KULLIYYAH OF INFORMATION AND COMMUNICATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

FINAL YEAR PROJECT REPORT

IDECIDE: DEGREE PROGRAM RECOMMENDATION SYSTEM

SUHIB AHMAD ABDULLA HASAN AHMED
1423183

SUPERVISED BY
Dr. AMIR 'AATIEFF BIN AMIR HUSSIN

AUG 2020
SEMESTER 2 2019/2020

FINAL YEAR PROJECT REPORT

i*Decide*: Degree Program Recommendation System

by

SUHIB AHMAD ABDULLA HASAN AHMED
1423183

SUPERVISED BY
Dr. AMIR 'AATIEFF BIN AMIR HUSSIN

In partial fulfillment of the requirement for the
Bachelor of Computer Science
Department of Computer Science
Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

AUG 2020
Semester 2 2019/2020

ACKNOWLEDGEMENTS

Praise and thanks to Allah first and foremost whose blessing enabled me to accomplish this project. I wish to express my deepest appreciation to my supervisor DR. AMIR 'AATIEFF BIN AMIR HUSSIN and FYP coordinator DR. NORZARIYAH BINTI YAHYA for relentless guidance, helpful suggestion, close supervision and moral encouragement to complete this task. Special thanks to my parents, my lecturers, and classmates I have had. My sincerely thanks to all those whom directly or indirectly helped me to complete this project.

ABSTRACT

This project - titled as iDecide - involves analyzing and designing a prototype to provide an online recommendation system that would help students choosing the most fitting degree program to join. It is built to help them to decide what major program is best for them based on certain data required to be entered and analyzed. The recommendation process is machine-learning based. The system provides a new approach to help students deciding, as the only common approaches now are about reading a lot of related articles and written resources or to consult some other individuals who might have an opinion about the issue.

This system is to be classified as artificial intelligence project, works basing on data science tools and a machine learning algorithm. For prototyping purpose, the project was scoped down to cover IIUM Kulliyyahs (faculties) of applied sciences (ICT, Engineering and Architecture). The software would only consider some of the personal factors in the analysis, excluding other personal factors and all the environmental and external influencers. This application would be as a seed for a universal software, which would cover every major program provided around the world and could be more complex by including all relevant factors before recommending a degree program. This universal

version could overcome humans' ability to advice and recommend on the decision of choosing a degree program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Background

One requirement for people to advance in their communities is to graduate from a high education institute, as it would allow them to start a modern career, impower their selves and climb up the chain of food. Numbers show that those who got a degree and a higher certificate do have the highest employment rate comparing to the those who did not finish a bachelor program (U.S. Department of Commerce, 2018). They also show that graduates are wealthier - by far - than those who are not graduates (Wolla, S. A. & Sullivan, J., 2017). Same goes for nations, as they need to educate their people so they could advance as a whole and compete with other nations around the world (Hanushek, E. A. & Wößmann, L., 2007). As the embranchment of knowledge fields and the deepening of specializations is an undergoing endless process, the matter of choosing the right and most-fitting major will always be highly sensitive, because a wrong study field choice would result a loss in the one's time, money and governments' efforts and facilities. Numbers do show that there is a problem in the current processes of selection a degree major program and several solutions did not overcome the problem.

Recommendation systems are relatively new products generated by the evolution of computing and the applications of artificial intelligence field. They are basically algorithms aimed at suggesting certain elements / items for users to adopt. They managed to contribute to solve a lot of choosing questions in

many fields of daily life like which items to purchase or which nations to immigrate to.

1.2 Problem Statement

It is noticeable that many students finish high school and pick their degree program based on fault approaches. Many statistics show the unsatisfactory of graduates for the major decision they made. According to 'Inside High Ed' survey (Fain, P, 2017), %40 of students regrets their major choice and say that they would choose different one if they got the chance. The end result of this disappointment that %80 of U.S. college students decide a major change at least once (Ramos, Y, 2013), or to drop what was studied in the bachelor and start an irrelevant career like %73 of graduates in the U.S. (Plumer, B, 2013).

There is a big number of causes for that fault pickup process, like the awareness and responsibility level of the students at the time of making the decision, ill-advice, the blindness about the available major choices and many more reasons. But one main concern here is the question of how relevant the study field is to one's fields of interests, previous experiences and already acquired knowledge.

1.3 Project Objectives

- Explore previous related works and research to determine the most fitting approach for the problem.

- Construct a prototype of computer system to help students while picking their college majors.

1.4 Project Significance

As far as concerned, there is not any automated system that provides such a service. It is provided by humans, and mostly with face-to-face meetings. Which do cost a large amount of money, consume long time, could require travelling from one place to another to meet or could also end to a wrong choice. It is also available through exploring a huge flow of information shared online about this problem, including personal opinions, related reflections and sharing of successful persons' stories. But all of that mostly causes confusion, get the decision-making process to be harder, and do not help to answer the question of what major to pick.

Being an online-based system means that it would give an ability to thousand or millions of people around the world to use. It also means that its commitment to go through a structured process before coming to a result is tight, as a programmed software do not have a choice of breaking the process like humans. A machine-based software also means that it has a capability of processing and learning faster than people. So, with time and a large number of system users, it could overcome the professional consultants and give a more educated guess of what bachelor programs will be most fitting to a user.

