Informatics Institute of Technology

In Collaboration With

University of Westminster, UK



*University of Westminster, Coat of Arms*

Generalized Abstractive Text Summarization Using Optimized Transformers

Software Requirements Specification (SRS)

Mr. Nazhim Kalam

w1761265 / 2019281

Supervised by

Mr. Torin Wirasingha

September 2022

This Project Proposal is submitted in partial fulfilment of the requirements for

the BSc (Hons) Computer Science degree at

the University of Westminster.

Table of Contents

[List of Figures ii](#_Toc119965928)

[List of Tables ii](#_Toc119965929)

[1. CHAPTER OVERVIEW 1](#_Toc119965930)

[2. RICH PICTURE 1](#_Toc119965931)

[3. STAKEHOLDER ANALYSIS 2](#_Toc119965932)

[3.1 Stakeholder Onion Model 2](#_Toc119965933)

[3.2 Stakeholder Viewpoints 3](#_Toc119965934)

[4. REQUIREMENT ELICITION METHODOLOGIES 4](#_Toc119965935)

[5. ANALYSIS OF DATA & PRESENTATION OF THE OUTCOME THROUGH ELICITATION METHODOLOGIES 5](#_Toc119965936)

[5.1 Literature Review 5](#_Toc119965937)

[5.2 Survey 6](#_Toc119965938)

[5.3 Interviews 9](#_Toc119965939)

[5.4 Prototyping 9](#_Toc119965940)

[5.5 Brainstorming 9](#_Toc119965941)

[6. SUMMARY OF FINDINGS 10](#_Toc119965942)

[7. CONTEXT DIAGRAM 10](#_Toc119965943)

[8. USE CASE DIAGRAM 11](#_Toc119965944)

[9. USE CASE DESCRIPTIONS 11](#_Toc119965945)

[10. REQUIREMENTS 13](#_Toc119965946)

[10.1 Functional Requirements 13](#_Toc119965947)

[10.2 Non-functional Requirements 14](#_Toc119965948)

[11. CHAPTER SUMMARY 14](#_Toc119965949)

[REFERENCES I](#_Toc119965950)

[APPENDIX A – CONCEPT MAP II](#_Toc119965951)

# List of Figures

[Figure 12.1 - Prototype Feature Diagram (Self-composed) 13](#_Toc117550682)

[Figure 13.1 - Gantt Chart 16](#_Toc117550683)

[Figure 13.2 - Model development flow (Self-composed) 21](#_Toc117550684)

# List of Tables

[Table 5.1 - Related work in abstractive text summarization 3](#_Toc117584436)

[Table 11.1 - Research Objectives 9](#_Toc117584437)

[Table 13.1 - Research Methodology 13](#_Toc117584438)

[Table 13.2 - Deliverables and Dates 17](#_Toc117584439)

[Table 13.3 - Risk Mitigation Plan 19](#_Toc117584440)

**Acronyms**

|  |  |
| --- | --- |
| AI | Artificial Intelligence. |
| DL | Deep Learning |
| ML | Machine Learning |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# CHAPTER OVERVIEW

