Decision Support System for Determining: Right Education Career Choice

Ankit Mundra[1], Ashutosh Soni[2], Sunil Kumar Sharma[3], Pardeep Kumar[4] and Durg Singh Chauhan[5]

[1,2,3,4] *Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat, Disst. Solan, Himachal Pradesh, India.*
[5] *GLA University, Mathura, Uttar Pradesh, India.*
e-mail: [1] *ankitmundra8891@gmail.com*

Abstract. In today's education environment, where it is very hard to choose your career option in a right way. Currently there is an increasing interest in Machine learning and its role in educational systems, making this as a new growing research area. Machine learning practices can provide significant contribution by providing support to users for opting the right education domain to shape their career. In this paper, we have proposed education decision support system model, which compromises the components i.e. user interface, inference engine and knowledge base. Our model supports active machine learning type techniques to provide a knowledge discovery pattern, thus it also includes the knowledge of several domain experts. Our proposed model generates the output results (career choice) by taking several past studies, records, student's information and current education environment into consideration with the help of respective domain experts. Further, we have evaluated the performance of our model by considering several test cases.

Keywords: Expert system, Rule based engine, Knowledge base.

1. Introduction

Career decision making is the most important process of individual's career development [1]. Today, for right career choice students and their parents go to the councillors or academicians for getting their help by the experiences. Sometime this strategy may be helpful but most of the time due to lack of knowledge and ability about the student, right career option cannot be suggested [2]. However, the wrong career choice or the inappropriate career choice affects the future of students and as well as the development of the nation.

According to the survey of Ministry of Human Resource Development of India, Higher education is the powerful tool to build knowledge-based society of the 21st Century [3]. The survey report shows that during the last four decades the number of students enrolled in higher studies has grown up to 12 times (Table 1).

Figure 1 shows the rate of the subjects opted by the students for their higher studies. It clearly defines that students are more and more interested to take engineering and arts field as the career option.

Table 1. Student enrollment over last 4 decades [3].

Years	Student Enroll (in Millions)
1970–80	2.0
1980–90	2.8
1990–2000	4.9
2000–10	12.5
2010	25.9

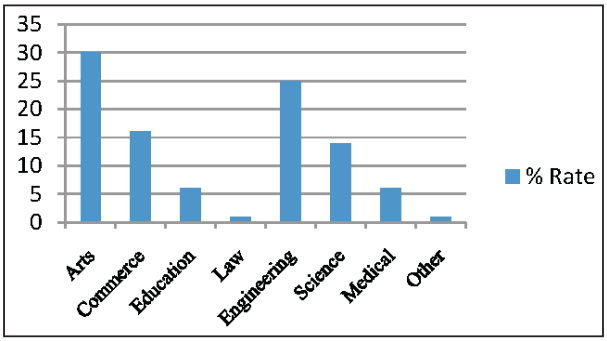

Figure 1. Enrollment by field of study (2011–12).

But, various other steams and fields are available which also have significant exposure and less competition. It is required to suggest those career options (other than arts, engineering) to the students so that they can shape their career in a non-traditional or less competitive environment and also related to their ability or interest.

In order to provide a right advice to the students we have proposed a model i.e. Career Advice Model *(CAM)* which is based or machine learning techniques and used a rule based decision support system to generate the expert knowledge.

1. User Interface: This is the first component of *CAM* through this user interact with the system. This consist of two modules one for user's personal skills and information processing and second for the personality analysis.
2. Inference Engine: This is the information processing component of *CAM* and having rule based decision support capability.
3. Knowledge base: This is the core component of *CAM* which consists of several knowledge sources for supporting the inference engine so that new knowledge is discovered from the previous knowledge patterns.

The layout of this paper is as follows: First section introduces the need of decision support system in higher education career option choices and brief about proposed model (*CAM*). Further, section second elaborated the proposed model and its working along with system design. Section third shows the JAVA based interfaces of *CAM* and evaluates the performance of *CAM* using certain test cases. Lastly, section fourth concludes this paper followed by references.