The proposed system could be available for a small amount of payment to be used online, provide automatically generated list of degree recommendations based on a structured analysis process. In other words, the recommendation system will try to give more accurate results than human consultation, for lower cost and fast enough to help students considering and deciding.

TABLE 1.1 COMPRESSION BETWEEN PROPOSED SYSTEM AND EXISTED REPLACEMENTS

Comparison Point	Consultants	Online Browsing	Proposed System
Physical Availability	Mostly available in big cities only	Any place with computer and internet connection	Any place with computer and internet connection
Time Availability	Need to make an appointment	Anytime	Anytime
Time Consuming	Medium - High	High	Low
Accuracy	Various / Depends on the consultant experience	Week	High
Cost	High charge	Free of charge	Low charge

1.5 Project Scope

To be constructed software is aiming to help about-to-graduate and fresh graduates of high-school who want to join a university for a degree program. It will only include the majors offered by IIUM Kulliyyahs (faculties) of ICT, engineering and architecture. The factors to be considered are some personal factors, excluding of nationality, gender and race. Also, without entertaining all

environmental factors like living place, economic status, market needs and study cost.

1.6 Project Stages

TABLE 1.2 PROJECT STAGES: FYP 1

Task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Responsible
Obtain supervisor consent form															Student Supervisor
Supervision and Seminar															Coordinator Student Supervisor
Submission of FYP Synopsis															Student
Implementation; Progress evaluation															Student
FYP Showcase															Coordinator Student Supervisor
Evaluation															Coordinator Supervisor

Table 1.3 PROJECT STAGES: FYP 2

Task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Responsible
Supervisor Consultation																	Student & Supervisor
Gather Requirements																	Student
Supervisor Discussion																	Student & Supervisor
Develop the software																	Student
Test the software																	Student
Submit technical report																	Coordinator & Student
FYP 2 Show case and Evolution																	Coordinator & Examiner & Student & Supervisor

CHAPTER TWO

REVIEW OF PREVIOUS WORKS

2.1 Introduction

Following the project objectives, the papers included cover researches about recommendation systems and their approaches, recommendation systems and machine learning algorithms and recommendation systems for educational uses.

2.2 Previous Works

TABLE 2.1 TALBE OF PREVIOUS WORKS INCLUDED

No	Year	Authors	Research Title
1	2015	Isinkaye, F.O. & Folajimi, Y.O. & Ojokoh, B.A.	Recommendation systems: Principles, methods and evaluation
2	2017	Portugal, Ivens & Alencar, Paulo & Cowan, Donald	The Use of Machine Learning Algorithms in Recommender Systems: A Systematic Review
3	2012	Sunita B Aher & Lobo L.M.R.J.	Best Combination of Machine Learning Algorithms for Course Recommendation System in E-learning

4	2002	Zhang, Tong & Iyengar, Vijay S.	Recommender Systems Using Linear Classifiers
5	2010	Gershman, A. & Meisels, A. & Lüke, K.-H. & Rokach, L. & Schclar, A. & Sturm, A.	A Decision Tree Based Recommender System
6	2016	Ali, Shereen H & El Desouky, Ali I & Saleh, Ahmed I	A New Profile Learning Model for Recommendation System based on Machine Learning Technique
7	2018	Reddy, Y. Subba & Govindarajulu, P.	College Recommender system using student' preferences/voting: A system development with empirical study

2.3 Discussion on Previous Works

Based on Isinkaye and Folajimi definition of recommender systems, it could be driven that using those systems is one possible way to solve the problem at hand (choosing degree major). And by exploring the two basic approaches (content-based and collaborative) and their mechanism, my problem could be solved by any of them. Content-based filter would need collecting personal information from the user, profile him/her and match the profile with the extensive described majors in the system. Collaborative filters require categorizing users based on personal info and to recall the pre-found right choices for that category (Isinkaye & Folajimi, 2015).

Portugal, Alencar and Cowan supported the definitions, terms and categories listed. They determined that machine learning (ML) -by definition- could do the job of career & major consultants to recommend, and numbered the ML types (supervised, unsupervised, semi-supervised and reinforcement). Which type to choose for this project is depend on the data to be collected and its validity to be a training set or not. By exploring their reviewed previous researches, they indicated that the research work on RSs and ML types has been growing in the recent decade, opening the chances for practitioners to implicate the found results upon real applications. Those searches focused on several ML approaches: ensemble, support vector machine, clustering, kernel method and decision trees. Which qualify those methods to be applied more than the remaining not-enough searched methods. This review also cleared the significant role of bigdata science for ML algorithms' performance. The more of collecting, preprocessing, and structuring work for data, the more accurate results to be gotten. Lastly, the paper discussed the performance indicators for ML algorithms, and showed that top three are: precision, recall and F-measure. Those testers were resulted to be more relevant to Bayesian & decision tree than the other algorithms (Portugal & Alencar & Cowan, 2017).

