

# Fuzzy-Based Recommendation System for University Major Selection

Shaima Alghamdi, Nada Alzhrani and Haneen Algethami <sup>a</sup>

Department of Computer Science, College of Computers and Information Technology, Taif University, Taif, Saudi Arabia

**Keywords:** Fuzzy Expert System, Recommendation System Decision Support System, Major Selection.

**Abstract:** The decision of choosing a university major is one of the most important decisions in every adult life. To make a suitable decision, a student needs an expert opinion, time, and effort. Therefore, a decision-making system should be developed in order to help prospective students to increase their educational outcome and productivity. In Saudi Arabia, each university requires specific criteria in order to accept students. These criteria are made based on two factors: 1) the outcome of student's qualification exams and 2) overall high school grades. The student must take these calculations into consideration when selecting a major. Thus, in this paper, a Fuzzy-Based Recommendation System (FRS) is proposed to aid students in choosing a suitable major. This system designed using Fuzzy Expert System (FES). Additionally, a cluster-based preferences technique is implemented to obtain the student's preferred majors, using distance measurement. The system has been tested on fifteen prospective students to measure its accessibility. Results showed that students are stratified by the suggested majors that fell in line with their preferences.

## 1 INTRODUCTION

Most high school students are uncertain of which university major to choose after they finish their study. The decision-making process can be overwhelming. Hence, they usually look for guidance and support by reaching out to teachers, relatives and colleagues. These people are regarded as experts. However, the problem lies in finding an expert available to help the students. Also, some experts' opinions tend to be subjective to their own experience, without taking into consideration the student preferences. Hence, the decision support system (DSS) are more effective to help students in making life-affecting decisions (Turban, 1995).

In Saudi Arabia, there is a need for a DSS for the university major selection problem. Universities require qualification criteria in order to register students. In addition to the student's previous performance, preferences should be considered during the process to provide a more realistic result. Finally, some majors might share courses, which make the decision of choosing one major over the other is a difficult task.

In this paper, a fuzzy-based recommendation system is used to suggest a list of majors for the student. Recommender Systems (RSs) provide

suggestions for items to be of use to a user (Ricci, 2011). In this paper, the knowledge-based recommendation system (KBRS) is used, in which the system uses inference technique to find a relationship between the items and the user (Burke, 2000).

Fuzzy logic is used to handle uncertainty rising from similarities between the majors. Fuzzy logic can provide an effective means for conflict resolution of multiple criteria and better assessment of options (Burke, 2000).

The specific objectives of this paper are:

- To identify the features that contribute to maximizing student satisfaction on the major choice.
- To propose an Intelligent Decision Support System (IDSS) to aid students in the decision-making process
- To evaluate user satisfaction on the system performance.

In what follows, Section II reviews related work. Section III describes the problem definition. Section IV proposed methodology used in this paper. Section V presents the evaluation study and discusses the obtained results. The paper is then concluded in Section VI.

<sup>a</sup> <https://orcid.org/0000-0002-7582-4480>

## 2 LITERATURE REVIEW

Recommendation systems (RSs) were found to be a helpful tool (Yera, 2017). It can help students to find, organize, and use resources that match their individual goals, interests, and current knowledge (Al-Badarenah et al, 2016). Nevertheless, recent studies show different approaches for managing uncertainty in recommender systems, such as Bayesian approaches (Luis M. de Campos, 2008), Markov models (Nachiketa Sahoo, 2012), fuzzy approaches (Azene Zenebe, 2009), genetic algorithms (Holland, 1992), or neural networks (Lehr, 1990).

Recommendation systems have played an important role in education. One of these systems is a Markov Chain Collaborative Filtering Model for Course Enrollment Recommendations by (Elham S.Khorasani, 2016). Another recommendation system is course recommender system using association rules by (Narimel Bendakir, 2006). Another recommendation system is designed by (Desi Purwanti Kusumaningrum, 2017) entitled Recommendation System for Major University Determination Based on Student's Profile and Interest.

