RECOMMENDATION SYSTEM IN SELECTING COURSE OF PUBLIC UNIVERSITY IN MALAYSIA USING K-NEAREST NEIGHBOUR

THOI WEN BIN

BACHELOR OF COMPUTER SCIENCE

UNIVERSITI MALAYSIA PAHANG

UNIVERSITI MALAYSIA PAHANG

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name : THOI WEN BIN

Date of Birth : 19/05/1995

Title : Recommendation System in Selecting Course of Public University in Malaysia Using K-Nearest Neighbour

Academic Session : 2018/2019 Semester 1

I declare that this thesis is classified as:

☐	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997)*
☐	RESTRICTED	(Contains restricted information as specified by the organization where research was done)*
☑	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Universiti Malaysia Pahang reserves the following rights:

1. The Thesis is the Property of Universiti Malaysia Pahang
2. The Library of Universiti Malaysia Pahang has the right to make copies of the thesis for the purpose of research only.
3. The Library has the right to make copies of the thesis for academic exchange.

Certified by:

(Student's Signature)

950519-14-7025
New IC/Passport Number
Date: 7/1/2019

(Supervisor's Signature)

Junaida bt. Sulaiman
Name of Supervisor
Date: 10/1/2019

NOTE : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach a thesis declaration letter.

Universiti Malaysia PAHANG
Engineering • Technology • Creativity

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor in Computer Science (Computer Systems & Networking) with Honours.

(Supervisor's Signature)

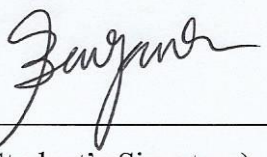
Full Name : DR. JUNAIDA BT SULAIMAN

Position : SENIOR LECTURER
FACULTY OF COMPUTER SYSTEMS
& SOFTWARE ENGINEERING
UNIVERSITI MALAYSIA PAHANG
[illegible] TUN RAZAK, 26300 GAMBANG, KUANTAN
TEL: [illegible] 2255 FAKS: 09-549 2144

Date : 10/1/2019

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang or any other institutions.



(Student's Signature)

Full Name : THOI WEN BIN

ID Number : CA15037

Date : 25 December 2018

RECOMMENDATION SYSTEM IN SELECTING COURSE OF PUBLIC UNIVERSITY IN MALAYSIA USING K-NEAREST NEIGHBOUR

THOI WEN BIN

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Computer Science (Computer Systems & Networking) with Honours

Faculty of Computer Systems & Software Engineering

UNIVERSITI MALAYSIA PAHANG

DECEMBER 2018

ACKNOWLEDGEMENT

Firstly, I would like to express my greatest gratitude to Dr. Junaida Binti Sulaiman, my supervisor who give me support and guidance throughout the research. It is a great honour to do this research under her supervision. During this period, she provided a lot of useful advices to improve my research.

I must express my appreciation to my family and friends as they support me mentally and physically. Their motivation kept me in completing the thesis in time. They also give me a lot of suggestions and ideas to continue my research.

ABSTRAK

Internet mula berkembang dengan kelajuan yang besar pada tahun-tahun kebelakangan ini. Perniagaan telah menemui banyak pelanggan dan pendapatan di internet. Terdapat maklumat dan item yang besar, dan ia menjadi berlebihan. Sistem cadangan dibangunkan untuk mengatasi masalah ini. Kini, pelajar lepasan sekolah khususnya untuk pelajar STPM dan Matrikulasi di Malaysia menghadapi masalah dengan memilih kursus universiti awam yang sesuai. Terdapat 20 universiti awam yang menawarkan 893 kursus sarjana muda di Malaysia. Mereka perlu memohon kursus-kursus yang berminat melalui UPU dalam talian dalam tarikh akhir yang ditetapkan. Di Malaysia, hanya iMASCU yang merupakan sistem cadangan mudah yang menunjukkan dan menyemak kelayakan kursus berdasarkan kelayakan pengguna. Sistem cadangan berasaskan K-Nearest Neighbour akan dilaksanakan untuk menyelesaikan masalah ini. Objektif projek ini adalah mengkaji algoritma dan teknik semasa dalam sistem cadangan untuk memilih kursus; untuk melaksanakan K-Nearest Neighbour dalam sistem cadangan; dan untuk menilai prestasisi sistem cadangan. Dataset projek ini adalah program ijazah yang ditawarkan oleh universiti awam Malaysia. Java akan digunakan untuk melaksanakan projek ini. Data dikumpulkan dari soal selidik yang dibuat kepada pelajar universiti awam. Terdapat enam input dan satu output sasaran dalam latihan dan ujian. Hasilnya menunjukkan bahawa nilai k yang dikehendaki adalah satu dan k-Nearest Neighbor boleh dilaksanakan di pemilihan kursus universiti awam Malaysia.

ABSTRACT

Internet began growing up with immense speed in these recent years. Business field had discovered many customers and income in the internet. There are huge information and items, and it became overload. Recommendation systems is developed to overcome this problem. Nowadays, freshly school leavers especially for STPM and Matriculation college students in Malaysia have trouble with selecting a suitable course of public university. There are 20 public universities offering 893 bachelor courses in Malaysia. They need to apply their interested courses through UPU online within stated deadline. In Malaysia, there is only iMASCU which is the recommendation system that show and check the qualification of courses based on the qualification of user. K-Nearest Neighbour based recommendation system is implemented to solve this problem. The objectives of this project are to study the current algorithm and technique in recommendation systems for selecting courses; to implement k-Nearest Neighbour in the recommendation system; and to evaluate the application of the recommendation system. The data is collected from survey made to current public university students. There are six inputs and one target output in the training and testing sets. The result showed that the desired k value for the data is one and k-Nearest Neighbour can be implemented in the Malaysia public university course selection.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$\sum$	SUMMATION
$\in$	ELEMENT OF
$\emptyset$	EMPTY SET

LIST OF ABBREVIATIONS

K-NN	K-Nearest Neighbour

CHAPTER 1

INTRODUCTION

1.1 Introduction

Internet began growing up with immense speed in these recent years. Business field have discovered many customers and income in the internet. So that, there are many online shop exists in the internet like Amazon, Lazada and many more. Today, there are huge information and items, and become overload. The seller is very difficult to find out what the user is looking for. This problem can be solved by search engines but only partially solved. However, some search engines not able provide personalization of information. So that, developers found a solution to solve the problem which is recommendation systems (Zhang et al., 2015).

Recommendation system is a software tool or technique that gives recommendations of items for user. Recent year, recommendation systems is being used in many field like movies, music, tourism and others. The purpose of recommendation system is to provide some suggestion to the user who have less experience or ability to assess over many choices. For example, Amazon.com implement recommendation systems to personalize the online store for their customer, so that their customer can receive different suggestion via this recommendation systems. Based on the user's preferences, recommendation system will guess the items that most suitable for user. In education field, there are many recommendation systems appeared in recent years such as Course Selecting Recommendation System (Al-Badarenah & Alsakran, 2016), E-Learning Recommendation System (Tan & Guo, 2008), and Reading Material Recommendation System (Hsu, Hwang, & Chang, 2010).

Nowadays, freshly school leavers especially for STPM and Matriculation college students in Malaysia have trouble with selecting a suitable course of public university. According UPU Online, there are 20 local universities offering 893 bachelor courses in

Malaysia. These fresh school leavers need to review all the courses before apply suitable courses. They will apply through UPU online within stated deadline.

1.2 Problem Statement

Recommendation system had several weaknesses such as the lack of data, changing of user preference, scalability, and privacy (Jain, Grover, Thakur, & Choudhary, 2015). Recommendation system need data to analyse, so that users can only get a good recommendation when the recommendation systems have a huge data set to analyse. On the other hand, user will change their preferences easily because of their purposes or any other factor. So that, recommendation systems will not recommend a suitable item for the user when user changing their preferences.

Course selecting recommendation system had been implemented in other country like Jordan, Korea and others. In Malaysia, there is a recommendation system called iMASCU to show the qualification of courses. iMASCU is a simple recommendation system that show and check the qualification of every courses and apply university course in UPU Online (Shahar, 2018). In other word, iMASCU only show the course that can applied or based on their STPM or Matriculation College result only and not consider their interest and preference. Therefore, the recommendation of iMASCU given will be not suitable for the student that have excellent exam result because he will get many recommendations from the recommendation system and the trouble of reviewing course is not reduce.

This project implement a recommendation system for selecting university course to reduce these problem of STPM students by giving them a suitable suggestion. This project implement k-Nearest Neighbour based recommendation system based on student survey response.

1.3 Objectives

The aim for this project is to develop a recommendation system in selecting course of local university in Malaysia. To achieve the aim of this project, there are some objectives that need to be consider. These objectives are as below,

i) To study the current algorithm and technique in recommendation systems for selecting university courses.

ii) To implement k-Nearest Neighbour in the recommendation system.

iii) To evaluate the application of recommendation system for selecting university course.

1.4 Scope

The scope of this project is divided in two categories, which are:

i) Dataset

The selection of courses is for degree level. The data of degree program is program of Malaysia public university that get from UPU Online. The data of public university also get from UPU Online.

The option used for secondary school and matriculation college subject include Malay Language, English Language, Chinese Language, Science, Mathematics, History, Geography, Biology, Physics, Chemistry, Additional Mathematics, Information Technology, Economics, Business Studies, Accounting and other subjects are categorised as others.

Other than that, choices of hobby include chatting, computer game, cooking, dance, eat, internet, listen to music, play music instrument, reading, singing, sleep, sport, watch drama, movie, and other hobbies are categorised as others.

For the academic discipline of the first choice of university course, the university courses are categorised into two groups, which are Science, Technology, Engineering and Mathematics (STEM), and non-STEM.

There are 73 set of data collected from university students in Malaysia and the 73 set of data is divided into 2 for training testing. 30 set of data used for training process and 43 set of data used for testing process.

ii) Language

The language for implementation of recommendation systems is Java programming language. NetBeans IDE 8.2 and WEKA 3.8 is used in this project.

1.5 Report Organization

This report consists of three chapters. Chapter 1 discuss on introduction of this project. Chapter 1 includes the introduction, problem statement, objectives, and scope.

Chapter 2 discuss about the literature review. This chapter explain the recommendation system, the technique used by recommendation system and existing recommendation systems.

Chapter 3 discuss about the methodology. Chapter 3 includes the framework and process flow of recommendation system that use in this project.

Chapter 4 discuss about the result and discussion of the research. Chapter 4 explain the training process and testing process of the experiment in this project.

Chapter 5 discuss about the conclusion. Chapter 5 first explain the overall of the research, follow by description of the achievement of objective that had achieve, then research constraint and future work had been discussed.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter explain recommendation system and its algorithm technique. Other than that, this chapter also explain the existing recommended system. Next section explains the overview about the recommendation system.

2.2 Recommendation System

Figure 2.1 shows the overview of recommendation system that derived from Poole, Mackworth, Baker, Szkuta, Sanabria, Berry, Linoff, Bogers, Van Den Bosch, Lane, Park, Kim, Choi, and Peterson. Recommendation system can be divided into many categories, the most common are content-based filtering and collaborative filtering. Many algorithms to implement these two type of recommendation systems. Bayesian Classifiers, Artificial Neural Network, and Decision Tree can use to implement content-based filtering recommendation system such as iMASCU. On the other hand, Item-Based Collaborative Filtering, Pearson Correlation Coefficient, and K-Nearest Neighbour can use to implement collaborative filtering recommendation system such as Reddit, Youtube, Last.fm, Amazon.com, Facebook, and Genius. The algorithms used, and existing application explain in the following section.

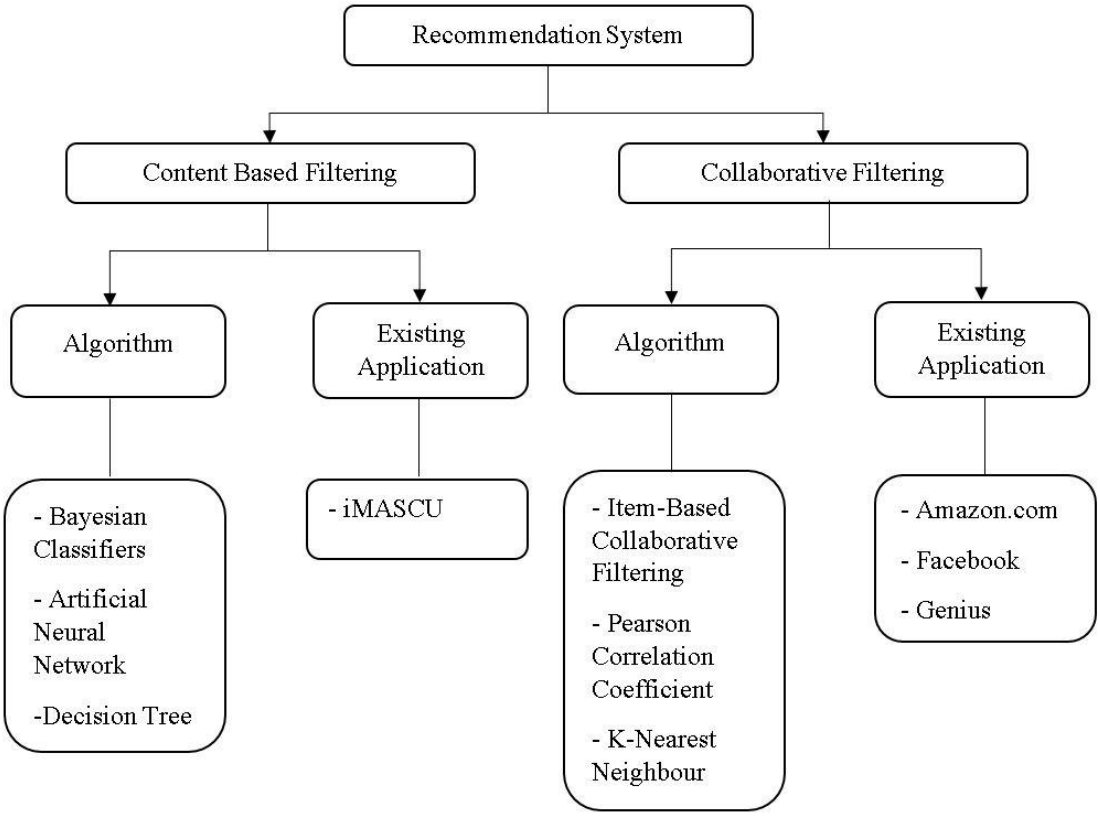


Figure 2.1 Overview of Recommendation System

2.3 Content-Based Filtering

Content-based filtering identify the preference of user at the beginning of the system and the interest is based on how the user rate the item. When create a profile, recommendation system will let user fill in a survey form and get the initial information which is the interest of the user. Items that already rated positively by the user and the item that didn't rated will be compared by content-based filtering. After that, content-based filtering will look for the similarity among users. The items with the most like the items that rated positively will be the one to recommend to the new user (Zhang et al., 2015). In content-based filtering, there are many algorithms can be implemented such as Bayesian Classifiers, Artificial Neural Network, and Decision Tree.

2.3.1 Algorithm Used in Content-Based Filtering

2.3.1.1 Bayesian Classifiers

Bayesian classifier is a probabilistic method for the classification. Bayesian classifier is to predict the values of features for members of a class based on the role of

that class. If user know the class, Bayesian Classifier can guess the values of the other features. If it does not know the class, Bayesian Classifier can use to predict the class based on feature values given. Batesian Classifier can build a probabilistic model of the features and uses that model to predict the classification of a new sample (Poole & Mackworth, 2010).

The simplest of Bayesian Classifier is Naïve Bayesian Classifier. The Naïve Bayesian Classifier independently assume that the input is conditionally independent from other classification given. Belief network will reflect the independence of the Naïve Bayesian Classifier where the features that has no parents, and classification is the only parent of the input.