According to Sunita and Lobo (2012), Pattern analysis is the process of analyzing big segments of data in order to find the implicit data behind the explicit numbers / texts / categories ...etc. They studied the recommendation process in Moodle, the online education systems. This process is meant to recommend the best course for a user to take, by collecting every user activity on the website. This data of activities is to be stored in MYSQL system, to be

explored to recognize the user pattern, using apriori association rule. They found that a combination of classification, clustering and association rule minding is the best for their course recommendation system case. Which would also be used for this project of program recommendations it have the similarity of education topic involved.

Based on Zhang's and Iyengar's paper, to have a ranked list of recommendations, the system needs to generate a specific figure / value which indicates how accurate and correct an option is to choose. They used 3 different datasets to test and compare recommendation systems based on model-based recommendation system, with linear classification models. Model-based cooperative filtering is way weaker than model based, and within model-based, linear is better than decision tree in cases of text comparison and classification (Zhang & Iyengar, 2002).

One advantage of using decision tree is the output of multiple choices instead of only one, it is also having the ability to deal with several data forms. It is shortcoming is the need to construct huge number of trees in a system. So, in cases of online shopping system, decision tree if not an efficient approach to follow as it is always et new items. But in cases of limited lists to check, it is useful as it would o through every option and prioritize the recommendations list. But the paper introduced a solution for this problem, by recursive construction the tree to return weighted list of options. But there is another problem, they found that as the tree become deeper, the performance falls down (Gershman & Meisels & Lüke & Rokach & Schclar & Sturm, 2010).

After determining the basic definitions and types of the recommendation systems and machine learning. Ali, El Desouky and Salehproposed a profile learning model, in which user is required to give some personal data, preferences to produce a list. Then a merged multi-class classifier is used to cottach the profile into a domain hypothesis. By using thousands of profiles to test and evaluate, this MMC showed high accuracy and low error possibility using many evolution inindeeer (Ali & El Desouky & Saleh, 2016).

Reddy's and Govindarajulu's work is about reviewing and developing an existing recommendation system that helps students to choose the best college (from a list of colleges exist in one region in India) based on their preferences. It has the similarity with my project of profiling both parties involved to find the best match, which are colleges and students in the paper, and in my project, the parties are degree programs and students. The systems to be developed through the research is not a machine learning based system, so there is not any automated process of evolving and developing the recommendation process, it all need to be done by system admin. It used an improved version of R-tree data structure algorithm. The paper focused on developing the processing time need to give a recommendation. Which is an important factor when there are huge number of choices to be explored. So, in my case, it could be reference for future work of including big number of majors in the recommendation system (Reddy & Govindarajulu, 2018).

CHAPTER THREE

METHODOLOGY

3.1 Background

The system is proposed to be a web application to ease the access and use of the system, as it will be machine-algorithm based and the system will use provided training set (data with correct answers).

After uploading the system online and setting up, its working follow will be to collect some personal information of the user through a form. Then the system will analyze given profile, start the recommendation process, and finally provide a list of top degree program choices.

3.2 Software Devolvement Cycle

FIGURE 3.1 SOFTWARE DEVELOPMENT CYCLE

As the proposed system is a machine-learning based one, the work of reducing the results error and enhancing the accuracy fall on the system itself

as it is programmed to self-learn from inputs. But like any other system, and beside the tradition work of maintaining the system to perform better in sense of technical issues, the system would need to have a cycle of development even for the recommendation process. The process would be enhanced by adding more new relevant decision factors, or even remove the factors resulted to be irrelevant by the software self-analysis.

For the software to learn from own experience and develop the recommendation process and results' accuracy, it will need to get a feedback for every recommendation process made. So, it will store each result with a reference number, which will be given to user in order to use it later to feedback and evaluate.

3.3 Functioning Requirements

The main tool of analysis to be used is to get datasets. By an online search, there was no suitable datasets to be used for this project. So, while applying the system, it was supported by a dataset built by own surveys. Datasets were based on data collection through two surveys, as the software is going to need a big number of responds, which makes other methods like interviews quite difficult to be done.

First survey was filled by those who finished their undergraduate and started working and those who were about to finish their UG studies and planning to start a career. Second survey was for those who finished undergraduate and were doing their postgraduate studies and those who were about to finish their bachelor and going to do a Master program. Both surveys identified how many people were not working or completing studies in the same

field of science and decided to switch fields. They also offered the dataset for the system to analyze.

To ensure the commitment of users to give feedback, a promise for a partial refund when they give feedback would be made. Also, to make sure that the feedback would not be given just to get the refund and it would not be made inaccurate and in a careless sense, there could be a minimum pass of time required before a user could refill the feedback form. Which means that there will be a timestamp for each recorded recommendation, and a date to pass before allowing user to feedback and get the refund.

3.3.1 Datasets

Below is the list of factors included for the survey to collect and build dataset, and the same factors were the base for the system to analyze and base recommendation results on.

