasing a part, customers prefer looking into the customer reviews. There are many open sources available where people have expressed their experiences with the products. One of the sources is Facebook. People who have the access to those sources will go through the comments and have an idea about the product. But it is not easy to go through thousands of comments and get an idea. Most of the time people will go through first couple of comments.

Below diagrams show the data we gathered using an online survey we conducted. Our main motive through the survey was to identify our target audience, the size of the audience and also their expectations.

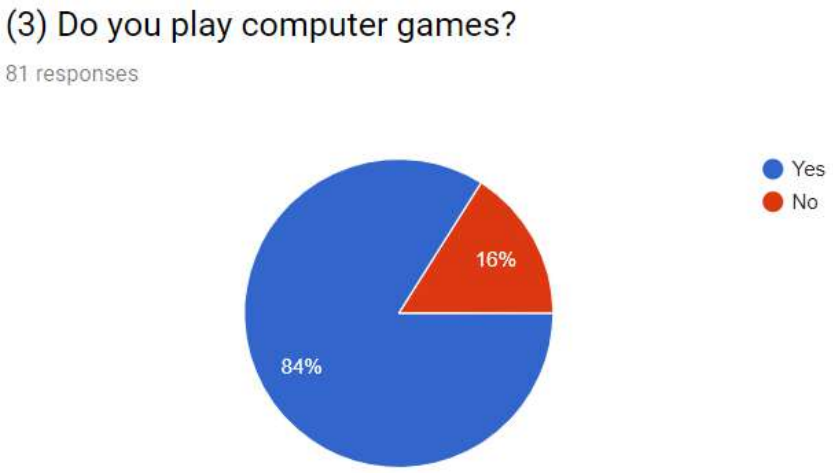

Figure1. 1 Survey Result displaying the percentage of gaming community.

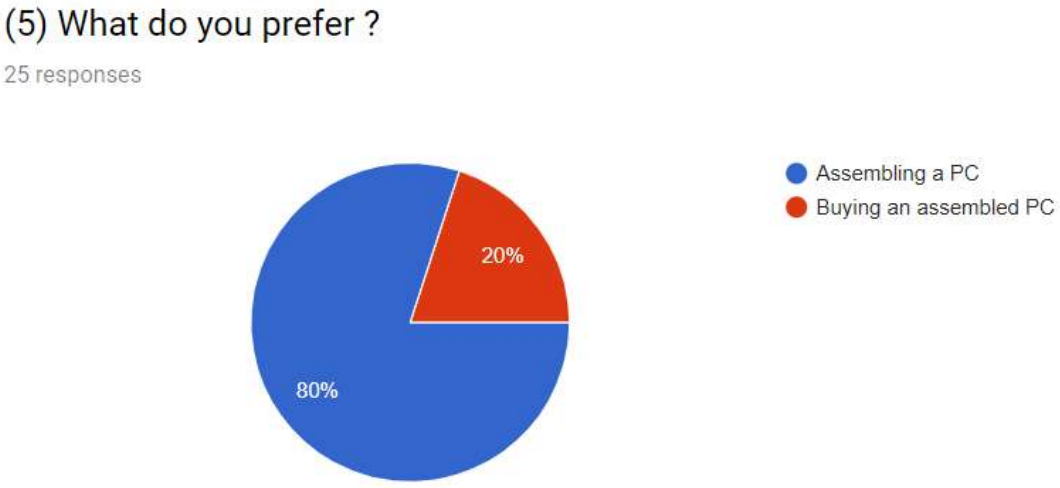

Figure1. 2 Survey Results displaying the preference of assembling PC

1.1 Background and Literature Survey

1.1.1 Background Context

The concept of having an online assistant for assembling PC's considering the requirements and the limitations is one of the challenging topics that many researches had their interests on. This is an area that is spread worldwide. Because computers and laptops are becoming a need rather than want in their lives. There were many researchers conducted to what we are building. They have used many techniques, concepts, models that is useful for us as well. Some of the research problems aren't related to our area yet the concepts they have used is very useful for us.

1.1.2 Literature Review

- Sentiment Analysis

Sentiment analysis is one of the main areas that is becoming popular in the society. This method is very useful in many of the areas. This is also known as opinion mining, sentiment mining and sentiment extraction [1]. According to Zeenia Singla sentiment analysis can be considered as a "computational study of extract subjective information from a text". Online reviews are becoming very important since they have become a measurement in quality of businesses. According to Andreea Salinca we should use a large data set for analysis and Yelp Data set provides a large review database. Also, according to this article to improve the efficiency of analysis they use two feature extraction methods and four machine learning models for automatic review analysis. The customer review analysis helps the manufacturers to identify the unrealized potential as well. Therefore, it is useful not only to the customers but also for manufacturers. Online reviews of e-commerce giants like Amazon, Flipkart has large review bases. Thus, the Big Data commerce came into role with them. These parameters help in taking profitable and accurate decisions for a business [2].

To proceed in the review analysis, they have used multiple approaches. In summarizing text Minquin Hu and Bing Lu divided the process into three tasks [3]. Namely,
Mining comments based on product features.

Identifying the opinion sentence in each review and decide whether it is positive or negative.

Summarize the results.

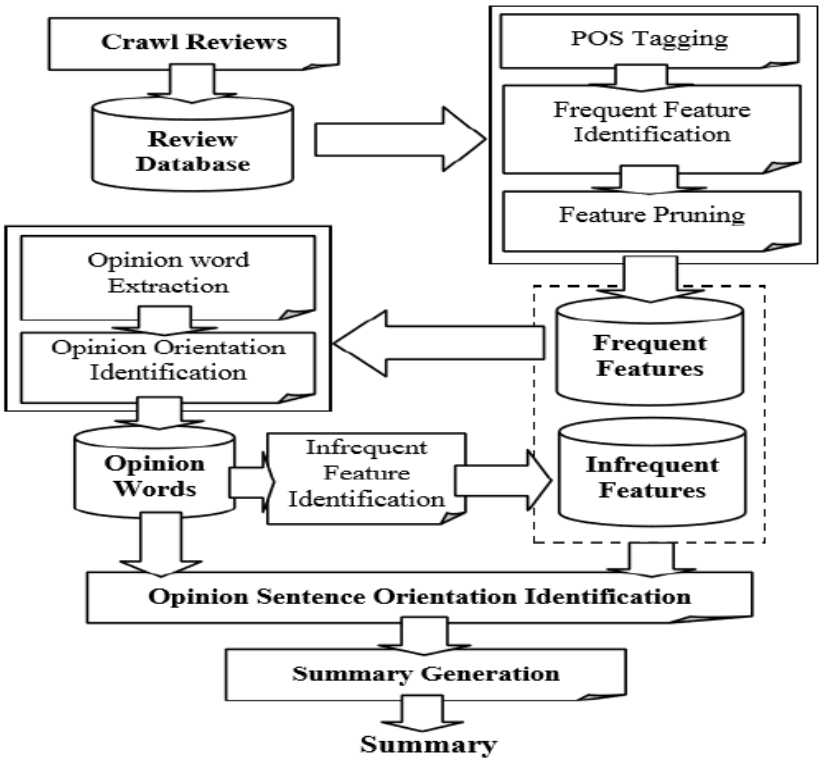

Figure2. 1- Model for Feature-based opinion summarization [3]

Part of Speech (POS) Tagger is one of the common and efficient methodologies that is used for sentiment analysis. This method identifies the nouns, adjectives, adverbs and verbs in a sentence. Below is one of the Machine learning process that is used for an efficient method to analyze the comments [4].