According Bayes' rule, formula given as follow.

$$\mathrm{P}(\mathrm{C} = \mathrm{c}\,|D) = \frac{P(C = c)P(D\,|C = c)}{P(D)} \qquad \text{Equation 1}$$

where *D* is classified as relevant if the probability that *D* belongs to class *c* given that is contains or does not contain specific terms is larger than the probability that *D* does not belong to class *c* given the features of *D* (Poole & Mackworth, 2010).

2.3.1.2 Artificial Neural Network

Artificial Neural Network is a back-propagation network that learn from the data that are layered, connected, feed-forward neural network (Baker, 2001). The process is from a input layer enter a hidden layer, after that enter to a output layer. Whole process only carries out in one direction only. Every unit that in first layer will connect to every unit that in the next layer in one direction only. Multiple hidden layers may contain in a back-propagation network. Mapping from one data space to another can be learnt by Artificial Neural Network with high generalization capability (Szkuta & Sanabria, 1999). Figure 2.2 (Han, Kamber, & Pei, 2012) shows the example of the process of Artificial Neural Network.

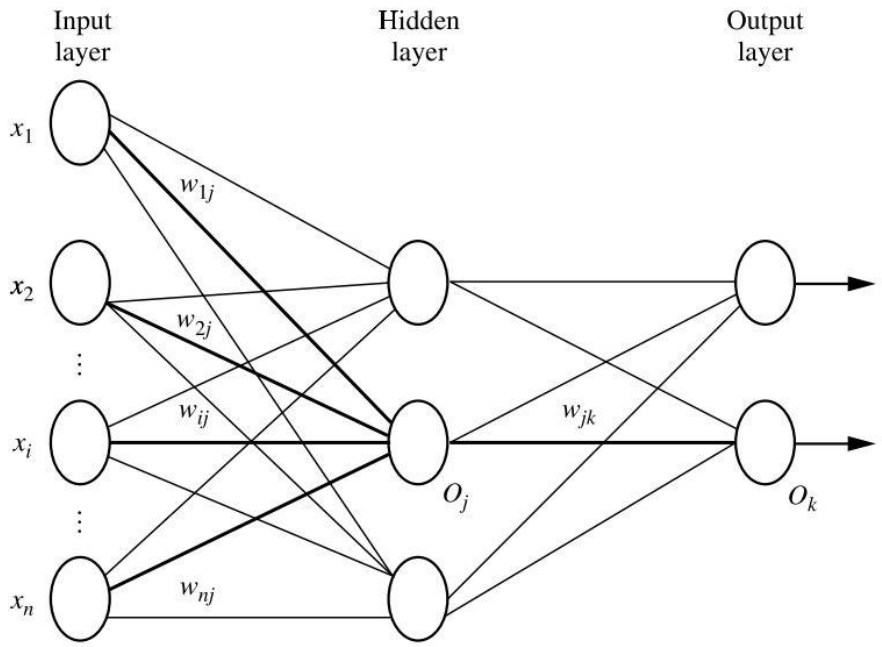


Figure 2.2 Example of Artificial Neural Network Process

2.3.1.3 Decision Tree

Decision tree is a tool that use tree-like graph of decisions and their possible consequences. Decision tree is one of the way to display algorithm that contain conditional control statement only.

To build an effective decision tree model, the data in the model set must fit the time frame (Berry & Linoff, 1999). Figure 2.3 (Yoon Ho Choa Soung Hie Kim, 2002) below shows the modelling time frame.

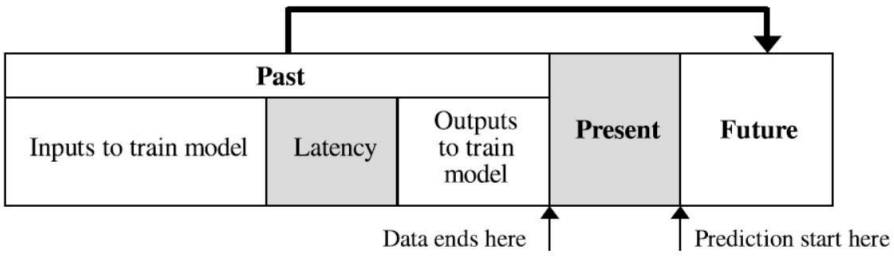


Figure 2.3 Modelling Time Frame

The modelling time frame from Figure 2.3 has three important components which is past, current, and future. Past consists of history data that has been gathered and handled. Present is the model build time. Future will be a prediction occur. Since the future prediction can be done through the past, the past is divided into three periods which

are the period of input data used by distant past; the period of output is determined; and a period of present latency. Based on the values of the independent variables, dependant variable of a new case can be assigned to a class in the score set by decision tree (Yoon Ho Choa Soung Hie Kim, 2002).

2.3.2 Existing Application of Content-Based Filtering Based Recommendation System

2.3.2.1 iMASCU

iMASCU is a simple recommendation system that show and check the qualification of every courses and apply university course in UPU Online (Shahar, 2018). In iMASCU, the details of students will be requested. For example, Identity Card Number, Name, Email, Phone Number and the most important the result of SPM or STPM. iMASCU will then show the qualification of every course to the students, so that students will know what courses is applicable, and what coursed is not applicable. Figure 2.4 (Shahar, 2018) below is the interface of iMASCU.

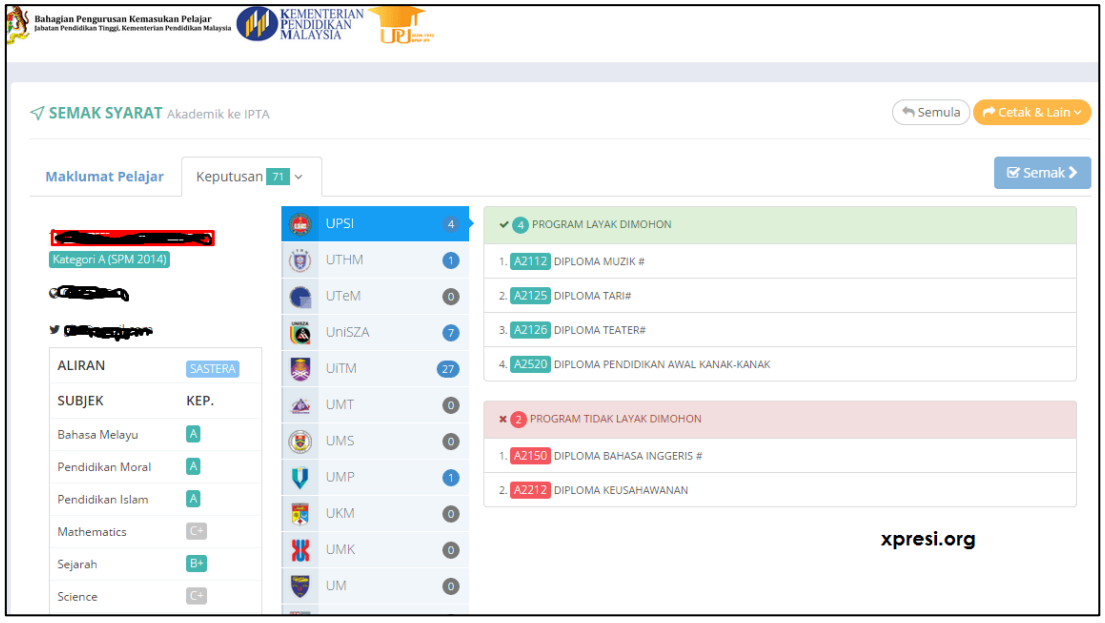


Figure 2.4 Interface of iMASCU

Although it reduces few option for students, but if the students get straight A's for SPM or get 4.0 for STPM, the number of option will be increase to maximum. So that iMASCU only can show the qualification of courses, and not suggest the most suitable courses for STPM students.

2.4 Collaborative Filtering

Collaborative filtering has become one of the most researched techniques in recent years. Collaborative filtering is a domain-independent prediction technique for the object that not likely described by metadata. Collaborative filtering will gather the preferences of user and recommend user a recommendation by finding the similarity of the other users that have similar preference. User with same preference or characteristic will build a group called neighbourhood. (Isinkaye, Folajimi, & Ojokoh, 2015). In collaborative filtering, several algorithms can be implemented such as Item-Based Collaborative Filtering, Pearson Correlation Coefficient, and K-Nearest Neighbour.

2.4.1 Algorithm Used in Collaborative Filtering

2.4.1.1 Item-Based Collaborative Filtering

Item-based collaborative filtering as known is a collaborative filtering for recommendation system to identify the similarity between rated items. User with similar items rated will build a neighbourhood based on the preference of the user. Based on the user preference, recommendation system will give recommendations to user with the items that rated in neighbourhood (Bogers & Van Den Bosch, 2009).

Item-based collaborative filtering is invented to overcome the problems of user-based collaborative filtering. Users are the main role in the user-based collaborative filtering recommendation system. User will similar preference will join into a group. Recommendation system will give recommendations based on the rated items by other users that are from the same group. User-based collaborative filtering had some trouble like recommendation system had bad performance when system contain too many items, and preference of user changed frequently will make recommendation system need to be recomputed (Sarwar, Karypis, Konstan, & Riedl, 2001).

Item-based collaborative filtering solves the troubles stated in recommendation system that have contain more users than items. Item-based collaborative filtering use rating distributions is based on item, not user. With system contains more users than items, more ratings of each item than user, so the average rating of items will not change

frequently. Therefore, rating distribution in the item-based collaborative filtering is more stable, recommendation system will not be recomputed more frequently.

For the example that show in Figure 2.5 (Zhang et al., 2015), Item A is rated by user A, user B, user C; and Item B is rated by user A, user D, user C. The similarity of Item A and Item B is both items are rated by user A and user C. After the similarity matching, Recommendation System recommend Item C, Item D, and Item E to user that same group with user A, user B, user C, and user D. This is because, Item C is rated by user A and user C; Item D is rated by user C; and Item E is rated by user B and user C.

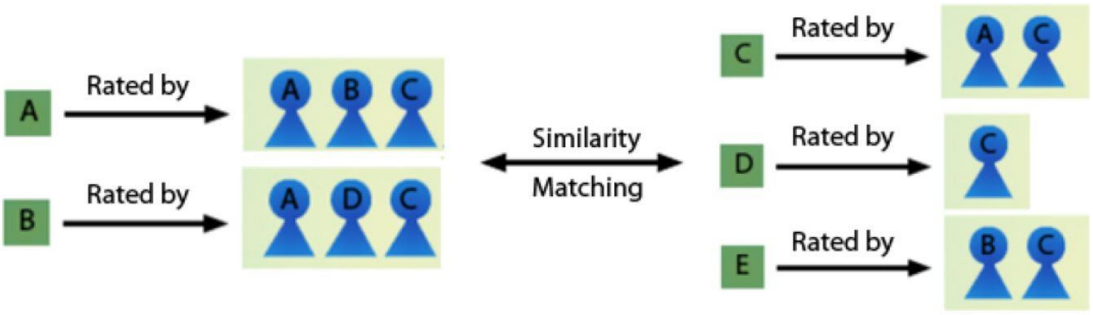


Figure 2.5 Example of Item-Based Collaborative Filtering

2.4.1.2 Pearson Correlation Coefficient

Pearson correlation coefficient is a measurement of linear relationship between two variables. Pearson correlation coefficient has a value between 1 and −1, where 1 is positive linear correlation, 0 is no linear correlation, and −1 is negative linear correlation(Lane, 2013). Table 2.1 shows the rating of five users of five restaurants. " √ " represent that the positive rated of restaurant by user, however "×" represent that the negative rated of restaurant by user.

Table 2.1 Rating of five users of five restaurants

	User B	User D	User A	User C	User E
Restaurant A	√	√	×	√	×
Restaurant B	√	√	√	√	√
Restaurant C	×	×	√	√	√
Restaurant D	√	√	×	×	×
Restaurant E	×	×	√	√	?

From example in Table 2.1, to guess the rating of Restaurant E given by User E Restaurant E, similarity between User E and other users need to be determined. An approach used to identify the correlation between User E and other users. Average of weighted of the recommendations form some users can be determined. User's rating that given by the weight will be found by correlation between USER E and the other users. Pearson *r* is a measure of correlation. Let R_{ij} be the rating of user *i* on document *j*. The formula of correlation between user *x* and user *y* is given by:

$$r(x,y) = \frac{\sum_{d \in document}(R_x, d - \bar{R}_x)(R_y, d - \bar{R}_y)}{\sqrt{\sum_{d \in document}(R_x, d - \bar{R}_x)^2 \sum_{d \in document}(R_y, d - \bar{R}_y)^2}} \quad \text{Equation 2}$$

Where $\bar{R}_x$ is the mean value of ratings by user *x*. From the Table 2.1, the result of correlation calculation between User and other users result will be User A is -0.577, User B is –0.577, User C is 1.0, User D is 0.577. Therefore, the weighted average of the production of rating for Restaurant E given by each user's and the correlation between User E and other users is 0.682. Pearson correlation coefficient will predict that User E will give Restaurant E a positive rating based on the rating of other users and the calculation of Pearson correlation coefficient.

2.4.1.3 K-Nearest Neighbour

In the past 10 years, most of the researches are using k-Nearest Neighbour to deploy their recommendation system (Park, Kim, Choi, & Kim, 2012). K-Nearest Neghbour(K-NN) is one of the simplest method for solving classification problems. Fix and Hodges first introduce K-NN in year1951 (Peterson, 2009). K-NN is not popular

when first introduced, but now it has become widely used in pattern recognition and classification since 1960s (Han et al., 2012).

K-NN defined as learning by analogy, it compares a specific test data with a set of training data that are similar to it. Classification is based on the result of their closest neighbours. The name K-Nearest Neighbour, the K indicates the number of neighbours taken to determine the result.

K-NN usually apply Euclidean distance to determine the similarity between two testers. Given formula as below:

$$\mathrm{d}(x, y) = \sqrt{\sum_{k=4}^{n} (x_k - y_k)^2} \qquad \text{Equation 3}$$

where *n* is the number of attributes; *k* is k^{th} components of data; *x* is referred to user 1; and *y* is referred to user 2.

There is another formula to predicted rating for user and item, given formula as below:

$$p_{ij} = \bar{r}_i + \frac{1}{k} \sum_{k=1}^{n} w(i,p)(\bar{r}_{pj} - \bar{r}_p) \qquad \text{Equation 4}$$

where $\bar{r}_I$ is the average rating for user *i*, and *k* is a normalizing factor to guarantee value of weights sum to 1; *w(i,p)* is the correlation between user *i* and *p*. *w(i,p)* is the Pearson correlation coefficient between user *i* and *p*.

After calculating the similarity, the K value of K-Nearest Neighbour need to determine. K value usually is odd number like 1, 3, 5, and 7. If K = 1, 1-NN approach is referred to the nearest neighbour. If K value is more than 1, for example K = 3, 3-NN approach will choose the 3 nearest neighbour and choose the majority of the same result in the 3 nearest neighbour to be recommended to the user, The K value also can be determined by testing the error rate of the classification. Many researchers use the range of 1 to 5 in their K-NN recommendation system (Adeniyi, Wei, & Yongquan, 2016).

2.4.2 Existing Application of Collaborative Filtering Based Recommendation System

2.4.2.1 Facebook

Facebook is well-known online social media in the world. Facebook applied recommendation system is to scale the massive data sets, and to overcome issues that appear because of certain data such as skewed items. The technique that Facebook applied in recommendation is collaborative filtering. At Facebook, collaborative filtering helps user find out the item that most relevant to them. This include the page, group, event and more. Collaborative filtering provides a good recommendation come from the user that have similar interests. Figure 2.6 (Facebook, 2015) at below is the phase of Apache Giraph which is a recommendation system that written by Facebook (Facebook, 2015).

