



الأكاديمية العربية للعلوم والتكنولوجيا والنقل البحري
Arab Academy for Science, Technology & Maritime Transport

Project Name :

Course Booking

Doctor's advisors:

Dr. Ahmed Salem

Dr. Fahima El Maghraby

Project Team:

| Name | ID |
|-----------------------------|----------|
| Muath Hisham Ismail | 19106673 |
| Abdalkareem Mounir Assi | 19106795 |
| Abdelrahman Hany | 19108297 |
| Ahmed Ashraf | 19104771 |
| Karim Aly Ibrahim Elmoghazy | 19106009 |

AKNOWLEDGEMENT

The path that led us to where we are today was not easy, but we got here eventually. We take this opportunity to thank Allah SWT for guiding and giving us the strength, as well as providing us with valuable resources, both people and information, that helped us complete this task in the way you will see.

We also would like to express our sincere gratitude to both our supervisors Dr. Ahmed Salem and Dr. Fahima Maghraby for their help and guidance throughout the process of this project. The knowledge they provided greatly helped guide us to the correct progress path.

Last but not least, we are extremely grateful to our parents for their support, love, sacrifices and prayers that gave us the strength from the beginning of our college experience till we reached the end of the road. We truly wouldn't have reached where we are today without them.

DECLARATION

We certify that this project that we are submitting for the *Arab Academy for Science, Technology and Maritime Transport - College of Computing and Information Technology*, is completely our own work that we went to extreme measures to complete at this level. This project is completely original and was not taken from the work of anybody else, hence no copyright laws were breached. We also would like to state that information taken from other sources has been mentioned in the references at the end.

ABSTRACT

Hot Course is a platform that helps users to access educational centers that are reliable and suitable for them without wasting their time and effort by searching for a suitable educational center for them and helps educational centers that have not gained their share of fame to reach users and offer their courses to them.

Moodle TextRank was used to summarize books for users so that the user does not waste time reading the entire book if he wants.

Table of Contents

| | |
|--|-----------|
| 1 Literature Review to Auto Summarization | 7 |
| 1.1 Introduction..... | 7 |
| 1.2 Types of Text Summarization | 8 |
| 1.2.1 Abstractive summarization..... | 8 |
| 1.2.2 Extractive summarization..... | 8 |
| 1.3 Comparison between NLTK and TextRank | 9 |
| 1.3.1 Advantage to NLTK | 9 |
| 1.3.2 Disadvantage to NLTK..... | 9 |
| 1.3.3 Advantage to TextRank..... | 10 |
| 1.3.4 Disadvantage to TextRank | 11 |
| 1.4 TextRank-Based Summarization..... | 11 |
| 1.5 Similarity function..... | 12 |
| 1.6 TextRank model..... | 12 |
| 1.7 Conclusion | 14 |
| 2 Literature Review to course booking | 15 |
| 2.1 Advantages:..... | 16 |
| 2.2 Limitations: | 16 |
| 3 Market and Business Analysis | 17 |
| 1 SWOT | 17 |
| 2 4Ps | 17 |
| 3 Survey | 18 |
| 4 SRS document..... | 19 |
| 1 Introduction..... | 19 |
| 1.1 Purpose..... | 19 |
| 1.2 Intended Audience and Reading Suggestions | 19 |
| 1.3 Intended Use | 19 |
| 1.4 Project Scope..... | 19 |
| 1.5 Definitions..... | 20 |
| 2 System Description | 20 |
| 2.1 Methodology | 20 |
| 2.2 SYSTEM ANALYSIS & DESIGN | 22 |

| | |
|--|-----------|
| 2.2.1 use case diagram | 22 |
| 2.2.2 Sequence Diagram..... | 23 |
| 2.2.3 ERD Diagram | 25 |
| 2.3 Assumptions and Dependencies..... | 25 |
| 2.4 Intended Technologies | 26 |
| 3 System Requirements..... | 26 |
| 3.1 Functional Requirements | 26 |
| 3.2 External Interface Requirements..... | 26 |
| 3.2.1 User Interfaces..... | 26 |
| 3.2.2 Hardware Interfaces | 28 |
| 3.2.3 Software Interfaces | 28 |
| 3.2.4 Communications Interfaces..... | 28 |
| 3.3 System features | 28 |
| 3.3.1 Auto summarization | 28 |
| 3.3.2 signup | 29 |
| 3.4 Nonfunctional requirements | 30 |
| 4 References | 31 |

1 Literature Review to Auto Summarization

1.1 Introduction

The amount of textual material available on the Internet and in other libraries is growing enormously every day. With data growing so much at once and containing irrelevant content and noise, consuming information has become a costly and time-consuming task.

Text summarization is a method of summarizing data. A manual text summarization process is arguably an effective way to preserve the meaning of text. However, this is a time-consuming task. [1]

The process of condensing a portion of a text into a shorter version, minimizing the size of the original text while retaining important informational aspects and substantive meaning, is called an abstract. Figure 1 shows the summary task in a simple way. An abstract is a reductive transformation of a source text into a summary text by extraction.

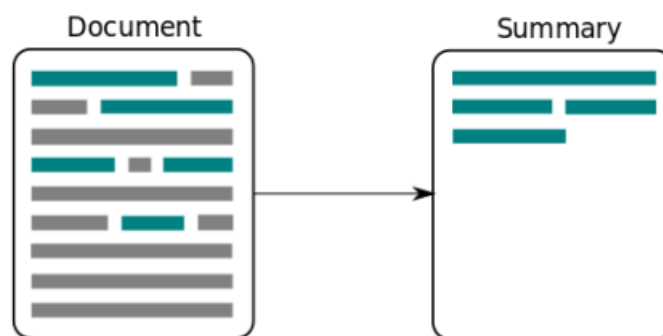


Fig. 1: generating summary from input document

According to another definition, "Automatic summaries are coherent texts generated by software that contain a substantial amount of relevant information from the source text. Its compression ratio τ is less than one third of the original document length. [2]

This article explores the area of text summarization. Understand how the TextRank algorithm works and implement it in Python. Buckle up and this is going to be a fun ride. [3]

1.2 Types of Text Summarization

There is two main type of text summarization:

1.2.1 Abstractive summarization

The adjective 'abstractive' is used to denote a summary that is not a mere selection of a few existing passages or sentences extracted from the source, but a compressed paraphrasing of the main contents of the document, potentially using vocabulary unseen in the source document.