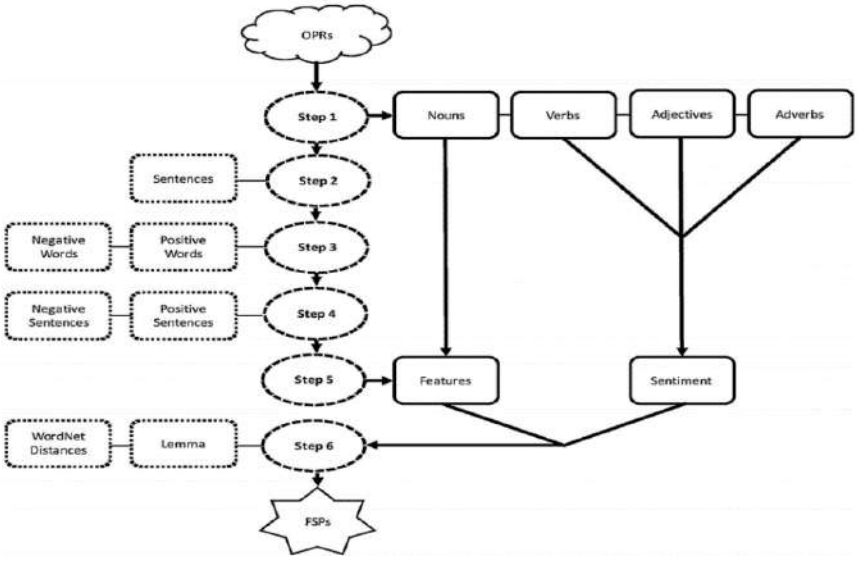

Figure2. 2 Cognition process [4]

By considering the readings it was clear that similar concepts have being used in different areas. Even though some concepts are from different fields the mechanisms that are used can be incorporated in our research as well. Because these mechanisms. Techniques and concepts produce similar outputs as “TechRing”. Considering the requirements and knowledge we gathered from the articles there are no prevailing online assistants that provides assembling plans and budget. Thus we believe this assistant will be one time solution for many of the problems faced by people in assembling and finding suitable parts.

1.2 Research Gap and Problem.

Through the data we gathered from research articles, observations, surveys and online sources it was clear that most of the sites have seldom approaches to provide solutions to the customers. With the rising demands for computers and laptops necessity of a platform which can provide solutions for customer problems is a must.

There are sites that are built for selling PC parts. Some sites display products/parts that belong to one vendor and some sites do display products from multiple vendors. There are some vendors who do not have online platforms as well. Therefore, this is a disadvantage for both the customer and the vendor. The customers might lose a chance of purchasing a product for a much cheaper price than expected. Because same product might be available for lesser price with another vendor. Figure 1.7 shows that majority of people prefer having an option to compare the prices.

The vendor might not be able to reach the customers who cannot visit them physically when they are not available online. One of the main reason local vendors are reluctant to reach the online market is due to the lack of technological knowledge and they feel safe within their comfort zone. But when we spoke with them, it was clear that they also want to join the online market if someone can provide the required guidance.

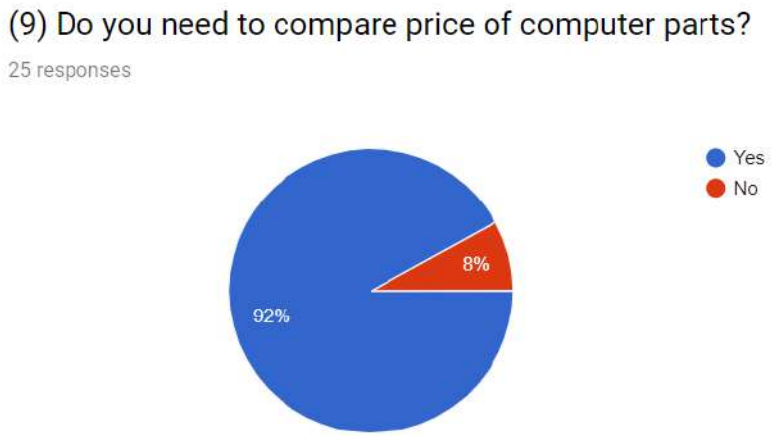

Figure2. 3- Survey results of computer parts price comparing

Through the survey we conducted we got to know that people use different methods to find the compatibility of parts. Mostly, online resources are widely used. They have to access these sources separately and get the required details. Below figure show the most common methods people use in order to access the necessary details.

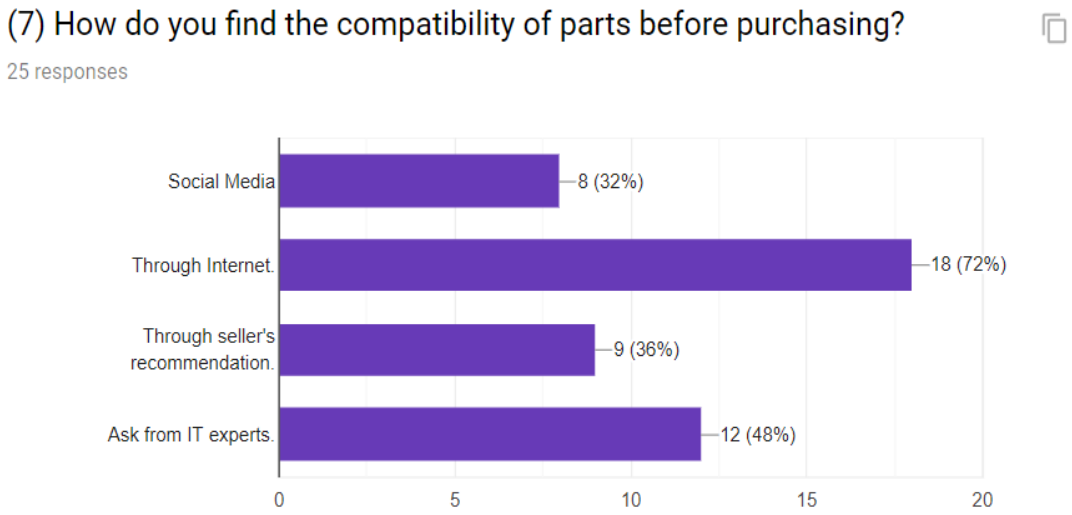

Figure2. 4- Survey result of how people check information about PC parts

There are variety of Laptops in the market. When selecting a Laptop these options confuses the customer because they have many functionalities incorporated with one another. Most of the sites which sell laptops have the option where people can compare two or more laptops. But none of those sites recommend the best laptop from them. So the user is left with the compared functions and choose the best one as they believe. Through the survey we were able to gather factors users mainly consider before purchasing a

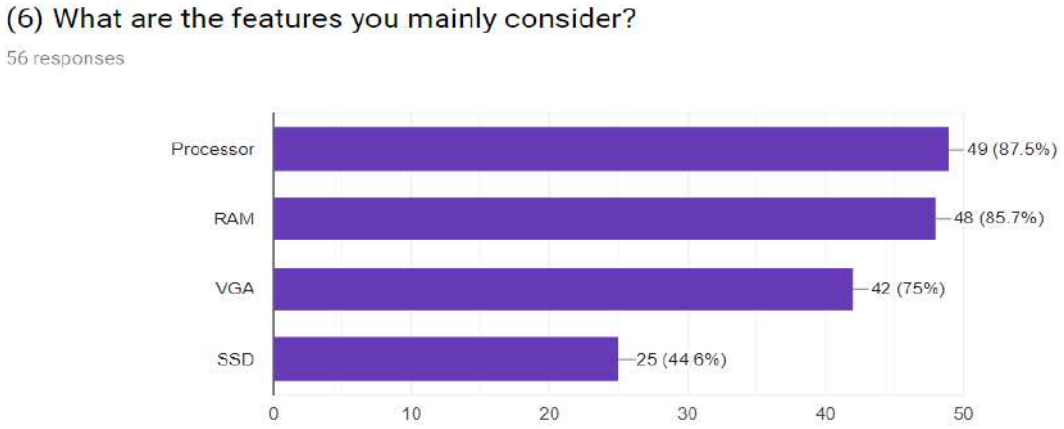

Figure2. 5-Survey results of features user mainly consider in a laptop

By going through many of the research articles several approaches were conducted in order to analyze the comments made in the social media. Reason for us to consider this area is customers look into the customer reviews done for products before purchasing. Therefore, this area is important for PC parts and Laptops. Most of the sites either display the comments that were placed by their previous customers or there are Facebook pages that contain customer reviews for products. Customers have limited access for an analyzed customer overview for the products (E.g.: Number of people satisfied with the product).