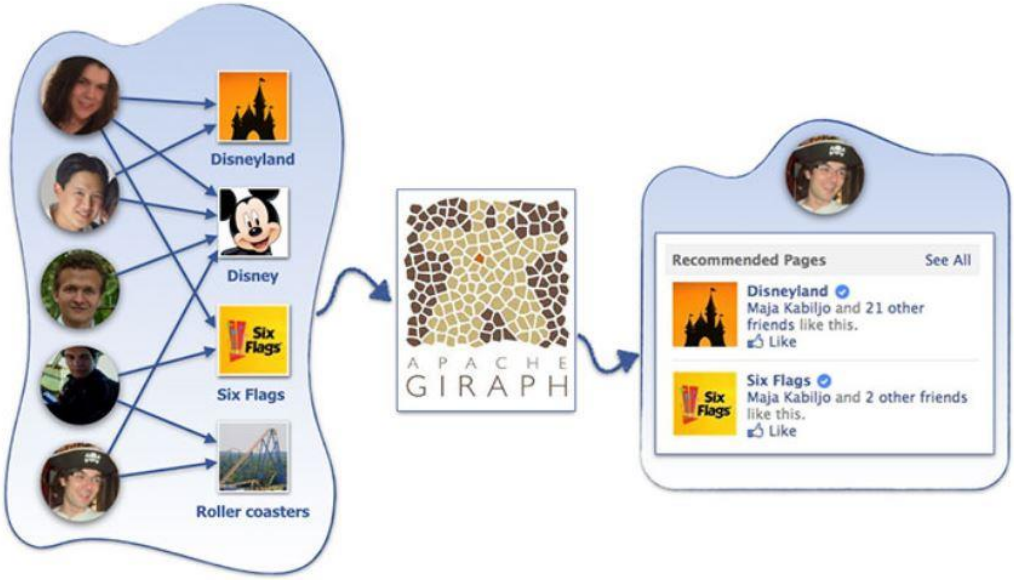


Figure 2.6 Facebook Recommendation System

Based on Figure 2.5, the interested page of user is “Disney” and “Roller coasters”. So that, Facebook will provide a list of recommended pages to the user which are “Disneyland” and “Six Flags” based on the other user that also have interest in “Disney” and also “Roller coasters”.

2.4.2.2 Amazon.com

Amazon is the largest online shopping website in the world based on revenue and market capitalization. Recommender system that applied in their website is using their interest of customer as the input to produce recommendation items. Other application only uses the items that the user purchase and rate to show the user interests. But, Amazon use recommendation system to personalize the online store for their users.

The technique that applied in website of Amazon is item-based collaborative filtering. Item-based collaborative filtering unlike the other traditional collaborative filtering, item-based collaborative filtering scales the number of users and items independently. This technique give suggestion in real-time, scope of huge data sets, and produce a high-quality suggestion (Linden, Smith, & York, 2003)

2.4.2.3 Genius

Genius is a well-known music recommendation system that developed by Apple Inc. Genius use the Gracenote MusicID service to fingerprint songs in music library, it is to identify the songs' details in the Genius' database. Although the technique of recommendation system applied in Genius is a trade-secret of Apple Inc but Genius appear to use collaborative filtering to compare the song's metadata.

When Genius is first initialized, it will analyses user's music library and compiles all the collaborative filtering data necessary to build playlist from the library based on seed song given. If we choose a seed song that is very atypical of the style of the artist or album, Genius will recommend music that represent the more common aspects of the artist. For example, using a country-folk ballad by a rock metal band as a seed song, Genius will recommend the thrash-medal songs. Genius is expected to perform well when recommend song based on a popular seed song but have to suffer when based on a less popular seed song (Barrington, Oda, & Lanckriet, 2009).

2.5 Comparison Table

In Table 2.2 describe the summary of the two type of recommendation system which are content based filtering and collaborative filtering.

Table 2.2 Summary of Type of Recommendation System

	Content Based Filtering	Collaborative Filtering
Description	Identify the preference of user at the beginning of the system and the interest is based on how the user rate the item.	Gather the preferences of user and recommend user a recommendation by finding the similarity of the other users that have similar preference.
Advantage	Use items' content to predict user's preference (Park et al., 2012).	Predict item through similar user patterns (Park et al., 2012).
Disadvantage	Impossible to predict the totally distinct types of items (Park et al., 2012).	Recommendations difficult for the users with distinct tastes. This situation is called black sheep or grey sheep (Park et al., 2012).
Algorithm that can use	Bayesian Classifiers, Artificial Neural Network and Decision Tree.	Item-Based Collaborative Filtering, Pearson Correlation Coefficient and K-Nearest Neighbour.
Existing Application	iMascu.	Amazon, Facebook and Genius.

2.6 Conclusion

This chapter explained recommendation system and its algorithm technique. Other than that, this chapter also explained the existing recommended system. Recommendation system can be divided into many categories, the most common are content-based filtering and collaborative filtering. Many algorithms to implement these two type of recommendation systems. Bayesian Classifiers, Artificial Neural Network, and Decision Tree can use to implement content-based filtering recommendation system such as iMASCU. On the other hand, Item-Based Collaborative Filtering, Pearson Correlation Coefficient, and K-Nearest Neighbour can use to implement collaborative filtering recommendation system such as Amazon.com, Facebook, and Genius.

In next chapter will discuss about the methodology used to conduct this project.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter will discuss about the methodology used to conduct this project. Next section will explain briefly about the overview framework and process flow of recommendation system.

3.2 Recommendation System Framework

Figure 3.1 (Wang, Yu, Feng, & Wang, 2014) shows the recommendation system framework of this project. From Figure 3.1, the new user will give their preference by filling a set of question. After that, system will extract data into user preference based on the question given. After that, system will calculate the similarity between new user and old users by using Euclidean Distance. The result of the calculation will be collected to identify which user is the nearest to the new user. After that, K-Nearest Neighbour is used to determine, *K* number of user will be considered to identify the recommendation. The majority of result of *K* user number will be recommend to the new user (Wang et al., 2014).

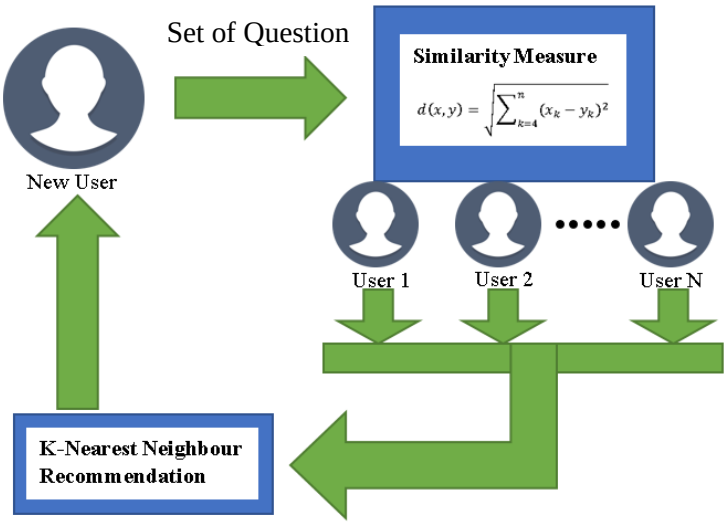


Figure 3.1 Framework of Recommendation System

3.3 Development Of Recommendation System

Figure 3.2 (Bhumichitr, Channarukul, Saejiem, Jiamthapthaksin, & Nongpong, 2017) shows the process flow of development recommendation system. There is some process flow in the development, which are data gathering, data pre-processing and prediction.

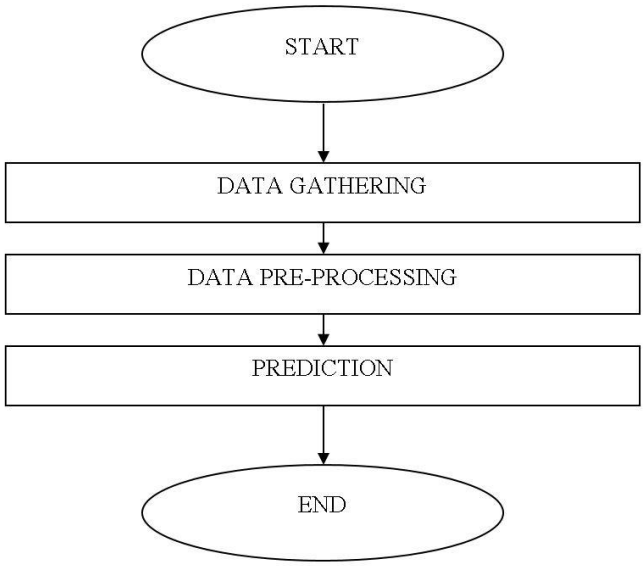


Figure 3.2 Process Flow of Recommendation System

3.3.1 Data Gathering

Data can be defined as a set of objects and their attributes, where an attribute is defined as a characteristic of an object. Data of user will be gather at this phase by filling a set of question to identify their preference, and result. The data will be divide into few categories like the example given in Table 3.1. The data that in the example of Table 3.1 is fiction data.

Table 3.1 Example of Data Gathered

USER	PREFERED SUBJECT	PREFERED FIELD	PREFERED UNIVERSITY	PREVIOUS RESULT	COURSE CODE
USER 1	BIOLOGY	SCIENCE	UMP	3.0	JG44
USER 2	ACCOUNT	BUSINESS	UUM	3.5	UE02
USER 3	PHYSICS	SCIENCE	UM	3.25	MS11

These data will be ready to process in the next process which is data pre-processing.

3.3.2 Data Pre-Processing

Data typically needs to be pre-processed in order to let machine learning technique use to analysis (Bogers & Van Den Bosch, 2009). Data that from the questionnaire will be extracted in this phase. The data will be divide into few categories like the example given in Table 3.1. After that, the similarity of old user and new user will be measured in this phase. The similarity measurement will be discussed in the following section.

3.3.2.1 Similarity Measurement

In this project, the Euclidean Distance formula to calculate the similarity is as follow:

$$d(x,y) = \sqrt{\sum_{k=4}^{n} (x_k - y_k)^2}$$

where n is the number of attributes; k is k^{th} components of data; x is referred to user 1; and y is referred to user 2. Based on the example in Table 3.1, the data is processed as in Table 3.2.

Table 3.2 Example Data Set when New User Register

USER	PREFERED SUBJECT	PREFERED FIELD	PREFERED UNIVERSITY	PREVIOUS RESULT	COURSE CODE
USER 1	BIOLOGY	SCIENCE	UMP	3.0	JG44
USER 2	ACCOUNT	BUSINESS	UUM	3.5	UE02
USER 3	PHYSICS	SCIENCE	UM	3.25	MS11
USER 4	BIOLOGY	SCIENCE	UTM	3.5	?

From Table 3.2, the difference between new user and old user can be calculated by comparing two values of each component which is $x_k - y_k$ in the formula. If two values are different, the result of $x_k - y_k$ is 1. If two values are same, the result of $x_k - y_k$ is 0. For result, the difference will be taken as usual, for example result of USER 1 is 3.0 and USER 4 is 3.5, so the difference is -0.5. The outcome of calculation is shown in Table 3.3

Table 3.3 Expected Outcome of Calculation from Example Data Set

RANK	USER	Distance to USER 4
1	USER 1	1.118033989
2	USER 3	1.436140662
3	USER 2	1.732050808

As result from Table 3.3, USER 1 is the nearest to the USER 4; and USER 2 is the farthest to USER 4. In another word, USER 1 have most similarity with USER 4; and USER 2 is the least similarity with USER 4.

3.3.3 Prediction

In this phase, recommendation system will predict the result based on the calculation from data pre-processing phase. After the calculation of the formula given is obtained, recommendation system will give recommendation using k-Nearest Neighbour. From Table 3.3, the k value is 1, which is k = 1, so that 1-NN approach will pick the user

who nearest to USER 4, so the recommendation of USER 1 will be recommended to USER 4, which is JG44. If k value is 3, which is k =3, 3-NN approach will compare the result 3 nearest distance to user 4. USER 1 is JG44, USER 2 is UE02, and USER 3 is MS11, because of three results is different and there are no majority result, so system will choose the nearest neighbour which is the result of USER 1, JG44.

3.4 Evaluation

To make sure the quality of recommendation given by the recommendation system, two standard information retrieval measures is used which are precision and recall. Precision is the fraction of recommended items that is relevant to the user. Recall is the fraction of relevant items that are part of recommended items. The formula is given as follow:

$$\text{Precision} = \frac{Number\ of\ relevant\ recommendation}{Total\ number\ of\ recommendation}$$

$$\text{Recall} = \frac{Number\ of\ relevant\ recommendation}{Total\ number\ of\ useful\ recommendation}$$

The highest of these two values will be 1, means the system have a best performance in giving a relevant recommendation to user. However, the lowest value is 0, means the system have a worst performance in giving a relevant recommendation to user. Precision is more significant than recall, because high quality recommendation is more important than a large number of recommendations in recommendation systems (Isinkaye et al., 2015).

3.5 Gantt Chart

Refer to appendix A.

CHAPTER 4

RESULT AND DISCUSSION

4.1 Introduction

This chapter discuss about the implementation of the project as mentioned in Chapter 3. NetBeans IDE 8.2 and WEKA 3.8 were used in this project. Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. The implementation of the project is divided into two parts which are training process and testing process. In training process, selection of K value for the data had been done. In testing process, the desired of k value that selected in training process is used in testing process. On the other hand, recommendation system is implemented using K-Nearest Neighbour in this project.

4.2 Experiment Set Up

The dataset that used in this experiment is collected from a survey and filled in by public university students. There are total of 73 respondents' data has been collected in this survey. The example of survey can refer to appendix F. The example of raw data collected from survey are shown in Figure 4.1 and other data can refer to appendix B.

Gender	Which subject you liked the most in secondary school/Matriculation college?	What is your hobby?	Which type of personality best suits you?	Which of the following university is the first choice for continuing your undergraduate study ?	You have chosen Universiti Kebangsaan Malaysia (UKM). Please choose your first-choice undergraduate program.
Female	ECONOMIC	SPORT	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Islam Antarabangsa Malaysia (UIAM)	ECONOMICS - SARJANA MUDA EKONOMI (KEPUJIAN)
Female	add math	badminton	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Kebangsaan Malaysia (UKM)	STATISTICS - IJAZAH SARJANAMUDA SAINS DENGAN KEPUJIAN (SAINS AKTUARI)
Female	Biology	Read novel	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Kebangsaan Malaysia (UKM)	APPLIED SCIENCE - IJAZAH SARJANAMUDA SAINS FORENSIK DENGAN KEPUJIAN
Female	Chemistry	watch Tv	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Kebangsaan Malaysia (UKM)	DENTAL STUDIES - DOKTOR PERGIGIAN
Female	Chinese	Listening to music	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Kebangsaan Malaysia (UKM)	APPLIED SCIENCE - IJAZAH SARJANAMUDA SAINS DENGAN KEPUJIAN (BIOTEKNOLOGI DENGAN PENGURUSAN)

Figure 4.1 Raw Data That Collected from Survey Form

These data were divided into two groups for training and testing. The data used for training process can refer to appendix D, while data used for testing process can refer to appendix E. Figure 4.2 show the main interface of system to allow user to proceed to training process or testing process.

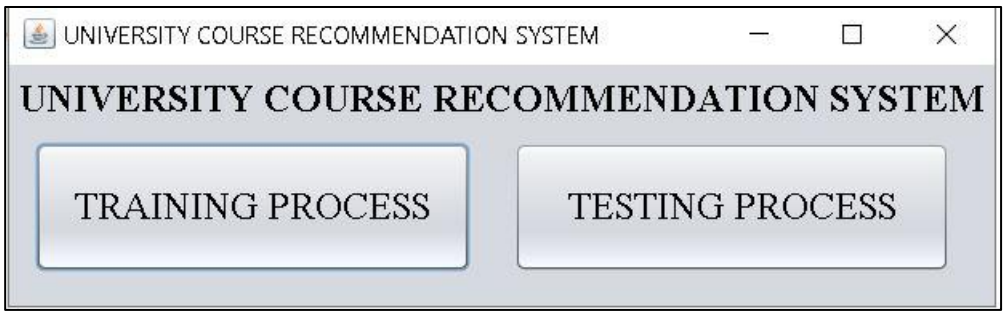


Figure 4.2 Main Interface of University Course Recommendation System

4.2.1 Training Process

In training process, the WEKA API were used to determine which k value is suitable for this dataset. WEKA jar file was imported to the NetBeans's project source library. Figure 4.3 show the interface of training process.