There are also, number of studies that have addressed the major selection problem. One of these studies is a Prototype Rule-based Expert System with an Object-Oriented Database for University Undergraduate was proposed by (Ahmar, 2012). The study highlighted the importance of using an expert system supported by an object-oriented database. Also, it used Kappa-PC expert system development environment, which supports rule-based reasoning, object-oriented modelling, list processing, and graphical user interface construction components. This ES has three major components that are: 1) Knowledgebase; 2) Inference engine; 3) User interface.

Another study is the Decision Support System for Major Selection Vocational High School (VHS) using Fuzzy Logic Android-Based was proposed by (Salaki, 2015). It is a DSS to aid the student in the decision-making process, based on the score of acceptance exams to specify the appropriate VHS major for the student. The DSS consists of three main parts: 1) Information system; 2) DSS, which has three subsystems Database subsystem, Model subsystem, Dialog Subsystem; 3) Fuzzy Inference System.

In the previously presented literature, fuzzy logic was used to deal with uncertainty in relative problems. Additionally, database systems were also used to store data. Finally, a graphical user interface

was also used to retrieve online information. Hence, the same components are used in developing the FRS for the university major selection problem presented in this paper. Even though the previous studies have used fuzzy expert systems to solve this problem, it is worthy to develop an efficient IDSS for tackling real-world major selection, for students applying to Taif University at Saudi Arabia. The intended contribution focuses on the use of fuzzy logic to improve the performance of knowledge-based recommender systems. The combination harnesses its power with the fuzzy expert system.

## 3 PROBLEM STATEMENT

The major selection problem aims to maximise student satisfaction on their major choice to minimise the number of ungraduated students.

There are two tracks in high school, science and art. Each track has specific majors. Students from each track can apply only to those majors. However, the science track has more options than the art track. For example, a student applying for mathematics must be from the science track. On the other hand, a student applying for linguistics can be from science or art track. Thus, the high school track affects the direction of the result of the system.

Universities require three qualification criteria to accept students. third-year high school percentage (HSA), and percentages of two tests: 1) General Ability Test (GAT) and 2) Achievement test (AT).

The proposed recommendation system suggests a list of suitable majors based on the student's overall percentage and the student's preferences. The percentage is calculated based on the GAT, AT and HSA values.

In Taif University (TU), there are ten colleges, with each college having several majors to choose from and different calculation scheme. The student's overall percentage to be accepted in medicine and pharmacy colleges are calculated as shown in equation (1), where HSA and AT must be greater than or equal 75%. The student's overall percentage to be accepted in engineering, computers and information technology and applied medical sciences colleges is calculated also as shown in equation (1), however HSA and AT must be greater than or equal 70%. The student's overall percentage value for the science college is calculated as shown in equation (2). The student's overall percentage value for the art, education, shari'a, and business administration colleges is calculated as shown in equation (3). Note that  $\alpha = 0.3$ ,  $\beta = 0.4$  and  $\gamma = 0.5$ .

$$P_{M|E} = (HSA \times \alpha) + (GATS \times \alpha) + (AT \times \beta) \quad (1)$$

$$P_S = (HSA \times \beta) + (GATS \times \alpha) + (AT \times \alpha) \quad (2)$$

$$P_A = (HSA \times \gamma) + (GATS \times \gamma) \quad (3)$$

### 3.1 Data Collection

Prospective students find it difficult to select a university major. If they are not satisfied with their selection, they might change their major during their four-year degree program. In order to understand the factors affecting their decisions, two surveys were conducted. The first survey targeted high school students. The second survey targeted university students to give their insights after spending a year in a specific major.

The survey had 239 prospective participants and 392 university participants. In both surveys, a high percentage of students agreed on the difficulty of choosing a university major. In addition, more than half of the participants recommend the need for a system to help them in the decision-making process.

## 4 METHODOLOGY

### 4.1 Fuzzy Expert System Implementation

The aim of this step is to define a set of available majors for each student, i.e. the majors where the student's overall percentage matches the major requirements.