# RICH PICTURE

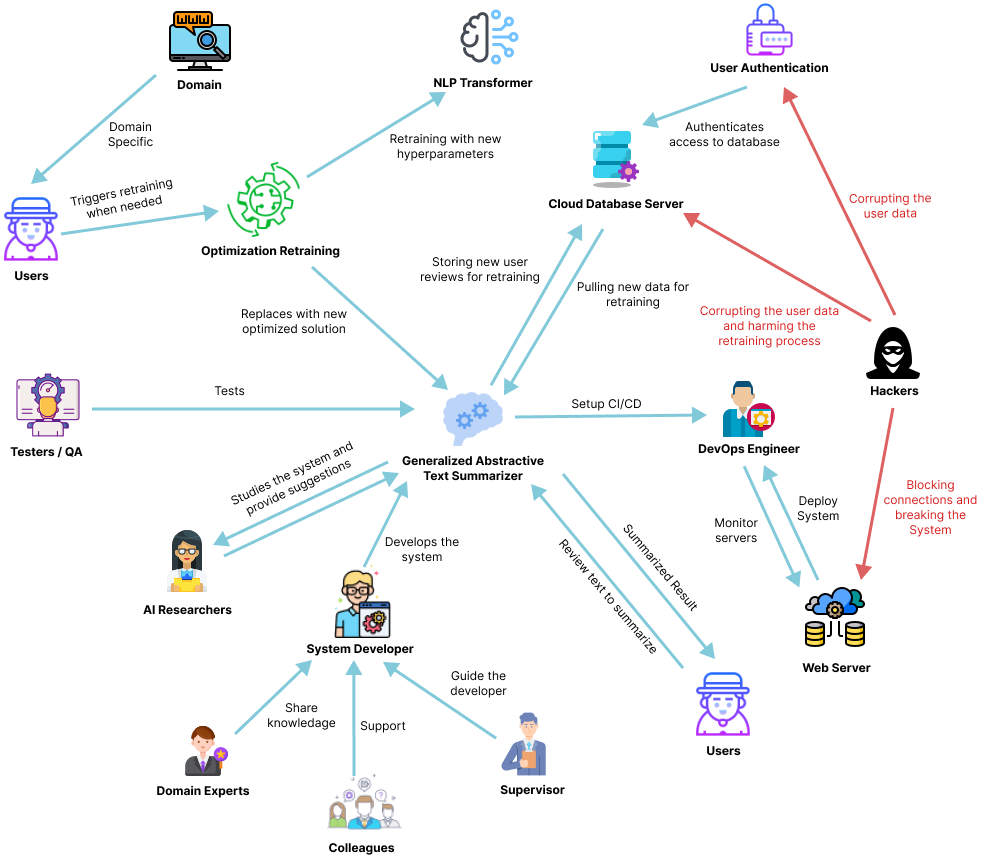
****

Figure 2.1 – Rich Picture Diagram (*self-composed*)

# STAKEHOLDER ANALYSIS

## 3.1 Stakeholder Onion Model

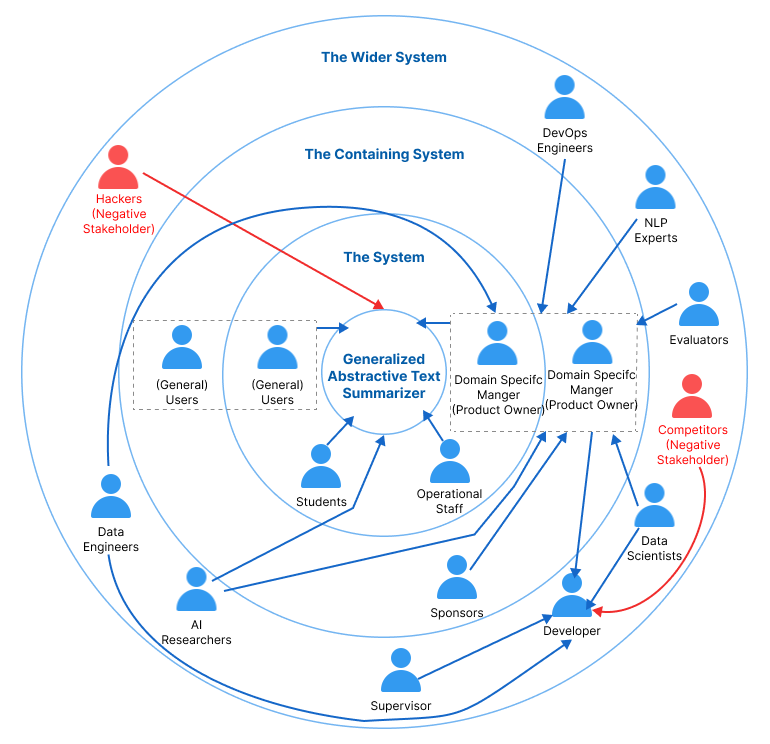


Figure 3.1 – Stakeholder Onion Model (*self-composed*)

## 3.2 Stakeholder Viewpoints

|  |  |  |
| --- | --- | --- |
| **Stakeholder** | **Role** | **Benefits/ Role Description** |
| Developer | Financial Contributor | Works on developing the system |
| Sponsors | Funds to market the system and allows developers to advance the system with time. |
| Data Scientists | Quality Control Regulator | Provides performance enhancements for the models and algorithms used in data science. |
| Data Engineers | Gives guidance on potential data that may be used to generate the best suggestions possible. |
| AI Researchers | Conduct research in the specified area to enhance and implement reliable text summarizing models. |
| NLP Experts | Offers specialized guidance and insights on the field  knowledge, to enhance the functionality of the system. |
| Students | Functional beneficiary | Understanding how the generalization works along with hyperparameter retraining with respect to the domain. |
| Domain Specific Users | System Owner and Operational | Inputs the text reviews for abstractive summarization and handles the need for model retraining when need to improve performance with the previously used inputs as new data for the model. |
| General Users | General Usage | General users (Not Domain Specific) will be using the general abstractive summarization model without any specific model assigned to them and hyperparameter retraining unless they want to. |
| Operational Staff | Operations support | Ensures that the system is up and functioning while responding to user requests and problems. |
| DevOps Engineers | Product Deployment & Maintenance | Makes ensuring the system is up and running in the cloud and is serving users without being throttled |
| Hackers | Negative Stakeholder | May manipulate the review data stored in the database which will affect the retraining process. |
| Competitors | May build competing systems that may outperform the existing system. |
| Evaluators | Quality Inspector | Checks to see if the system is ready for production use and puts it through its paces. |