TABLE 3.1 TABLE OF FIELDS AND EVALUTION POINTS AND RANGES

Fields	Interests Evaluation (Not Interested – Very Interested)	Experience Evaluation (Did not Experience – Highly Experienced)
Graphic Designing / Video Editing	1 – 5	1 – 5
Web Design / Development	1 – 5	1 – 5

Coding / Programming	1 – 5	1 – 5
Building Computer Networks	1 – 5	1 – 5
Hacking / White Hacking	1 – 5	1 – 5
Manual Drawing / Digital Drawing	1 – 5	1 – 5
Sketching Structures / Buildings	1 – 5	1 – 5
Planning Districts / Cities	1 – 5	1 – 5
Designing Internal Furniture / Apartment Finishing of paints, floors ... etc	1 – 5	1 – 5
Building Robots / Remote Controlled Items	1 – 5	1 – 5
Chemical Creations & Reactions	1 – 5	1 – 5

TABLE 3.2 TABLE OF HIGH SCHOOL SCORES

Subject	ICT / Computer	Chemistry	Math	Physics	Arts
Evolution	1 – 5 (Worst Score – Best Score)				

3.4 Logic Design

Use-case Diagram

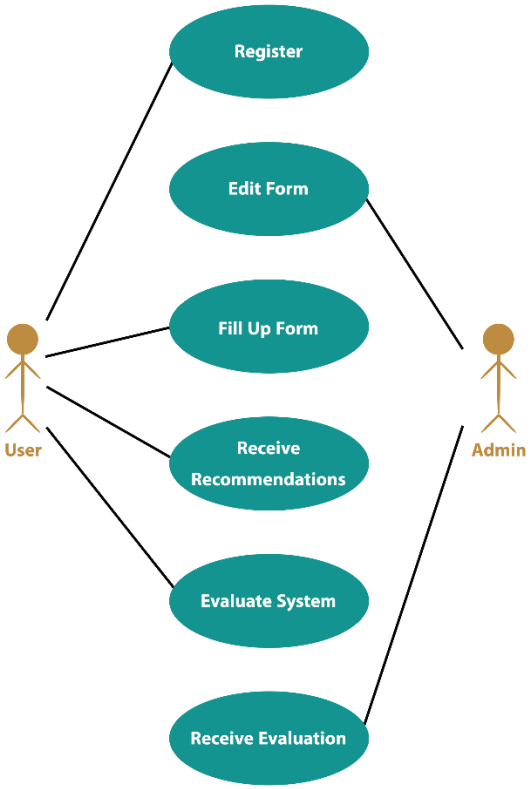

FIGURE 3.2 USE-CASE DIAGRAM

Sequence Diagram