Abstractive summarization has shown the most promise towards addressing issues in extracting important information from the text documents, but Abstractive generation may produce sentences not seen in the original Copyright c 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). [4]

Trying to understand the content of the text and then providing a summary based on that, which may or may not have the same sentences as present in the original test.

Abstractive summarization tries to create its own sentences and is definitely a step towards more human-like summaries.

The best method of abstractive is NLTK.

1.2.2 Extractive summarization

Extracted summaries, as their name suggests, contain formulations and phrases extracted from the original passage of text. There are various extraction techniques, but the most common and simplest is simply extracting sentences in the correct order from a passage of text. In general, the first few sentences play a greater role in the

meaning of the passage and are often extracted but sentences later in the passage may also be extracted.

Note that in the research community, text passages are often rewritten into summaries by humans, so gold-annotated summaries are usually abstract in nature. Abstractive Summary's supervised learning process is therefore straightforward because the goals are clearly defined. However, the sentence-by-sentence extract summarization may perform additional steps. [5]

Picking out sentences from the text that can best represent its summary.

It's more about learning to understand the importance of each sentence and their relations with each other rather than trying to understand the content of the text.

The best method of extractive is TextRank.

1.3 Comparison between NLTK and TextRank

1.3.1 Advantage to NLTK

1. Wide range of tools and algorithms: NLTK comes with a number of tools and algorithms for tasks like tokenization, part-of-speech tagging, named entity recognition, and more that can help you save time and energy when working on natural language processing projects.
2. Easy to use: With NLTK, natural language processing in Python is simple for beginners to get started because of its user-friendly API and extensive documentation.

1.3.2 Disadvantage to NLTK

1. Performance: NLTK is implemented in Python, which is not a particularly fast programming language. This can make NLTK somewhat slower than other toolkits that are implemented in faster languages such as C or C++.

2. Limited support for some languages: NLTK was developed primarily for working with English text, and while it does include support for other languages, this support may not be as extensive as for English.
3. Some features may be outdated: NLTK is an actively developed toolkit, but some of the algorithms and features included in the toolkit may be based on older research and may not be as up-to-date as more recent approaches.
4. May require more development time: Because NLTK is a general-purpose toolkit, you may need to spend more time developing and fine-tuning your natural language processing system compared to using a more specialized toolkit that is tailored to a specific task.

1.3.3 Advantage to TextRank

1. It is simple to implement: TextRank is relatively easy to implement, as it does not require a lot of preprocessing or training data.
2. It is unsupervised: TextRank is an unsupervised learning algorithm, which means that it does not require any labeled data to work. This makes it useful for tasks where labeled data is not available.
3. It can handle large datasets: TextRank is efficient and can handle large datasets, making it suitable for use on long documents or large collections of documents.
4. It can handle different languages: TextRank has been successfully applied to a variety of languages, including English, Arabic, Spanish, and Chinese.
5. It can extract keyphrases: TextRank is good at extracting keyphrases, which are short phrases that capture the main points of a document. This can be useful for tasks such as keyword extraction or document summarization.

1.3.4 Disadvantage to TextRank

1. It relies on the quality of the input data: TextRank works by building a graph of the relationships between words or phrases in the text. If the input data is of poor quality, such as being poorly written or having a lot of noise.
2. It can be computationally expensive: Building the graph of relationships between words or phrases can be computationally intensive, especially for large datasets.

1.4 TextRank-Based Summarization

In the proposed method, the probability of moving from set A to set B is equal to the pairwise similarity of the sets. The modified TextRank proposed in this article uses the intuition behind the PageRank algorithm to rank sentences. This allows you to select the most important sentences from your input text document. PageRank links important websites to other important websites. Similarly, our approach assumes that key phrases are (as well) linked to other key phrases in the input document. In this model, we assigned an index to each sentence according to the input sequence of sentences in the document. The index value is heavily used to order the summarized sentences correctly to make the summary meaningful. In the next step, each sentence is tokenized into a series of words.

Represent each sentence in a given text document as a word vector to define the similarity between sentences. [6]

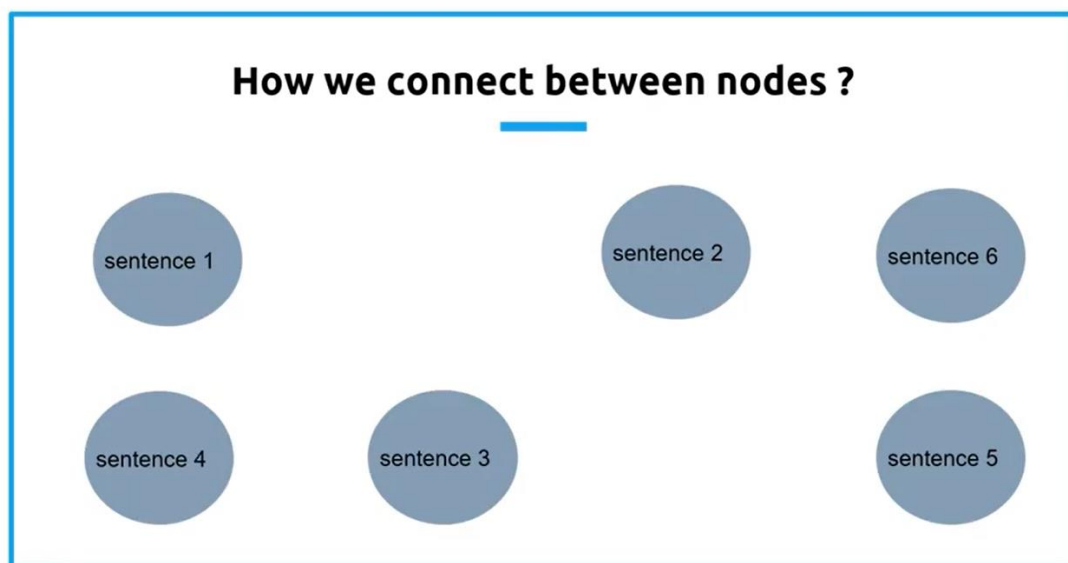
1.5 Similarity function

The weights between nodes based on this similarity function. The number of common words between two sentences divided by the length of each to avoid promoting long sentences