# REQUIREMENT ELICITION METHODOLOGIES

There were several requirement elicitation approaches used to collect needs for the creation of the research project. The approaches selected for this were literature review, survey, and prototype. The following is a discussion of the rationales behind selecting the mentioned requirement elicitation approaches.

|  |  |
| --- | --- |
| **Method** | **Description** |
| Literature Review | To determine research gaps in the chosen domain of interest and the intended topic of study at the project's outset, the author conducted a thorough literature analysis. Current systems were researched together with comparable technologies that might be applied to the existing systems that were referenced in literature in order to discover research gaps available in technologies that can be used. |
| Survey | A questionnaire was utilized as a survey instrument to obtain requirements and opinions from possible users of the suggested system. The author will benefit from this sort of poll in understanding people's thought processes and expectations for the prototype. It will also enable the author to explain whether or not the targeted users will benefit from the suggested solution. |
| Interviews | Interviews were performed to gain expert insight into domain-specific requirements and to determine the best method to address the issue at hand while adding to the body of knowledge through research. Interviews were determined to be the greatest source of information because the field is new and the technical expertise needed is very precise. Additionally, this technique allowed for the qualitative evaluation of the suggested system, allowing for the identification of any shortcomings or difficulties that could need to be resolved during prototyping. |
| Prototyping | The project was chosen to follow the Agile Software Development Life-cycle, thus prototyping would allow the author to test and evaluate the prototype while iteratively trying out several alternative implementations to find any potential areas for improvement. |
| Brainstorming | Whether you're attempting to come up with a broad subject before you start your research, you're trying to focus more specifically, or you're determining what evidence to use for a particular paragraph, brainstorming is a useful technique to produce ideas at every step of the process. In order to assess the system for personally, the author has a number of brainstorming sessions with his colleagues at various project stages. |

# ANALYSIS OF DATA & PRESENTATION OF THE OUTCOME THROUGH ELICITATION METHODOLOGIES

The data analysis from the requirement elicitation methods that were selected are shown below.

## **5.1 Literature Review**

Table 5.1 – Requirement Analysis from literature review findings

|  |  |
| --- | --- |
| **Finding** | **Citation** |
| In the completion of the literature review on the existing work done, it was identified that abstractive text summarization systems for customer reviews helps users to make better and quicker decisions on their actions let it be on buying products or watching a movie, user review summarization proves to save time for customers. | (Boorugu, Ramesh and Madhavi, 2019) |
| When exploring technologies that can be applied to achieve the required outcome, it was clear that traditional machine learning and deep learning approaches were only used for abstractive text summarization in the domain of movie reviews. Leaving the usage of advanced deep learning approaches such as Transformers untouched for this domain. | (Khan et al., 2020) |
| It was identified that transformer optimization has not been looked into when working with transformers in abstractive text summarization domain in general and not specific to the movie domain. | (Gupta et al., 2021) |
| Dataset related to working with model generalized has been used previously and is recommended to be used if researchers are willing to work with the idea of generalization for the domain of abstractive text summarization. | (Kouris, Alexandridis and Stafylopatis, 2019) |

