ESCORT – UNIVERSITY STUDENTS GUIDANCE SYSTEM BASED ON NATURAL LANGUAGE PROCESSING

2022-179

Project Proposal Report

Abarna U

B.Sc. (Hons) Degree in Information Technology

(Specialization in Software Engineering)

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

January 2022

ESCORT – UNIVERSITY STUDENTS GUIDANCE SYSTEM BASED ON NATURAL LANGUAGE PROCESSING

2022-179

Project Proposal Report

B.Sc. (Hons) Degree in Information Technology

(Specialization in Software Engineering)

Department of Information Technology

Sri Lanka Institute of Information Technology
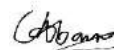
Sri Lanka

January 2022

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.

2022.02.09
…………………………
(Date)

…………………………………..
(Abarna.U)

The above candidate is carrying out research for the undergraduate Dissertation under my supervision

…………………………
(Date)

………………………………
(Signature of the supervisor)

Abstract

Being in a good mental state is very important to university students because university life impacts students' career and happiness. Many university students are having issues related to psychology, nowadays. There are a lot of reasons for the psychological issues. Several students work as part time employees and interns. Workload in studies and workplace, family problems, misunderstandings, physical diseases, and fear are some examples for the causes of psychological issues such as stress, hostility, depression, hopelessness, and job control. Covid 19 also leads to many psychological issues these days. Especially it can be observed that the number of psychological issues and suicides increased during the lockdown period. Further, many university students have difficulties in communication because of lack of language skills. Several students are not feeling comfortable to express the issues in person to psychiatrists, councilors, family, and friends. Also, some prefer their mother tongs to explain the issues. Therefore, there is a need for a software system to provide solutions for psychological issues in preferred language of university students. The proposed system will recommend some activities and exercises to university students as a solution for their issues, using Natural Language Processing.

Keywords: psychology, stress, Covid 19, hostility, Natural Language Processing

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Background

Natural Language Processing (NLP) which is a component of Artificial Intelligence, is known as the ability of a program to know human language including spoken and written [1]. A review of literature from the past thirty years establishes psychological distress as both a long and current issue of university students worldwide. Poorer academic outcomes and problematic health behaviors are joined to students' distress, and these wider implications conjointly highlight the requirement for acceptable policies and services to support students throughout what is clearly a difficult time. Additional review known varied socio demographic, situational and educational factors as potential bases of students' distress. Undoubtedly, the strain of the university way are inherently stressful. However experiencing these as distressing is not inevitable [2]. During the COVID-19 natural event, studies deciding the problems and psychological distresses practiced by university students had been conducted. University students had additional negative feelings, psychological symptoms, and economic/social issues than the general population due to educational delays, economic effects of the pandemic, and its effects on their standard of living [3]. The scholars also practiced serious mental issues like depression, anxiety, alexithymia, stress, and posttraumatic stress disorder [4][5]. Due to all these issues practiced throughout the natural event, Grubicet an emphasized that educational establishments ought to particularly arrange interventions to develop header skills for college students by taking some precautions and characteristic the deficiencies in instructional make reduce negative educational and psychosocial outcomes. The main reason for conducting this research is to provide better solution for psychological issues of university students without communication issues.

1.2. Literature review

The expertise of elevated levels of psychological issues is a biggest problem for several university students. However, it is any combined by its association with alternative problematic health behaviors and underperformance, and analysis demonstrating the links between these problems and psychological distress is given [2][14]. In addition to students' retrospective self-ratings of impairment, the impacts of psychological issues in universities outcomes have conjointly been incontestable lengthwise [8]. For instance, students' levels of psychological issues at mid-course have foreseen decreases in test performance from first to second year, even when dominant for previous performance [9].