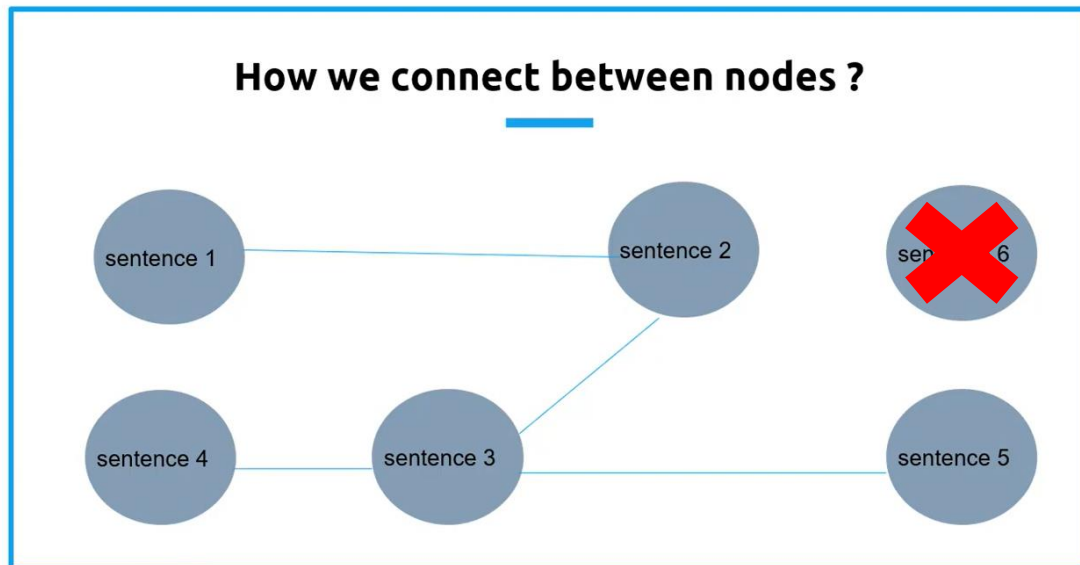
$$Sim(S_i, S_j) = \frac{|\{w_k | w_k \in S_i \& w_k \in S_j\}|}{\log(|S_i|) + \log(|S_j|)}$$

1.6 TextRank model

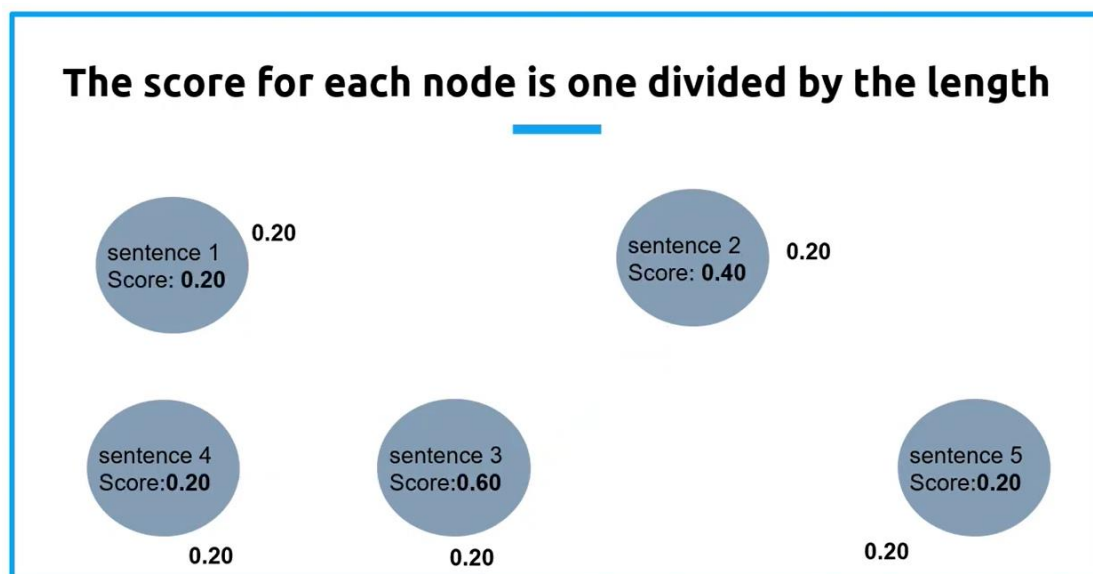
At first, we separated the paragraph into sentences and put each one of them in a shape of a small circle.



Then, we connect between the similar sentence that is related to each other and removes the sentences that are not connected.



Next, we get the value of each sentence according to the number of edges outside it, each edge taking a value (number of sentences/1).



The sentences are displayed according to the number of sentences desired.

1.7 Conclusion

- TextRank models any document as an undirect graph using sentences as nodes.
- A function to compute the similarity of sentences is needed to build edges in between.
- This function is used to weigh the graph edges, the higher the similarity between sentences the more important the edge between them.
- We can say that we are more likely to go from one sentence to another if they are very similar.