Functions	PCPartPicker.com	NewEgg.com	Noteb.com	TechRing
Select Compatible PC-Parts	✓			✓
Build PC according to a System requirement Of a Game				✓
Display prices and compare of different vendors	✓			✓
Analyze Comments and display rating based on that				✓
Recommends assemble plan according to budget	✓	✓		✓
Recommend the best PC part in a price range				✓
Notify Price Drops to user	✓	✓		✓
Laptop comparison			✓	✓
Recommending the best laptop				✓

Table 1. 1 - Comparison of current available systems with "TechRing".

1.3 Research Objectives

Main Objective

Main objective is to provide an online assistance for people who wants to assemble their PC's by themselves for different purposes and also to empower the users with information related to purchasing laptops and PC parts. The system analyses the requirements through extracting data using web scraping and provide the user with information.

Specific Objective

Display an analyzed overview of the customer reviews on the particular product.

Customers who have purchased the parts will post their experiences in social media platforms. But these platforms do not provide an overview of the comments.

We will access those comments and analyze them. After that, these comments will be categorized as positive expression and negative expressions. Through an algorithm a count of those categorized comments will be taken. Finally, we will display a summary of the customer feedback in a rating format.

2 METHODOLOGY

2.1 Methodology

Before purchasing any product one of the common practice people follow is to check the previous customer experience with the product. Earlier people use to talk with the consumers and get their experiences. But now there are online platforms where people from all over the world can share their experiences. One of the most common platforms that are used is the Social Media Platform. For example platforms such as YouTube, Facebook, Twitter. This has become the most common platform used by the consumers since it easy to access and post their experiences. Customer feedback is one aspect that is considered by the PC part customers. From the survey, we conducted it was proved that this segment of consumers also prefers the content in the social media content before purchasing the PC parts.

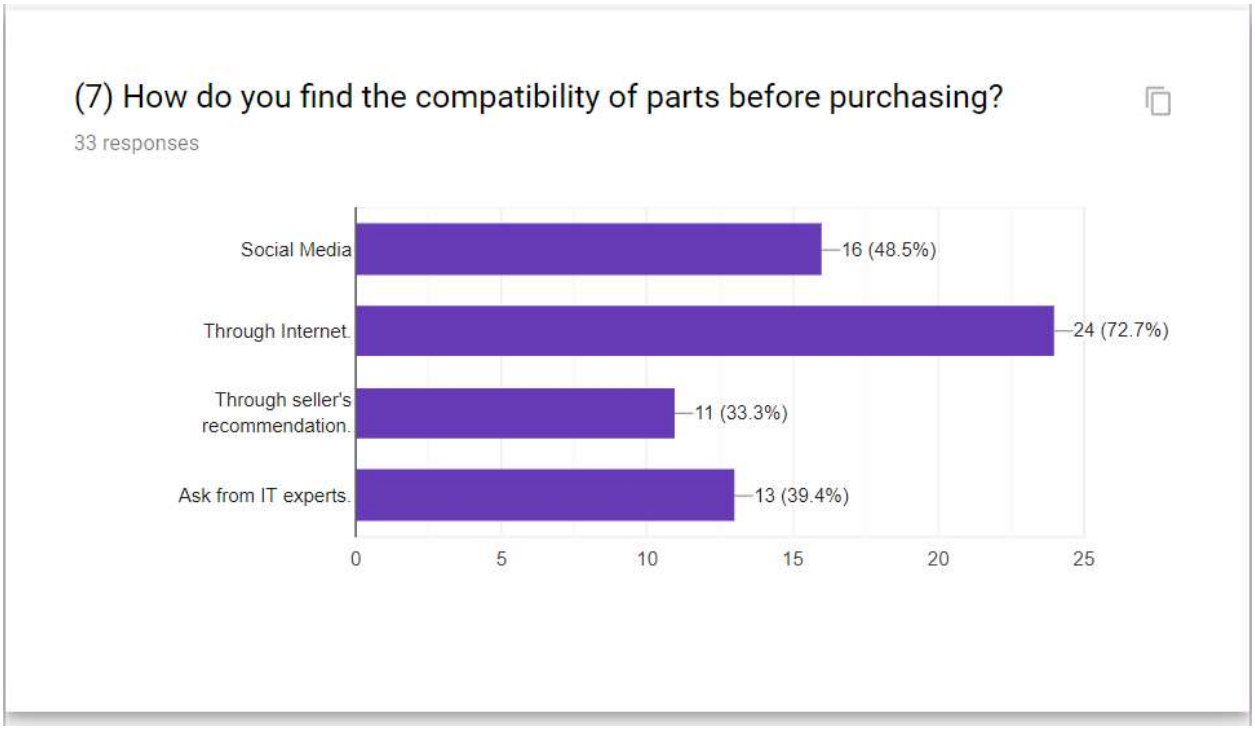

Figure3. 1-Survey Result displaying the social media percentage.

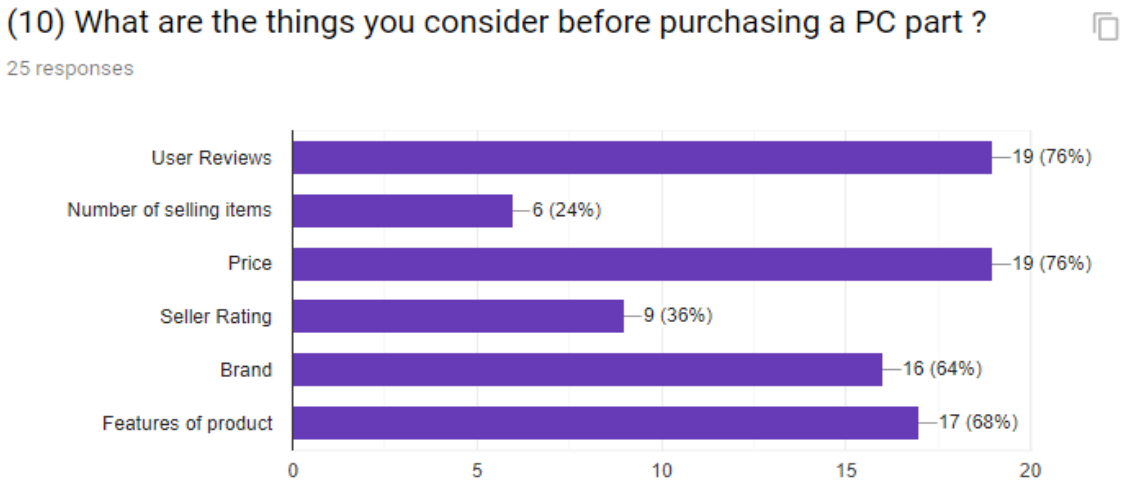

Figure3. 2-Survey Result displaying the percentage of people preferring user reviews.