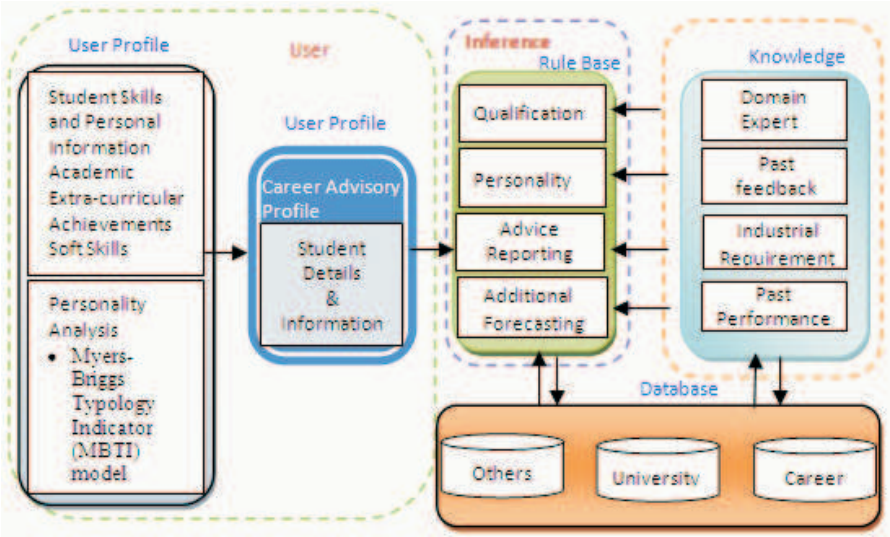

Figure 2. *CAM* architecture.

2. Proposed Approach

In this section we have elaborated the proposed *CAM* approach and its working process for generating the advice report to the user for his/her career guidance. As discussed in section 1 *CAM* approach start with the user interface module by taking the user information and skills into consideration. Then through inference engine and communicating with knowledge base *CAM* generates the advice report to user. Figure 2 shows the architecture of *CAM* approach which consist three main modules (section 2).

2.1 *Module 1*

This is the user interface module and consists of three parts that are described below.

2.1.1 Student skills and personal information

This module is to acquire student's attribute that are academic abilities, personal qualities, skills which a student require to inform system. In this phase, data regarded to student academics and skill is collected for assessment. Data is collected under two categories. First category is academic abilities. this covers three domains i.e. i) reading literacy; ii) mathematical literacy and iii) scientific literacy [7]. Reading literacy is taken in broad sense. Reading literacy stresses the interactive environment of reading and the constructive nature of comprehension. In constructing meaning, the reader uses various processes, skills, and strategies to promote, monitor and uphold understanding. The mathematical literacy domain is related with the ability of students to draw upon their mathematical competencies to meet the challenges of the future. Scientific literacy denotes the degree to which student is inclined towards physics, chemistry and biology. Assessment of student in these domains is performed through asking academic records of the student in high school and senior secondary. Apart from records, various tests are conducted in series

to retrieve details of the student needed to analyse. The outcome of this process is the basic profile of knowledge and academic skills.

Second category is of personal skills, hobbies and personal preferences. Since motive of our approach is not only to advice one primary career choice but it comprehensively guide user to choose secondary task including training, part time job, Social activities participation. Our approach emphasis on mapping secondary task with users personal skills.

2.1.2 Personality analysis

Personality of person is description of person's behavioral and attitudinal response patterns or combination of features and attributes that define the person as being individual that falls into certain category. In CAM, Myers Briggs Type Indicator (MBIT) model is employed to identify the personality type of the person [8]. This model demonstrates that we have an natural favorite choice for using our mind in one of two different ways further divided into four different category. First one is "orientation to world" that may be Extraverted or Introverted. Second is "take in information "that may be sensing or intuition. Third is "make decisions" that may be thinking or feeling. Fourth is "take info or decide" that may be perceiving or judging [8]. The working of module is as follow:

i) A test is organized for student to appear in structured according to MBIT model.
ii) Inference engine accepts the test data as input along with personality test knowledge base. Using fuzzy logic, inference engine generates the personality type of the user.