2 Literature Review to course booking

E-learning has gotten better and better over time, and has brought about major changes in how we learn and who can achieve that learning. Advances in information technology (IT) have led to improvements in many areas, such as education. Education is growing rapidly, spurring the adoption of e-learning, which is a direct result of the integration of education and technology and is recognized as a powerful medium of learning (Al-Fraihat, Joy, & Sinclair, 2017 Year). E-Learning is widely accepted in all educational sectors and academic institutions. In recent years, the application of e-learning for teaching and learning has increased significantly, especially after lockdowns. Professionals who want to be educated usually find it difficult to follow the courses they are enrolled in due to time zones and schedule differences. eLearning has solved this problem because everyone has the knowledge and teaches them what they want, regardless of time difference or location. Self-paced learning allows students to learn at their own pace. However, due to lack of time or motivation, I am unable to complete the entire course.

Hot Course is an e-learning platform that provides educational content in various categories and levels. We will provide a new form of self-paced education that will change the way education is for working adults and those who work short hours. With automatic summaries, Hot-Course cuts course length by more than half, giving students only the essential information they need, eliminating time-consuming fill-out tasks that don't add valuable new knowledge.

| | Hot-Course | Udemy | Coursera |
|--------------------------------------|-------------------|---------------|---|
| Who can become an instructor? | anyone | anyone | Only experts and professional instructors |
| Community forum support | supported | Not supported | supported |
| Course summarization | supported | Not supported | Not supported |

2.1 Advantages:

- Courses are summarized
- courses are offered in both English and Arabic
- flexible learning schedule
- Courses can be bought individually
- Easy to enroll as an educational center
- Availability of different levels of content
- Students get instant, lifetime access to the registered course
- Past courses' reviews and rating
- Posts on courses with latest updates and discussion forums

2.2 Limitations:

- There is a possibility that the course you searched for is not available
- There is a possibility that the cycle that was searched for will be repeated when more than one educational center

3 Market and Business Analysis

1 SWOT

| S | W | O | T |
|--|---|--|---|
| <ul style="list-style-type: none">❖ We have Facilities for payment (cash, credit card).❖ Simple design.❖ Great competitive advantage in the Middle East Market.❖ Serving many student in short time❖ Flexibility in learning | <ul style="list-style-type: none">❖ other competitors that share the same market with me.❖ I' m a new business in the market so no one know me❖ Absence of trainer❖ Technical difficulties | <ul style="list-style-type: none">❖ Middle East Market is in big need For this system❖ Introduce an services more than competitors like (auto smmirization) | <ul style="list-style-type: none">❖ Economic problems.❖ People don' t know how to use this system.❖ Increase student expectation❖ Internet connection problems |

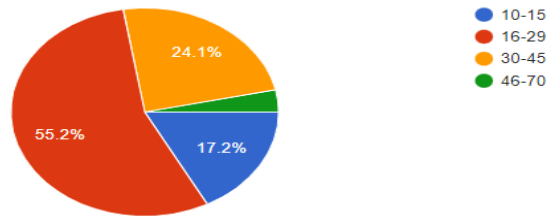
2 4Ps

| Product | Price | Promotion | Place |
|---|---|---|--|
| <ul style="list-style-type: none">❖ Courses❖ Auto-summarized books and texts❖ UX/UI | <ul style="list-style-type: none">❖ Discounts❖ Depends on every author❖ Offers❖ Bundling | <ul style="list-style-type: none">❖ Direct marketing❖ Push strategy❖ Social media advertising | <ul style="list-style-type: none">❖ online |

3 Survey

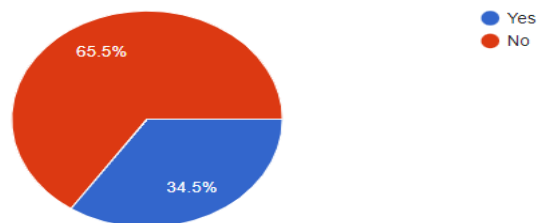
How old are you ?

29 responses



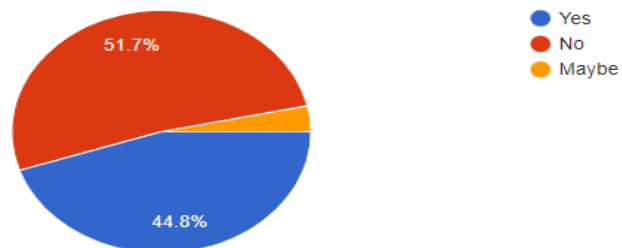
Have you heard about this idea before ?

29 responses



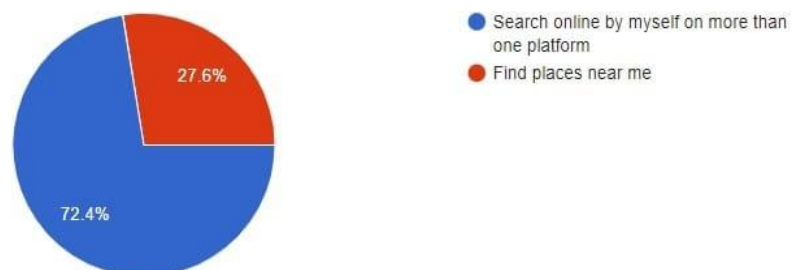
Do you like to use the feature of summarizing texts to read books?

29 responses



What are your preferred research methods to take a course?

29 responses



4 SRS document

1 Introduction

1.1 Purpose

The program allows easy access to distinguished educational companies and the best courses at the best prices and discounts. Also, the user can summarize the large articles to make them smaller and easier to read to save his time and give him the same informational value. You can also see the latest publications of your favorite companies and comment on them.

1.2 Intended Audience and Reading Suggestions

We target students who need courses and companies that need these students.

Readers who need to summarize large books and articles.

1.3 Intended Use

For easy access to good educational courses without cost or effort, and I can summarize large books and articles.

1.4 Project Scope

The purpose of the program is to help students get access to the best educational companies, and we have thousands of educational companies in all fields and all over the world.

We also have artificial intelligence to summarize large articles and books.