Since there billions of users in the social media platforms there are many customer reviews and feedback available. For people to get an idea about the experiences they need to go through these comments and posts. But in reality, people do not have enough time to go through these content one by one. Most of the people skim through the first comments they encounter and get an idea. But they prefer having an analyzed overview of the over role customer experiences.

Thus considering this aspect “TechRing” has a specialized feature for this aspect. In here “TechRing” will look into the customer reviews one by one. Analyze each customer review and detect the positivity or negativity of each comment. Then a count will be taken for each sector. This analyzed counts will be then displayed in a rating format. Through this “TechRing”’ allows the customers to have an idea about the previous customer experiences at a glance saving the time they will have to spend in reading the customer reviews. For this purpose reviews submitted in the YouTube will be extracted real time once the user enter name of the product in the search bar. Real time data extraction is used since the new comments get added every minute.

2.2 Research Strategy

This section in the research is focusing on analyzing the consumer feedbacks. There are other existing platforms that display customer comments as it's or allow the customers to display their comments. But no such platform that will analyses and display the summary results. Results of this analysis will not only be used for displaying the ratings, but also used for the price optimization process in displaying the products in the grid. Initially we needed to collect set of data (customer comments) in order to train our system. By this step we our system is capable of making accurate analysis of the comments.

2.3 Data Collection Method

Initially a sample data set of 540 comments was used to train the system. This set consisted of comments posted in the Amazon site.

Data collection process will be conducted Real Time. This is practiced in order to track the newly added comments. For this we will be using the social media platform, You Tube. Once the user enter the name in the search bar, it will be broken in to key words. Then it will search the videos related to these key words. Then using the Python Libraries such as Beautiful Soup and Selenium comments will be extracted real time. These extracted data will be then stored in a CSV file and will be sent to analysis. Below is a display of comments in a CSV file. Each time comment extraction function is called this CSV file will be created with the new added comments.

	A	B
1	username	comment
2	Fuck Google	Then what about DDR5 ?
3	Natanel	Alright, lets see that 256gb stick.
4	Eastyy	would love to see a 256gb ram disk
5	Goatie89	This was very informative.
6	Shadowfury	True True!!!
7	Kit Anilapde	Is there laptops in the market right now using ddr4 ram?
8	judyM johnson	thankyou
9	Dilan Rukmal	Thank you Mr. Mike Mohney and Kingston Technology...
10	Sabre Gold	What if
11	LutfoR Rahman	I'm always happy Kingston technology used feel better
12	PizzaLover	Im sticking to ddr3 thanks, none of the improvements are work changing a motherboards/cpu as well as spending another 200 bucks on ram
13	Santi Carrizo	kingston ddr4 8GB 2133 CL15 es una buena RAM?
14	Alok Adhikary	I love Kingston products
15	eufilip2 Kappa	I can add this to my pc ? or is only for leptops / macbooks
16	Dominic Torretto	what
17	Ritik Singh	is 8gb ram supported for hp pavilion au084tx
18	moteados	is Jessalyn Gilsig on 0:11 ???
19	David Nmt	hi I'm going to buy a "Kingston KVR21N15S8-8 DDR4 2133MHz CL15 Single Channel Desktop RAM - 8GB" for my lenovo y700 laptop is it compatible?
20	wisnu yogapraditya	Okay, how do i download them?
21	Rajiv key's	Hi I am using macbook pro (2Ã—8gb)ram ddr3 can i upgrade to ddr4 (2Ã—16gb)ram..my macbook pro model is A-1278
22	N Shen	Great Video! Very informative.
23	ZvXGamer	Fvfvdv
24	Saikat Chakraborty	i want 2133 mhz ddr4 ram for my laptop. From where i get this??
25	Aon Lazio	wating for DDR10
26	.::Let's Play::.	Can you please tell me How do I know that the RAM installed in my PC is manufactured by Kingston? because a seller promised a Kingston RAM DDR4 in t
27	Rupjyoti Mekar	When ddr4 ram prices wil go down? I want a 8gb ram, But this price stopping me to it. Do something so we can buy ram at cheap price.
28	DonDiego256	What

Figure3. 3- CSV File format storing the comments.

2.4 **Customer Feedback Analysis**

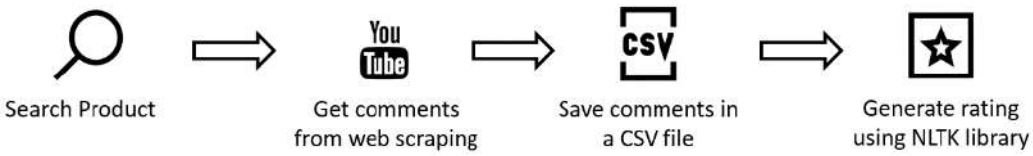

Figure3. 4 Customer feedback analysis process

This function starts when the user enter a name in the search bar. Once the user enter the name it will be broken down to key words. Using these key words the video related to the name will be searched automatically by our platform. For this purpose key words entered is used as the Search Tag. The path to access YouTube is a predefined path. This will run whenever the user enter a name.

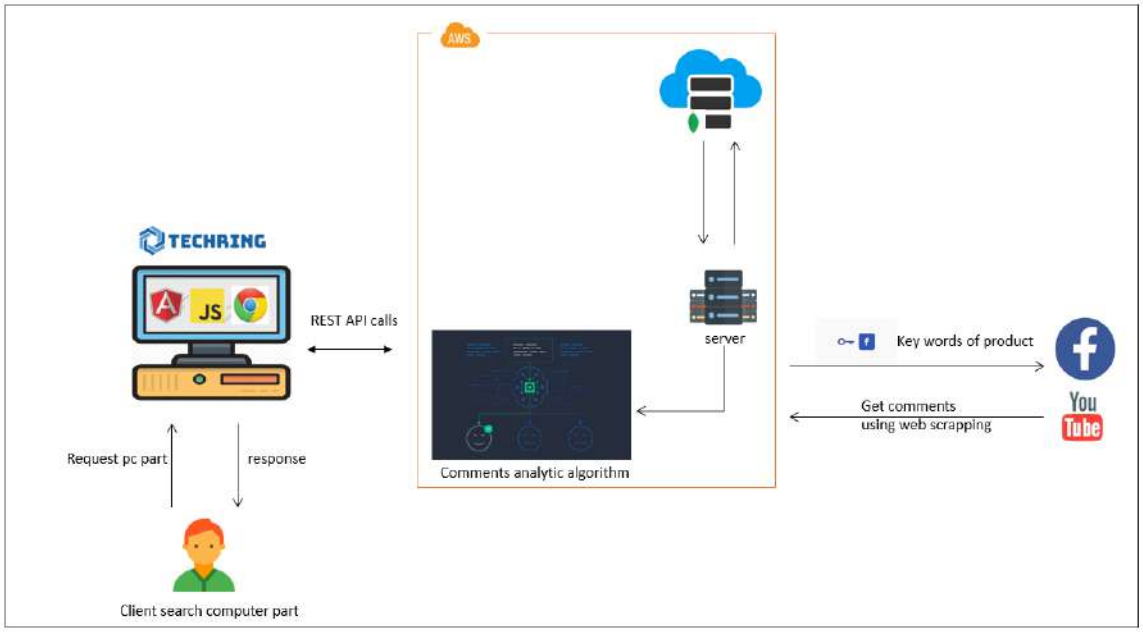

Figure3. 5- High-level architecture diagram for comment analysis.