2.1.3 User profile

Every user accessing CAM has Career Advisory Profile (CAP). To start interaction with CAM, user registers him/herself. Registration process creates the user profile where user gives the required information. Having registered once, user can access the profile anytime. User personality type, after appearing in personality test, is stored in the CAP. Similarly Student's academic record, skills along with the result of test conducted in the student skills and personal information phase.

2.2 *Module 2*

This module consist of Inference engine which is responsible for generating the advice report by applying the rules and taking knowledge base into consideration.

As we have discussed expert system in our approach is rule-based which refer knowledge in the form of rule for problem solving. Rules have been represented as relations, recommendations, directives, strategies and heuristics [9]. We have included all five essential members of the expert system development team: domain expert, knowledge engineer, programmer and end users [9]. Modular approach has been incorporated in design of inference engine that make sure that changes in the system can be embedded with minimum effort.

Functionality of the inference engine has been divided into four modules: Qualification forecasting, personality forecasting, Advice reporting, Additional forecasting shown in Figure 1.

2.2.1 Qualification forecasting

Qualification forecasting module has been associated in inference engine to predict future performance using Heteroscedastic models of linear regression. Our approach anticipate to facilitate career option for both job and higher studies defined as sperate categories.

Taking heuristics and assumption in consideration, we have assigned a threshold value to each subject in which a category includes one or more subject.

- The eligibility criteria for the student weather he/she eligible for the certain subject is checked on the basis of the threshold (shown below) associated with the subject .As student makes the choice of the career, all the subject fall into the category are fetched.

$$\text{If Student\_GPA} \geq \text{Subject\_Threshold} \tag{1}$$

  Then Student is eligible for Subject and Proceed Further.
- Each subject falling into the career option will have the some impact. Subject has been assigned some weight. Our approach has embodied soft skills, required for subject, of the student that is recorded in database after test conducted in Student skills and Personals Record module. So the impact of each subject on career category is evaluated by using following formula (equation 2).

$$\text{Subject}_{\text{Impact Factor}} = \text{Subject\_Predicted Value} * \text{Subject\_weight} + \text{Subject\_SoftSkill} \tag{2}$$

- Career option category assigns certain Impact factor to subjects that come within the career option. So the calculated Impact factor of the subject is check against assigned Impact factor as follows:

$$\text{If Subject}_{\text{Impact factor}} \geq \text{Assigned}_{\text{Impact factor}} \tag{3}$$

  Then proceed else Send Advisory Report to concerned Student.

2.2.2 Personality forecasting

Personality analysis is a part of inference engine that provides personality type of the student. Personality forecasting module uses this information to match the appropriate personality type with the industrial and academic requirement. Qualification forecasting module obeys eligibility criteria that emphasizes on non interactive attributes i.e. academic record. In order to make career advice process more close to what student really is, we utilizes expert system concept to cast career advice to person matching with his/her personality.

2.2.3 Advice reporting

Advice reporting module is the part of the inference engine that takes the input from the student skills and personal information module, qualification forecasting and addition forecasting module to generate advice for student. Advice reporting process is of two type i.e.

a. Career advice

Career advice report (shown in Figure 3) is governed by the inference rule. Motive of the report is to suggest the career option or approving student's opted career choice. Report also insures that if user has not gained the approval from the qualification forecasting then its advices student to improve on the subjects whose impact factor has value below required value assigned to the subject. Format of advisory report is given below. Career advice report has three sections. Section one is suggested career option. This section names the career option suitable for the concerned student. It also includes the reason why an option is opted for. At last, it has like/dislike column that student can opt to like or dislike option. Section one is improvement section listing the entire subject that student needs to improvise on. Last section is Student's feedback and Suggestions.

Advise report		
Student Name - Registration Number - Report Generation time & date -		
Suggested Career Option		
Option List	Reason	Like /dislike
Improvement Section		
Subject List	Improvisation %	
Student's Feedback and Suggestions		

Figure 3. Advice report format.