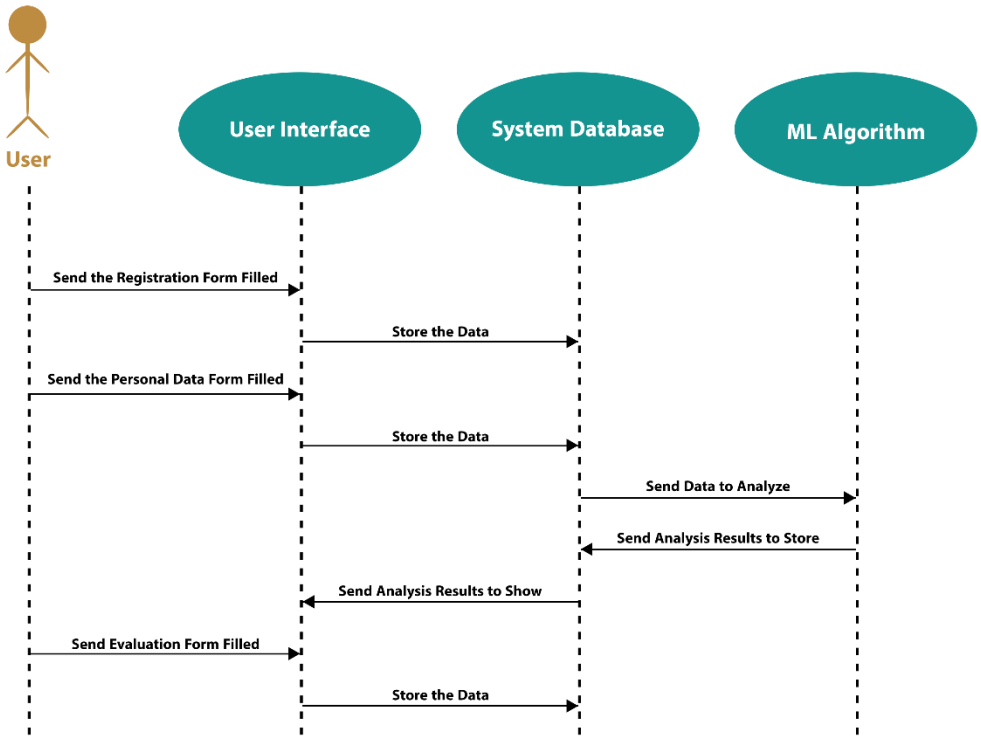

FIGURE 3.3 SEQUENCE DIAGRAM

Class Diagram

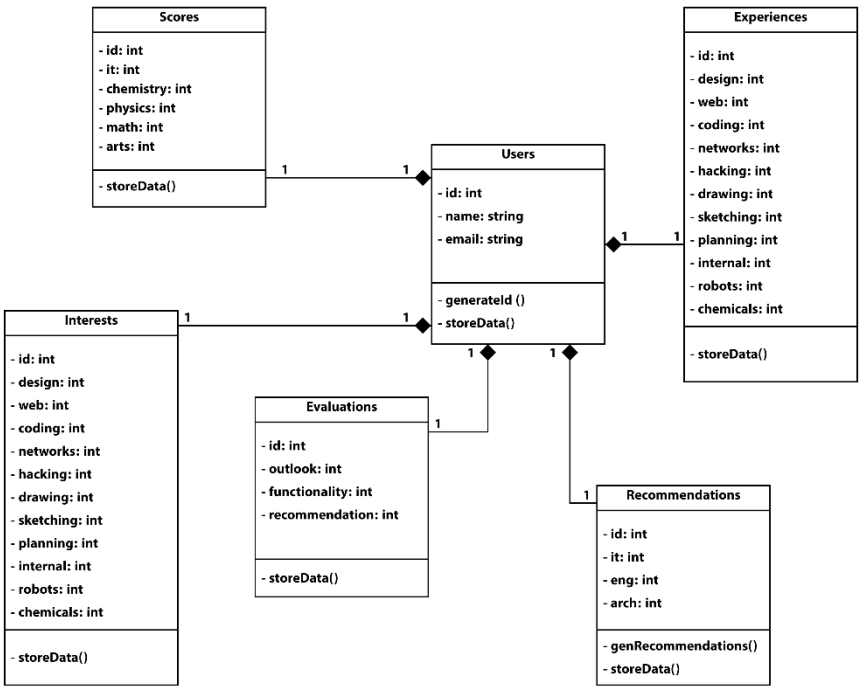

FIGURE 3.4 CLASS DIAGRAM

3.5 Functioning Flow

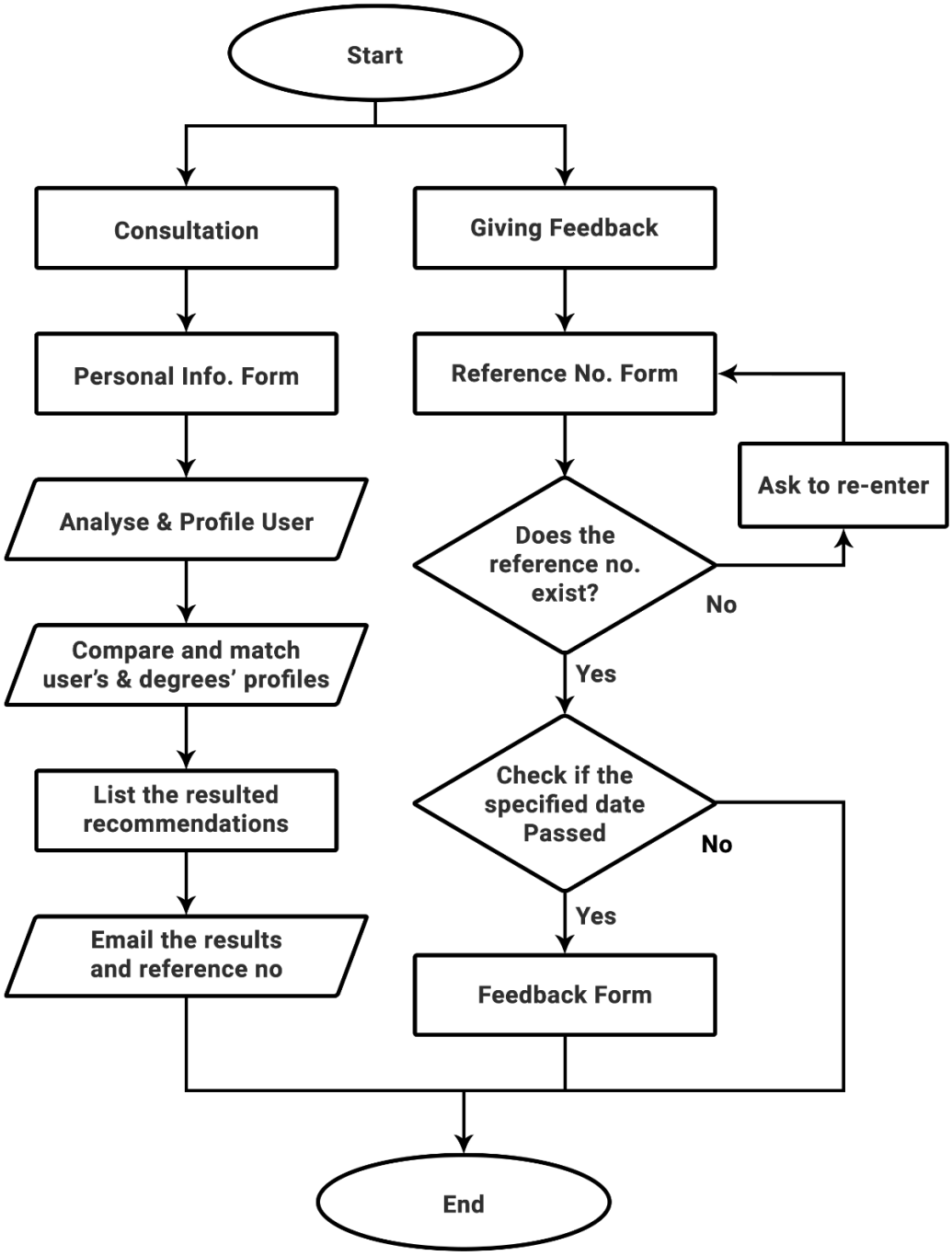

FIGURE 3.5 FUNCTIONING FLOW CHART

CHAPTER FOUR

ANALYSIS OF RESULTS AND OUTCOME

4.1 Background

The system was developed using Laravel back-end framework as it is one of the most developed and market-share growing frameworks for web applications. It was also adopted as it does offer a suitable API package called Larcombee to build a recommendations system. Larcombee is a Recombee API package for Laravel, which is an artificial Intelligence powered recommendation system service with an intuitive RESTful API & SDKs tailored by data scientists.

Additionally, as the front-end of the application is simple and do not require high specifications, there was no front-end framework used in this project, only responsive HTML and CSS coding.