First, the high school track must be identified. the overall percentage is calculated as explained in Section III by using HSA, AT and GAT values. For a science student,  $P_{M|E}$ ,  $P_S$ , and  $P_A$  values are computed, since she/he can enroll in both science and art majors. On the other hand, only  $P_A$  value is computed for art students, as they are only allowed to enroll in art majors. If the user is a science student, the following step is to ask her/him to choose their preferred track in the university. If the student chooses the art track; the system only deals with  $P_A$ , i.e. for the art section. If the science student chooses the science track; the system handles three values, i.e.  $P_{M|E}$ ,  $P_S$ , and  $P_A$ .

The values are then passed to the FES to determine the applicable majors for the student, with respect to the previous criteria.

### 4.1.1 Fuzzy Logic Process

This process consists of a number of steps as follows:

1. Identify the linguistic variables and values, as presented in Table 1 (Jang J. S., 1997).

Table 1: Linguistic variables.

Type	Linguistic variable	Linguistic value
Input	Science, Engineering, Medicine, Art	High, Medium, or Low
Output	Science_Major, Engineering_Major, Medicine_Major, Art_Major	High, Medium, or Low

2. Identify fuzzy sets and their corresponding membership functions. The antecedent fuzzy sets represent the overall percentage constraints imposed by the university. For each college, the student's overall percentage is classified into one of the three fuzzy sets. This definition is used further in the rule evaluation, in order to ensure that the colleges and majors are within the student's range. Figures 1, 2, 3 and 4 present the antecedent fuzzy sets. In these figures, the x-axis represents the student's overall percentage values and the y-axis represent the corresponding membership values. The overall percentage values (i.e. the x-axis) are driven from Taif University enrollment data from last year.

Table 2: Antecedent fuzzy sets and ranges.

Medicine and Pharmacy fuzzy sets and their ranges		
Fuzzy set	Range	$\mu=1$
Low	[80.00, 88.00]	80.00
Medium	[85.00, 91.00]	87.00
High	[89.00, 100.00]	100.00
Science fuzzy sets and their ranges		
Fuzzy set	Range	$\mu=1$
Low	[70.00, 78.00]	70.00
Medium	[75.00, 88.00]	81.00
High	[85.00, 100.00]	100.00
Engineering fuzzy sets and their ranges		
Fuzzy set	Range	$\mu=1$
Low	[75.00, 80.00]	75.00
Medium	[78.00, 90.00]	84.00
High	[88.00, 100.00]	100.00
Art fuzzy sets and their ranges		
Fuzzy set	Range	$\mu=1$
Low	[70.00, 78.00]	70.00
Medium	[75.00, 88.00]	81.00
High	[85.00, 100.00]	100.00

They depend on 1) the average of all student who were accepted last year and 2) the number of available seats in each major. These values are uncertain and can change every year. The ranges provided by the

Table 3: Consequent fuzzy sets for each applicable major.

Science Major fuzzy sets			
Applicable Majors	$\mu=1$	Range	Fuzzy set
Biology- Zoology – Microbiology- Biotechnology- Food Science - Chemistry – Physics – Mathematics Accounting	00.00	[00.00, 11.00]	High
Biology- Zoology – Microbiology- Biotechnology- Food Science - Chemistry – Physics – Mathematics Accounting	15.00	[10.00, 21.00]	Medium
Biology- Zoology –Chemistry – Physics –Mathematics	30.00	[20.00, 30.00]	Low
Engineering Major fuzzy sets			
Applicable Majors	$\mu=1$	Range	Fuzzy set
Computer Science – Information Technology –Interior Design – Industrial Engineering – Architectural Engineering – Computer Engineering – Radiology –Nursing – Physical Therapy – Laboratories	30.00	[30.00, 41.00]	High
Computer Science – Information Technology –Interior Design – Industrial Engineering – Architectural Engineering – Computer Engineering – Radiology –Nursing Laboratories	45.50	[40.00, 51.00]	Medium
Computer Science – Information Technology	60.00	[50.00, 60.00]	Low
Medicine Major fuzzy sets			
Applicable Majors	$\mu=1$	Range	Fuzzy set
Medicine - Pharmacy	30.00	[60.00, 71.00]	High
None	45.50	[70.00, 81.00]	Medium
None	60.00	[80.00, 90.00]	Low
Art Major fuzzy sets			
Applicable Majors	$\mu=1$	Range	Fuzzy set
All art majors	90.00	[90.00, 101.00]	High
Marketing- Management – English- Arabic- Media and communication science- Early childhood – Sports – Graphical Design – Fabric Design and Fashion - Shari'a - alqara'at – Art –Psychology – Economics -Systems - Sciences of Quran - Islamic Culture	105.00	[100.00, 111.00]	Medium
English	120.00	[110.00, 120.00]	Low