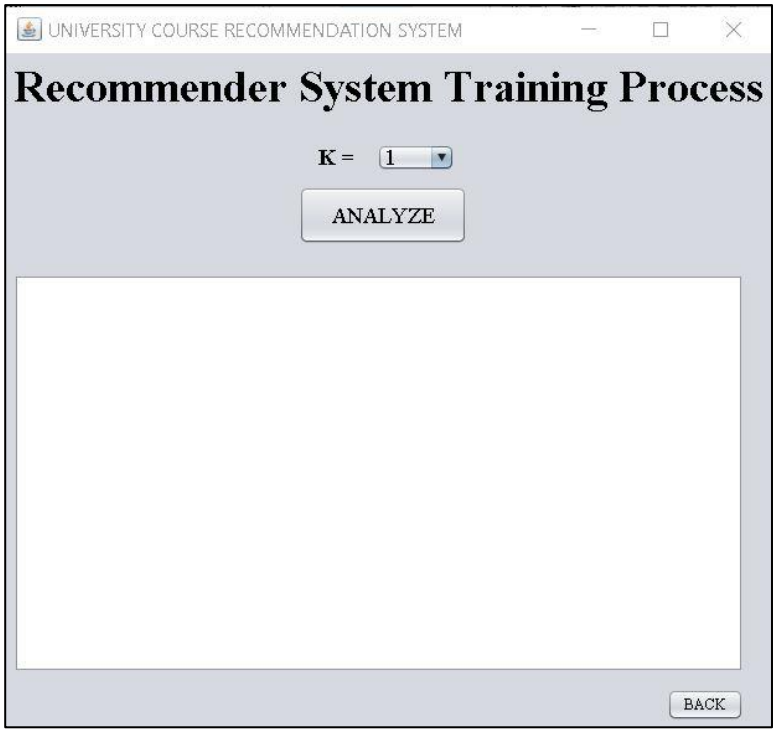


Figure 4.3 Interface of Training Process

After user choose the K value from 1 to 10 and click the ANALZE button to show the classification result of the K value. K value was only set from 1 to 10 because of the small dataset used. The K-Nearest Neighbour classification result of K value were shown as Figure 4.4.

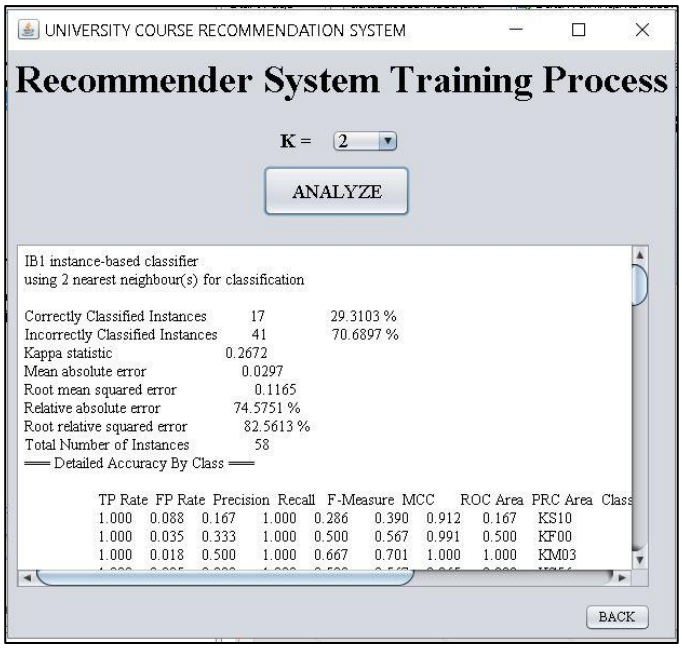


Figure 4.4 Training for K Value

In Figure 4.4, user can view the statistical summary and measurement and confusion matrix of the K value chosen. Correctly classified instances show the accuracy predicted value while incorrectly classified instances show the inaccuracy of predicted value. Kappa statistic show how closely the instances classified by Weka matched the data labelled as ground truth, controlling the accuracy of random classifier as measured by the expected accuracy. Mean absolute error measure the average magnitude of the error in a set of forecasts without considering their direction. Root mean squared error show the quadratic scoring rule which measure the average magnitude of the error. Relative absolute error is the total absolute error which same kind of normalization, errors are normalized by the error of the simple predictor that predicts average values. Root relative squared error refer to error is made relative to what it would have been if a simple predictor had been used.

Three measurements used to evaluate the K value which are the correctly classified instances, mean absolute error and root mean squared error (Srivastava, 2018). The K value decided by the highest correctly classified instances, lowest mean absolute error and lowest root mean squared error. The result of K value training from 1 to 10 is shown in Table 4.1.

Table 4.1 Result of K Value in Training Process

K Value	Correctly Classified Instances (%)	Mean Absolute Error	Root Mean Squared Error
1	100	0.0183	0.0647
2	29.3103	0.0297	0.1165
3	24.1379	0.033	0.1259
4	24.1379	0.0346	0.1297
5	18.9655	0.0362	0.1336
6	17.2414	0.0368	0.1349
7	13.7931	0.037	0.1354
8	10.3448	0.0372	0.1359
9	10.3448	0.0375	0.1365
10	12.069	0.0376	0.1364

Based on the result in Table 4.1, the K value chosen was one for this dataset and used for testing process. The reason of K value chosen was one because of its accuracy and low error. In Figure 4.5, the code is used for calling the WEKA API in java programming.

```
private void btnANAActionPerformed(java.awt.event.ActionEvent evt) {
    try
    {
        ArffLoader loader = new ArffLoader();
        loader.setSource(new File("C:/Users/Pc/Documents/TRAINING PROCESS.arff"));
        Instances data = loader.getDataSet();
        //Load the Data Training arff file

        data.setClassIndex(data.numAttributes()-1);

        Classifier ibk = new IBk(Integer.parseInt(cbNumber.getSelectedItem().toString()));
        ibk.buildClassifier(data);

        Evaluation eval = new Evaluation(data);
        eval.evaluateModel(ibk, data);

        taTrainingResult.setText(ibk + eval.toSummaryString() + eval.toClassDetailsString() + eval.toMatrixString());

    }
    catch (Exception ex) {
        Logger.getLogger(DataTrainingInterface.class.getName()).log(Level.SEVERE, null, ex);
    }
```

Figure 4.5 Coding of Call the WEKA API

4.2.2 Testing Process

In testing process, the desired of k value that selected in training process is used in testing process. The model obtain from the training process is implemented using testing data. The remaining 15 data is used in testing process. The testing result is shown in Figure 4.6 when K is one.

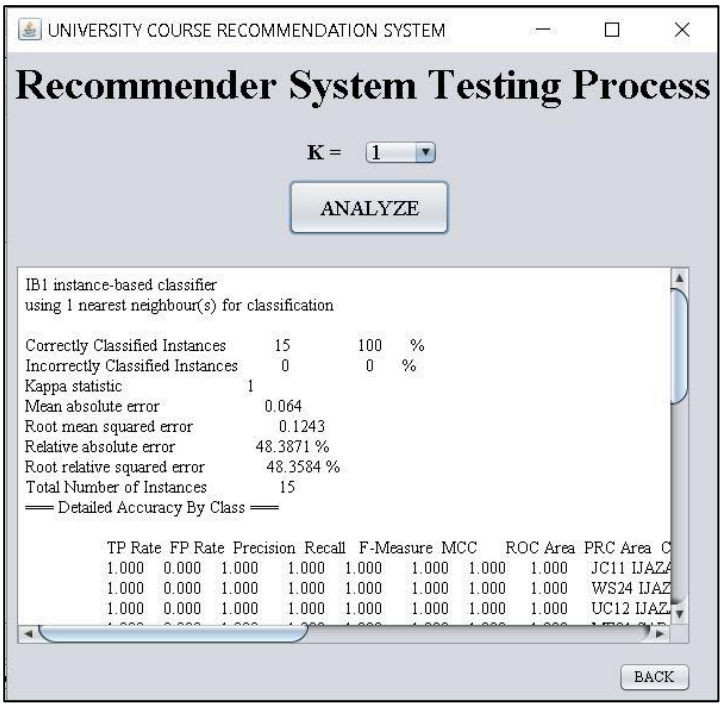


Figure 4.6 Testing for Training ModelFigure4 Testing for Training Model

Based on Figure 4.6, the correctly classified instances are 100% which same with training model when k value is one. This is because the closest point to any training data point is itself. Hence, the correctly classified instances are always get 100% accuracy

with k value is one. The mean absolute error and root squared error for testing model is low which are 0.064 and 0.1243 respectively.

4.2.3 Recommendation System

Recommendation system for university course selection is implemented in this project using the desired k value that selected from training and testing model which is one.

Figure 4.7 show the interface of the recommendation system. After user choose the options of all question, the system recommends a university course to the user. In Figure 4.8, the user had to choose the option of gender, favourite secondary school/Matriculation college subject, hobby, personality, university for continuing undergraduate study and the academic discipline. After user click SUBMIT button, the system made a recommended course to the user.

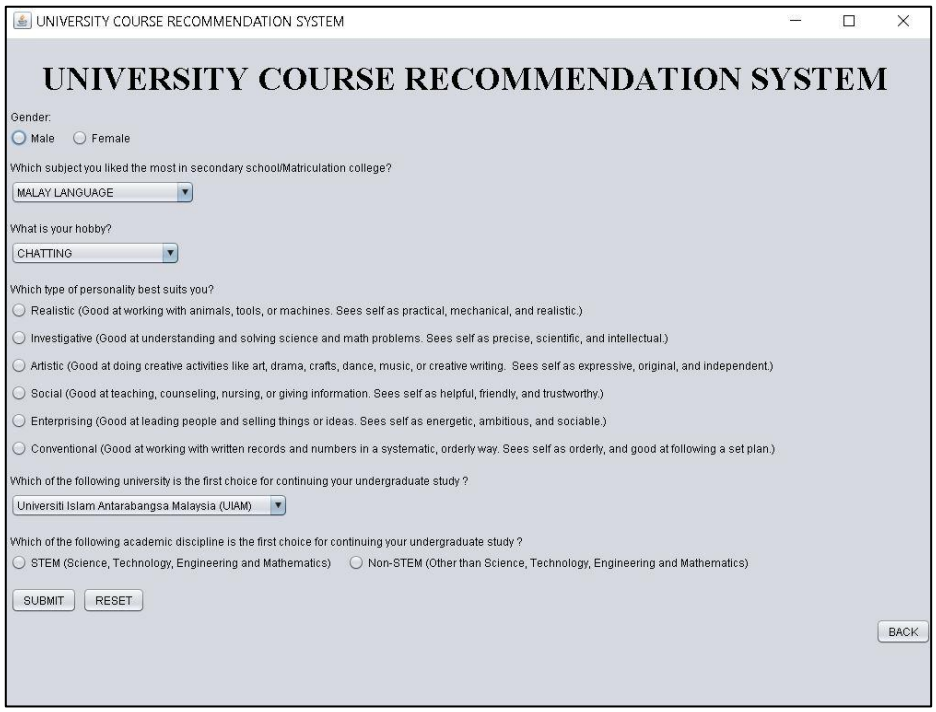


Figure 4.7 Interface of Recommendation System

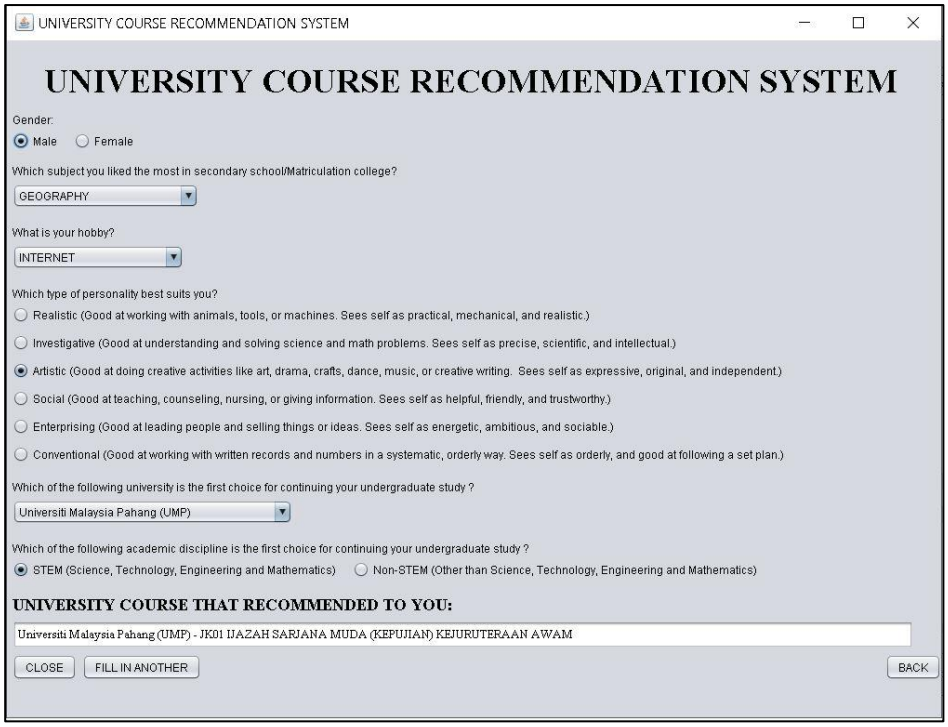


Figure 4.8 Example of System Recommend User a Course

The code shown in

Figure 4.9, Figure 4.10, Figure 4.11, Figure 4.12 and Figure 4.13 were implemented in the project to build recommendation system. The code in

Figure 4.9 is connect to the database and code in Figure 4.10 is to retrieve the data from the database.

Then, the data of the database will compare with the input data of the simple system and calculate the Euclidean Distance in Figure 4.11. After that, the Euclidean Distance of each instance will be listed and sorted from the nearest to farthest in Figure 4.12. The coding in Figure 4.13, the system will list down the nearest neighbour.

```
public databaseconnect(){
    try{

    conn=DriverManager.getConnection("jdbc:derby://localhost:1527/PSMDATA","Benjamin","12345678");
    st=conn.createStatement();

   }
   catch(Exception e){
    System.out.println(e);
   }
}
```

Figure 4.9 Coding for Connect the Database

```
//Retrieve data from database
int Username = rs.getInt("USERNAME");
String GenderD = rs.getString("GENDER");
String SubjectD = rs.getString("SUBJECT");
String PersonalityD = rs.getString("PERSONALITY");
String UniversityD = rs.getString("UNIVERSITY");
String STEMD = rs.getString("STEM");
String FirstChoiceD = rs.getString("FIRSTCHOICE");
String HobbyD = rs.getString("HOBBY");
```

Figure 4.10 Coding for Retrieve Data from Database

```
//Compare Survey Data with Database Data
if (GenderF.equals(GenderD)){resultcg = 0.0;}
else {resultcg = 1.0;}
if (SubjectF.equals(SubjectD)){resultcs = 0.0;}
else {resultcs = 1.0;}
if (PersonalityF.equals(PersonalityD)){resultcp = 0.0;}
else {resultcp = 1.0;}
if (UniversityF.equals(UniversityD)){resultcu = 0.0;}
else {resultcu = 1.0;}
if (STEMF.equals(STEMD)){resultcS = 0.0;}
else {resultcS = 1.0;}
if (HobbyF.equals(HobbyD)){resultch = 0.0;}
else {resultch = 1.0;}


//Calculate Euclidean Distance
dist = Math.sqrt(Math.pow(resultcg+resultcs+resultch+resultcp+resultcu+resultcS, 2));
```

Figure 4.11 Coding for Compare Two Type of Data and Calculate Euclidean Distance

```
public class sortdistance implements Comparator<distanceresult> {

    @Override
    public int compare(distanceresult o1, distanceresult o2) {
        return o1.distance < o2.distance ? -1 : o1.distance == o2.distance ? 0 : 1;
    }
```

Figure 4.12 Coding for Sort the Euclidean Distance

```
int k=1; // 1 of nearest neighbour
String[] ss;
ss = new String[k];
```

Figure 4.13 Coding for List Down the Nearest Neighbour

4.3 Conclusion

The implementation of the project has been conducted successfully. The implementation of the project is divided into two parts which are training process and testing process. In training process, analysis and determination of K value for the dataset had been done. Based on the result discussed, the K value chosen was one and used for testing process. The reason of K value chosen was one because of its accuracy and low error. In testing process, data had been tested by a recommendation system based on the K value chosen from training process.