There are many YouTube videos available for each keyword. From all the videos the one with the highest views will be chosen. Then it will scroll down to the video and extract the comments posted by the people related to the product. These extracted comments will be saved in a CSV File. This file will be recreated every time user search for a product.

This CSV file will be then used for analysis. Each comment holds three characteristics with themselves. They are,

- Polarity: Positivity or Negativity
- Subject: What is it about?
- Opinion holder: Person or the entity who express the opinion.

In the analysis I focus on the Polarity of the comment. Before the analysis the data set needs to be preprocessed. This will be done in two segments using the Natural Language Toolkit in python. Because then we can reduce the observation space of the comment. The segments are,

Uniform Spelling.

For human beings the two words, “Good” and “good” carry the same meaning. But for the system this appears as two different words due to the capitalization in the first letter. Therefore the first thing that will be done is convert all words in to the lower case.

Removing Special characters.

Special characters like “! @,?, < “ doesn’t contribute for sentiment analysis. Therefore these characters will be removed in the preprocessing.

Tokenization.

Next step is the tokenization. This process breakdown the sentence into smaller part called **Tokens.** These tokens are important for finding the pattern in the texts. This form the base of stemming and lemmatization which can be considered as cleaning process of textual data. In tokenization there are two options. Namely,

Word Tokenization.

Sentence Tokenization.

In my function I used Word Tokenization. In word tokenization it spilt each comment into word sets. Final output of word tokenization is can be converted in to a Data Frame which will be easily understood by the machine learning algorithms. For this step text cleaning methods such as lemmatization and stemming is added. Thus it will remove the punctuation and numerical values.

Analysis

In order to conduct the comment analysis I used the Natural Language Toolkit in python. In there they have in built python libraries for sentiment analysis. In these libraries they have used the model Recurrent Neural Network. RNN is capable of using the sequential information in a text. Since a sentence has grammatical structure and an order, each word

depend on the previous word. Thus if the neural network needs to know the order the words came in. RNN is an extended version of the Neural Network. In this model the same task is repeated for each word and RNN captures and store the memory of each which have being calculated previously. Below diagram display the functioning of a neural network. Using this model accuracy of analyzing the comment was high.

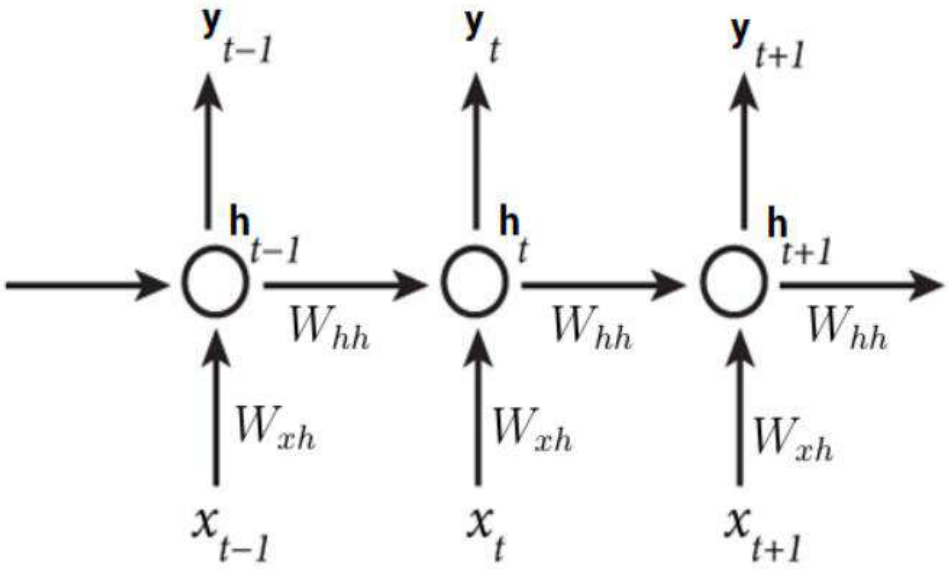

Figure3. 6-RNN architecture.

Through the RNN model each comment will be analyzed and categorized as either **positive** or **negative.**

Then a count of each will be taken. Number of positive comments in the CSV file and vice versa. Finally the counts will be displayed in a Rating format for each product.

```
In [1]: import nltk

In [2]: # nltk.download('vader_lexicon')

In [3]: from nltk.sentiment.vader import SentimentIntensityAnalyzer

In [4]: sid = SentimentIntensityAnalyzer()

In [5]: positve_comment = "PC part is good qulity and high speed, recommend, awsome, best!!!"
        sid.polarity_scores(positve_comment)
Out[5]: {'neg': 0.0, 'neu': 0.433, 'pos': 0.567, 'compound': 0.8879}

In [6]: negative_comment = "PC part is bad and low qulity not recommend"
        sid.polarity_scores(negative_comment)
Out[6]: {'neg': 0.562, 'neu': 0.438, 'pos': 0.0, 'compound': -0.7724}

In [7]: import pandas as pd

        #Ignore warnings
        import warnings
        warnings.filterwarnings("ignore")

In [8]: df = pd.read_csv('fb_comments2.csv')

In [9]: df.head()

Out[9]:
```

	date_and_time	Comment
0	2017-09-03T00:00:00.000Z	Love it
1	2017-06-06T00:00:00.000Z	Great Tablet
2	2018-04-20T00:00:00.000Z	My K Fire
3	2017-11-02T17:33:31.000Z	Confused
4	2018-04-24T00:00:00.000Z	Great, simple, and resonably priced tablet.

Figure3. 7- Coding segment for positive/negative comment analysis.

```
In [15]: from nltk.sentiment.vader import SentimentIntensityAnalyzer

In [16]: sid = SentimentIntensityAnalyzer()

In [17]: df['score'] = df ['Comment'].apply(lambda Comment:sid.polarity_scores(Comment))

In [18]: df['compound'] = df['scores'].apply(lambda d:d['compound'])

In [19]: df['comp_score'] = df['compound'].apply(lambda score: 'pos' if score >=0 else 'neg' )

In [20]: df.head()

Out[20]:
```

	date_and_time	Comment	scores	score	compound	comp_score
0	2017-09-03T00:00:00.000Z	Love it	{'neg': 0.0, 'neu': 0.192, 'pos': 0.808, 'comp...	{'neg': 0.0, 'neu': 0.192, 'pos': 0.808, 'comp...	0.6369	pos
1	2017-06-06T00:00:00.000Z	Great Tablet	{'neg': 0.0, 'neu': 0.196, 'pos': 0.804, 'comp...	{'neg': 0.0, 'neu': 0.196, 'pos': 0.804, 'comp...	0.6249	pos
2	2018-04-20T00:00:00.000Z	My K Fire	{'neg': 0.706, 'neu': 0.294, 'pos': 0.0, 'comp...	{'neg': 0.706, 'neu': 0.294, 'pos': 0.0, 'comp...	-0.3400	neg
3	2017-11-02T17:33:31.000Z	Confused	{'neg': 1.0, 'neu': 0.0, 'pos': 0.0, 'compound...	{'neg': 1.0, 'neu': 0.0, 'pos': 0.0, 'compound...	-0.3182	neg
4	2018-04-24T00:00:00.000Z	Great, simple, and resonably priced tablet.	{'neg': 0.0, 'neu': 0.549, 'pos': 0.451, 'comp...	{'neg': 0.0, 'neu': 0.549, 'pos': 0.451, 'comp...	0.6249	pos

```
In [21]: df['compound'] = df['scores'].apply(lambda d:d['compound'])
```

Figure3. 8: Coding segment for taking a count of each segment.