university represents the medium fuzzy set, where any value above that range is considered high and any value below is considered low.

Table 2 summaries all the antecedent fuzzy sets. It consists of the fuzzy sets and their corresponding ranges. The range consists of the average values starting from 70%, i.e. the lower bound for enrollment in TU, is 100%.

The consequent fuzzy sets will specify a range for each value (Medicine, Engineering and Computers and Information Technology, Science, Art). The ranges determine the applicable majors for the student. Generally, the ranges in the consequent part start from 0 to 120. Figures 5, 6, 7 and 8 present the consequent fuzzy sets. Values overlapping is minimized to help in the elimination process to provide the student with the available majors only.

Table 3 displays the consequent fuzzy sets. It illustrates the applicable majors for each fuzzy set. The system returns a value that is used to determine the set of applicable majors for the student based on her/his overall percentage. For example, in Engineering\_Major fuzzy sets if the value of the consequent is 35, then the applicable majors for the student is only computer science and information

technology. This value is used to retrieve the available majors for the student from the database, which stores the fuzzy sets ranges based on the university majors. A triangular membership function is used, as shown in equation (4), due to its suitability of the overall percentage value. This function has four parameters: 1) average value (AVG), 2) Lower Bound (LB), 3) Membership Function (MP), and 4) Upper Bound (UB) (Jang J. S., 1997).

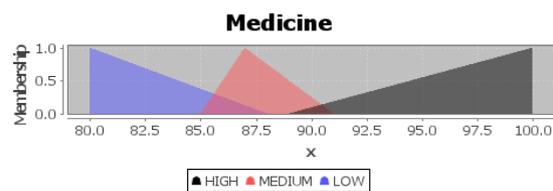


Figure 1: Medicine fuzzy set.

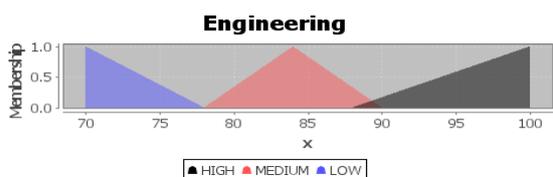


Figure 2: Engineering fuzzy set.

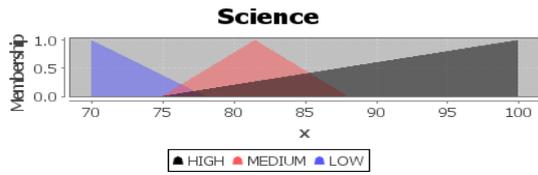


Figure 3: Science fuzzy set.

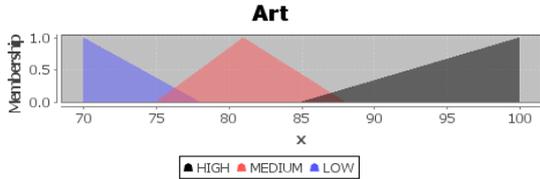


Figure 4: Art fuzzy set.

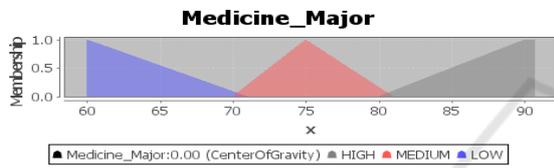


Figure 5: Medicine\_Major fuzzy sets.