Many of the university students prefers a computerized system to share the issues in psychology they have. Some students shares the issues with family, relatives and friends. Very less students prefer to consult psychiatrist and counsellors and some do not share the issues which they have with others.
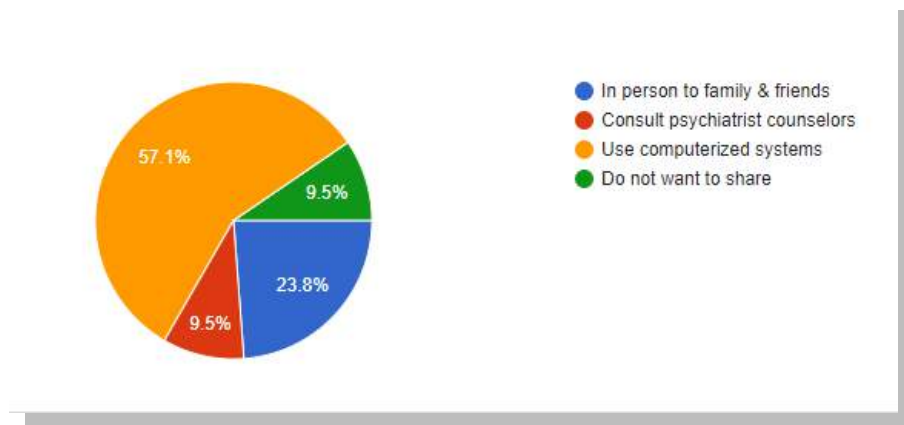

Figure 1.2.1. Summary of the responses for methods of sharing psychological issues

Most of the university students thinks a system for recommending solution to psychological issues, will be very useful to them because students have no time to consult psychiatrist, go to counselling for treatments because students are far away from home and continues university academic activities.
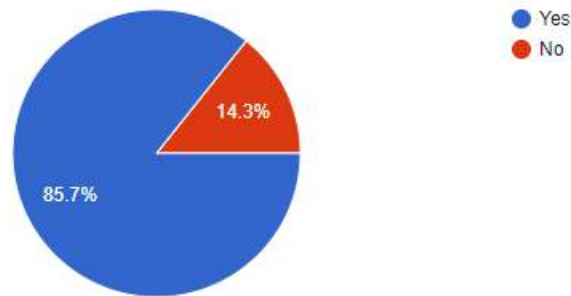
Figure 1.2.2. Summary of the responses for methods of sharing psychological issues

Lack of communication is one of the important issues to university students. Some students are not fluent in English and some are unable to describe the issues clearly in English. All of the students prefer their mother tong to describe the issues in psychology which they have. None of the students suggest any other languages instead of their mother tongs to describe psychological issues.

Figure 1.2.3. Summary of responses for using mother tongs to share psychological issues

If you answer is No, Which language do you prefer?

0 responses

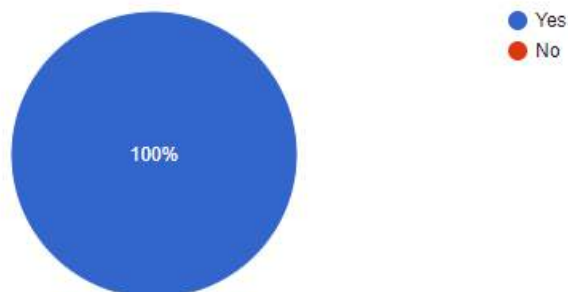No responses yet for this question.

Figure 1.2.4. Summary of responses for other languages instead of mother tongs to share psychological issues