## 

## **5.2 Survey**

Table 5.2 – Requirement Analysis from questionnaire

|  |  |
| --- | --- |
| **Question** | Have you ever realized that reading lengthy reviews takes a significant amount of time? |
| **Aim of question** | To determine whether the audience as a whole considers reading lengthy reviews to be a time-consuming activity. |
| **Findings & Conclusion**  It can be concluded that a large part of the audience (more than 90% of the audience) finds that’s reading lengthy reviews is a time-consuming hassle which also proves that they would appreciate if there would be a quicker approach for this problem, like a summarization. This also concludes to see a positive correlation from the results which was expected from the author of the project. | |
| **Question** | Do you believe that developing a generic system for all domains would be a wise course of action? |
| **Aim of question** | Ensuring that developing a generic system would be beneficial in all domains |
| **Findings & Conclusion**  It can be concluded that most of the participants (more than 90% of the audience) agrees that developing a generalized system which can adapt to the domain as they use, is beneficial and worth the effort to process with the project research. This also concludes to see a positive correlation from the results which was expected from the author of the project | |
| **Question** | Who do you think will most benefit from this system? |
| **Aim of question** | Getting to know about the thoughts of the participants about to whom the system would mostly benefit from? |
| **Findings & Conclusion**  It can be concluded that a majority of the participants (more than 60%) finds that this system will benefit the movie, restaurant, tourist, hotel, ecommerce domains (these domains were considered since they are mostly interacted with the users on a daily bases and uses customer reviews for their domain as a part of their business) as well as the general public users. | |
| **Question** | How much do you think that this system would benefit you? |
| **Aim of question** | Getting to know how much the system would benefit the general participants which are NOT domain specific |
| **Findings & Conclusion**  From the statistics graph, it can be concluded that roughly 75% of the audience finds that the system would benefit them for their general work or needs given that its not domain specific to them, which is a positively correlated result from the achieved statistics. | |
| **Question** | How much do you think that this system would benefit businesses? |
| **Aim of question** | Getting to know from the participants as to how much the system would benefit businesses/domains in solving this problem. |
| **Findings & Conclusion**  From the statistics graph, it can be concluded that roughly 84% of the audience finds that the system would benefit the businesses, which is a positively correlated results from the achieved statistics and that’s what the author expected to achieve. | |
| **Question** | Before making a reservation or booking a movie or a hotel, do you read the customer reviews? |
| **Aim of question** | Getting an idea from the audience if in general they give importance to customer/user reviews to any domain before consuming their product or services. |
| **Findings & Conclusion**  It can be concluded that a majority of the participants (more than 95% of the audience) agrees that they value and read customer reviews before they consume one’s product or service. Therefore, making customer reviews a major contributing factor for business growth. | |
| **Question** | How much you think customer reviews are important with respect to any domain? |
| **Aim of question** | Getting an idea from the audience to see how much they value customer reviews. |
| **Findings & Conclusion**  From the statistics graph, it can be concluded that roughly 90% of the audience finds that customer/user reviews are very important irrelevant to the domain, which is a positively correlated results from the achieved statistics and that’s what the author expected to achieve. | |
| **Question** | Which additional features would you want to see in this system. |
| **Aim of question** | To identify the systems non-functional requirements which could potentially improve the system. |
| **Findings & Conclusion**  The majority of participant responses were concerned with classifying the review text's sentiment after it had been summarized and with managing a list of review uploads so as to add filtering for the summarized review text based on the sentiment, whether it was positive or negative, along with the sentiment score. | |

## 

## **5.3 Interviews**

In order to get opinions of the technical domain expertise, interviews were conducted with the experts & researchers from the respective fields. Experts & researchers in ML and Abstractive text summarization systems were chosen to be interviewed in order to establish project requirements.

Table 5.3 – Requirement Analysis from interviews

|  |  |  |
| --- | --- | --- |
| **Interviewer** | **Designation** | **Findings** |
| Mrs. Krishnakripa Jayakumar | Lecturer & NLP researcher | The interviewee considered that BERT, GPT-2, RoBERTa are few of the top tier transformer architecture. Regarding hyperparameter tuning it told that manual implementation takes a lot of time but their tools to save time like Raytune, Optuna. It’s also discussed that working with a single domain for the time frame of the project is a good scope but creating a generalized solution will be challenging. For evaluation metrics ROUGE and BLEU can be used but ROUGE is more preferred. |
| Mrs. Nelum Weerakoon | Technical Architect & NLP researcher |  |
| Mr. Dinuka Piyadigama | Software Engineer |  |
| Prof. Nihal | Lecturer & ML Expert |  |

## **5.4 Prototyping**

Several challenges and requirements emerged though the iterative prototyping. Since this project problem domain was related to the field of movies, finding the dataset with the expected metadata was challenging, however through multiple research paper reviewing the author was able to find the datasets but the amount of data present was massive (around 8million records), the author had to break it down to multiple dataset portions to work with. However, finding the dataset for model generalization wasn’t challenge but preprocessing the data was a major task since the dataset was massive and also contained noisy text data, which needed a lot of cleaning.