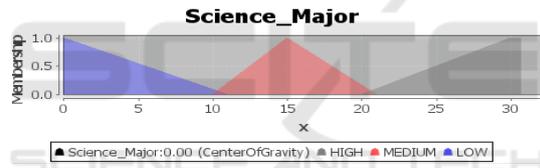


Figure 6: Science\_Major fuzzy sets.

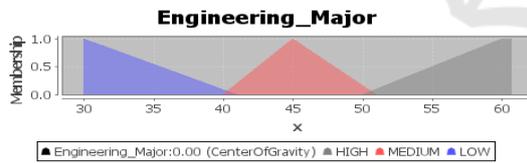


Figure 7: Engineering\_Major fuzzy sets.

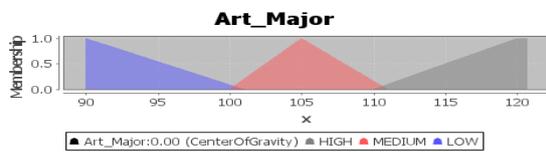


Figure 8: Art\_Major fuzzy set.

$$\mu(AVG, LB, MP, UB) = \begin{cases} 0, & \text{If } AVG \leq LB \\ \frac{AVG - LB}{MP - LB}, & \text{If } AVG \in (LB, MP] \\ \frac{UB - AVG}{UB - MP}, & \text{If } AVG \in (MP, UB) \\ 0, & \text{If } AVG \geq UB \end{cases} \quad (4)$$

### 3 INFERENCE MECHANISM

Table 4: System Knowledgebase.

Fuzzy Rules			
IF Science IS	Low	THEN Science_Major IS	Low
	Medium		Medium
	High		High
IF Engineering IS	Low	THEN Engineering_Major IS	Low
	Medium		Medium
	High		High
IF Medicine IS	Low	THEN Medicine_Major IS	Low
	Medium		Medium
	High		High
IF Art IS	Low	THEN Art_Major IS	Low
	Medium		Medium
	High		High

Table 4 presents the knowledge base of the system as IF-THEN rules. Rule evaluation executes the rules based on the student’s input. Antecedent value is used in the evaluation of the consequent part. Rule aggregation combines all the fuzzy sets that resulted from firing the rules in the last stage. The combined fuzzy sets are in the same universe of discourse. This combination is used as an input for the defuzzifier.

### 4 DEFUZZIFICATION

The algorithm used to defuzzify the values is the center of gravity algorithm (COA), as shown in equation (5). Centre-of-Area algorithm is commonly used in the defuzzification process and there are many studies that have used it such as (Maranate, 2014), (X.Y. Djam, 2011), and (Enes Erkan, 2016). The COA, as shown in equation 6, is defined for a finite universe of discourse (Jang J.-S. R., 1997). The x is the value in the universe of discourse, and  $\mu(x)$  is the corresponding membership value.

$$COA = \frac{\sum x \mu(x)}{\sum \mu(x)} \quad (5)$$

#### 4.1 Cluster-based Preferences

The aim of this step is to cluster preferred majors into groups, where majors in the same group are more

similar than the majors in different groups. This arrangement can help the student to 1) choose set majors that they prefer and 2) give the student a chance to see other majors in case the student's average does not qualify to enter the major. The cluster-based preferences use distance measurements to calculate the similarity. To do so, the following steps are applied:

- Majors are divided into regions, based on the content similarity or the work field. Each region has a number of keywords and associated to one or more question, as shown in Table 5. Regions and questions are formed and validated by the university's faculty members of each major.
- Student's answers are recorded through an online questionnaire to define their preferences. This process aims to measure the student's interest in each region on a scale from one to five, where five is the maximum score. If the Group has more than one question, the average score is calculated based on the student's answers.
- The difference between the maximum score and the user score of a current question is calculated, as shown in equation (7).

$$Score = \frac{\sum_{i=1}^N score_i}{N} \quad (6)$$

$$Difference = Max_{score} - Score \quad (7)$$

In equation (6),  $N$  is the number of questions for the region. Note the minimum distance is the closest to the preferred region.

- Record majors with the minimum distance region score, i.e. preferred majors.