CHAPTER 5

CONCLUSION

5.1 Introduction

This research was about the implementation of k nearest neighbour in university course recommendation system. The objective of the research was to study the current algorithm and technique in recommendation systems for selecting university courses, to implement k-Nearest Neighbour in the recommendation system, and to evaluate the application of recommendation system for selecting university course. The expected output of this research was to determine the suitable k value for recommendation system. There were six techniques and four systems discussed in Chapter 2. Each technique was explained, and a summary was made in the end of the chapter. K nearest neighbour was used in this research because of most of the researches are using k-Nearest Neighbour to deploy their recommendation system in the past 10 years (Park, Kim, Choi, & Kim, 2012). K-NN is one of the simplest method for solving classification problems. The dataset that used in this research is collected from a survey and filled in by public university students. There are total of 73 respondents' data has been collected in this survey as shown in appendix B. These data were divided into two groups for training and testing. The result was analysed in Chapter 4. Based on the experiment, the K value chosen is one.

5.2 Achievement of Objective

Based on the result presented in Chapter 4, the objectives have been achieve. The description of each achievement is as follow.

i) To study the current algorithm and technique in recommendation systems for selecting university courses.

Some of the current algorithm and technique in recommendation systems for selecting university courses were discussed in Chapter 2. The algorithm and technique are Bayesian Classifiers, Artificial Neural Network, Decision Tree, Item-Based Collaborative Filtering, Pearson Correlation Coefficient and K-Nearest Neighbour. On the other hand, current recommendation systems also being studied in Chapter 2 which are iMASCU, Facebook, Amazon.com and Genius.

ii) To implement k-Nearest Neighbour in the recommendation system.

Based on the result discussed in Chapter 4, the K value chosen was one and used for testing process. The result of training and testing model that discussed in Chapter 4 show that k value is one is most suitable for university course selection, and k-Nearest Neighbour can be implemented in the recommendation of Malaysia public university course selection.

iii) To evaluate the application of recommendation system for selecting university course.

Based on the result discussed in Chapter 4, the K value chosen was one and used for testing process. The reason of K value chosen was one because of its accuracy and low error rate. In testing process, result.

5.3 Research Constraint

There are some constraints in this research.

i) Type of data used.

The dataset that used in this research is collected from a survey and filled in by public university students and stored in excel file. WEKA import the file with the type of *.arff* file type. WEKA cannot import data type other than *.arff* file so that type of data is limited, and if another data need to add into the dataset. The file need to convert to *.arff* file again.

ii) Size of the dataset is small.

The amount of the data that collected from survey is very less. So, it makes the dataset small. The dataset contains of too much classes, with a small amount of data, some evaluations may not very accurate.

5.4 Future Work

There are several enhancements that can be carried out for future improvement for this project.

i) Use another source of dataset.

Dataset other than collected from survey form can be used in the research to let the data more fit to the algorithm. For example, can request data from the Bahagian Pengurusan Kemasukan Pelajar Jabatan Pendidikan Tinggi, Kementerian Pendidikan Malaysia.

ii) Use larger dataset.

Dataset other than collected from survey form can be used in the research to get more data. So that, the result and evaluation of recommendation system is more accurate.

REFERENCES

Adeniyi, D. A., Wei, Z., & Yongquan, Y. (2016). Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. *Applied Computing and Informatics*, *12*(1), 90–108. https://doi.org/10.1016/j.aci.2014.10.001

Al-Badarenah, A., & Alsakran, J. (2016). An Automated Recommender System for Course Selection. *International Journal of Advanced Computer Science and Applications*, *7*(3), 166–175. https://doi.org/10.14569/IJACSA.2016.070323

Baker, F. B. (2001). *The Basics of Item Response Theory. Second Edition.* (Second Edi). Washington: ERIC Clearinghouse on Assessment and Evalution, College Park, MD. Retrieved from https://files.eric.ed.gov/fulltext/ED458219.pdf

Barrington, L., Oda, R., & Lanckriet, G. (2009). Smarter than genius? human evaluation of music recommender systems. *... Symposium on Music ...*, (Ismir), 357–362. https://doi.org/10.1007/978-3-540-72079-9_9

Berry, M. J. A., & Linoff, G. S. (1999). *Mastering Data Mining - The Art and Science of Customer Relationship Management*.

Bhumichitr, K., Channarukul, S., Saejiem, N., Jiamthapthaksin, R., & Nongpong, K. (2017). Recommender Systems for university elective course recommendation. *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 1–5. https://doi.org/10.1109/JCSSE.2017.8025933

Bogers, T., & Van Den Bosch, A. (2009). *Collaborative and content-based filtering for item recommendation on social bookmarking websites*. *CEUR Workshop Proceedings* (Vol. 532). https://doi.org/10.1007/978-0-387-85820-3

Facebook, I. (2015). Recommending items to more than a billion people | Engineering Blog | Facebook Code. Retrieved from https://code.facebook.com/posts/861999383875667/recommending-items-to-more-than-a-billion-people/

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. *San Francisco, CA, itd: Morgan Kaufmann*. https://doi.org/10.1016/B978-0-12-381479-1.00001-0

Hsu, C. K., Hwang, G. J., & Chang, C. K. (2010). Development of a reading material recommendation system based on a knowledge engineering approach. *Computers and Education*, *55*(1), 76–83. https://doi.org/10.1016/j.compedu.2009.12.004

Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, *16*(3), 261–273. https://doi.org/10.1016/j.eij.2015.06.005

Jain, S., Grover, A., Thakur, P. S., & Choudhary, S. K. (2015). Trends, problems and solutions of recommender system. *International Conference on Computing, Communication and Automation, ICCCA 2015*, 955–958. https://doi.org/10.1109/CCAA.2015.7148534

Lane, D. (2013). Values of the Pearson Correlation. *Online Statistics Education: An Interactive Multimedia Course of Study*. Retrieved from http://onlinestatbook.com/2/introduction/variables.html

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, *7*(1), 76–80. https://doi.org/10.1109/MIC.2003.1167344

Park, D. H., Kim, H. K., Choi, I. Y., & Kim, J. K. (2012). A literature review and classification of recommender systems research. *Expert Systems with Applications*, *39*(11), 10059–10072. https://doi.org/10.1016/j.eswa.2012.02.038

Peterson, L. E. (2009). K-nearest neighbor - Scholarpedia. https://doi.org/doi:10.4249/scholarpedia.1883

Poole, D. L., & Mackworth, A. K. (2010). Artificial Intelligence - Foundations of Computational Agents. https://doi.org/10.1016/j.artint.2010.12.004

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-Based Collaborative Filtering Recommendation, (JANUARY). https://doi.org/10.1145/371920.372071

Shahar, F. M. (2018). Business | New Straits Times | Malaysia General Business Sports and Lifestyle News. Retrieved from https://www.nst.com.my/business

Srivastava, T. (2018). Introduction to KNN, K-Nearest Neighbors _ Simplified. Retrieved December 4, 2018, from https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/

Szkuta, B., & Sanabria, L. (1999). Electricity price short-term forecasting using artificial neural networks. *IEEE Transactions on Power*, *14*(3). Retrieved from http://ieeexplore.ieee.org/abstract/document/780895/

Tan, H., & Guo, J. (2008). E-Learning Recommendation System, 430–433. https://doi.org/10.1109/CSSE.2008.305

Wang, Z., Yu, X., Feng, N., & Wang, Z. (2014). An improved collaborative movie recommendation system using computational intelligence. *Journal of Visual Languages and Computing*, *25*(6), 667–675. https://doi.org/10.1016/j.jvlc.2014.09.011

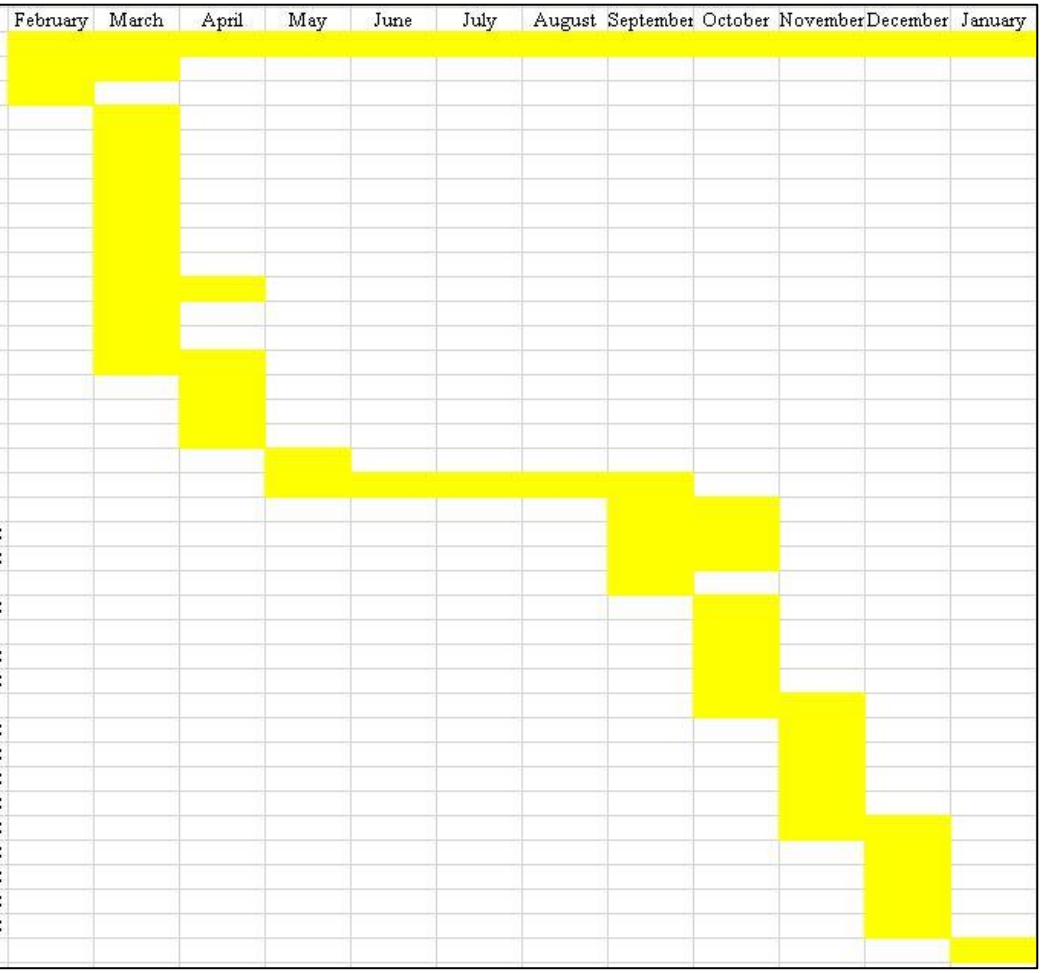
Yoon Ho Choa Soung Hie Kim, J. K. K. (2002). A personalized recommender system based on

web usage mining and decision tree induction. *Expert Systems with Applications*, *23*, 329–342. https://doi.org/10.1016/S0957-4174(02)00052-0

Zhang, L., Bu, Z., Wu, Z., Cao, J., Wu, D., Lu, J., … Biradar, S. (2015). A Personalized E-Learning Based on Recommender System. *Other Conferences*, *1*(2), 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

APPENDIX A
GANTT CHART

Task Name	Duration	Start	Finish
PSM Planning	325d	15/2/2018	5/1/2019
Introduction	24d	15/2/2018	10/3/2018
Identify the requirement, problem and objective	5d	15/2/2018	19/2/2018
Submission Chapter 1	1d	8/3/2018	8/3/2018
Correction on Chapter 1	3d	8/3/2018	10/3/2018
Literature Review	14d	11/3/2018	24/3/2018
Review Existing Technique	14d	11/3/2018	24/3/2018
Review Existing System	14d	11/3/2018	24/3/2018
Submission Chapter 2	1d	15/3/2018	15/3/2018
Correction on Chapter 2	3d	15/3/2018	17/3/2018
Methodology	31d	25/3/2018	24/4/2018
Design Flowchart	1d	25/3/2018	25/3/2018
Design Framework	1d	26/3/2018	26/3/2018
Gather information about the technique applicati	10d	27/3/2018	5/4/2018
Implement technique	5d	6/4/2018	10/4/2018
Submission of Chapter 3	1d	12/4/2018	12/4/2018
Correction on Chapter 1, 2 and 3	13d	12/4/2018	24/4/2018
Submission PSM 1	1d	4/5/2018	4/5/2018
Presentation PSM 1	1d	15/5/2018	15/5/2018
Study on PSM2 Implementation	120d	16/5/2018	12/9/2018
Result and Discussion	30d	13/9/2018	11/10/2018
System coding	30d	13/9/2018	11/10/2018
Collect data from survey form	16d	15/9/2018	30/9/2018
Develp model for training and testing	11d	1/10/2018	11/10/2018
Training for K value	6d	2/10/2018	7/10/2018
Testing K value	5d	7/10/2018	11/10/2018
Submission Chapter 4	1d	12/10/2018	12/10/2018
Correction on Chapter 4	25d	13/10/2018	6/11/2018
Conclusion	10d	7/11/2018	16/11/2018
Research constraint	5d	7/11/2018	11/11/2018
Future Work	4d	12/11/2018	15/11/2018
Submission on Chapter 5	1d	16/11/2018	16/11/2018
Correction on Chapter 4 and 5	25d	17/11/2018	11/12/2018
Submission PSM 2	1d	12/12/2018	12/12/2018
Presentation PSM 2	1d	18/12/2018	18/12/2018
Refine PSM documents	7d	20/12/2018	28/12/2018
Complete research thesis	1d	29/12/2018	29/12/2018
Submit thesis to faculty	1d	5/1/2019	5/1/2019



APPENDIX B
RAW DATA

Gender	Which subject you liked the most in secondary school/Matriculation college?	What is your hobby?	Which type of personality best suits you?	Which of the following university is the first choice for continuing your undergraduate study ?	You have chosen Universiti Kebangsaan Malaysia (UKM). Please choose your first-choice undergraduate program.
Female	ECONOMIC	SPORT	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Islam Antarabangsa Malaysia (UIAM)	ECONOMICS - SARJANA MUDA EKONOMI (KEPUJIAN)
Female	add math	badminton	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Kebangsaan Malaysia (UKM)	STATISTICS - IJAZAH SARJANAMUDA SAINS DENGAN KEPUJIAN (SAINS AKTUARI)
Female	Biology	Read novel	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Kebangsaan Malaysia (UKM)	APPLIED SCIENCE - IJAZAH SARJANAMUDA SAINS FORENSIK DENGAN KEPUJIAN
Female	Chemistry	watch Tv	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Kebangsaan Malaysia (UKM)	DENTAL STUDIES - DOKTOR PERGIGIAN
Female	Chinese	Listening to music	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Kebangsaan Malaysia (UKM)	APPLIED SCIENCE - IJAZAH SARJANAMUDA SAINS DENGAN KEPUJIAN (BIOTEKNOLOGI DENGAN PENGURUSAN)