Since manual hyperparameter tuning was a very time-consuming task, a research for hyperparameter tuning libraries were experimented and “Optuna” worked for our needs when it comes to model automatic optimized retraining. New data entered by the domain user of the application has been designed to be stored and used for retraining their model. At least 5 top tier transformer architectures had to be explored with for choosing the best architecture.

## **5.5 Brainstorming**

The author engaged in brainstorming across various project phases. These were carried out both with the authors' colleagues and supervisors as well as through a self-analysis process.

Table 5.4 – Requirement Analysis from brainstorming

|  |
| --- |
| **Purpose** |
| Able to figure out several other research gaps/ limitations which can be fit into the current project domain in order to increase the magnitude of research effort. |
| **Findings** |
| Multiple ideas were brought up as the result of the brainstorming session. The concept of creating a performance adaptive generalization model was brought up by the authors supervisor, along with several other approaches to increase the performance of the system exponentially such like making use of the new data from the domain users for retraining and combine all data with the common domain for retraining since the data count increases with respect to the common domain user. |

# SUMMARY OF FINDINGS

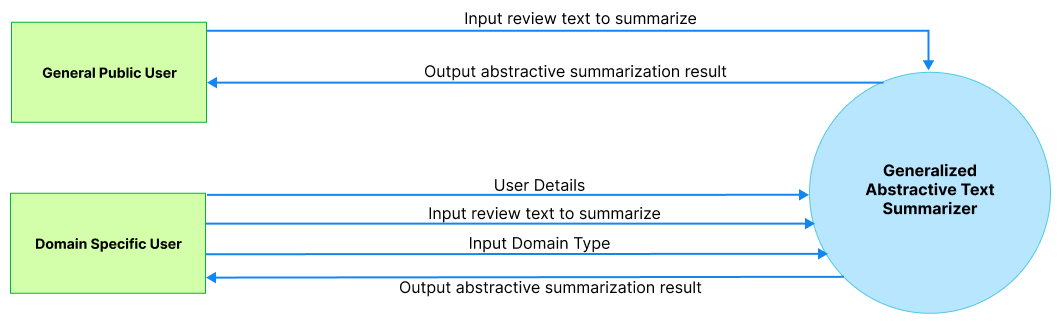
Table 5.3 – Summary of Findings

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Id** | **Finding** | **Literature Review** | **Survey** | **Interview** | **Brainstorming** | **Prototyping** |
| 1 | The proposed system would benefit businesses (domain specific users) and general users (not domain specific) |  | ✓ |  | ✓ |  |
| 2 | For the movie domain the limit of abstractive text summarization can be further pushed using optimized transformers to increase performance this being the existing limitation | ✓ |  | ✓ |  |  |
| 3 | Its clear that customer/user reviews are valued and reviewed mostly by a vast majority of the audience before they consume or use any product or service (applies to any domain) | ✓ | ✓ |  | ✓ |  |
| 4 | Its clear that users spend lot of time review long reviews and they would like it being short to save time and make quicker decisions. | ✓ | ✓ |  | ✓ |  |
| 5 | Hyperparameter tuning is one way to increase the performance of the system and it can be done both manually or by automated tools like Raytune, Optuna etc.… | ✓ |  | ✓ |  | ✓ |
| 6 | Data preprocessing for the domain of Movies and Generalization is requires a lot of effort since the datasets are mostly raw data difficult to find specially in the case of movie reviews (with expected metadata) | ✓ |  |  |  | ✓ |
| 7 | Additional features such as sentimental and sentimental score of the review summary is mostly expected from the users. |  | ✓ |  |  |  |
| 8 | Creating a hybrid transformer model to further increase the performance is a suggested improved. |  |  | ✓ | ✓ |  |
| 9 | It’s clear on what are the suitable evaluation metrics to be used for abstractive text summarization. | ✓ |  | ✓ |  |  |
| 10 | It’s clear on what the top tier transformer architecture that could be explored. | ✓ |  | ✓ |  |  |
| 11 | Making use of larger new data for retraining for a specific domain, from companies/businesses who uses data which are of the same domain. (e.g.: - 50 different restaurants data can be combined for retraining give that the domain is “Restaurants”) |  |  |  | ✓ |  |