#### 4.2 Elimination Process

The aim is to ensure that the suggested majors align with the student's overall percentage, with respect to the university constraints, while satisfying her/his preferences. Thus, the final majors are computed based on the stored values. Where, for each student, the suggested majors are in the intersection area between two sets: 1) the applicable majors and 2) the preferred majors. i.e., Final Majors = Applicable Majors  $\cap$  Preferred Majors

### 5 SYSTEM EVALUATION

This stage is the first software development stage. Hence, a pre-alpha version of the system was

released to test for the system accessibility in a high school in Taif, Saudi Arabia. The sample consisted of twelve high school students. Results showed that 66% were strongly pleased with the system and 54% were pleased with suggestions provided by the system as shown in figures 9. A positive feedback was received from the students, where each student was led to a suitable major that fell in line with their preferences. Accordingly, the system can help in increasing the student's satisfaction by giving each student the chance to succeed in the suggested major.

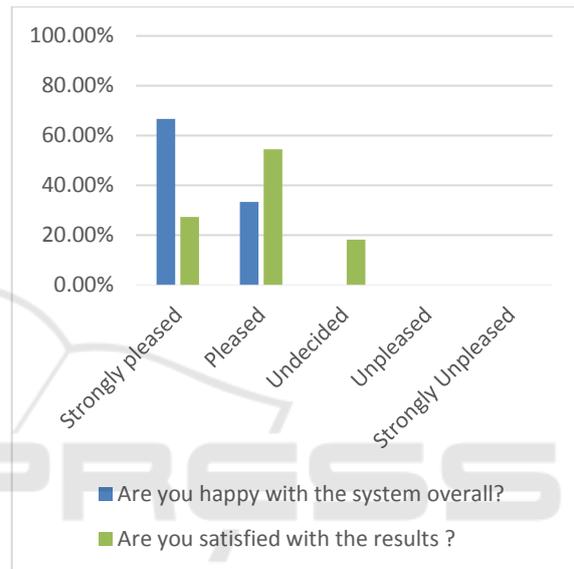


Figure 9: High School Students' Feedback.

### 6 CONCLUSIONS

This paper focused on designing a Fuzzy Recommendation System (FRS) that aided in students' decision in choosing their university major. The more satisfied are the students about their majors, the more productive they can be. However, measuring student's preferences and how it can relate specifically to the student's interests can raise the question of system accuracy. This problem can be solved by developing a detailed scale, with the help experts, to measure the student's preferences.

In future work, a comparison to existing methods must be conducted. Also, the system's accuracy and performance must be tested.

Table 5: Preferences regions based on similarities.

The Science Path Regions			
Region	Majors	Similarity	Keywords
One	Medicine	Content	Anatomy
Two	Medicine – Microbiology	Content	Immune System – Diseases
Three	Medicine –Nursing – Physical Therapy	Working field	Medical Care
Four	Chemistry – Food Science and Nutrition	Content	Food Science
Five	Pharmacy – Biotechnology – Chemistry	Content	Formulation – Drugs
Six	Radiology – Physics	Content	Radiation
Seven	Mathematics	--	Calculus – Numbers
Eight	Accounting	--	Finances – management
Nine	Interior Design – Architectural Engineering	Content	Design –Building (Interior /Exterior)
Ten	Computer Science – Information Technology – Computer Engineering	Content / Work filed	Programming/computers
Eleven	Laboratory – Microbiology –Chemistry	Content	Labs – substance
Twelve	Biology – Microbiology -Zoology	Content	Living Organisms
Thirteen	Physics	--	Natural Laws
Fourteen	Industrial engineering	---	Assembly, Numbers
The Art Path Regions			
Region	Majors	Similarity	Keywords
One	English – Arabic	Languages	Languages
Two	Arabic – Al-Shari'a	Content	Arabic
Three	Al-Shari'a - Systems - Islamic Culture	Content	Religion
Four	Al-qara'at , Sciences of Quran	Content	Al-Quran
Five	Al-Shari'a – Laws	Content	Islamic Law
Six	Economic and Finance-Management-Marketing- Management Information System	Content	Administration
Seven	Graphical Design	--	Computer -Design
Eight	Fabric Design and Fashion	--	Design – Fashion
Nine	Media and Communication Science	--	Media
Ten	Art	--	Art
Eleven	Sports	--	Sport
Twelve	Psychology – Early childhood	Work field	Behaviour – psychology

