

PAPER • OPEN ACCESS

A fuzzy-based career recommender system for senior high school students in K to 12 education

To cite this article: M C B Natividad *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **482** 012025

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

A fuzzy-based career recommender system for senior high school students in K to 12 education

M C B Natividad¹, B D Gerardo² and R P Medina³

¹Technological Institute of the Philippines, Quezon City, Philippines

²West Visayas State University, Iloilo City, Iloilo, Philippines

³Technological Institute of the Philippines, Quezon City, Philippines

¹marvecheska@gmail.com

²bobby.gerardo@gmail.com

³ruji_p_medina@yahoo.com

Abstract. In this paper, a proposed career recommender system was presented which aim to help not only the guidance counselor but most especially the senior high school students to guide them in considering numerous factors associated to their decision on what career they will pursue. In dealing with choosing the student attributes from numerous factors, a feature selection technique is appropriate to use to remove irrelevant features that affect the performance of the proposed fuzzy-based system. In this paper, different filter methods are used to select the best attributes then these attributes are used as crisp inputs. The result of the experiment shows reasonable result for making decisions. It is concluded that the proposed career recommender system for the students is very timely and will be one of the significant research work in the new era of education system in the Philippines.

1. Introduction

Life is full of difficult choices like the senior high school students where they need to face a tough decision as to what appropriate career will lead them to a bright future. That is why the education department of the Philippines as they continuously realizing their commitment to give quality education to the nation is now implementing the K to 12 education program that one primary reason is to lengthen the cycle of primary and secondary education from 10 years to 12 years (1). Aside from this, one feature of K-12 is to prepare the students and empower them to confidently join the labor market, so that by the time they graduate, they are already on their legal age and equipped with sufficient skills (2), thus, the Department of Education included career plans in their enhanced curriculum.

Studying career choice is important in employment especially in developing countries (3) like the Philippines as such, course preferences of graduates has a strong relationship with the labor market (4). But nowadays, students are more exposed with various factors that are associated with their career choices which need to be considered like interest in the course (5), situational factors, parental influence, school attended, gender, prestige attachment, and passion which are listed in the study of (6) (7). The ability of the learners to identify their preferred career choice, family, and teachers have significant role to influence the student (8) as well as environmental factors like counselor or other members of the family aside from parents, teachers, and friends, opportunity, their own personality traits (9), and capacity(10).



That is why career counseling has a very important role to help the students to be more aware (11) and not to just pick any course that they want(12) and to address this issue, it is only high time to develop a career recommender system to help the students in determining career tracks suited for them using fuzzy logic since this approach deals with linguistic variables which are words that are less precise than numbers just like in human reasoning.

Since in this study, different factors must first to identify and in dealing with this diverse and vast amount of student data, issues in computational time and complexity must also consider to produce a quality prediction model (13) and to address this, the application of feature selection is needed (14) since it is used in pre-processing step of data by selecting appropriate subset of features as relevant student attributes before they will be used in a career recommender system.

2. Literature Review

2.1. Feature selection

Feature selection techniques are effectively used together with many classification techniques in educational setting like in predicting a student performance model because the learning efficiency and performance accuracy are enhanced especially that the complexity of the learned results reduced (15). Consequently, it gives several benefits such as the removal of irrelevant features (16) and with these limited number of features, still, promising results were achieved (17). In addition, by using feature selection in identifying significant factors as well as the students' characteristics (18) that influence the studying performance of student (19) and to use in a warning system, is very helpful to the learners as well as to the teachers in evaluating academic performance for improvement (20) since this is the primary goal of all educational organizations (21).

There are three feature selection methods namely: filter, wrapper and embedded method, wherein in this study filter methods are used because they are generally functioning in pre-processing as features are selected based on the characteristic of the data (22) as it can easily cope with classification tasks in feature spaces of large dimensionality as found out by (23) in their paper.

2.2. Fuzzy Logic

Fuzzy logic was introduced by Lotfi Zadeh in 1965 as an approach to computing based on degrees of truth rather than the usual true or false of 1 or 0 of Boolean logic. It handles uncertain, imprecise and vague data which are encountered in most real-life problems and make use of them as basis for building fuzzy inference systems(24).

In recommender systems, especially in taking preferences of users, uncertainty cannot be ignored. But, with fuzzy logic, this issue could be handled to support recommender systems in giving accurate recommendation (25) with the effective and efficient result as presented in (26), even in career counseling (27) (28), as well as to analyze students' academic performances (29).

3. Methodology

3.1. Data Collection

The student data were obtained from senior high school students of the two schools division in the province of Isabela where their demographic data are gathered through a structured questionnaire which these attributes were adopted from different literature and with prior approval of the school heads the students' class marks were also collected. From the collected student data, several attributes were drawn as shown in Table 1.