3.1 Commercialization Aspect of the Project

The target market for this platform is consisting of two segments. That is the PC parts and Laptop seekers and the PC parts and Laptop vendors. Currently "TechRing" is focusing only the local market. The seekers category is then sub divided in to two categories as

- Gaming population.
- Student population.

From the total Sri Lankan population, that is 21.44 Million, 18% of the population represent the school community and 14% represent the Gaming population in Sri Lanka. Thereby our user base will be approximately 32% of the total Sri Lankan population.

Considering the vendor sector, this is considered as an industry in Sri Lanka. Governmental Industrial Statics conducted in 2016 stated that this is placed 26th among the industries in Sri Lanka. By the year 2016 there are 32 establishment under this category. Under these establishments there are 3226 people involved. This set of people will be our target vendors that needs to be onboard.

Initially we will be using a cost effective Market Plan. Thus the most effective marketing platform is the Digital Marketing. We will be using the Facebook, Instagram and YouTube for our marketing purpose. Below is a diagram of our Business Canvas that we have designed.

<table>
<tr>
<td rowspan="2">Key Partners
• Local PC parts Vendors
• Local Laptop Vendors
• E-Commerce Web sites
• Facebook

Key Resources
• Online Platform
• Registered Users
• Local Vendors
• YouTube Data</td>
<td rowspan="2">Key Activities
• Provide vendors for each product
• Provide assembling plans
• Price Optimization
• Display updated and analyzed customer reviews</td>
<td rowspan="2">Value Proposition
• Assembling plan customized according to game preference of customers
• Convenience – One stop for both PC parts and laptop requirements
• Speedy and on time assistance
• Trustworthy Suggestions. Notify customer about the price drops
• Real time data extraction</td>
<td rowspan="2">Customer Segments
• We will be dealing with a Mass market. Thus our customer base will be ranging from school students to professionals.
• Both expert and Non-expert person will be using this platform.
• Our target age range will be 15yrs to 60yrs or above</td>
<td>Customer Relationships
• Automated Service
• Customer Requirements are detected and then the system will provide necessary responses.
E.g. – Provide product Suggestions

• Customized Assistance
• Solutions provided will be based on each customer requirements
E.g. – Games specified Assembling Sequences</td>
</tr>
<tr>
<td>Channels
• Our main purpose is to direct the buyer to the best seller. Product delivery is the vendor's responsibility
• We will market our services initially using social media platforms such as Facebook Instagram and YouTube</td>
</tr>
<tr>
<td>Cost Structure
• Platform development costs
• Vendor acquisition cost
• Marketing the platform and services</td>
<td colspan="3">Revenue Streams
• Initially, we plan to build our revenue streams through customers
• The registered users are allowed o have one assembly plan build for free. This free test sample is given for them to identify our service. But for other plans, they need to make a payment and afterwards, only we provide them with the assembly plan.
• Later on, with our development rate and increased publicity we will ask the new vendors to pay a registration fee to enroll with our platform.
• In future we plan to advertise their products if there are any special promotions campaigns conducted.</td>
</tr>
</table>

Table 2. 1 - Business Canvas.

3.2 Testing and Implementation

- Unit Testing

The objective of the unit testing is to isolate a section of code and verify its correctness. In the SDLC unit testing is the first level of testing before doing the integration testing. [1] Our product "TechRing" has different sub units so we have to test the each sector individually before integration with other unit/ components. We divide our sub components into 4 main units which are logically separated units and over role there are 6 sub units implemented.

Each module is separately tested which will be explained in the test cases.

We can easily test the units and make sure they produce the expected output. Each member is responsible for their key area. They need to conduct the unit testing before the integration conducted

- Integration Testing

After the unit testing we integrate each unit and begin the integration testing. Integration testing is designed as a type of testing modules integrated locally and test the product as a group.

The four main components of our system are developed by four members in our group. Some components depend on the other component output so we have to test those parts separately together and then integrate. For example in price optimization they results of customer comments is also taken to considerations. Thus these two modules will be integrated first before integrating them separately into the system. These tests ensure the communication of data between the components of the system.

- System Testing

One of the black box testing method. After integrated all the components we have to do the system test to check whether the system functionalities meet the expected output. These testing enhanced the user user's experience with the application. These test done before introduce to the market.

Test Case No	Test Case - 01
Description	Scrape the product user comments from the YouTube
Pre-Conditions	1. System should be internet connection access to the application
Test Procedure	1. Visit the "TechRing" search the product 2. Click the product to view more details 3. After analyze the comments rating display bottom of the product image
Input	Product Name
Expected Output	1. Customer feedback rating should be calculated and display

Table 2. 2 -Test Case for comment Analysis.

Implementation

I. Scrape user comments/feedbacks from the YouTube videos.
Selenium is used to scrape the user feedback from YouTube.
After scraping the comments will be stored in CSV file.

II. Analysis of user feedbacks to generate the ratings
Python NLTK Library is used to analyze the Positivity or Negativity of the comments.

3 RESULTS AND DISCUSSION

3.1 Results

"TechRing" main focus is to provide what is BEST for the customers. Thus Customer feedback is a fact that is highly influencing in determination of the ideal hardware option. Therefore incorporating this feature in "TechRing" enhances the capability to find the best option. It has a high accuracy in analysis since I have used the RNN to for feedback analysis. According to the analyzed results the ratings will be displayed as shown in the figure.

It is one of the necessity to look into the customer feedback on the products. Yet most people do not have the time to go through these comments in order to get an idea of the experiences. Thus the feature introduced in "TechRing" helps the users to get an idea about the user experiences at glance giving them enough time to find another option if there is high negative feedback for the product.

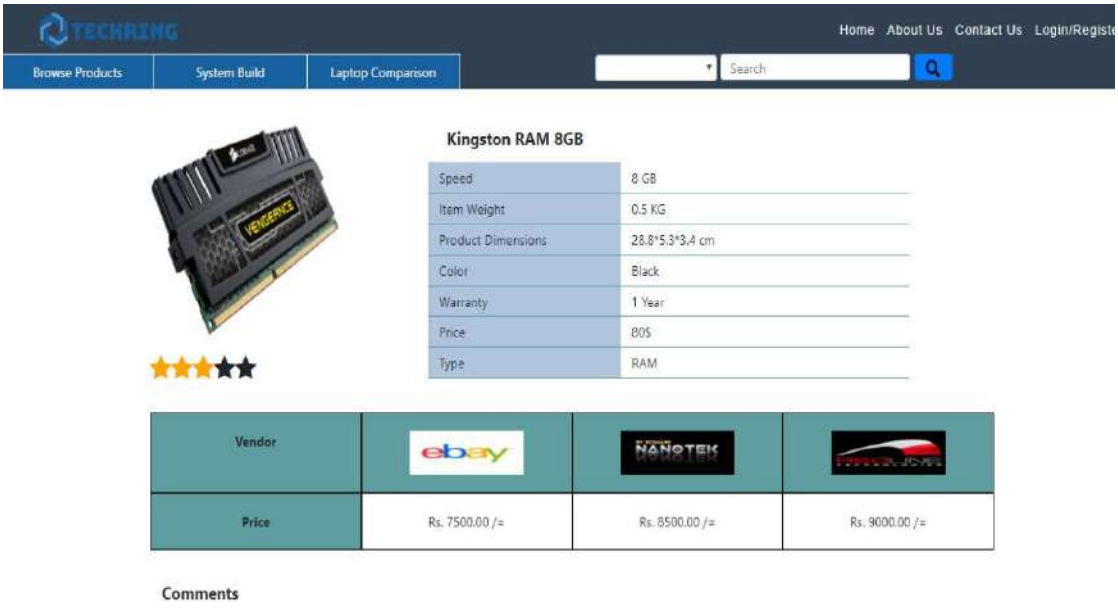

Figure4. 1: Interface displaying the ratings for each product.