Almost all the university students are unable to frequently visit psychiatrists and counsellors because of residence, academic and other reasons. The treatments which will be given by the consultants are long process and there are less chance to manage within the heavy schedule of university students. Simple activities such as cycling, painting and swimming, simple exercise including breathing exercises are mostly preferred by the university students as these solutions are manageable within the rush.
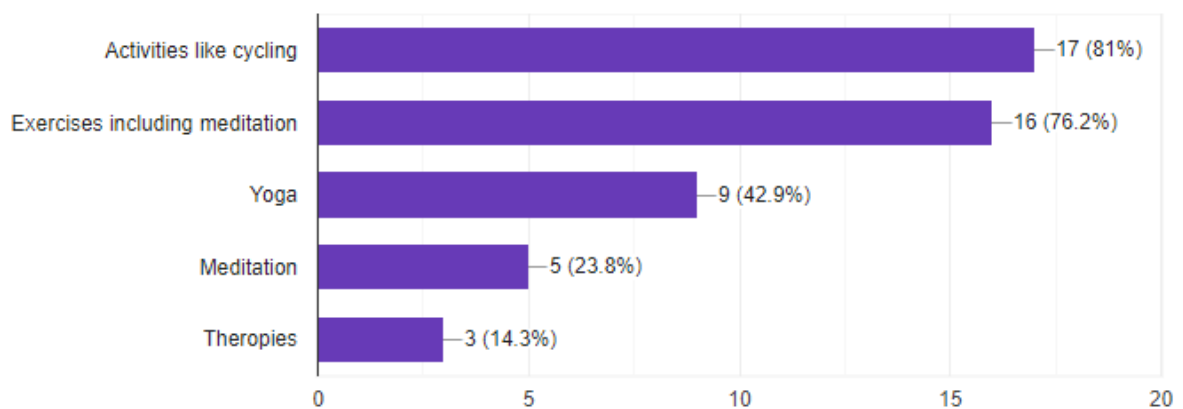

Figure 1.2.5. Summary of responses for solutions to psychological issues

1.3. Research Gap and Research Problem

1.3.1.   Research Gap

Psychology related problems are common nowadays. Thus, there is a demand for solutions for psychological issues. Even though there are some other applications available, people face some difficulties because everyone is not familiar with English to express what exactly they feel. Also, those available applications are not specific for university students as designed for common users [10].

Further, unlike directly providing the exact problem and finding solutions, this application gives chance to find out the solution simply by describing how they feel if the university students are unable to diagnose the psychological issues which they have [11]. Also, it will recommend solution as simple activities and exercise which can be manageable with their studies and work. This proposed system will be an easily accessible and user-friendly web application for university students.

Features	Champions mind	Sintelly	Positvt	Escort
Designed for university students	✖	✖	✖	✔
Language selection	✖	✖	✖	✔
Solution for known issues	✖	✔	✔	✔
Solution for unknown issues	✖	✖	✖	✔
Simple solution	✔	✖	✔	✔

Comparison Table

1.3.2. Research Problem

The entrance to the university marks an amount of transition for tykes. Through this transition, students face new challenges, like creating freelance selections regarding their lives and studies, adjusting to the tutorial demands of an ill-structured learning environment, and interacting with a various vary of latest individuals. Additionally, several students should, typically for the primary time, leave their homes and distance themselves from their support networks [12]. These challenges will have an effect on the psychological state and well-being of upper education students. Also, the possibilities of experiencing common psychological issues, like depression, anxiety, and stress, increase throughout adolescence and reach a peak in early adulthood around age twenty five [14] that makes university students a very vulnerable population.

2. Objectives

2.1. Main Objective

Main Objective is to provide a better solution for the issues in psychology, to the university students in preferred language.

2.2. Sub Objectives

- ❖ Implement a Web Application for all university students.
- ❖ Provide a useful and reliable system to university students in their preferred language.
- ❖ Make the system more efficient.
- ❖ Achieve a better knowledge in Natural Language Processing.

3. Methodology

NLP allows computers to know natural language as humans do. Whether or not the language is spoken or written, Natural Language Processing uses artificial intelligence to require real world input, process it, and be of it in an exceedingly manner a computer will perceive [1]. Even as humans have completely different sensors like ears to listen to and eyes to envision computers have programs to browse and microphones to gather audio. And even as humans have a brain to method that input, computers have a program to method their several inputs. At some purpose in process, the input is regenerate to code that the computer will perceive [13].