Additional Report		
Student Name - Registration Number - Report Generation time & date -		
Option List	Reason of Suggestion	Acceptance
Student's Feedback and Suggestions		

Figure 4. Advice report format.

b. Additional report

Additional report (shown in Figure 4) is special reports that retrieve its data through inference rule from Addition Forecasting Module which is discussed letter. As mentioned before, our approach does not only suggest primary career options but some additional options as per user demand. If a student demand any part time option like part time jobs, training, addition report suggest these options. these suggested option are chosen in such a way that there time of occurrence in future does not conflict with Primary suggested career option like higher study or job that student has already chosen. This is take care by the scheduler in Additional Forecasting Module.

2.2.4 Additional forecasting

Additional forecasting module has been designed for additional help of the student. Since last decade, evident changes have been noticed in choice of career by students. They are not satisfied with single career choice. In today's competitive environment, student tends to opt multiple options, utilizing time and exploiting talent, skills and opportunity leading to better and safe future. Understanding these requirements, we have proposed additional forecasting module that takes care of student extra demands. These extra and optional career options are denoted by Secondary career choices. Module works in following manner:

- Once Primary career option is chosen for student then Secondary option comes under consideration.
- In this module we fetched student skills set form database that are ignored or not undertaken in Primary career option choices. We also took student's hobbies, personal preferences to build up options for secondary career choices.
- Secondary career options are two types first is training. Our approach gives proper support to choose right training option that student can bind him/her with no conciliation. Second option is part time jobs. By help of inference rule, guided by forward chaining, module derive the part time options that lives up to the capability of student judged by test conducted in Student Skills and Personal Information module.
- All the evaluated available option are checked to see that there time of occurrence should not conflict with Primary career options.

2.3 *Module 3*

It consists of knowledge base based on four parts i.e. domain expert knowledge; past feedback collection; industrial requirements and demand; past performance of career choices. We have explained each part in detail here below.

2.3.1 Domain expert knowledge

As our model supports Active Machine Learning type, it has taken the knowledge of the several domain experts and their experiences into consideration. For this we have contacted the domain/field experts and gathered their experiences and feedback about career choices. Each time when a new user demands the career advice the inference engine contact to domain expert knowledge base and perform the forecasting accordingly.

2.3.2 Past feedback collection

We have collected the feedback about each career choice from the students who have previously taken the subjects. This feedback is taken on the basis of certain question sets and students answers those question on the behalf of their experience. This helps the inference engine for the forecasting.

2.3.3 Industrial requirements

In our *CAM* approach, in order to support the user in a more efficient way we have gathered the industrial requirements and demands and accordingly our system suggests the career choice to user. Further, in our knowledge base the industrial requirements are updated regularly and collaborate with user profile.

2.3.4 Past performances of career choice

In order to advice a right choice *CAM* maintains a knowledge base from past performance of available career choices. This performance is measure according to the Demands and job success rate which is offered by respective career domains during past years.

Further in this module consist of additional databases i.e. career option database which stores the information of available career option in higher education; university database which contains the information of the respective universities which offers the available courses; other database contains the information regarding past students, additional subjects offered by career domains etc.

2.4 *Rule Based Active Learning*

CAM is Rule-Based expert system but it is also an active learner, enriching its own knowledge base to suggest career option precisely. We have mentioned that Domain expert are consulted and their advice is stored as the knowledge in knowledge base. As a active learner, *CAM* learns from the Process of decision of Domain expert.

For incorporating active learning, we have deployed Instance-based learning method which is statistical method. Instance-based is lazy learning algorithm, hence required less computation time. IB methods [8,9] works on the principle that instances having similar properties and attributes will present in close proximity. IB saves all training instances in conceptual description that will be finally saved in the

knowledge-base. We have chosen IBL3 method of instance based learning that emphasize on saving significantly good instances.