Time take for comment analysis vary with the number comments that needs to be extracted. It will take an average time of 123 seconds to extract set of comments ranging from 100 to 1000.

Will be taken in order to extract the comments from YouTube. We tested the time taken to extract some sample sets of comments.

Video	**Number of Comments.**	**Time Taken.**
Sample - 1	150 comments	64 Seconds
Sample - 2	500 comments	125 Seconds
Sample - 3	700 comments	180 Seconds

Table 3. 1 – YouTube Comments crawler time comparison

3.2 Research Findings.

In order to conduct sentiment analysis there are number of tools, techniques and algorithms available. From all the Natural Language Processing tools available it is said that Natural Language ToolKit is the most usable and is considered as the mother of all NLP's. For the sentiment analysis I have also used the NLTK [5]. Thus the accuracy of the results are guaranteed.

Various NLP Libraries

NLP Library	Description
NLTK	This is one of the most usable and mother of all NLP libraries.
spaCy	This is completely optimized and highly accurate library widely used in deep learning
Stanford CoreNLP Python	For client-server based architecture this is a good library in NLTK. This is written in JAVA, but it provides modularity to use it in Python.
TextBlob	This is an NLP library which works in Pyhton2 and python3. This is used for processing textual data and provide mainly all type of operation in the form of API.
Gensim	Genism is a robust open source NLP library support in python. This library is highly efficient and scalable.
Pattern	It is a light-weighted NLP module. This is generally used in Web-mining, crawling or such type of spidering task. p
Polyglot	For massive multilingual applications, Polyglot is best suitable NLP library. Feature extraction in the way on Identity and Entity.
PyNLPl	PyNLPl also was known as 'Pineapple' and supports Python. It provides a parser for many data format like FoLiA/Giza/Moses/ARPA/Timbl/CQL.
Vocabulary	This library is best to get Semantic type information from the given text.

Figure5. 1: NLTK ranking.

One of the most common social media platform used by the society is Facebook. Earlier the rules of Facebook allowed to access the Facebook data by a Third party. But now they have tightened the rules and regulations. This resulted for us to use another social media platform to retrieve the customer feedbacks like YouTube and Twitter.

3.3 Discussion

Main motive of the "TechRing" is to be the ideal online assistant when it comes to searching the hardware components. This is to be one stop for both your PC parts and Laptop requirements. Any person with internet and average IT knowledge can access our platform and we will be readily available to satisfy your hardware requirements.

In order to identify the tool or technique to be used for sentiment analysis we did a prior analysis on the existing tools. For analysis of the techniques we contacted the experts, our lecturers, refereed documents, YouTube tutorials, and blogs. From the analysis we came

to a conclusion that we will be using NLTK as the analysis tool since it generate accurate results and it is also easy to use.

4 CONCLUSION

"TechRing" is a web based online assistant designed to provide solutions for your hardware requirements. This provides customized services especially to the Gaming Society. One of the major service that is provided by "TechRing" is providing hardware options matching the user requirements. In finding the matching requirements there many factors that affect like, compatibility matching, budget limitations, Vendor options and also the previous customer experiences.

In this document it focuses on the customer feedback analysis. Since people are curious of the customer feedbacks before purchasing product we decided to include this feature to our platform. Hus an analyzed overview of the customer feedbacks available in the social media platforms, especially YouTube will be taken in to considerations. From this feature it saves the time users need to spend in search and reading comments and posts posted related to the product that you are looking for. Because "TechRing" takes the burden off your shoulders and analyses these comments for you. This method used in "TechRing" is proved to have the best accuracy in sentiment analysis so far from the available tool for this purpose. Because it is not limited only to check the positive or negative words in the comments but it also checks and validate the relationship of one word has with another. Therefore it will be extracting the actual meaning of the post. Moreover, this platform will be having up to date information since it will be using real time data extraction.

5 REFERENCES

[1] F. N. Leo Rizky Julian, "THE USE OF WEB SCRAPING IN COMPUTER PARTS AND ASSEMBLY PRICE COMPARISON," 2015.

[2] S.-T. Tan, "Multimedia Based PC Assembly Learning Tool," 1996.

[3] N. D. Udi Boker, "Comparing Computaional Power," 2015.

[4] W. Hou, X. Li, Y. Jin and J. Wu, "A Study of Intelligent Decision-Making System Based on Neural Networks and Expert System," 2013.

[5] P. J. B. a. R. C. Jain, "Three-dimensional object," vol. 1, 1985.

[6] R. T. Chin and C. R. Dyer, " “Model-based recognition," vol. 18, 1986.

[7] a. R. B. F. Solina, " “Recovery of parametric models from range images : the case of superquadrics with global deformation"," Vols. vol:1, vol 2, 1990.

[8] K. N. Kirithika B, "Comparison of Intel processor with AMD processor with Green Computing," 2013.

[9] C. Kocas, "Online price competition within and between Heterogeneous Retailer Groups," 2004.

[10] R. H. Jianxia Chen, "A price comparison system based on Lucene," April 2013.

[11] A. Zalozhnev, "The ICT Products Prices and Quantities".

[12] A. Salinca, "Business reviews classification using sentiment analysis.," 2016.

[13] S. R. S. J. Zeenia Singla, "Statistical and Sentiment Analysis of consumer product reviews," 2017.

[14] B. L. Minqing Hu, "Mining and Summarizing Customer reviews.".

[15] A. L. Robert Ireland, "Application of data analytics for product design: Sentiment Analysis of online product reviews," 2018.

[16] Y. N. A. O. I. O. BabolaT. Issac, "Assemblin a Desktop Computer System with In-Bult Uninterrupted Power Supply.," 2017.

[17] Y. 2. LinghuiLiu1, "ApplicationofAgileMethodintheEnterprise WebsiteBackstageManagementSystem," 2012.

[18] A. J. M. Kamaljeet Kaur, "Applying Agile Methodologies in Industry Projects: Benefits and Challenges," 2015.

[19] J. D. A. a. O. B. Shvetha Soundararajan, "A Methodology for Assessing Agile Software Development Methods," 2012.

[20] [Online]. Available: http://jaspervanderhoek.com/wp/methodology/agile/pursuing-a-fully-agile-software-lifecycle/ [image].
S. I.V, "Analysing Sentiments with NLTK," December 2016. [Online]. Available: https://opensourceforu.com/2016/12/analysing-sentiments-nltk/.