3.2. Pre-processing of Data

The dataset was divided into three sets as sixty percent (60%) of the population were consider as the train set with 429 students, and the remaining forty percent (40%) has been equally divided to validation set as 144 students and test set consisting of 143 students, respectively, with a total of 716

students. All unnecessary spaces are removed and wrong input texts were edited before it converted to a .csv file and Weka 3.8.0 as an open source software and data mining tool was used for feature selection using five filtering methods which are Correlation-based, GainRatio, InfoGain, Relief and Symmetrical, respectively.

Table 1. Description of Student Attributes

Attribute Name	Description	Attribute Name	Description
Gender	Student’s gender	Parent_income	Parent’s monthly income
Age	Age of the student	Personality	Perception and Interest
Track Strands	SHS track enrolled Current strands	Parents_influence Relatives	Influence of Parents Influence of Relatives
Father_Occupation	Father’s Occupation	Peers	Peer Influence
Father_educ_attainment	Father’s Education	Socio_economic_status	Socio-Economic status of the family
Mother_Occupation	Mother’s Occupation	Proximity	Proximity of School
Mother_educ_attainment	Mother’s Education	Job_opportunities	Prestige of career
Final_Grade	Cumulative Grade Point Average		

3.3. System architecture

The system architecture of the proposed system as shown in figure 1 presented data from students and other factors associated to their decision as filtered using feature selection techniques to select the best attributes and prepare those data as crisp inputs before inputting in a fuzzy inference system. To make the data become fuzzy inputs, fuzzification process takes place as these inputs transform into grades of membership for linguistic terms of fuzzy sets. Rules now in the fuzzy inference will be evaluated and combined all these rules using fuzzy set operations to obtain a final result and defuzzified to obtain the final crisp output as the appropriate recommendation which the advisor as counselors will give to the student.

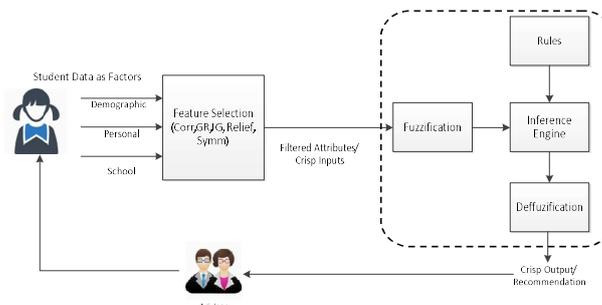


Figure 1. System architecture of the proposed career recommendation system

4. Experimental Results

The selected attribute from the different feature selection techniques were, National Career Assessment Examination (NCAE) result, strands, final grade and interest of the students and these attributes used as variables in developing the fuzzy inference engine in using Matlab as shown in figure 2 consisting of four input variables and one output as result.

The four input variables in figure 3 show the different degree of membership with their respective values and the rule viewer as shown in figure 4 presented 72 rules for the fuzzy model.

The fuzzy inference engine was then evaluated using evalfis and the result was shown in figure 5 where the defuzzification method used is centroid. Here, the four input data are 3 as strands input, 5 for interest, 1 for ncae and 89 as grade which the result as ans is equal to 88.9565 meaning that the student has 88.95% chance of pursuing their chosen career as related to their current track.