There are 2 main phases in natural language processing:
1. Data preprocessing
2. Algorithm development

1. Data preprocessing
Data preprocessing involves getting ready and "cleaning" text information for machines to be ready to analyze it. Preprocessing puts information in executable type associated highlights options within the text that an algorithmic program will work with [8][9]. There are many ways that this could be done, including:
- ❖ Tokenization – Breakdown text into smaller units.
- ❖ Stop word removal – Remove common words from text and make it unique to receive more information.
- ❖ Lemmatization and stemming – Reduce words to root forms to process.
- ❖ Part-of-speech tagging - Mark words based on the part of speech including nouns, verbs and adjectives.

2. Algorithm development

Once the data has been preprocessed, an algorithmic program is developed to process it [1]. There are many alternative natural language process algorithms, however 2 main varieties are usually used:

- ❖ Rules-based system - this technique uses fastidiously designed linguistic rules. This approach was used timely within the development of natural language process, and remains used.

- ❖ Machine learning-based system - Machine learning algorithms use statistical methods. They learn to perform tasks supported training data they are fed, and change their methods as a lot of information is processed. Using a combination of machine learning, deep learning and neural networks, natural language process algorithms hone their own rules through perennial process and learning [12].
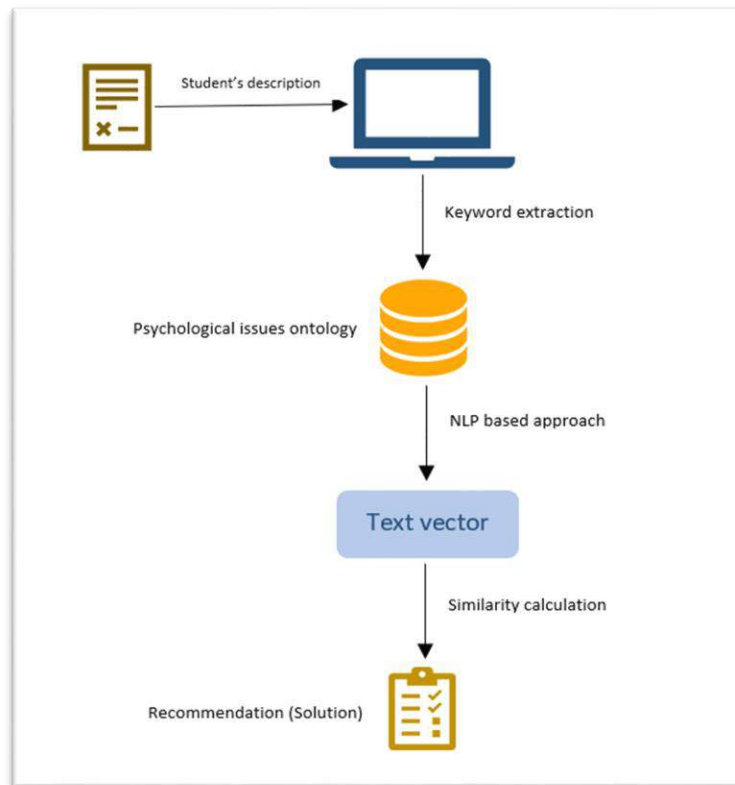
Figure 3.1

Overall Architecture Diagram for Psychology Management

Students' Description:

After selecting the language, university students are instructed to describe the psychological issues which they have. Or else students can describe the feelings as a paragraph in the given area for description.

Keyword extraction:

Keyword extraction is used to identify the important words.

1. Load dataset
2. Identify text which will going to analyze
3. List all the stop words
4. Preprocess for getting cleaned and normalized text
5. Extract most frequent keyword
6. List the top keywords

NLP based approach:

1. Text classification – to predict tags and categories
2. Word sequence – to predict next and previous words in language modelling
3. Text meaning – to represent word vectors and embedding
4. Sequence to sequence – to find and use hidden representation of text

Similarity calculation:

Cosine similarity is the cosine of the angle between two vectors, which gives the angular distance between the vectors.