6 APPENDICES

6.1 Survey Results.

- **Survey**

https://docs.google.com/forms/d/1kcaZ96I1M7lUrWDbo9UqmquWT8q9dlo-mg8O4J8HT68/edit#responses

- **Survey Results**

https://drive.google.com/file/d/1COKVjgpRApagijZNzkqG1XN5q-IMzPJn/view?usp=sharing

6.2 Appendix A: Use case Diagram.

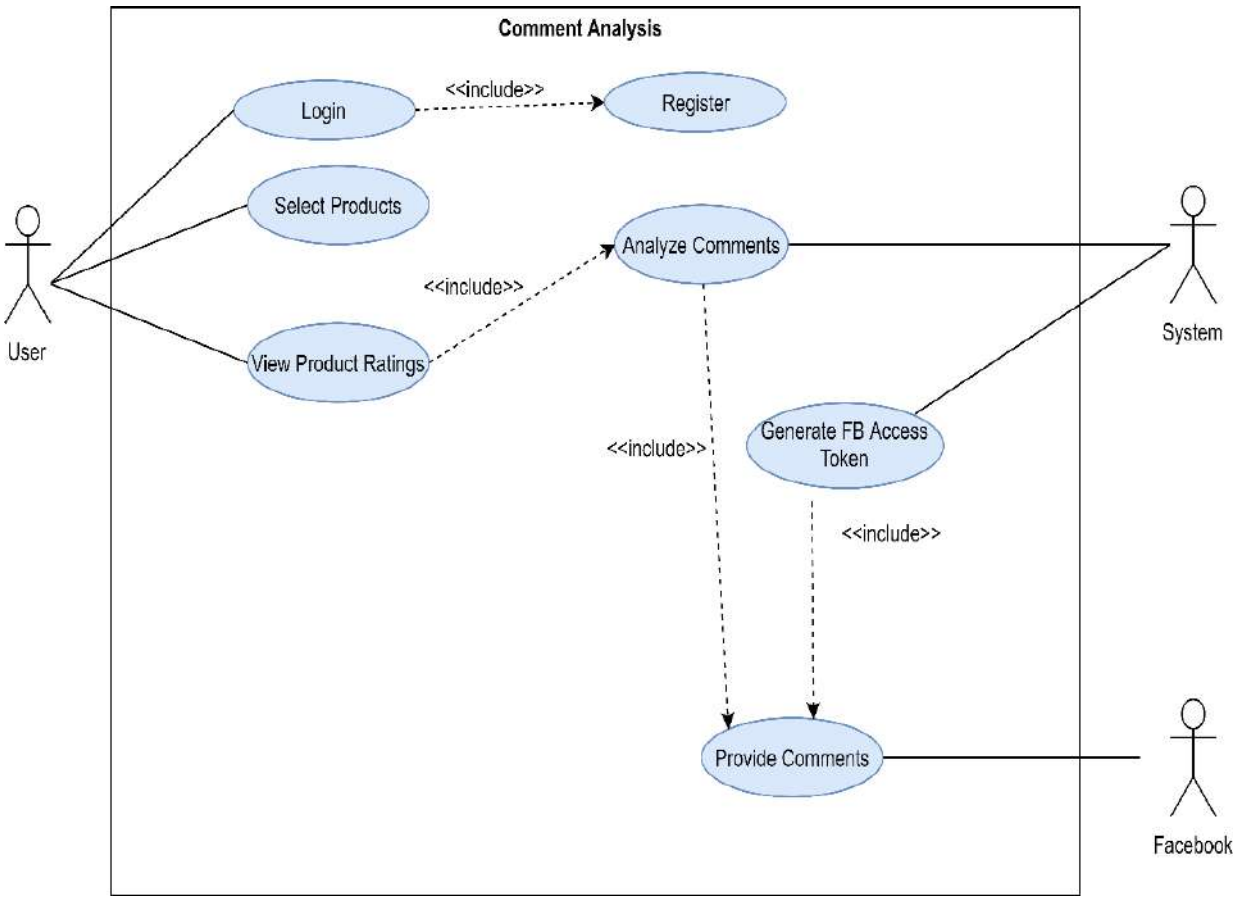

Figure 6. 1 -Use case Diagram for Comment Analysis

6.3 Appendix B: Activity Diagram.

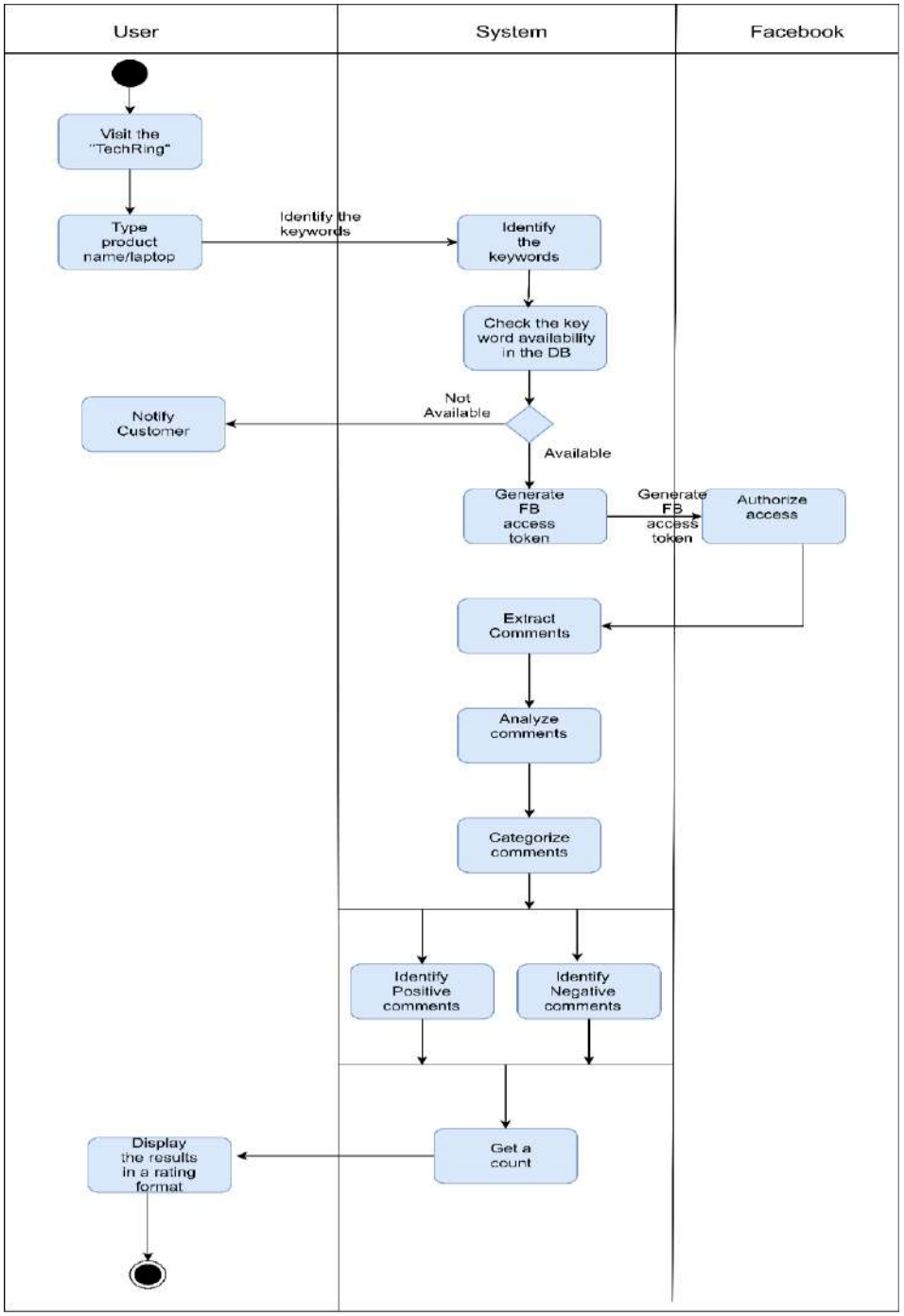

Figure 6. 2- Activity diagram for Comment Analysis.