Female	Maths	Play music instrument	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Kebangsaan Malaysia (UKM)	COMPUTER SCIENCE - IJAZAH SARJANAMUDA KEJURUTERAAN PERISIAN DENGAN KEPUJIAN (PEMBANGUNAN SISTEM MULTIMEDIA)
Female	Science Computer	Cooking	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Kebangsaan Malaysia (UKM)	COMPUTER SCIENCE - IJAZAH SARJANAMUDA SAINS KOMPUTER DENGAN KEPUJIAN
Male	Biology	Listen to music	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Kebangsaan Malaysia (UKM)	MEDICAL SERVICES - IJAZAH SARJANAMUDA SAINS BIOPERUBATAN DENGAN KEPUJIAN
Male	Kemahiran Hidup	Play Game	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Kebangsaan Malaysia (UKM)	PHARMACY - IJAZAH SARJANAMUDA FARMASI DENGAN KEPUJIAN
Female	Account	Basketball	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Malaya (UM)	ACCOUNTING AND TAXATION - SARJANA MUDA PERAKAUNAN
Female	Biology	Discover new thing	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Malaya (UM)	LAW - SARJANA MUDA UNDANG-UNDANG
Female	Mathematics	Reading	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Malaya (UM)	CIVIL ENGINEERING - SARJANA MUDA KEJURUTERAAN AWAM
Female	Mathematics T	Sleep	Investigative (Good at understanding and solving science and	Universiti Malaya (UM)	COMPUTER SCIENCE - SARJANA MUDA SAINS

			math problems. Sees self as precise, scientific, and intellectual.)		KOMPUTER (SISTEM MAKLUMAT)
Male	Accounting	Chating	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Malaya (UM)	ACCOUNTING AND TAXATION - SARJANA MUDA PERAKAUNAN
Male	Add maths	Badminton	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Malaya (UM)	MEDICINE - SARJANA MUDA PERUBATAN DAN SARJANA MUDA PEMBEDAHAN
Male	Business	Swimming	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Malaya (UM)	FINANCE, BANKING, INSURANCE - SARJANA MUDA KEWANGAN
Male	Chemistry	Playing video games	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Malaya (UM)	DENTAL STUDIES - SARJANA MUDA PEMBEDAHAN PERGIGIAN
Male	Chemistry	esports	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Malaya (UM)	COMPUTER SCIENCE - SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)
Female	BAHASA MELAYU	WATCH DRAMA	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Malaysia Kelantan (UMK)	BUSSINES AND ADMINISTRATION (BROAD PROGRAMMES) - IJAZAH SARJANA MUDA PENTADBIRAN PERNIAGAAN (PERBANKAN DAN KEWANGAN ISLAM) DENGAN KEPUJIAN

Female	Add Math	Eat	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (TEKNOLOGI GRAFIK & MULTIMEDIA) DENGAN KEPUJIAN
Female	Bahasa Melayu	Badminton	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Malaysia Pahang (UMP)	CIVIL ENGINEERING - IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM
Female	Chinese	eating	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN
Female	Math	Play badminton	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN
Female	Mathematic	Listening to music	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	Mathematics	Reading	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Malaysia Pahang (UMP)	CHEMICAL AND PROCESS - IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN KIMIA
Female	Mathematics	Watch movie	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Malaysia Pahang (UMP)	APPLIED SCIENCE - IJAZAH SARJANA MUDA

					SAINS GUNAAN (KEPUJIAN) KIMIA INDUSTRI
Female	Physics	Playing games	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	Pjk	Play	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Malaysia Pahang (UMP)	CIVIL ENGINEERING - IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM
Male	Accounting	Dance, play drum	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	Biology	Browsing the web	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	Chemistry	Sport	Enterprising (Good at leading people and selling things or ideas. Sees self as energetic, ambitious, and sociable.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	IT	Reading	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN

Male	Math	Badminton	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Malaysia Pahang (UMP)	CIVIL ENGINEERING - IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM
Male	Math	playing game	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	mathematics	watching movie and programming	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Malaysia Pahang (UMP)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	PHYSIC	SPORT	Enterprising (Good at leading people and selling things or ideas. Sees self as energetic, ambitious, and sociable.)	Universiti Malaysia Perlis (UniMAP)	ELECTRICITY AND ENERGY - IJAZAH SARJANA MUDA KEJURUTERAAN ELEKTRIK DENGAN KEPUJIAN
Female	mathematics	listening music	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Malaysia Sabah (UMS)	CIVIL ENGINEERING - IJAZAH SARJANA MUDA KEJURUTERAAN DENGAN KEPUJIAN (KEJURUTERAAN AWAM)
Female	Science	Reading novel	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Malaysia Sarawak (UNIMAS)	PHYSICAL SCIENCE (BROAD PROGRAMMES) - IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (SAINS KOGNITIF)
Male	Mathematics	Reading novel	Investigative (Good at understanding and solving science and	Universiti Malaysia Sarawak (UNIMAS)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER DENGAN

			math problems. Sees self as precise, scientific, and intellectual.)		KEPUJIAN (SAINS KOMPUTAN)
Male	MUET	Playing badminton	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Malaysia Sarawak (UNIMAS)	PHYSICAL SCIENCE (BROAD PROGRAMMES) - IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (SAINS KOGNITIF)
Male	SCIENCE	READIN G	Enterprising (Good at leading people and selling things or ideas. Sees self as energetic, ambitious, and sociable.)	Universiti Malaysia Terengganu (UMT)	BIOLOGY AND BIOCHEMISTRY - SARJANA MUDA SAINS (SAINS BIOLOGI)
Female	Bahasa Cina	Spend time with family	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Pendidikan Sultan Idris (UPSI)	TRAINING FOR PRE-SCHOOL TEACHERS - IJAZAH SARJANA MUDA PENDIDIKAN (PENDIDIKAN AWAL KANAK-KANAK) DENGAN KEPUJIAN
Male	MATHEMAT IC	COMPU TER GAME	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Pertahanan Nasional Malaysia (UPNM)	CIVIL ENGINEERING - IJAZAH SARJANA MUDA KEJURUTERAAN AWAM
Female	Account	Game	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Putra Malaysia (UPM)	ACCOUNTING AND TAXATION - BACELOR PERAKAUNAN
Female	Biology	Dancing	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Putra Malaysia (UPM)	FOOD PROCESSING - BACELOR SAINS DAN TEKNOLOGI MAKANAN
Female	biology	listening to music	Conventional (Good at working with written records and numbers in a	Universiti Putra Malaysia (UPM)	VETERINARY - DOKTOR PERUBATAN VETERINAR

			systematic, orderly way. Sees self as orderly, and good at following a set plan.)		
Male	Chemistry	Reading	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Putra Malaysia (UPM)	FOOD PROCESSING - BACELOR SAINS (PENGURUSAN PERKHIDMATAN MAKANAN)
Male	Chinese	Play game	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Putra Malaysia (UPM)	COMPUTER SCIENCE - BACELOR SAINS KOMPUTER (SISTEM KOMPUTER)
Male	BIOLOGY	SPORT	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Sains Islam Malaysia (USIM)	APPLIED SCIENCE - SARJANA MUDA SAINS DENGAN KEPUJIAN (BIOTEKNOLOGI MAKANAN)
Female	chemistry	play computer game	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Sains Malaysia (USM)	CHEMISTRY - IJAZAH SARJANA MUDA SAINS (KEPUJIAN) (KIMIA)
Female	Mathematics	Listening music	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Sains Malaysia (USM)	NURSING AND CARING - IJAZAH SARJANA MUDA SAINS KESIHATAN (KEPUJIAN) (KEJURURAWATAN)
Female	Mathematics	Dancing	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Sains Malaysia (USM)	ARCHITECTURE AND TOWN PLANNING - IJAZAH SARJANA MUDA SAINS (PERUMAHAN, BANGUNAN DAN PERANCANGAN) (KEPUJIAN) (SENIBINA)
Female	Mathematics	Listen to music	Investigative (Good at understanding and solving science and	Universiti Sains Malaysia (USM)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA

			math problems. Sees self as precise, scientific, and intellectual.)		SAINS KOMPUTER (KEPUJIAN)
Female	Mathematics (T)	swimming, badminton	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Sains Malaysia (USM)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEPUJIAN)
Female	PHYSIC	WATCHING MOVIE	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Sains Malaysia (USM)	CIVIL ENGINEERING - IJAZAH SARJANA MUDA KEJURUTERAAN (KEPUJIAN) (KEJURUTERAAN AWAM)
Male	Chemistry	Gaming	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Sains Malaysia (USM)	CHEMICAL AND PROCESS - IJAZAH SARJANA MUDA KEJURUTERAAN (KEPUJIAN) (KEJURUTERAAN KIMIA)
Female	ACCOUNT	READING	Enterprising (Good at leading people and selling things or ideas. Sees self as energetic, ambitious, and sociable.)	Universiti Sultan Zainal Abidin (UniSZA)	FINANCE, BANKING, INSURANCE - IJAZAH SARJANA MUDA PERAKAUNAN DENGAN KEPUJIAN
Male	Physical Education	Singing	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Teknikal Malaysia Melaka (UTeM)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (KEPINTARAN BUATAN) DENGAN KEPUJIAN
Male	physics	computer games	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Teknikal Malaysia Melaka (UTeM)	COMPUTER SCIENCE - IJAZAH SARJANA MUDA SAINS KOMPUTER (MEDIA INTERAKTIF) DENGAN KEPUJIAN

Female	Chemistry	Swimming	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Teknologi Malaysia (UTM)	CHEMICAL AND PROCESS - SARJANA MUDA KEJURUTERAAN (KIMIA)
Female	Mathematics	Watching videos	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Teknologi Malaysia (UTM)	CHEMICAL AND PROCESS - SARJANA MUDA KEJURUTERAAN PROSES KIMIA
Female	Maths	Music, games n movies	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Teknologi Malaysia (UTM)	MOTOR VEHICLE, SHIPS AND AIRCRAFT - SARJANA MUDA KEJURUTERAAN (MEKANIKAL - AUTOMOTIF)
Female	Physics	Sports	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Teknologi Malaysia (UTM)	CHEMICAL AND PROCESS - SARJANA MUDA KEJURUTERAAN (KIMIA - GAS)
Male	English	Reading	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Teknologi Malaysia (UTM)	COMPUTER SCIENCE - SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)
Male	Mathematics	Sleep	Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)	Universiti Teknologi Malaysia (UTM)	COMPUTER SCIENCE - SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)
Male	Maths	Badminton	Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)	Universiti Teknologi Malaysia (UTM)	MOTOR VEHICLE, SHIPS AND AIRCRAFT - SARJANA MUDA KEJURUTERAAN (MEKANIKAL - AUTOMOTIF)
Male	physics	watching drama	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as	Universiti Teknologi Malaysia (UTM)	COMPUTER SCIENCE - SARJANA MUDA SAINS KOMPUTER (BIOINFORMATIK)

			orderly, and good at following a set plan.)		
Female	CHEMISTRY	LISTEN MUSIC	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Teknologi MARA (UiTM)	APPLIED SCIENCE - SARJANA MUDA SAINS (KEPUJIAN) KIMIA GUNAAN
Female	Science	Singing, dancing, play mobile games	Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)	Universiti Tun Hussein Onn Malaysia (UTHM)	COMPUTER SCIENCE - SARJANA MUDA SAINS KOMPUTER (PENGKOMPUTERAN MULTIMEDIA) DENGAN KEPUJIAN
Male	Maths	Sports	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Tun Hussein Onn Malaysia (UTHM)	COMPUTER SCIENCE - SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	Economy	Listen music	Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)	Universiti Utara Malaysia (UUM)	TRAINING FOR TEACHERS WITH SUBJECT SPECIALISATION - SARJANA MUDA PENDIDIKAN DENGAN KEPUJIAN (BIMBINGAN DAN KAUNSELING)
Female	Mathematics	watching drama	Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)	Universiti Utara Malaysia (UUM)	ECONOMICS - SARJANA MUDA SAINS EKONOMI DENGAN KEPUJIAN
Female	Mathematics	Playing Basketball	Enterprising (Good at leading people and selling things or ideas. Sees self as energetic, ambitious, and sociable.)	Universiti Utara Malaysia (UUM)	AUDIO-VISUAL TECHNIQUES AND MEDIA PRODUCTION - IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (MULTIMEDIA)

APPENDIX C
DATA AFTER PREPROCESSING

GENDER	SUBJECT	PERSONALITY	HOBBY	UNIVERSITY	FIELD	COURSE
Female	CHINESE LANGUAGE	Social	OTHERS	Universiti Pendidikan Sultan Idris (UPSI)	NON STEM	AT19 IJAZAH SARJANA MUDA PENDIDIKAN (PENDIDIKAN AWAL KANAK-KANAK) DENGAN KEPUJIAN
Male	MATHEMATICS	Conventional	SPORT	Universiti Tun Hussein Onn Malaysia (UTHM)	STEM	BC10 SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	SCIENCE	Artistic	SINGING	Universiti Tun Hussein Onn Malaysia (UTHM)	STEM	BC24 SARJANA MUDA SAINS KOMPUTER (PENGKOMPUTERAN MULTIMEDIA) DENGAN KEPUJIAN
Male	PHYSICS	Realistic	COMPUTER GAME	Universiti Teknikal Malaysia Melaka (UTeM)	STEM	CC13 IJAZAH SARJANA MUDA SAINS KOMPUTER (MEDIA INTERAKTIF) DENGAN KEPUJIAN
Male	OTHERS	Artistic	SINGING	Universiti Teknikal Malaysia Melaka (UTeM)	STEM	CC25 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEPINTARAN BUATAN) DENGAN KEPUJIAN
Female	ACCOUNTING	Enterprising	READING	Universiti Sultan Zainal Abidin (UniSZA)	NON STEM	DE02 IJAZAH SARJANA MUDA PERAKAUNAN DENGAN KEPUJIAN
Female	CHEMISTRY	Social	LISTEN TO MUSIC	Universiti Teknologi MARA (UiTM)	STEM	ES59 SARJANA MUDA SAINS (KEPUJIAN) KIMIA GUNAAN

Male	SCIENCE	Enterprising	READING	Universiti Malaysia Terengganu (UMT)	STEM	GS02 SARJANA MUDA SAINS (SAINS BIOLOGI)
Female	MATHEMATICS	Realistic	LISTEN TO MUSIC	Universiti Malaysia Sabah (UMS)	STEM	HK01 IJAZAH SARJANA MUDA KEJURUTERAAN DENGAN KEPUJIAN (KEJURUTERAAN AWAM)
Female	PHYSICS	Artistic	COMPUTER GAME	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	ACCOUNTING	Artistic	DANCE	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	MATHEMATICS	Conventional	COMPUTER GAME	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	BIOLOGY	Investigative	INTERNET	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	MATHEMATICS	Social	LISTEN TO MUSIC	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	CHEMISTRY	Enterprising	SPORT	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN

Male	MATHEMATICS	Investigative	WATCH DRAMA, MOVIE	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	MATHEMATICS	Artistic	SPORT	Universiti Malaysia Pahang (UMP)	STEM	JC11 IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN
Female	CHINESE LANGUAGE	Conventional	EAT	Universiti Malaysia Pahang (UMP)	STEM	JC11 IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN
Male	INFORMATION TECHNOLOGY	Social	READING	Universiti Malaysia Pahang (UMP)	STEM	JC11 IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN
Female	ADDITIONAL MATHEMATICSS	Social	EAT	Universiti Malaysia Pahang (UMP)	STEM	JC24 IJAZAH SARJANA MUDA SAINS KOMPUTER (TEKNOLOGI GRAFIK & MULTIMEDIA) DENGAN KEPUJIAN
Female	MATHEMATICS	Social	WATCH DRAMA, MOVIE	Universiti Malaysia Pahang (UMP)	STEM	JG04 IJAZAH SARJANA MUDA SAINS GUNAAN (KEPUJIAN) KIMIA INDUSTRI
Female	OTHERS	Artistic	COMPUTER GAME	Universiti Malaysia Pahang (UMP)	STEM	JK01 IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM
Male	MATHEMATICS	Social	SPORT	Universiti Malaysia Pahang (UMP)	STEM	JK01 IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM
Female	MALAY LANGUAGE	Investigative	SPORT	Universiti Malaysia Pahang (UMP)	STEM	JK01 IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM