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}$$

Figure 3.2 Cosine formula

Recommendation (Solution):

System will recommend simple activates and exercises as a solution for the psychological issues which are manageable to university students.

4. Project Requirements

4.1. Hardware requirements

Laptop or desktop

Smart phone

Windows 10

4.2. Software Requirements

JavaScript

Python

Sklearn

NLTK and other Python libraries

TensorFlow

Anaconda IDE

Jupyterlab

4.3. Functional Requirements

Browse 'Escort'

Log into the system

Click 'Psychology' option in homepage

Select language

Describe the issue

Click 'Submit' button

Receive solution

4.4. Non–Functional Requirements

Accuracy

Speed

Portability

Security

Compatibility

Usability

Privacy

Reliability

Scalability

5. Description of Personal and Facilities

Component:

- ❖ Solution Recommendation System for Psychological issues using Natural Language Processing.

Tasks:

- ❖ Requirement and information gathering
- ❖ Requirement analysis
- ❖ System Designing
- ❖ Function Implementation
  - ✓ Describe issues or feelings as a paragraph.
  - ✓ Keyword extraction
  - ✓ Text Preprocessing
  - ✓ Postagging
  - ✓ Solution as result

    Activities and exercises
- ❖ Testing
  - ✓ Unit testing
  - ✓ Integration testing
  - ✓ System testing

6. Budget and Budget Justification

Preparation for report and printing – 1500.00

Internet usage for research – 5000.00

Domain name registration (annual) – 2500.00

Hosting (annual) - 15000

Other costs (Data Collection, Travelling Expenses) - 5000

Total – 29000

7. References

[1] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, State-of-the-Art Natural Language Processing, 2020.

[2] Jessica Sharp, Stephen Theiler, A Review of Psychological Distress Among University Students: Pervasiveness, Implications and Potential Points of Intervention, 2018.

[3] Duong V, Pham P, Yang T, Wang Y, Luo J, The ivory tower lost: how college students respond differently than the general public to the COVID-19 pandemic, 2020.

[4] Garcia R, Psychological effects of the COVID-19 outbreak and lockdown among students and workers of a Spanish university, Psychiatry Res, 2020.

[5] Emel Bahadir-Yilmaz, Arzu Yüksel, Emel Bahadir-Yilmaz, PhD, RN, Department of Psychiatric Nursing, Faculty of Health Sciences, Giresun University, Evaluation of the psychological problems experienced by university students during the COVID-19 outbreak and suggestions, 2021.

[6] Dada EG, Bassi JS, Chiroma H, Abdulhamid SM, Adetunmbi AO, Ajibuwa OE, Machine learning for email spam filtering: review, approaches and open research problems, 2019

[7] Oliva J, Serrano JI, del Castillo MD, Iglesias A, A methodology for the characterization and diagnosis of cognitive impairments-application to specific language impairment, Artif Intell Med 2014.

[8] M. Munikar, S. Shakya and A. Shrestha, Fine-grained Sentiment Classification using BERT, Artificial Intelligence for Transforming Business and Society (AITB), 2019.

[9] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio, Neural machine translation by jointly learning to align and translate, 2014.

[10] Aziliz Le Glaz1 , MD; Yannis Haralambous2 , PhD; Deok-Hee Kim-Dufor1 , PhD; Philippe Lenca2 , PhD; Romain Billot2 , PhD; Taylor C Ryan3 , MS; Jonathan Marsh4 , MA; Jordan DeVylder4 , MD, PhD; Michel Walter1,5 , MD, PhD; Sofian Berrouiguet1,2,5,6 , MD; Christophe Lemey 1,2,5 , MD, Machine Learning and Natural Language Processing in Mental Health: Systematic Review, 2021

[11] Tanya Nijhawan1 Girija Attigeri2* Ananthakrishna, Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Karnataka, India, Depertment of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Karnataka, India, Stress Detection using Natural Language Processing and Machine Learning over social Interactions, 2021