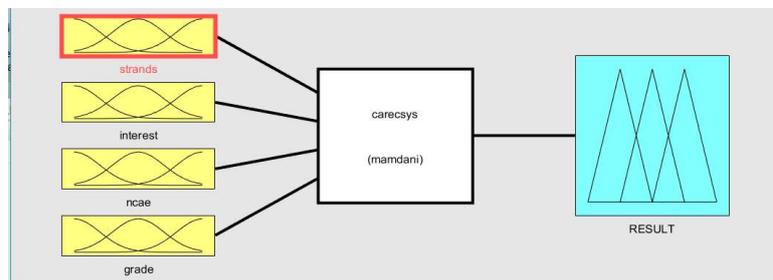


Figure 2. Fuzzy model with Input Variables

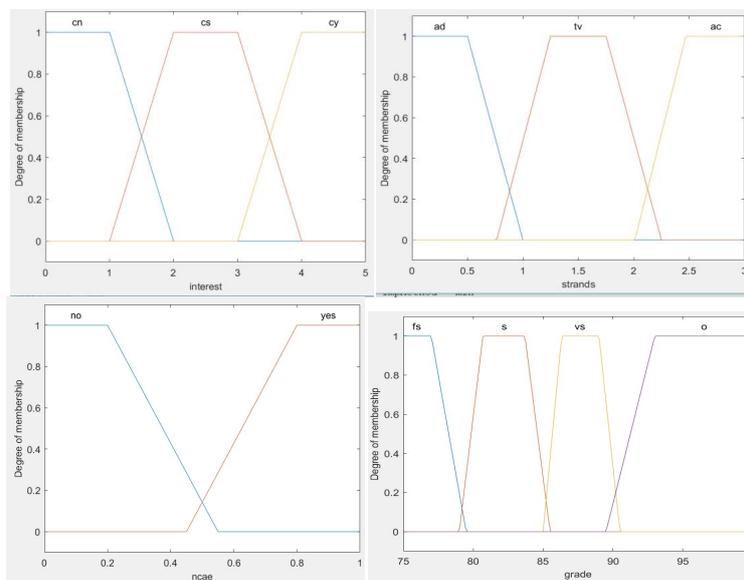


Figure 3. Degree of membership for four input variables

```

1. If (strands is ad) and (interest is cn) and (ncae is no) and (grade is fs) then (RESULT is very_low) (1)
2. If (strands is ad) and (interest is cn) and (ncae is no) and (grade is s) then (RESULT is very_low) (1)
3. If (strands is ad) and (interest is cn) and (ncae is no) and (grade is vs) then (RESULT is low) (1)
4. If (strands is ad) and (interest is cn) and (ncae is no) and (grade is o) then (RESULT is low) (1)
5. If (strands is ad) and (interest is cs) and (ncae is no) and (grade is fs) then (RESULT is very_low) (1)
6. If (strands is ad) and (interest is cs) and (ncae is no) and (grade is s) then (RESULT is very_low) (1)
7. If (strands is ad) and (interest is cs) and (ncae is no) and (grade is vs) then (RESULT is low) (1)
70. If (strands is ac) and (interest is cs) and (ncae is yes) and (grade is s) then (RESULT is high) (1)
71. If (strands is ac) and (interest is cs) and (ncae is yes) and (grade is vs) then (RESULT is high) (1)
72. If (strands is ac) and (interest is cs) and (ncae is yes) and (grade is o) then (RESULT is very_high) (1)
    
```

If strands is and interest is and ncae is and grade is Then RESULT is

not not not not not

Connection: or and Weight:

Figure 4. Fuzzy rules for input variables

```

>> evalfis([3 5 1 89], a)
ans =
    88.9565
fx >>
    
```

Figure 5. Result of evaluation of the fuzzy inference engine

4.1. Performance Evaluation

The root mean square error (RMSE) and mean absolute error (MAE) are used to measure the performance of the developed fuzzy model as shown in equation (1) and (2).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n e_i^2} \quad (1)$$

$$MAE = \frac{1}{N} \sum_{i=1}^n e_i \quad (2)$$

Table 2. Performance Evaluation of the developed model in the dataset

Dataset	MAE	RMSE
Training Set	2.251	4.482
Validation Set	.1927	3.954
Test Set	.9857	3.861

Table 2 shows that RMSE as a measure of predictive accuracy decreases the value from the training set to the test set which means that a smaller value indicates less difference between the estimated and actual values. Additionally, with the decrease of the RMSE, the predictive accuracy of the model improves.

5. Conclusion and Future Works

It is presented in this study a fuzzy based recommender system that would help senior high school students to consider factors associated to their career decision, but since numerous factors are needed to be considered, feature selection techniques are used to remove irrelevant student attributes so that the fuzzy inference model will generate reasonable results.

For further works, the proposed system needs to be evaluated for more test sets with the different set of criteria since the student attributes selected will be changed as the significant factors from real data will also be changed. It is concluded that a fuzzy based career recommender system is timely and will be useful for helping the senior high school students under the new education program which is the k to 12.

References

- [1] SEAMEO INNOTECH 2012 *K to 12 Toolkit*
- [2] Yap R H 2011 K to 12: the key to quality education? *Policy Brief PB-11-02* pp 1–12
- [3] Ng E S W, Gossett C W, Chinyoka S and Obasi I 2016 Public vs private sector employment An exploratory study of career choice among graduate management students in Botswana *Pers Rev.* **45(6)** pp 1367–85
- [4] Pitan O S and Adedeji S O 2015 Students' choice of courses : determining factors , sources of information , and relationship with the labour market demands in Nigeria *Africa Educ. Review* **11:3** pp 445-58 DOI: 10.1080/18146627.2014.934997
- [5] Laguador J M 2014 Examination of influence and intention towards Lyceum of the Philippines University and career choice of general engineering students *Int. J. Manag. Sci.* **3(11)** pp 847–55
- [6] Esther S U and Ayuk A A 2013 Templates that determines the choice of course selection in the university : a case study of University of Calabar Department of Sociology *J. Emerg. Trends Educ. Res. Policy. Stud.* **4(2)** pp 246–51
- [7] Eremie M D 2014 Comparative analysis of factors influencing career choices among senior secondary school students in rivers state Nigeria *Arab J. Bus Manag. Rev.* **4(4)** pp 20–5
- [8] Shumba A and Naong M 2012 Factors influencing students' career choice and aspirations in