Female	MATHEMATICS	Investigative	READING	Universiti Malaysia Pahang (UMP)	STEM	JK03 IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN KIMIA
Female	INFORMATION TECHNOLOGY	Social	COOKING	Universiti Kebangsaan Malaysia (UKM)	STEM	KC00 IJAZAH SARJANAMUDA SAINS KOMPUTER DENGAN KEPUJIAN
Female	MATHEMATICS	Artistic	PLAY MUSIC INSTRUMENT	Universiti Kebangsaan Malaysia (UKM)	STEM	KC09 IJAZAH SARJANAMUDA KEJURUTERAAN PERISIAN DENGAN KEPUJIAN (PEMBANGUNAN SISTEM MULTIMEDIA)
Female	CHEMISTRY	Realistic	WATCH DRAMA, MOVIE	Universiti Kebangsaan Malaysia (UKM)	STEM	KD00 DOKTOR PERGIGIAN
Male	OTHERS	Realistic	COMPUTER GAME	Universiti Kebangsaan Malaysia (UKM)	STEM	KF00 IJAZAH SARJANAMUDA FARMASI DENGAN KEPUJIAN
Male	BIOLOGY	Investigative	LISTEN TO MUSIC	Universiti Kebangsaan Malaysia (UKM)	STEM	KM03 IJAZAH SARJANAMUDA SAINS BIOPERUBATAN DENGAN KEPUJIAN
Female	BIOLOGY	Realistic	READING	Universiti Kebangsaan Malaysia (UKM)	STEM	KM11 IJAZAH SARJANAMUDA SAINS FORENSIK DENGAN KEPUJIAN
Female	ADDITIONAL MATHEMATICSS	Investigative	SPORT	Universiti Kebangsaan Malaysia (UKM)	STEM	KS10 IJAZAH SARJANAMUDA SAINS DENGAN KEPUJIAN (SAINS AKTUARI)
Female	CHINESE LANGUAGE	Realistic	LISTEN TO MUSIC	Universiti Kebangsaan Malaysia (UKM)	STEM	KS56 IJAZAH SARJANAMUDA SAINS DENGAN KEPUJIAN (BIOTEKNOLOGI DENGAN PENGURUSAN)

Female	MALAY LANGUAGE	Investigative	WATCH DRAMA, MOVIE	Universiti Malaysia Kelantan (UMK)	NON STEM	LE23 IJAZAH SARJANA MUDA PENTADBIRAN PERNIAGAAN (PERBANKAN DAN KEWANGAN ISLAM) DENGAN KEPUJIAN
Male	CHEMISTRY	Social	COMPUTER GAME	Universiti Malaya (UM)	STEM	MC04 SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)
Female	ADDITIONAL MATHEMATICSS	Investigative	SLEEP	Universiti Malaya (UM)	STEM	MC06 SARJANA MUDA SAINS KOMPUTER (SISTEM MAKLUMAT)
Male	CHEMISTRY	Conventional	COMPUTER GAME	Universiti Malaya (UM)	STEM	MD01 SARJANA MUDA PEMBEDAHAN PERGIGIAN
Male	ACCOUNTING	Realistic	CHATING	Universiti Malaya (UM)	NON STEM	ME01 SARJANA MUDA PERAKAUNAN
Female	ACCOUNTING	Realistic	SPORT	Universiti Malaya (UM)	NON STEM	ME01 SARJANA MUDA PERAKAUNAN
Male	BUSINESS STUDIES	Investigative	SPORT	Universiti Malaya (UM)	NON STEM	ME07 SARJANA MUDA KEWANGAN
Female	MATHEMATICS	Investigative	READING	Universiti Malaya (UM)	STEM	MK09 SARJANA MUDA KEJURUTERAAN AWAM
Female	BIOLOGY	Conventional	OTHERS	Universiti Malaya (UM)	NON STEM	ML01 SARJANA MUDA UNDANG-UNDANG
Male	ADDITIONAL MATHEMATICSS	Investigative	SPORT	Universiti Malaya (UM)	STEM	MM01 SARJANA MUDA PERUBATAN DAN SARJANA MUDA PEMBEDAHAN
Male	CHINESE LANGUAGE	Social	COMPUTER GAME	Universiti Putra Malaysia (UPM)	STEM	PC04 BACELOR SAINS KOMPUTER (SISTEM KOMPUTER)
Female	ACCOUNTING	Conventional	COMPUTER GAME	Universiti Putra Malaysia (UPM)	NON STEM	PE02 BACELOR PERAKAUNAN

Female	BIOLOGY	Realistic	DANCE	Universiti Putra Malaysia (UPM)	STEM	PG09 BACELOR SAINS DAN TEKNOLOGI MAKANAN
Male	CHEMISTRY	Social	READING	Universiti Putra Malaysia (UPM)	STEM	PG18 BACELOR SAINS (PENGURUSAN PERKHIDMATAN MAKANAN)
Female	BIOLOGY	Conventional	LISTEN TO MUSIC	Universiti Putra Malaysia (UPM)	STEM	PM01 DOKTOR PERUBATAN VETERINAR
Male	BIOLOGY	Realistic	SPORT	Universiti Sains Islam Malaysia (USIM)	STEM	QG07 SARJANA MUDA SAINS DENGAN KEPUJIAN (BIOTEKNOLOGI MAKANAN)
Male	PHYSICS	Enterprising	SPORT	Universiti Malaysia Perlis (UniMAP)	STEM	RK23 IJAZAH SARJANA MUDA KEJURUTERAAN ELEKTRIK DENGAN KEPUJIAN
Female	MATHEMATICS	Investigative	LISTEN TO MUSIC	Universiti Sains Malaysia (USM)	STEM	SC00 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEPUJIAN)
Female	ADDITIONAL MATHEMATICSS	Realistic	SPORT	Universiti Sains Malaysia (USM)	STEM	SC00 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEPUJIAN)
Female	MATHEMATICS	Artistic	DANCE	Universiti Sains Malaysia (USM)	STEM	SH03 IJAZAH SARJANA MUDA SAINS (PERUMAHAN, BANGUNAN DAN PERANCANGAN) (KEPUJIAN) (SENIBINA)
Female	PHYSICS	Investigative	WATCH DRAMA, MOVIE	Universiti Sains Malaysia (USM)	STEM	SK01 IJAZAH SARJANA MUDA KEJURUTERAAN (KEPUJIAN) (KEJURUTERAAN AWAM)
Male	CHEMISTRY	Investigative	COMPUTER GAME	Universiti Sains Malaysia (USM)	STEM	SK03 IJAZAH SARJANA MUDA KEJURUTERAAN (KEPUJIAN) (KEJURUTERAAN KIMIA)

Female	MATHEMATICS	Social	LISTEN TO MUSIC	Universiti Sains Malaysia (USM)	STEM	SM12 IJAZAH SARJANA MUDA SAINS KESIHATAN (KEPUJIAN) (KEJURURAWATAN)
Female	CHEMISTRY	Investigative	COMPUTER GAME	Universiti Sains Malaysia (USM)	STEM	SS28 IJAZAH SARJANA MUDA SAINS (KEPUJIAN) (KIMIA)
Male	MATHEMATICS	Investigative	SLEEP	Universiti Teknologi Malaysia (UTM)	STEM	TC10 SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)
Male	ENGLISH LANGUAGE	Realistic	READING	Universiti Teknologi Malaysia (UTM)	STEM	TC10 SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)
Male	PHYSICS	Conventional	WATCH DRAMA, MOVIE	Universiti Teknologi Malaysia (UTM)	STEM	TC24 SARJANA MUDA SAINS KOMPUTER (BIOINFORMATIK)
Female	CHEMISTRY	Artistic	SPORT	Universiti Teknologi Malaysia (UTM)	STEM	TK03 SARJANA MUDA KEJURUTERAAN (KIMIA)
Female	PHYSICS	Investigative	SPORT	Universiti Teknologi Malaysia (UTM)	STEM	TK30 SARJANA MUDA KEJURUTERAAN (KIMIA - GAS)
Male	MATHEMATICS	Realistic	SPORT	Universiti Teknologi Malaysia (UTM)	STEM	TK40 SARJANA MUDA KEJURUTERAAN (MEKANIKAL - AUTOMOTIF)
Female	MATHEMATICS	Realistic	LISTEN TO MUSIC	Universiti Teknologi Malaysia (UTM)	STEM	TK40 SARJANA MUDA KEJURUTERAAN (MEKANIKAL - AUTOMOTIF)
Female	MATHEMATICS	Artistic	WATCH DRAMA, MOVIE	Universiti Teknologi Malaysia (UTM)	STEM	TK79 SARJANA MUDA KEJURUTERAAN PROSES KIMIA

Female	MATHEMATICS	Enterprising	SPORT	Universiti Utara Malaysia (UUM)	NON STEM	UC12 IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (MULTIMEDIA)
Female	MATHEMATICS	Social	WATCH DRAMA, MOVIE	Universiti Utara Malaysia (UUM)	NON STEM	UE00 SARJANA MUDA SAINS EKONOMI DENGAN KEPUJIAN
Female	ECONOMICS	Conventional	LISTEN TO MUSIC	Universiti Utara Malaysia (UUM)	NON STEM	UT04 SARJANA MUDA PENDIDIKAN DENGAN KEPUJIAN (BIMBINGAN DAN KAUNSELING)
Male	MATHEMATICS	Investigative	READING	Universiti Malaysia Sarawak (UNIMAS)	STEM	WC00 IJAZAH SARJANA MUDA SAINS KOMPUTER DENGAN KEPUJIAN (SAINS KOMPUTAN)
Male	ENGLISH LANGUAGE	Conventional	SPORT	Universiti Malaysia Sarawak (UNIMAS)	STEM	WS24 IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (SAINS KOGNITIF)
Female	SCIENCE	Artistic	READING	Universiti Malaysia Sarawak (UNIMAS)	STEM	WS24 IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (SAINS KOGNITIF)
Female	ECONOMICS	Artistic	SPORT	Universiti Islam Antarabangsa Malaysia (UIAM)	NON STEM	YE00 SARJANA MUDA EKONOMI (KEPUJIAN)
Male	MATHEMATICS	Conventional	COMPUTER GAME	Universiti Pertahanan Nasional Malaysia (UPNM)	STEM	ZK01 IJAZAH SARJANA MUDA KEJURUTERAAN AWAM

APPENDIX D
DATA FOR TRAINING PROCESS

GENDER	SUBJECT	PERSONALITY	HOBBY	UNIVERSITY	FIELD	COURSE
Female	ADDITIONAL MATHEMATICSS	Investigative	SPORT	Universiti Kebangsaan Malaysia (UKM)	STEM	KS10 IJAZAH SARJANAMUDA SAINS DENGAN KEPUJIAN (SAINS AKTUARI)
Male	OTHERS	Realistic	COMPUTER GAME	Universiti Kebangsaan Malaysia (UKM)	STEM	KF00 IJAZAH SARJANAMUDA FARMASI DENGAN KEPUJIAN
Male	BIOLOGY	Investigative	LISTEN TO MUSIC	Universiti Kebangsaan Malaysia (UKM)	STEM	KM03 IJAZAH SARJANAMUDA SAINS BIOPERUBATAN DENGAN KEPUJIAN
Female	CHINESE LANGUAGE	Realistic	LISTEN TO MUSIC	Universiti Kebangsaan Malaysia (UKM)	STEM	KS56 IJAZAH SARJANAMUDA SAINS DENGAN KEPUJIAN (BIOTEKNOLOGI DENGAN PENGURUSAN)
Female	BIOLOGY	Realistic	READING	Universiti Kebangsaan Malaysia (UKM)	STEM	KM11 IJAZAH SARJANAMUDA SAINS FORENSIK DENGAN KEPUJIAN
Male	ADDITIONAL MATHEMATICSS	Investigative	SPORT	Universiti Malaya (UM)	STEM	MM01 SARJANA MUDA PERUBATAN DAN SARJANA MUDA PEMBEDAHAN
Female	ACCOUNTING	Realistic	SPORT	Universiti Malaya (UM)	NON STEM	ME01 SARJANA MUDA PERAKAUNAN
Male	CHEMISTRY	Conventional	COMPUTER GAME	Universiti Malaya (UM)	STEM	MD01 SARJANA MUDA PEMBEDAHAN PERGIGIAN
Male	CHEMISTRY	Social	COMPUTER GAME	Universiti Malaya (UM)	STEM	MC04 SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)

Female	MATHEMATICS	Investigative	READING	Universiti Malaya (UM)	STEM	MK09 SARJANA MUDA KEJURUTERAAN AWAM
Male	BUSINESS STUDIES	Investigative	SPORT	Universiti Malaya (UM)	NON STEM	ME07 SARJANA MUDA KEWANGAN
Male	MATHEMATICS	Social	SPORT	Universiti Malaysia Pahang (UMP)	STEM	JK01 IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM
Female	MALAY LANGUAGE	Investigative	SPORT	Universiti Malaysia Pahang (UMP)	STEM	JK01 IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM
Male	MATHEMATICS	Conventional	COMPUTER GAME	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	CHINESE LANGUAGE	Conventional	EAT	Universiti Malaysia Pahang (UMP)	STEM	JC11 IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN
Female	ADDITIONAL MATHEMATICSS	Social	EAT	Universiti Malaysia Pahang (UMP)	STEM	JC24 IJAZAH SARJANA MUDA SAINS KOMPUTER (TEKNOLOGI GRAFIK & MULTIMEDIA) DENGAN KEPUJIAN
Male	BIOLOGY	Investigative	INTERNET	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	MATHEMATICS	Social	LISTEN TO MUSIC	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN

Female	MATHEMATICS	Investigative	READING	Universiti Malaysia Pahang (UMP)	STEM	JK03 IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN KIMIA
Male	CHEMISTRY	Enterprising	SPORT	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Male	MATHEMATICS	Investigative	WATCH DRAMA, MOVIE	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	MATHEMATICS	Social	WATCH DRAMA, MOVIE	Universiti Malaysia Pahang (UMP)	STEM	JG04 IJAZAH SARJANA MUDA SAINS GUNAAN (KEPUJIAN) KIMIA INDUSTRI
Male	PHYSICS	Enterprising	SPORT	Universiti Malaysia Perlis (UniMAP)	STEM	RK23 IJAZAH SARJANA MUDA KEJURUTERAAN ELEKTRIK DENGAN KEPUJIAN
Female	SCIENCE	Artistic	READING	Universiti Malaysia Sarawak (UNIMAS)	STEM	WS24 IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (SAINS KOGNITIF)
Male	MATHEMATICS	Investigative	READING	Universiti Malaysia Sarawak (UNIMAS)	STEM	WC00 IJAZAH SARJANA MUDA SAINS KOMPUTER DENGAN KEPUJIAN (SAINS KOMPUTAN)
Female	CHINESE LANGUAGE	Social	OTHERS	Universiti Pendidikan Sultan Idris (UPSI)	NON STEM	AT19 IJAZAH SARJANA MUDA PENDIDIKAN (PENDIDIKAN AWAL KANAK-KANAK) DENGAN KEPUJIAN
Female	ACCOUNTING	Conventional	COMPUTER GAME	Universiti Putra Malaysia (UPM)	NON STEM	PE02 BACELOR PERAKAUNAN