1.5 Definitions

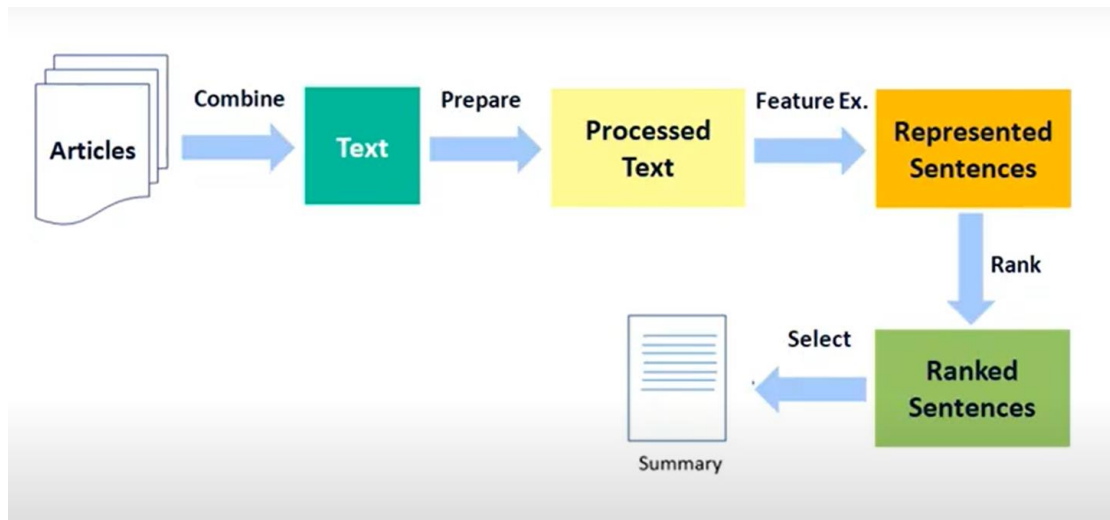
- Can auto-summarize large books and articles.
- Educational companies can join our world and upload their own courses and publications, and students can access excellent educational companies quickly without effort or cost, and they can see the opinions of others about educational companies.
- Users who write negative comments that can be resolved by educational companies can make a report on the comment, and the necessary measures will be taken.

2 System Description

2.1 Methodology

In the proposed method, the probability of moving from set A to set B is equal to the pairwise similarity of the sets. The modified TextRank proposed in this article uses the intuition behind the PageRank algorithm to rank sentences. This allows you to select the most important sentences from your input text document. PageRank links important websites to other important websites. Similarly, our approach assumes that key phrases are (as well) linked to other key phrases in the input document. In this model, we assigned an index to each sentence according to the input sequence of sentences in the document. The index value is heavily used to order the summarized sentences correctly to make the summary meaningful. In the next step, each sentence is tokenized into a series of words.

Represent each sentence in a given text document as a word vector to define the similarity between sentences.