In our approach, advices that Domain expert has suggested are known instances whose class label is given [10]. These instances are tagged with classification label. Classes have been made considering student personality, academic record, results of test conducted and soft skill. Advantage of having the active learning algorithm in addition to rule base system is following.

- After having the end result in the form of the advice and proposed carrier choice generated by the system or as the advice of the human expert, active learning algorithm captures and observe the pattern and the procedure of the decision to predict advice for the student which whose profile has the similarity.
- Active learning automates the process of the prediction with less commitment from the human expert. It saves the time as it is more likely to have the number of profile to have the similarities. Active learning can recognize what profile can be of same type and how close these profiles are with each other.

```
PA ← φ
PA ← {known Advices}
for each Testing sequence a in Training Set do
 /* testing sequence is record of the student seeking for career advice */
    for each Known Advice in PA do
        Resemblance (b) ← Relevance (a, b)
/* a shows record of student seeking for career advice.  b shows record  of
student that has been given advice   */
    if each b ∈ PA is acceptable
    then  bmax  ←acceptable  b with maximum Resemblance(b)
/* acceptable shows most similar student record that has been advised by the
Doman Expert */
    else
        i ← a random value between 1 and |PA|
        bmax ← some b ∈ PA that is i-th most similar to x
/* selection of the advised records that is most close to seeker student record */
   if class(a)=class(bmax)
    then classification ← right
    else
      classification ← wrong
/* checking weather predictable class is right or wrong */
    for each b in PA do
    if   Resemblance (b) ≥  Resemblance (bmax)
    then
        Change b's classification record
        if b's record is significantly poor
        then PA ← PA – {b}
/* CAM-IB3 discards the noisy saved instance with significantly poor
classifications */
```

Relevance(a,b)

return $\left(\sum_{i=0}^{n}|a_i - b_i|^2\right)^{1\backslash 2}$

3. Implementation and Evaluation

In this section, we have shown the JAVA module implementation of the *CAM* approach. We have created the JAVA forms for each *CAM* module. Further, in order to evaluate the performance of *CAM* we have

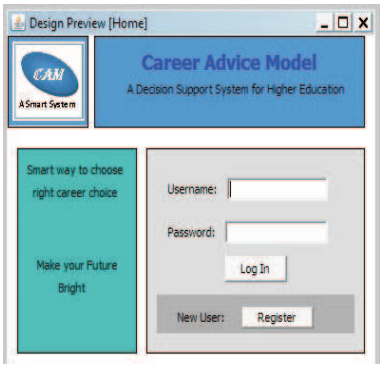

Figure 5. *CAM* home page.

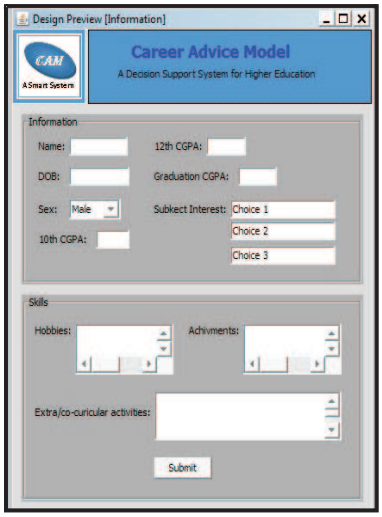

Figure 6. *CAM* home page.

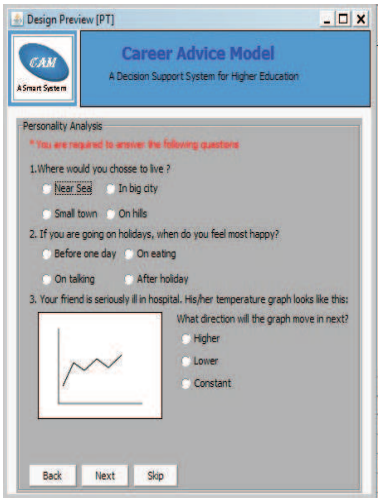

Figure 7. *CAM* home page.