4.2 Implementation

User Interface Design

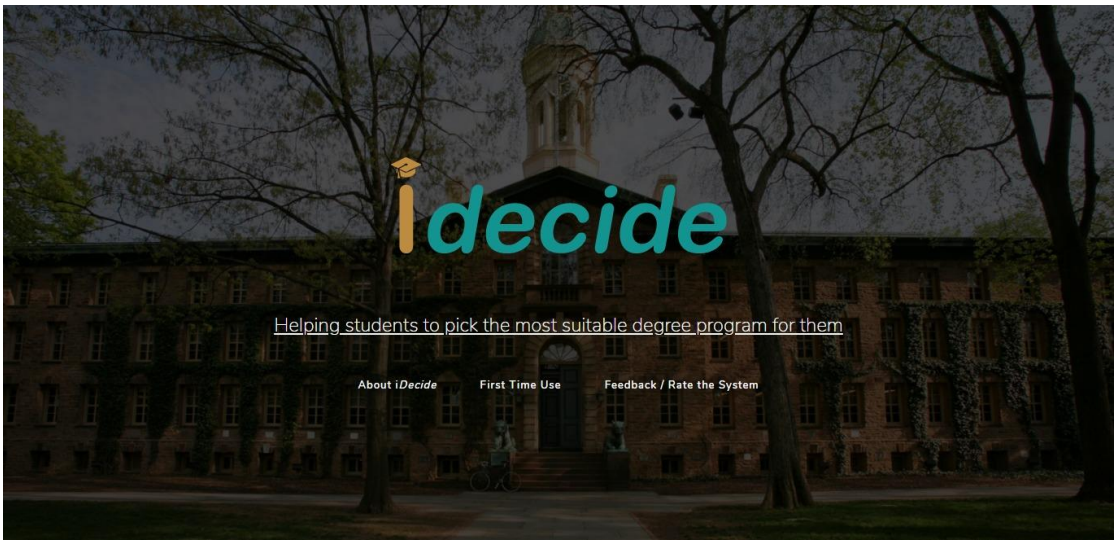

FIGURE 4.1 INITIAL DESIGN: WELCOME PAGE

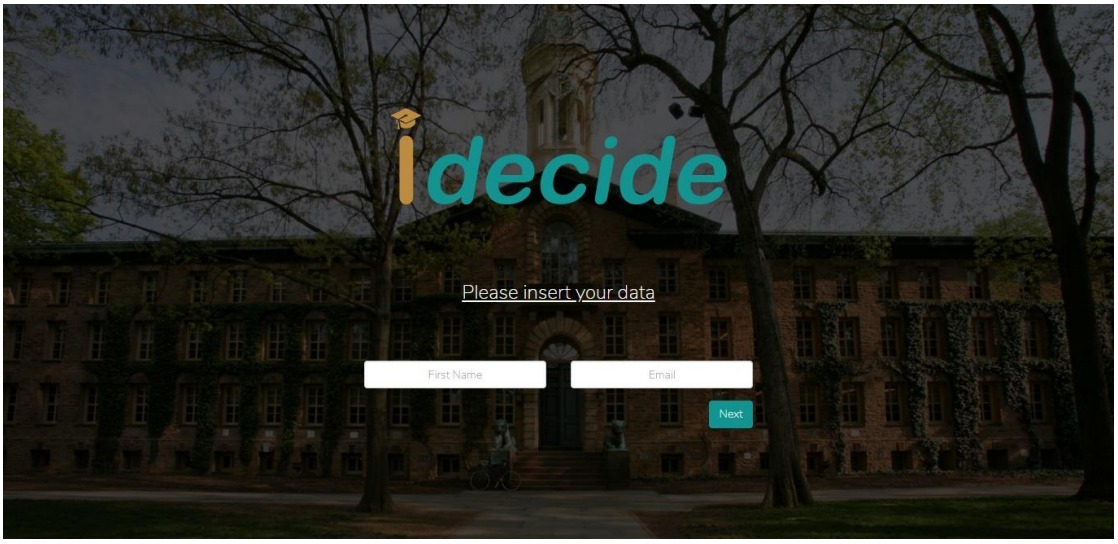

FIGURE 4.2 INITIAL DESIGN: REGISTRATION FORM

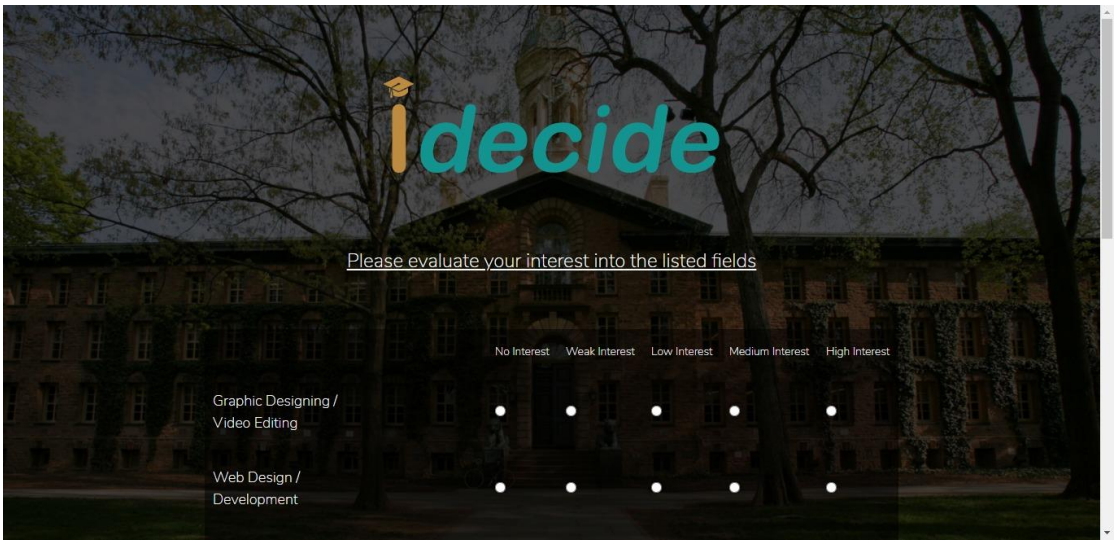

FIGURE 4.3 INITIAL DESIGN: PERSONAL DATA FORM

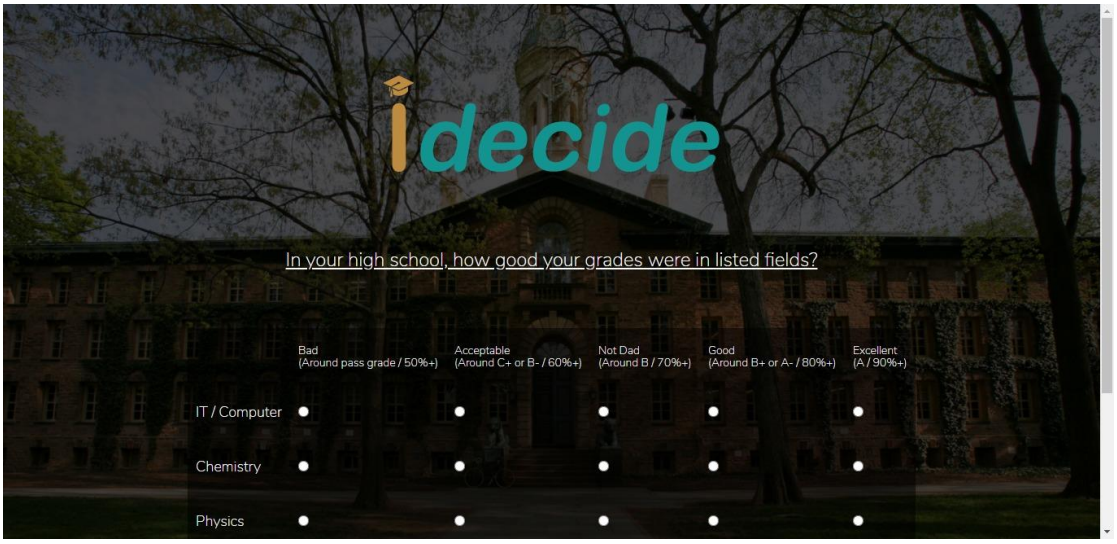

FIGURE 4.4 INITIAL DESIGN: HIGH SCHOOL SCORES FORM

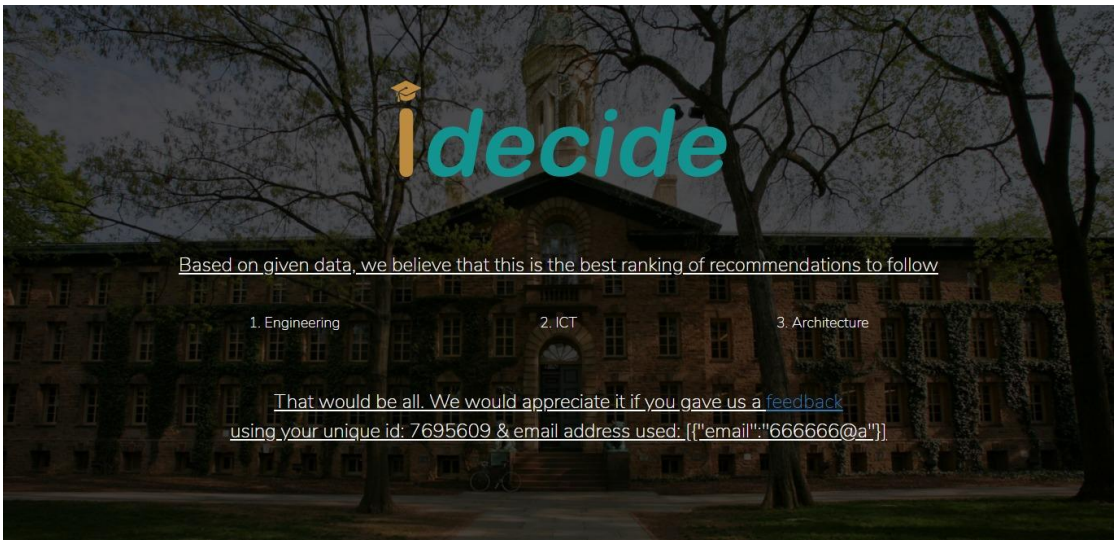

FIGURE 4.5 INITIAL DESIGN: RESULTED RECOMMENDATIONS PAGE

CHAPTER FIVE

CONCLUSION AND FUTURE WORK

The evaluation of test-group users stated that the recommendation system is functioning well, but it needs to be refined to get more accurate results. To have an extensive and more effective analysis, many more student-related factors need to be included in the recommendation process like student's financial status and market needs / trends. The system could be also extended to include factors like student's financial status and study cost. Which require an extensive and constant updatable database about the study costs for majors throughout nations. The project at hand was scooped down to bachelor levels only, and to have only three options of ICT, engineering, and Architecture. So, it could be widened to include all majors provided by all colleges and faculties within a university, a city or a whole state. It could also be upgraded to recommend for postgraduate study fields.