Female	BIOLOGY	Conventional	LISTEN TO MUSIC	Universiti Putra Malaysia (UPM)	STEM	PM01 DOKTOR PERUBATAN VETERINAR
Male	CHEMISTRY	Social	READING	Universiti Putra Malaysia (UPM)	STEM	PG18 BACELOR SAINS (PENGURUSAN PERKHIDMATAN MAKANAN)
Female	CHEMISTRY	Investigative	COMPUTER GAME	Universiti Sains Malaysia (USM)	STEM	SS28 IJAZAH SARJANA MUDA SAINS (KEPUJIAN) (KIMIA)
Male	CHEMISTRY	Investigative	COMPUTER GAME	Universiti Sains Malaysia (USM)	STEM	SK03 IJAZAH SARJANA MUDA KEJURUTERAAN (KEPUJIAN) (KEJURUTERAAN KIMIA)
Female	MATHEMATICS	Social	LISTEN TO MUSIC	Universiti Sains Malaysia (USM)	STEM	SM12 IJAZAH SARJANA MUDA SAINS KESIHATAN (KEPUJIAN) (KEJURURAWATAN)
Female	MATHEMATICS	Investigative	LISTEN TO MUSIC	Universiti Sains Malaysia (USM)	STEM	SC00 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEPUJIAN)
Female	ADDITIONAL MATHEMATICSS	Realistic	SPORT	Universiti Sains Malaysia (USM)	STEM	SC00 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEPUJIAN)
Male	PHYSICS	Realistic	COMPUTER GAME	Universiti Teknikal Malaysia Melaka (UTeM)	STEM	CC13 IJAZAH SARJANA MUDA SAINS KOMPUTER (MEDIA INTERAKTIF) DENGAN KEPUJIAN
Male	MATHEMATICS	Realistic	SPORT	Universiti Teknologi Malaysia (UTM)	STEM	TK40 SARJANA MUDA KEJURUTERAAN (MEKANIKAL - AUTOMOTIF)

Female	MATHEMATICS	Realistic	LISTEN TO MUSIC	Universiti Teknologi Malaysia (UTM)	STEM	TK40 SARJANA MUDA KEJURUTERAAN (MEKANIKAL - AUTOMOTIF)
Male	MATHEMATICS	Investigative	SLEEP	Universiti Teknologi Malaysia (UTM)	STEM	TC10 SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)
Female	PHYSICS	Investigative	SPORT	Universiti Teknologi Malaysia (UTM)	STEM	TK30 SARJANA MUDA KEJURUTERAAN (KIMIA - GAS)
Female	CHEMISTRY	Artistic	SPORT	Universiti Teknologi Malaysia (UTM)	STEM	TK03 SARJANA MUDA KEJURUTERAAN (KIMIA)
Female	CHEMISTRY	Social	LISTEN TO MUSIC	Universiti Teknologi MARA (UiTM)	STEM	ES59 SARJANA MUDA SAINS (KEPUJIAN) KIMIA GUNAAN
Male	MATHEMATICS	Conventional	SPORT	Universiti Tun Hussein Onn Malaysia (UTHM)	STEM	BC10 SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	ECONOMICS	Conventional	LISTEN TO MUSIC	Universiti Utara Malaysia (UUM)	NON STEM	UT04 SARJANA MUDA PENDIDIKAN DENGAN KEPUJIAN (BIMBINGAN DAN KAUNSELING)
Male	INFORMATION TECHNOLOGY	Social	READING	Universiti Malaysia Pahang (UMP)	STEM	JC11 IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN
Male	SCIENCE	Enterprising	READING	Universiti Malaysia Terengganu (UMT)	STEM	GS02 SARJANA MUDA SAINS (SAINS BIOLOGI)

Female	ACCOUNTING	Enterprising	READING	Universiti Sultan Zainal Abidin (UniSZA)	NON STEM	DE02 IJAZAH SARJANA MUDA PERAKAUNAN DENGAN KEPUJIAN
Male	ENGLISH LANGUAGE	Realistic	READING	Universiti Teknologi Malaysia (UTM)	STEM	TC10 SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN)
Male	OTHERS	Artistic	SINGING	Universiti Teknikal Malaysia Melaka (UTeM)	STEM	CC25 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEPINTARAN BUATAN) DENGAN KEPUJIAN
Female	SCIENCE	Artistic	SINGING	Universiti Tun Hussein Onn Malaysia (UTHM)	STEM	BC24 SARJANA MUDA SAINS KOMPUTER (PENGKOMPUTERAN MULTIMEDIA) DENGAN KEPUJIAN
Female	ADDITIONAL MATHEMATICSS	Investigative	SLEEP	Universiti Malaya (UM)	STEM	MC06 SARJANA MUDA SAINS KOMPUTER (SISTEM MAKLUMAT)
Female	ECONOMICS	Artistic	SPORT	Universiti Islam Antarabangsa Malaysia (UIAM)	NON STEM	YE00 SARJANA MUDA EKONOMI (KEPUJIAN)
Male	BIOLOGY	Realistic	SPORT	Universiti Sains Islam Malaysia (USIM)	STEM	QG07 SARJANA MUDA SAINS DENGAN KEPUJIAN (BIOTEKNOLOGI MAKANAN)
Female	CHEMISTRY	Realistic	WATCH DRAMA, MOVIE	Universiti Kebangsaan Malaysia (UKM)	STEM	KD00 DOKTOR PERGIGIAN
Female	MALAY LANGUAGE	Investigative	WATCH DRAMA, MOVIE	Universiti Malaysia Kelantan (UMK)	NON STEM	LE23 IJAZAH SARJANA MUDA PENTADBIRAN PERNIAGAAN (PERBANKAN

						DAN KEWANGAN ISLAM) DENGAN KEPUJIAN
Female	MATHEMATICS	Artistic	WATCH DRAMA, MOVIE	Universiti Teknologi Malaysia (UTM)	STEM	TK79 SARJANA MUDA KEJURUTERAAN PROSES KIMIA
Male	PHYSICS	Conventional	WATCH DRAMA, MOVIE	Universiti Teknologi Malaysia (UTM)	STEM	TC24 SARJANA MUDA SAINS KOMPUTER (BIOINFORMATIK)
Female	MATHEMATICS	Social	WATCH DRAMA, MOVIE	Universiti Utara Malaysia (UUM)	NON STEM	UE00 SARJANA MUDA SAINS EKONOMI DENGAN KEPUJIAN
Female	PHYSICS	Investigative	WATCH DRAMA, MOVIE	Universiti Sains Malaysia (USM)	STEM	SK01 IJAZAH SARJANA MUDA KEJURUTERAAN (KEPUJIAN) (KEJURUTERAAN AWAM)

APPENDIX E
DATA FOR TESTING PROCESS

GENDER	SUBJECT	PERSONALITY	HOBBY	UNIVERSITY	FIELD	COURSE
Female	MATHEMATICS	Artistic	SPORT	Universiti Malaysia Pahang (UMP)	STEM	JC11 IJAZAH SARJANA MUDA SAINS KOMPUTER (SISTEM KOMPUTER & RANGKAIAN) DENGAN KEPUJIAN
Male	ENGLISH LANGUAGE	Conventional	SPORT	Universiti Malaysia Sarawak (UNIMAS)	STEM	WS24 IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (SAINS KOGNITIF)
Female	MATHEMATICS	Enterprising	SPORT	Universiti Utara Malaysia (UUM)	NON STEM	UC12 IJAZAH SARJANA MUDA SAINS DENGAN KEPUJIAN (MULTIMEDIA)
Male	ACCOUNTING	Realistic	CHATING	Universiti Malaya (UM)	NON STEM	ME01 SARJANA MUDA PERAKAUNAN
Female	PHYSICS	Artistic	COMPUTER GAME	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	OTHERS	Artistic	COMPUTER GAME	Universiti Malaysia Pahang (UMP)	STEM	JK01 IJAZAH SARJANA MUDA (KEPUJIAN) KEJURUTERAAN AWAM
Male	MATHEMATICS	Conventional	COMPUTER GAME	Universiti Pertahanan Nasional Malaysia (UPNM)	STEM	ZK01 IJAZAH SARJANA MUDA KEJURUTERAAN AWAM

Male	CHINESE LANGUAGE	Social	COMPUTER GAME	Universiti Putra Malaysia (UPM)	STEM	PC04 BACELOR SAINS KOMPUTER (SISTEM KOMPUTER)
Female	INFORMATION TECHNOLOGY	Social	COOKING	Universiti Kebangsaan Malaysia (UKM)	STEM	KC00 IJAZAH SARJANAMUDA SAINS KOMPUTER DENGAN KEPUJIAN
Male	ACCOUNTING	Artistic	DANCE	Universiti Malaysia Pahang (UMP)	STEM	JC10 IJAZAH SARJANA MUDA SAINS KOMPUTER (KEJURUTERAAN PERISIAN) DENGAN KEPUJIAN
Female	BIOLOGY	Realistic	DANCE	Universiti Putra Malaysia (UPM)	STEM	PG09 BACELOR SAINS DAN TEKNOLOGI MAKANAN
Female	MATHEMATICS	Artistic	DANCE	Universiti Sains Malaysia (USM)	STEM	SH03 IJAZAH SARJANA MUDA SAINS (PERUMAHAN, BANGUNAN DAN PERANCANGAN) (KEPUJIAN) (SENIBINA)
Female	BIOLOGY	Conventional	OTHERS	Universiti Malaya (UM)	NON STEM	ML01 SARJANA MUDA UNDANG-UNDANG
Female	MATHEMATICS	Realistic	LISTEN TO MUSIC	Universiti Malaysia Sabah (UMS)	STEM	HK01 IJAZAH SARJANA MUDA KEJURUTERAAN DENGAN KEPUJIAN (KEJURUTERAAN AWAM)
Female	MATHEMATICS	Artistic	PLAY MUSIC INSTRUMENT	Universiti Kebangsaan Malaysia (UKM)	STEM	KC09 IJAZAH SARJANAMUDA KEJURUTERAAN PERISIAN DENGAN KEPUJIAN (PEMBANGUNAN SISTEM MULTIMEDIA)

APPENDIX F
SURVEY FORM

Survey on student preference when selecting undergraduate program through Unit Pengambilan University(UPU) online system (UPUOnline).

Dear all,

I am Thoi Wen Bin, fourth year student at Universiti Malaysia Pahang. I am currently conducting a research on implementation of machine learning algorithm in university course selection in Malaysia for my undergraduate project. The objective of the research is to develop a recommendation system in selecting undergraduate course in Malaysia based on user profile.

The survey contains 2 sections, and will take about 5-10 minutes to answer. All information and responses will be treated as strictly private and confidential, and anonymous for research purposes only.

Your coorperation is highly appreciated, and your feedback is valuable to me. Should you have any questions pertaining to this survey, please do not hesitate to contact me at 016-9541128 or email me at benjaminthoi0519@gmail.com

Thank you.

Regards,
Thoi Wen Bin, Undergraduate student, Faculty of Computer Systems & Software Engineering(FSKKP),
Universiti Malaysia Pahang(UMP)

NEXT

Never submit passwords through Google Forms.

Section 1: Demographic profile

Gender *

- ○ Male
- ○ Female

Which subject you liked the most in secondary school/Matriculation college? *

Fill in the blank with 1(one) answer.

Your answer

What is your hobby? *

Fill in the blank with one answer.

Your answer

Which type of personality best suits you? *

- ○ Investigative (Good at understanding and solving science and math problems. Sees self as precise, scientific, and intellectual.)
- ○ Conventional (Good at working with written records and numbers in a systematic, orderly way. Sees self as orderly, and good at following a set plan.)
- ○ Realistic (Good at working with animals, tools, or machines. Sees self as practical, mechanical, and realistic.)
- ○ Social (Good at teaching, counseling, nursing, or giving information. Sees self as helpful, friendly, and trustworthy.)
- ○ Enterprising (Good at leading people and selling things or ideas. Sees self as energetic, ambitious, and sociable.)
- ○ Artistic (Good at doing creative activities like art, drama, crafts, dance, music, or creative writing. Sees self as expressive, original, and independent.)

BACK NEXT

Section 2 : User preference on university and undergraduate program

Which of the following university is the first choice for continuing your undergraduate study ? *

- ○ Universiti Islam Antarabangsa Malaysia (UIAM)
- ○ Universiti Teknologi MARA (UiTM)
- ○ Universiti Kebangsaan Malaysia (UKM)
- ○ Universiti Malaya (UM)
- ○ Universiti Malaysia Kelantan (UMK)
- ○ Universiti Malaysia Pahang (UMP)
- ○ Universiti Malaysia Sabah (UMS)
- ○ Universiti Malaysia Terengganu (UMT)
- ○ Universiti Malaysia Perlis (UniMAP)
- ○ Universiti Malaysia Sarawak (UNIMAS)
- ○ Universiti Sultan Zainal Abidin (UniSZA)
- ○ Universiti Putra Malaysia (UPM)
- ○ Universiti Pertahanan Nasional Malaysia (UPNM)
- ○ Universiti Pendidikan Sultan Idris (UPSI)
- ○ Universiti Sains Islam Malaysia (USIM)
- ○ Universiti Sains Malaysia (USM)
- ○ Universiti Teknikal Malaysia Melaka (UTeM)
- ○ Universiti Tun Hussein Onn Malaysia (UTHM)
- ○ Universiti Teknologi Malaysia (UTM)
- ○ Universiti Utara Malaysia (UUM)

BACK NEXT

Section 2 : User preference on university and undergraduate program

You have chosen Universiti Putra Malaysia (UPM). Please choose your first-choice undergraduate program. *

- ACCOUNTING AND TAXATION - BACELOR PERAKAUNAN
- APPLIED SCIENCE - BACELOR SAINS (KEPUJIAN) BIOTEKNOLOGI
- ARCHITECTURE AND TOWN PLANNING - BACELOR SAINS SENI BINA
- ARCHITECTURE AND TOWN PLANNING - BACELOR SENI BINA LANDSKAP
- BIOLOGY AND BIOCHEMISTRY - BACELOR SAINS (KEPUJIAN) BIOKIMIA
- BIOLOGY AND BIOCHEMISTRY - BACELOR SAINS (KEPUJIAN) BIOLOGI SEL DAN MOLEKUL
- BIOLOGY AND BIOCHEMISTRY - BACELOR SAINS (KEPUJIAN) MIKROBIOLOGI
- BIOLOGY AND BIOCHEMISTRY - BACELOR SAINS (SAINS BIOPERUBATAN)
- BIOLOGY AND BIOCHEMISTRY - BACELOR SAINS BIOLOGI DENGAN KEPUJIAN
- CHEMICAL AND PROCESS - BACELOR KEJURUTERAAN (KIMIA)
- CHEMICAL AND PROCESS - BACELOR KEJURUTERAAN (PERTANIAN DAN BIOSISTEM)
- CHEMISTRY - BACELOR SAINS KIMIA DENGAN KEPUJIAN
- CHEMISTRY - BACELOR SAINS KIMIA PERINDUSTRIAN DENGAN KEPUJIAN
- CHEMISTRY - BACELOR SAINS KIMIA PETROLEUM DENGAN KEPUJIAN
- CIVIL ENGINEERING - BACELOR KEJURUTERAAN (AWAM)
- COMPUTER SCIENCE - BACELOR KEJURUTERAAN PERISIAN
- COMPUTER SCIENCE - BACELOR SAINS KOMPUTER (MULTIMEDIA)
- COMPUTER SCIENCE - BACELOR SAINS KOMPUTER (RANGKAIAN KOMPUTER)
- COMPUTER SCIENCE - BACELOR SAINS KOMPUTER (SISTEM KOMPUTER)
- CROP AND LIVESTOCK PRODUCTION - BACELOR PERTANIAN (SAINS TERNAKAN)

Survey on student preference when selecting undergraduate program through Unit Pengambilan University(UPU) online system (UPUOnline).

Thank you for your response.
If you have any questions pertaining to this survey, please do not hesitate to contact me at 016-9541128 or email me at benjaminthoi0519@gmail.com
Thank you.

Submit another response