- South Africa *J. Soc. Sci.* **33(2)** pp 169–78
- [9] Olamide S O and Oluwaseun S 2013 The factors determining the choice of career among secondary *Int. J. Engineering Sci.* **2(6)** pp 33–44
- [10] Gorad N, Zalte I, Nandi A and Nayak D 2017 Career counselling using data mining *Int. J. Innov. Res. Comput. Commun. Eng* **5(4)** pp 0–5
- [11] Goyal P and Kukreja T 2015 Narrowing awareness gap by using e-learning tools for counselling university entrants *Int. Conf. Adv. Comput. Eng. Appl.* **11(4)** pp 847–51
- [12] Durosaro I A and Nuhu M A 2012 An evaluation of the relevance of career choice to school subject selection among school going adolescents in ondo state *Asian J. Manag. Sci. Educ.* **1(2)** pp 140–5
- [13] Shahiri A M, Husain W, Aini N, and Rashid A 2017 A proposed framework on hybrid feature selection techniques for handling high dimensional educational data *AIP Conf. Proc.* **1891** 020130
- [14] Acharya A and Sinha D 2014 Application of feature selection methods in educational data mining *Int. J. Comput. Appl.* **103(2)** pp 34–38
- [15] Anuradha C and Velmurugan T 2016 Feature selection techniques to analyse student academic performance using Naïve Bayes student academic performance using Naïve *Conf. Proc. Small Mediu Bus* 2016 January 19 pp 345–50
- [16] Regha R S and Rani R U 2016 Optimization feature selection for classifying student in educational data mining *Int. J. Innov. Eng. Technol.* **7(4)** pp 490–6
- [17] Almayan H and Mayyan W 2016 Improving accuracy of students' final grade prediction model using PSO *6th Int. Conf. Inf. Commun. Manag.* 16 pp 35–9
- [18] Maharani K, Adji T B, Setiawan N A, and Hidayah I 2015 Comparison analysis of data mining methodology and student performance improvement influence factors in small data set *Conf. Proc. Int. Conf. Sci. Inf. Technol.* 15 pp 169–74
- [19] Saedudin R R and Sutoyo E 2017 Attribute selection on student performance dataset using maximum dependency attribute *5th Int. Conf. Electr. Electron. Inf. Eng.* 17 pp 176–9
- [20] Shanmugarajeshwari V and Lawrence R 2016 Analysis of students' performance evaluation using classification techniques *Int. Conf. Comp. Tech. & Intell. Data Engi.* p 16
- [21] Zaffar M, Iskander S and Hashmani M A 2018 A study of feature selection algorithms for predicting students academic performance *Int. J. Adv. Comput. Sci. Appl.* **9(5)** pp 541–9
- [22] Wang S, Tang J, Liu H and Lansing E 2016 Feature selection *Encycl. Mach. Learn. Data Min. ed Samut C et al* (New York City:Springer Science Business) 101–1 pp 1–9
- [23] Haindl M, Somol P and Ververidis D 2006 Feature selection based on mutual correlation *Prog. in Patt. Recog., Image Analand Appl.* (Springer Link)
- [24] Oladokun V O and Oyewole D I 2015 A Fuzzy inference based decision support system for solving the university-course admission choice problem *Int. J. Comput. Appl.* **112(3)** pp 1–7
- [25] Samia B, Allel H and Aicha A 2014 Handling preferences under uncertainty in recommender systems *2014 IEEE Int. Conf. Fuzzy. Syst.* 14 pp 2262–9 doi:10.1109/FUZZ-IEEE.2014.6891823
- [26] Ari E, Vatansever F and Uzun A 2009 Student professional orientation using fuzzy logic rules and quality *6th Res. Expert Conf. "Quality 2009"* pp 719–24
- [27] Gupta M V, Patil P, Deshpande S, Arisetty S and Asthana S 2017 FESCCO : Fuzzy expert system for career counselling *Int. J. Recent Innov. Trends Comput. Commun.* **5(12)** pp 239–43
- [28] Razak T R, Hashim M A, Noor N M, Hazwam I and Halim A 2014 *Career Path Recommendation System for UiTM Perlis Students Using Fuzzy Logic* 14 10.1109/ICIAS.2014.6869553
- [29] Banswal R and Madaan V 2016 SPACS : Students' performance analysis and counseling system using fuzzy logic and association rule mining *Int. J. Comput. Appl.* **134(3)** pp 12–7