If the mentioned student-related factors and various majors where included, the approach used in the project may not be the best to use, and a faster and more effective method should be adopted to reduce the necessary time and processes to recommend a major / study field. The software accuracy also needs to be enhanced by providing an enormous dataset to train, which may recall for methods different from online surveys to collect, as the collected records in used dataset were limited and did not help the accuracy to sharply rise and become at significant levels.

REFERENCES

Ali, S. H., El Desouky, A. I. & Saleh, A. I (2016). A New Profile Learning Model for Recommendation System based on Machine Learning Technique. Journal of Information Technology & Software Engineering, DOI: 10.4172/2165-7866.1000170.

Fain, P (2017). Second Thoughts About Higher Education Decisions. Inside higher ed. Retrieved from https://www.insidehighered.com/news/2017/06/01/survey-finds-regrets-among-most-former-college-students-belief-quality-their.

Gershman, A., Meisels, A., Lüke, K.-H., Rokach, L., Schclar, A. & Sturm, A. (2010). A Decision Tree Based Recommender System.

Hanushek, E. A. & Wößmann, L. (2007). Education Quality and Economic Growth. The International Bank for Reconstruction and Development - The World Bank. Retrived from https://siteresources.worldbank.org/EDUCATION/Resources/278200-1099079877269/547664-1099079934475/Edu_Quality_Economic_Growth.pdf.

Isinkaye, F.O., Folajimi, Y.O. & Ojokoh, B.A (2015). Recommendation systems: Principles, methods and evaluation. Egyptian Informatics Journal.