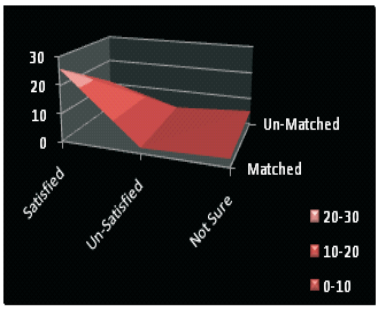

Figure 8. Test tesults.

implemented the test bed for 50 users and asked them to provide their valuable feedback. Here, Figure 5 shows the Login form for existing users and also provides the registration facility for new user. Existing users (who have already registered) have their username and password respectively.

Figure 6 shows the form which accepts user information and skills and makes it available for inference engine to process. Figure 7 shows, sample form for personality test conducted by *CAM* in order to perform personality analysis of the user. After, getting all the details *CAM* performs forecasting and the result of the same is generated in an advice report. Further, the format of advice reports has already been defined in previous section (Figure 3 and Figure 4).

Table 2. Comparison between previous approaches and *CAM*.

Functionality \ Approach	CGM[5]	iadvice[6]	CAM
Personality Identification	Yes	No	Yes
Fuzzy Logic	Yes	No	Yes
Rule-based expert system support	No	Yes	Yes
Active learning	No	No	Yes
Secondary career advices	No	No	Yes
Qualification forecasting	No	Yes	Yes

To check validation and acceptance in real world, we have conducted a evaluation process. Evaluation has been performed by taking 50 students in consideration. Motive of the Evaluation was to check the user's satisfaction level on the advisory report and addition report generated by *CAM*.

Result of the evaluation process has been shown in Figure 8. Advisory report's acceptance has been categorized in two levels. First one illustrates the level of satisfaction and second one shows the user's acceptance to choices of career given by CAM. Satisfaction levels are of three kinds: Satisfied, Non-Satisfied, Not Sure. Satisfied students are those who have accepted the advice of CAM.

User acceptance to CAM choice is of two type: Matched, Unmatched. Matched denotes that user whose career choice matches with the CAM choices. Unmatched denotes those whose career choice not matched with CAM choices. We have also done the comparison study between *CAM*, *i* advice and Career guidance Model (*CGM*).

4. Conclusion

In this paper we have proposed an approach which supports decision making in higher education system. *CAM* is a very useful approach for students who want to shape their career in a right way and egger to achieve success in the field which is made for them. We have also implemented the *CAM* approach in a JAVA based application and shown the effectiveness of our model.

References

[1] www.pearsonhighered.com/samplechapter/0131712772.pdf
[2] Patton, Wendy and Mary McMahon, "Career development and systems theory," Vol. 1, Sense Publishers, (2006).
[3] Ministry of Human Resource Development of India, Higher Education in India: Twelth five year plan (2012–2017) and Beyond, FICCI Higher education summit, (2012).
[4] Bresfelean, Vasile Paul and Nicolae Ghisoiu, "Higher education decision making and decision support systems," *WSEAS Transactions on advantages in engineering education*, Vol. 2, 43–52, (2009).
[5] Winston, Ojenge and Muchemi Lawrence, "Career guidance using expert system approach," *Strengthening the Role of ICT in Development*, 123, (2008).
[6] Hendahewa, Chathra, *et al.*, "Artificial intelligence approach to effective career guidance," *Sri Lanka Association for Artificial Intelligence*, (2006).
[7] Grabe and William, "Reading in a second language," Ernst Klett Sprachen, (2010).
[8] Briggs and Katharine C., "Myers-Briggs type indicator," Palo Alto, CA: Consulting Psychologists Press, (1976).
[9] Abraham and Ajith, "Rule-based expert systems," Handbook of measuring system design, (2005).
[10] Ankit Mundra, Bhagvan Krishna Gupta, Geetanjali Rathee, Nitin Rakesh and Vipin Tyagi, "Validated real time middleware for distributed cyber physical systems using HMM," International Journal of Distributed & Parallel Systems 4, no. 2, (2013).