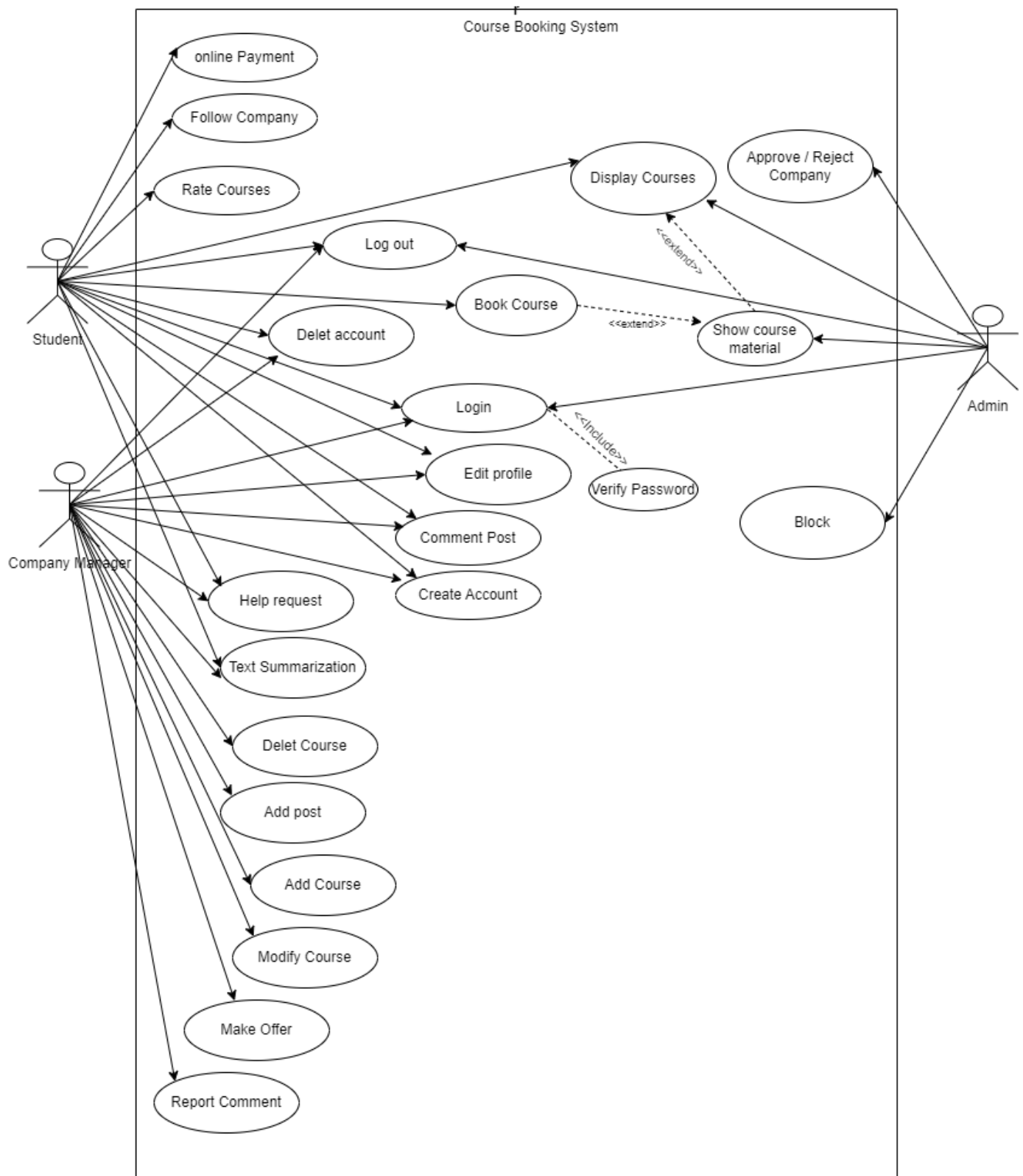


At first, combine the articles into a single text, then do the preparation and processing of text to get rid of sentences that do not contain meaning, then in feature Ex. Change the string to numbers with the method TextRank and represent it.

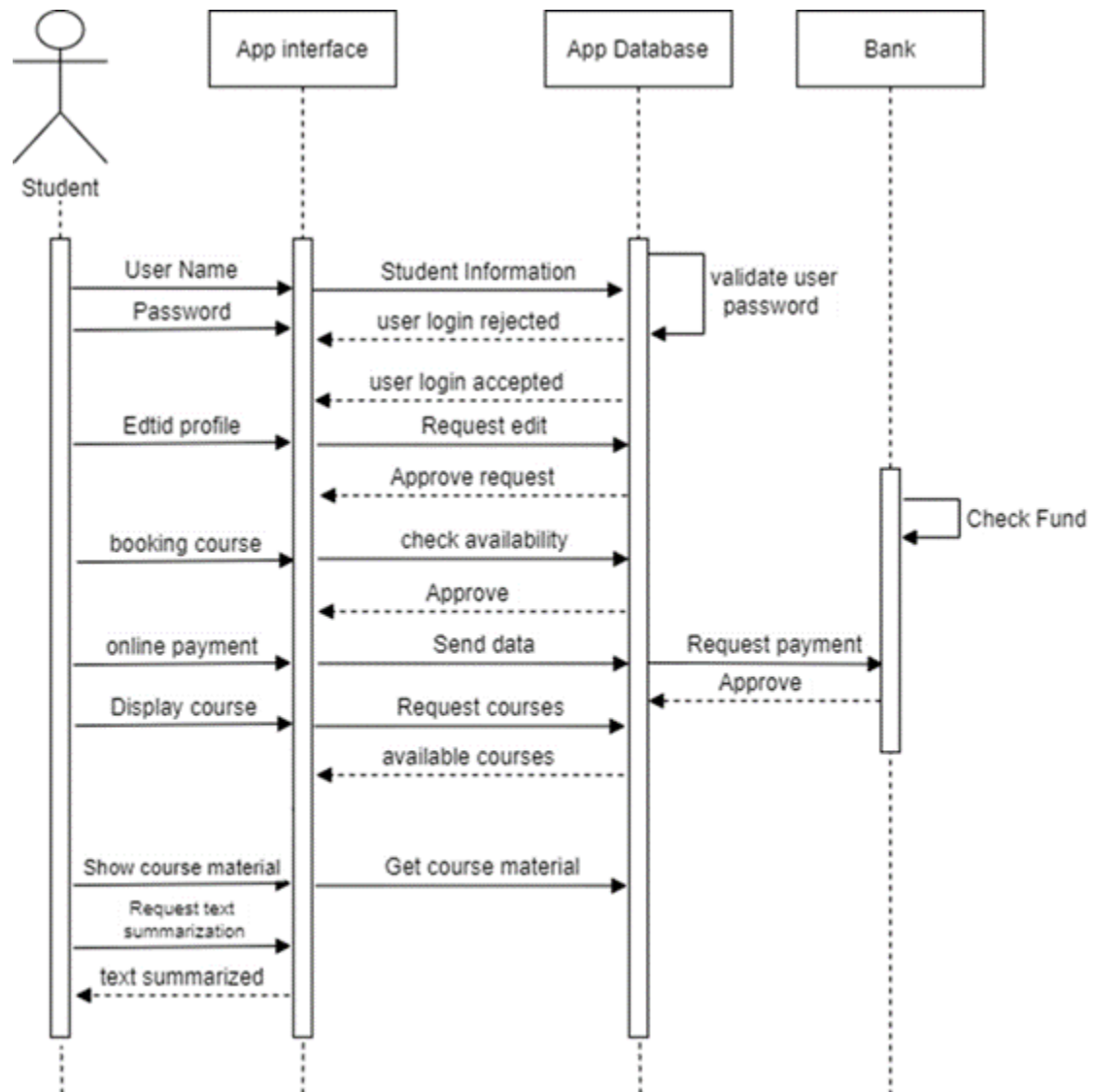
All sentences are represented in a specific way, then chose the important sentence and represent it.