# CONTEXT DIAGRAM

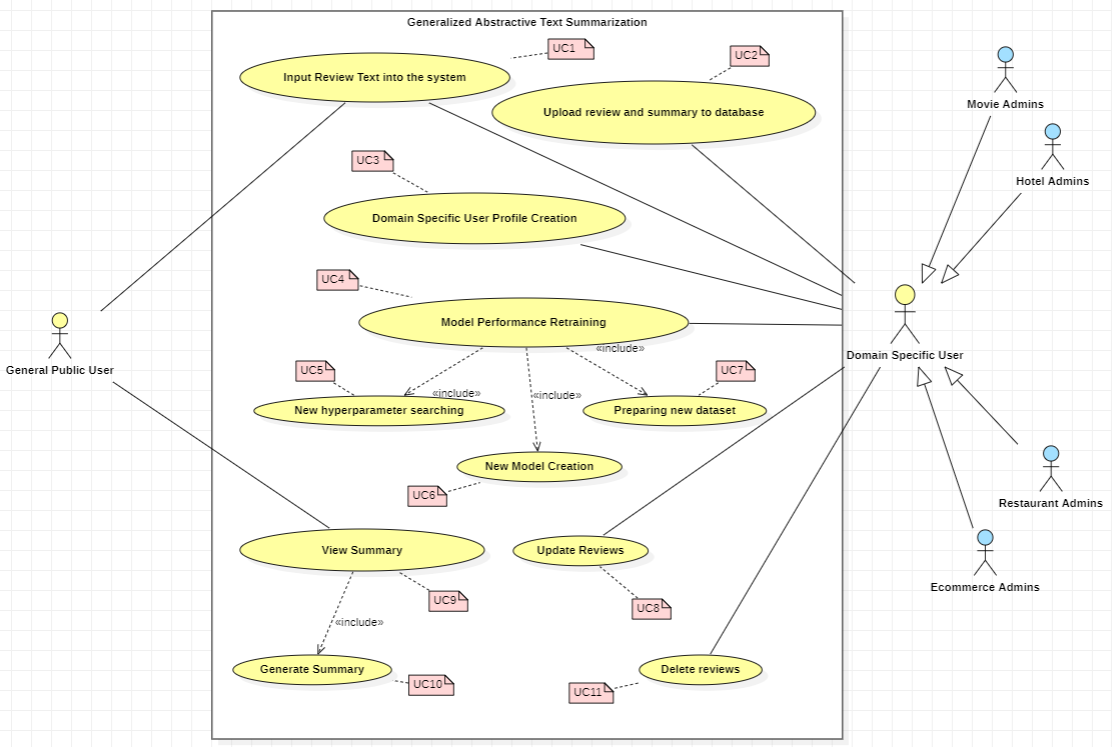
The boundaries and interactions of the system should be established before development. The graphic below shows how the system is situated.

Figure 7.1 – Context Diagram (self-composed)



# USE CASE DIAGRAM

Figure 8.1 – Use case Diagram (self-composed)



# USE CASE DESCRIPTIONS

Table 9.1 – Use case description UC:0X

|  |  |
| --- | --- |
| Use Case |  |
| Id |  |
| Description |  |
| Primary Actor |  |
| Supporting Actors (if any) |  |
| Stakeholders and Interests (if any) |  |
| Pre-Conditions |  |
| Post Conditions |  |
| Trigger |  |
| Main Success Scenario |  |
| Variations |  |

Table 9.2 – Use case description UC:0X

|  |  |
| --- | --- |
| Use Case |  |
| Id |  |
| Description |  |
| Primary Actor |  |
| Supporting Actors (if any) |  |
| Stakeholders and Interests (if any) |  |
| Pre-Conditions |  |
| Post Conditions |  |
| Trigger |  |
| Main Success Scenario |  |
| Variations |  |

# REQUIREMENTS

## 10.1 Functional Requirements

Based on the significance of the system demands, the MoSCoW approach was utilized to identify their priority levels.

Table 10.1 – Priority Levels

|  |  |
| --- | --- |
| **Priority Level** | **Description** |
| Must have (M) | The demand at this level is the fundamental functional requirement for a prototype, and it must be carried out. |
| Should have (S) | Although not strictly required for the anticipated prototype to function, important criteria do provide a lot of value. |
| Could have (C) | Optional, non-essential desirable needs are crucial to the project's scope. |
| Will not have (W) | Requirements that the system might not meet right now and that are not given first consideration. |

Table 10.2 – Functional requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **FR ID** | **Requirement** | **Priority Level** | **Use Case** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

## 10.2 Non-functional Requirements

Table 10.3 – Non-functional requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **NFR ID** | **Requirement** | **Priority Level** | **Use Case** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# CHAPTER SUMMARY

# REFERENCES

# APPENDIX A – CONCEPT MAP