[12] Himanshu Choudhary, Aditya Kumar Pathak, Rajiv Ratn Shah, Ponnurangam Kumaraguru, Neural Machine Translation for English-Tamil, 2018

[13] R A F A E L A. C A L V O1, D A V I D N. M I L N E1, M. S A Z Z A D H U S S A I N1,2 and H E L E N C H R I S T E N S E N3, School of Electrical and Information Engineering, The University of Sydney, Australia, Natural language processing in mental health applications using non-clinical texts, 2016
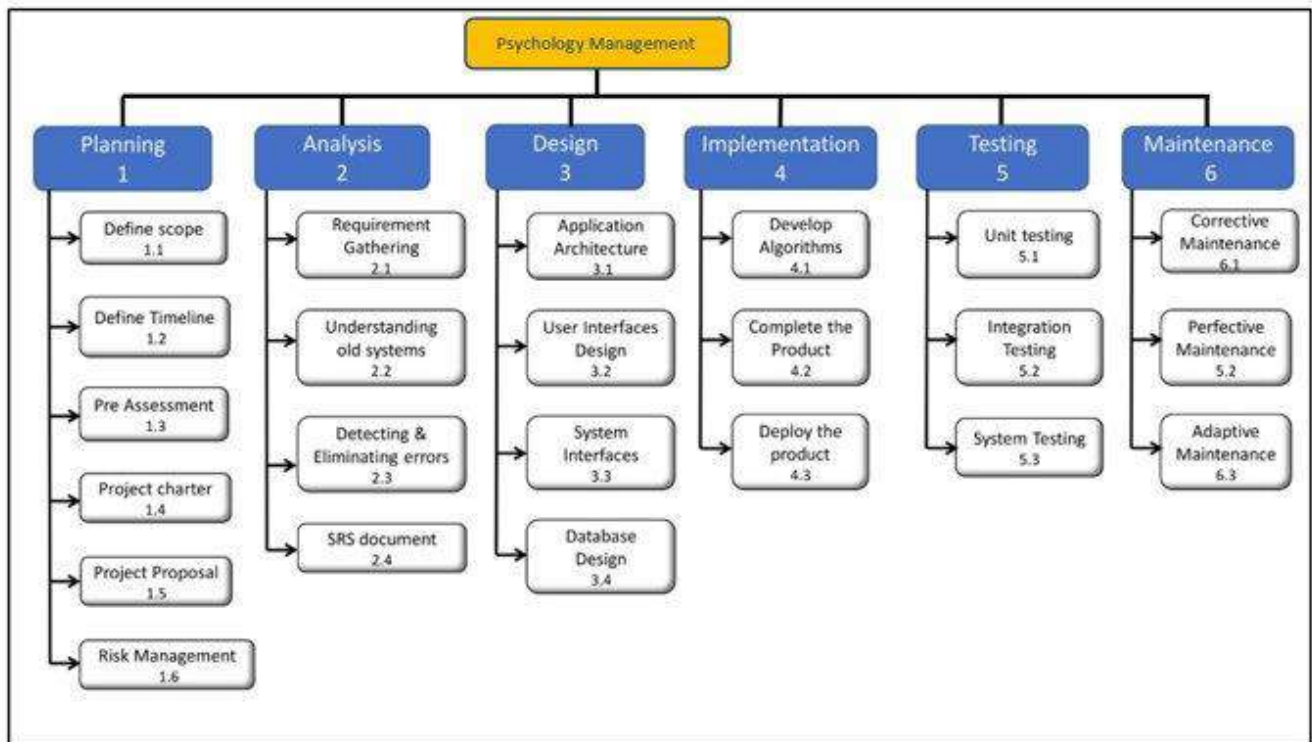
[14] Mitchell, S. L., Oakley, D. R., and Dunkle, J. H, White paper: a multidimensional understanding of effective university and college counseling center organizational structures, 2019.