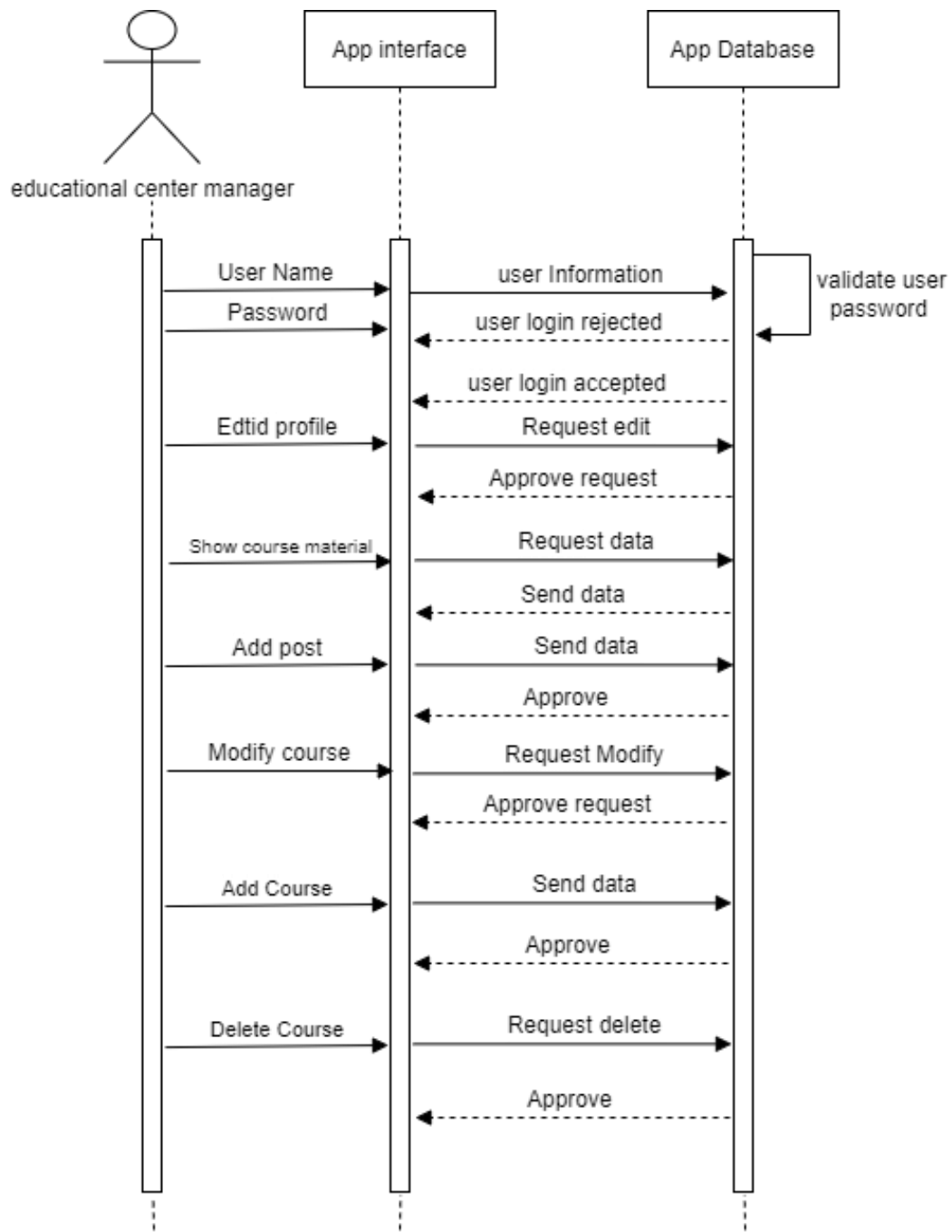
2.2 SYSTEM ANALYSIS & DESIGN

2.2.1 use case diagram

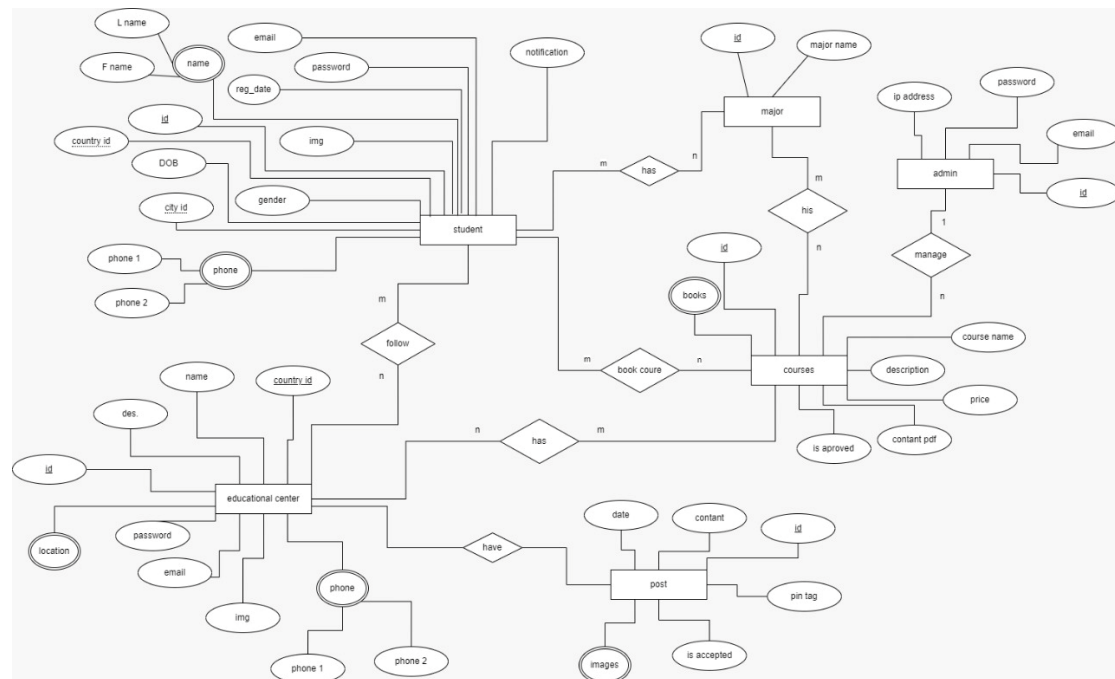


2.2.2 Sequence Diagram





2.2.3 ERD Diagram



2.3 Assumptions and Dependencies

2.3.1 Assumptions:

- there is a company's have access to login to the system and make an account to add courses and posts is already added by them
- the (id) for the company and student are auto generated by the system
- the help and support have answers for the popular questions by the people

2.3.2 Dependencies:

- To login the user needs to create an account
- To view courses available the user has to log in.
- To purchase a course a user has to choose the desired course and make a payment.