Plumer, B (2013). Only 27 percent of college grads have a job related to their major. Washington Post. Retrieved from https://www.washingtonpost.com/news/wonk/wp/2013/05/20/only-27-

percent-of-college-grads-have-a-job-related-to-their-major/?outputType=amp.

Portugal, I., Alencar, P. & Cowan, D (2017). The Use of Machine Learning Algorithms in Recommender Systems: A Systematic Review. Expert Systems With Applications Jornal.

Reddy, Y. S. & Govindarajulu, P (2018). College Recommender system using student' preferences/voting: A system development with empirical study. IJCSNS International Journal of Computer Science and Network Security, VOL.18 No.1, January 2018.

Ramos, Y. (2013). College students tend to change majors when they find the one they really love. Borderzine. Retrieved from https://borderzine.com/2013/03/college-students-tend-to-change-majors-when-they-find-the-one-they-really-love/.

Sunita B A. & Lobo L.M.R.J. (2012). Best Combination of Machine Learning Algorithms for Course Recommendation System in E-learning. International Journal of Computer Applications (0975 – 8887).

U.S. Department of Commerce (2018). Education Statistics 2018. Retrieved from https://nces.ed.gov/programs/coe/indicator_cbc.asp.

Wolla, S. A. & Sullivan, J. (2017). Education, Income, and Wealth. PAGE ONE Economics. Retrieved from https://files.stlouisfed.org/files/htdocs/publications/page1-econ/2017-01-03/education-income-and-wealth_SE.pdf.

Zhang, T. & Iyengar, V. S. (2002). Recommender Systems Using Linear Classifiers. Journal of Machine Learning Research 2, 313-334.