8. Appendices

Appendix A : System Diagram (Escort)
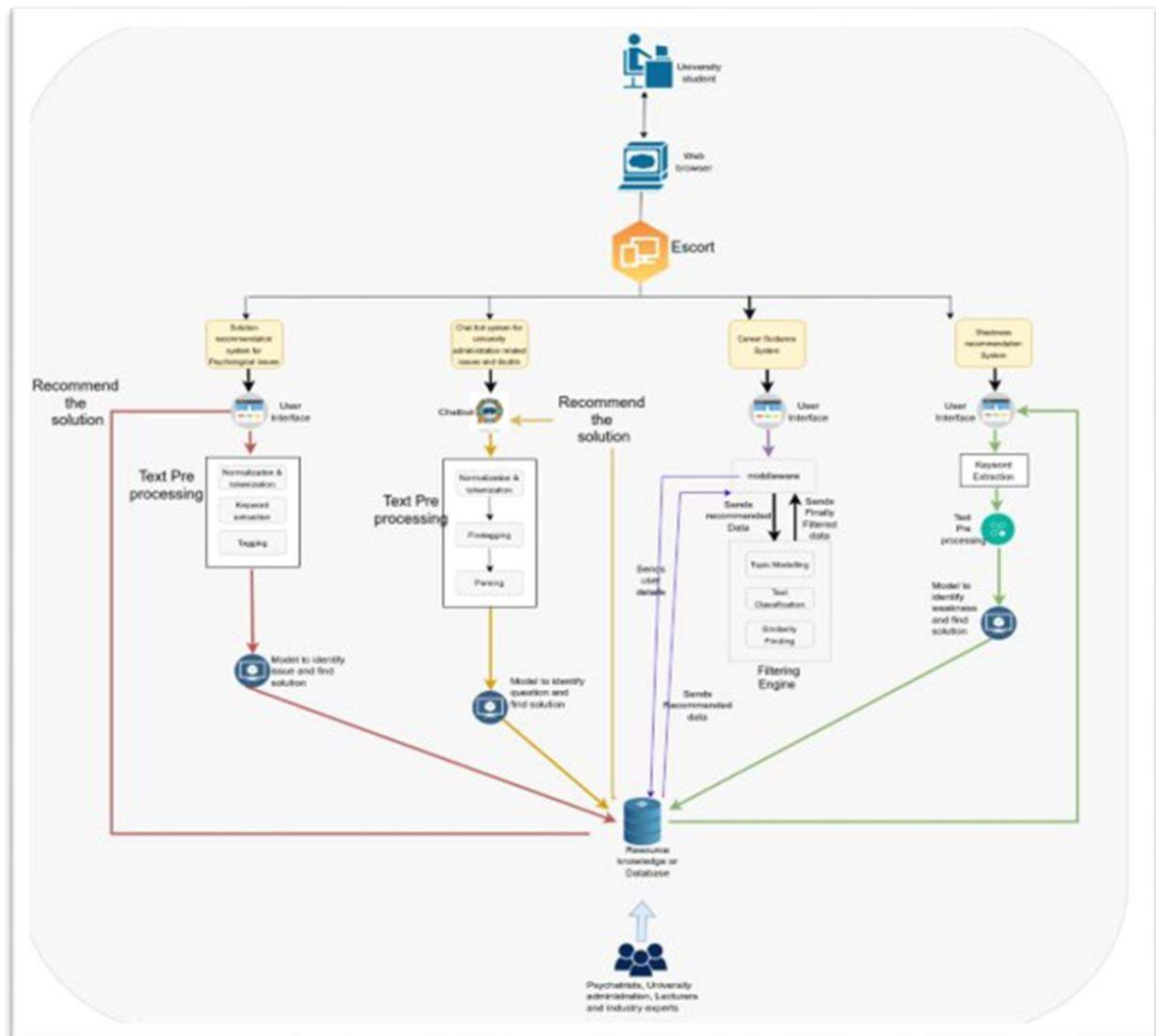
Appendix B : Work Breakdown Structure

Appendix C : Gantt Chart