2.4 Intended Technologies

Platform: flutter, Django

Tools: APIs, library summarizer from module summa

Languages: dart, python

3 System Requirements

3.1 Functional Requirements

- Users must be able to create their own accounts.
- Users must be able to book courses.
- Courses should be archived indefinitely so users can go back to previous chats.
- Users should be able to download files from courses or auto-summaries.

3.2 External Interface Requirements

3.2.1 User Interfaces

We have three main users:

3.2.1.1 Student

- **login page:** enter your email and password to log in, also you can click (forgot password) if you forgot your password.
- **create an account:** enter your name, gender, city, country, email, password, phone number, and date of birth, and upload a profile picture, then click next to go to the major's page.
- **major's page:** choose your favorite fields, then click (next) to go to the home page.
- **home page:** there are all posts from educational centers, you can react to posts and add comments to any post.
- **student courses:** in our system, the student can type in the search bar for a specific category, and also can view all courses that he is interested in.
- **book page:** in this page, the student will be able to view the name and picture of the educational center, all materials and the cost of the course,

and the rate of the course, then click (book) to choose the payment method.

- **help page:** in this page, we answer the most popular question, if the student has any other problem or question then he has to click the (contact us) button to fill the form and click (submit).
- **student profile:** in this page, the student can represent his profile picture, all his data, and all educational centers he follows, he can click to (edit) if he wants to edit any personal information.

3.2.1.2 educational center

- **login page:** enter your email and password to log in, also you can click (forgot password) if you forget your password.
- **create an account:** enter the educational center name, description, picture, all locations, email, password, phone numbers, and which payment method they accept, then click next to go to add a course.
- **add a course:** add name, description, cost, and material to the course and make an offer if they want
- **educational center courses:** in this page, show all the courses that the educational center published
- **help page:** in this page, we answer the most popular question, if the educational center has any other problem or question then they have to click the (contact us) button to fill the form and click (submit).
- **educational center profile:** in this page, the educational center will be able to represent the profile picture, and all the data, with the availability to (edit) if they want to edit any information and view the overall rate.
- **home page:** in this page, the educational center will be able to review all published posts and at any time they can edit, pin, or remove any post, and reply or report to any comments.

3.2.1.3 admin

- **statistic page:** in this page, he can show all statistics.

- **setting:** in this page, he can answer to all reports, and answer to help requests.
- **accept page:** in this page, he approves or rejects the new educational centers and the new courses.
- **home page:** in this page, he shows all posts and courses from the educational centers.

3.2.2 Hardware Interfaces

- Android
- iOS
- Windows

3.2.3 Software Interfaces

| Software used | Description |
|------------------|--|
| Operating system | We have chosen Windows, Android, and iOS operating systems for their best support and user-friendliness. |
| Database | To save the course record and student data, we have chosen the SQL + database. |
| flutter | To implement the project we have chosen flutter for its more interactive support. |

3.2.4 Communications Interfaces

This project supports all types of web browsers and android. We are using online payment.

3.3 System features

3.3.1 Auto summarization

3.3.1.1 Description and Priority

The auto-summarization feature provides a summary of the text that you provide, regardless of its size, and this has a very high priority because of its importance in saving readers' time.

3.3.1.2 Stimulus/Response Sequences

- He can see the old summaries he has summarized
- You can clear an old summary from history

3.3.1.3 Functional Requirements

- Users should be able to summarize large books and articles.
- Users should be able to download files after automatic summaries.

3.3.1.4 Other system features include:

Download: Download the book or text after summarizing

New summarize: to do new summarize as free cost

3.3.2 signup

3.3.2.1 Description and Priority

Providing a feature for the user to create a new account to save his data and refer to it at any time.

3.3.2.2 Stimulus/Response Sequences

- Can create an account with google or Facebook
- You can delete the account if you want

3.3.2.3 Functional Requirements

- Users should be able to create any number of private accounts.
- Users should be able to log in from anywhere on any type of device.

3.3.2.4 Other system features include:

major: choose your favorite fields

locations: choose all locations

3.4 Nonfunctional requirements

| | |
|--------------------------|---|
| Performance | <ul style="list-style-type: none">✓ Ensure that System features are highly responsive in a very short time so that the system works under high efficiency with an average time (3 seconds)✓ the system and database can handle pdf files✓ database able to meet the requirements of continuous data growth continuously to ensure fast response of data presentation |
| Safety and secure | <ul style="list-style-type: none">✓ If there is extensive damage to a wide portion of the database due to catastrophic failure, such as a disk crash, the recovery method restores a past copy of the database that was backed up to archival storage✓ user cannot use previously used password as new one✓ the system never save the payment methods information such as credit or debit cards✓ the system will always backup itself for easy restoration and data protection |
| Quality | <ul style="list-style-type: none">✓ the system should auto-summarize the input in an understandable way that includes all important points that were originally mentioned in the original text. |

4 References

- [1] A. A. L. K. N. S. C. D. M. H. M. K. M. F. MRIDHA, "A Survey of Automatic TextSummarization: Progress, Process andChallenges," *researchgate*, p. 29, 2021.
- [2] J. D. A. K. Y. Divakar Yadav, Automatic Text Summarization Methods: A Comprehensive Review, 2020.
- [3] P. Joshi, "An Introduction to Text Summarization using the TextRank Algorithm (with Python implementation)," 2022. [Online]. Available: <https://www.analyticsvidhya.com/blog/2018/11/introduction-text-summarization-textrank-python/>.
- [4] T. M. T. a. T. H. Ekaterina Zolotareva, Abstractive Text Summarization using Transfer Learning.
- [5] T. P. Alvin, "Introduction to Text Summarization with ROUGE Scores," 2022. [Online]. Available: <https://towardsdatascience.com/introduction-to-text-summarization-with-rouge-scores-84140c64b471>.
- [6] A. K. D. M. D. A. K. D. & A. S. Chirantana Mallick, "Graph-Based Text Summarization Using Modified TextRank," 2018. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-13-0514-6_14#Sec2.