## REFERENCES

- Ahmar, M. A. (2012). A Prototype Rule-based Expert System with an Object-Oriented Database for University Undergraduate Major Selection.
- Ali, M. S. (n.d.). A Neuro-Fuzzy Inference System For Student Modeling In Web-based tutoring system.
- Amer Al-Badarenah, J. A. (2016). An Automated Recommender System for Course Selection. *International Journal of Advanced Computer Science and Applications*.
- Azene Zenebe, A. F. (2009). Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems. *Fuzzy Sets and Systems*, pp. Volume 160, Issue 1, Pages 76-94.
- Burke, R. (2000). Knowledge-based recommender systems. *Encyclopedia of Library and Information Science*, 175-186.
- Desi Purwanti Kusumaningrum, N. A. ( 2017, April). Recommendation System for Major University Determination Based on Student's Profile and Interest. *Journal of Applied Intelligent System (e-ISSN: 2502-9401 / p-ISSN: 2503-0493)*, pp. Vol. 2 No. 1, pp. 21 – 28.
- Elham S.Khorasani, Z. Z. (2016). A Markov Chain Collaborative Filtering Model for Course Enrollment Recommendations. *IEEE International Conference on Big Data (Big Data)*, pp. 3484 - 3490.
- Enes Erkan, Ç. T. (2016). A Fuzzy Expert System for Risk Self-Assessment of Chronic Diseases. *IOSR Journal of Computer Engineering (IOSR-JCE)*, pp. PP 29- 33.
- Francesco Ricci, L. R. (2011). *Introduction to Recommender Systems*. New York: Springer Science+Business Media, LLC.
- Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press.
- Jang, J. S. (1997). Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence. *IEEE Transactions on automatic control*, pp. 42(10), 1482-1484.
- Jang, J.-S. R. (1997). *Neuro-fuzzy and soft computing: a computational approach to learning and Machine learning*. Ann Arbor: Prentice Hall, 1997.

- Lehr, B. W. (1990, September). 30 years of adaptive neural networks: perceptron, Madaline, and backpropagation. *in Proceedings of the IEEE*, pp. vol. 78, no. 9, pp. 1415-1442.
- Luis M. de Campos, J. M.-L. (2008). A collaborative recommender system based on probabilistic inference from fuzzy observations. *Fuzzy Sets and Systems*, pp. Volume 159, Issue 12, Pages 1554-1576.
- Maranate, T. (2014). Fuzzy FMEA Application for Variables in FactorAnalysis to Prioritize Suspected OSA Patients for sleep study. *The 3rd International Conference on Computer Engineering and Mathematical Sciences (ICCEMS)* .
- Nachiketa Sahoo, P. V. (2012, December ). A hidden Markov model for collaborative filtering. *MIS Q.* 36, pp. 1329-1356.
- Narimel Bendakir, E. A. (2006). Using Association Rules for Course Recommendation. *In Proceedings of the AAAI Workshop on Educational Data Mining.*, pp. pp. 31-40.
- Ronald R. Yager, L. A. (1992). *An Introduction to Fuzzy Logic Applications in Intelligent Systems*. Kluwer Academic Publishers.
- Salaki, C. R. (2015). Decision Support Systems Major Selection Vocational High School in Using Fuzzy Logic Android-Based.
- Turban, E. e. (1995). *Decision support and expert systems: management support systems*. University of Michigan: Prentice Hall.
- X.Y. Djam, G. M. (2011). A Fuzzy Expert System for the Management of Malaria. *International Journal of Pure and Applied Sciences and Technology*, pp. pp. 84-108.
- Yera, R. &. (2017). Fuzzy tools in recommender systems: A survey. *International Journal of Computational Intelligence Systems*, 10(1), 776-803.