



Myers-Briggs Type Indicator and Social Media Mining Based Career Counselling System

Mohd Sahid Khan¹, Dr Manuj Darbari²

¹Research Scholar, Nirwan University, Jaipur, Rajasthan, India.

²Professor, Department of Computer Science, School of Engineering, Babu Banarasi Das University, Lucknow, India.

Abstract – One of the dilemma of student life is choosing the right career. It is more so for those students who belong to underprivileged section and whose parents and the community itself is not generally aware of the various options available. Even if they are aware of it, they may not know how to choose the right one. We have proposed a Myers-Briggs Type Indicator (MBTI) and Social Media mining-based career counselling system. Explicit and Implicit inputs of the student is taken as input and the system recommends the best career option. The necessary information about the student's personality traits and interest is obtained through Social Media mining and traditional MBTI questioners.

Keywords: Recommender System, MBTI, NEP 2020, Facebook Mining, Career Counselling.

1. INTRODUCTION

National Education Policy – 2020 (NEP) of India has given "Flexibility" and "Extensive use of technology in education planning and management" as a part of its Fundamental Principles. Flexibility, so that learners have the ability to choose their learning trajectories and programs, and thereby choose their own paths in life according to their talent and interest. Another principle is extensive use of technology in education planning and management. In this age of Internet of everything, the available options and the explosion of information is major contributing factor in the dilemma of choosing the right career option. Thus,

using technology in career planning to achieve the degree of flexibility envisaged in NEP is need of the hours. Various expert systems, recommender systems have been developed in the past. In [1] Anusha et.al has developed a rule-based career advisory expert system named iAdvice taking into consideration past examination performance, student's preference and skills. In [2] Faliagka et.al implemented a system that model the candidate CV in HR-XML and rank the candidate based on Analytic Hierarchy Process(AHP). In [3] Saraswathi et.al develops a rule-based system which takes input from students and a knowledge-based system through pattern matching and jsoup java HTML parsing technique. In [4] Razak et.al measures student's skill strengths, abilities and personality facets and recommend them with possible career choices by using fuzzy logic approach. In [5] Abisoye et.al have developed a web-based career guidance using data collected from interviewing human counsellors and survey. A holistic approach in capturing the accurate input matrix of personnel information, academic performance, aptitude, personality trait, interest, influence of role model, parents and environment is required to recommend the best choice. In a survey in [5], it was found that 66% of the respondents realized that they have chosen a wrong career. This could be due to peer pressure, family and society pressure or more pertinent in India is due to non-availability of correct advice in the villages and remote areas. Thus, too many information and no information both ways necessitate an automatic online Artificial Intelligence based solution. The penetration of



Mobile and Social media in the remote villages should be exploited in enabling the masses towards our goal of India being a Developed country. The proposed system considers the holistic input approach where explicit and implicit inputs of the students is taken and recommended the best choice of career.

2.RELATED WORK

The various recommender and mining system have been proposed based on rule, business logic, fuzzy logic adaptive neuro fuzzy inference system. Some of them are given below: In [6] Raj Kishor Bisht proposed a framework in which student's capabilities is analyzed and jobs related to their capabilities are suggested. The student's marks in various subjects and interest in different domains have been taken into a fuzzy logic-based system. In [7] A. Lwayemi et.al utilize MBTI traits, mapping to suitable careers in a rule-based system based. In [8] Verma et.al utilized preference, fuzzy logic, and influence to integrate with Analysis Hierarchical Process (AHP) decision system. In [9] Ghalia et.al proposed a rule-based system for academic guidance to undergraduate students. In [10] Peker et.al proposed a Web Based Career Guidance System (WEB-CGS) using fuzzy logic. In [11] Gorad et.al designed an online system using data mining algorithm C5.0. The students were suggested a career option based on personality attributes. In [12] H. Takci et.al uses different data mining methods and classifiers for a system that will guide the student in selecting a career. The comparisons of various data mining algorithms, different data set results, and models used are also discussed. In [13] Gupta M V et.al suggested an AI based fuzzy expert system. In [14] Razak et.al explore the use of HFSs (Hierarchical Fuzzy System) for an expert system. In [15] Natividad et.al proposed career recommender system that uses different filter for selecting the best attributes, which are used as inputs. In [16] Faraaz Ansari et.al develops an AI based expert

system (chat-bot) for giving possible career opportunities to students. In [17] Tehseen Mehraj and Asifa Mehraj Baba scrutinized numerous contemporary career counselling platforms based on AI techniques. They found all the solutions having loopholes in one context or the other, lack of understanding of factors which guide the decision of career selection, and Non-existence of a proper mechanism to identify the personality traits of an individual and suggested solutions based on deep learning to make them more intelligent and efficient. In [18] Bikse et.al uses the Analytic Hierarchy Process (AHP) and was used to analyze the possible development scenarios of career education in Latvia. [19], proposed an Adaptive Neuro Fuzzy Inference System (ANFIS) to provide guidance and counselling for students. MATLAB was used to develop the ANFIS component. Java was used to design the Graphical User Interface (GUI), and to implement the logic behind the counselling module of the system. In [20] Hosam F et.al designed a Fuzzy logic-based system for students with interactive tools for selecting suitable colleges matching their educational skills. In [21] Comendador et.al implemented a fuzzy logic technique to develop a mobile based decision support tool called Generalized STEM-College Aptitude Test (GSTEM-CAT). The system give recommendation for appropriate university program to be enrolled suitable to the individual personality and knowledge. In [22] Ankush Daharwal et.al have developed a user-friendly website to give a test. The answers given by the user are used to store a classifier model for prediction. Already developed code for classifier prediction is appended with the input given by the user. Naïve Bayes classifier gives three recommendations based on the top three probabilities. The output is displayed to the user. The user can check the recommended Career and can choose the best among three choices. The Career with the highest probability is displayed first

then second and at last third. A feedback system was also included. In [23] Bhuvneswaran et.al develops an application that helps the Graduates from various streams, explore their interests and map their interest to the available professions. The application uses Machine learning algorithms to research the user’s interests and strengths through a series of Psychometric tests and list out the highest matching career options. In [24]Awoyelu et.al , proposed a system using the concept of fuzzy logic. Crisp sets are converted to fuzzy sets using fuzzy K- nearest neighbor algorithm method. The model was implemented in the MATLAB environment. In [25] Manoharan et.al proposes fuzzy logic approach for predicting users’ diversified interest and its categories by analyzing their implicit user profile. Depending on users’ interest categories, the viral news articles and their categories were determined and analyses through mining social media feeds–Facebook and Twitter. Furthermore, fresh news articles are retrieved from news feeds incorporated with retrieved viral news articles provided as recommendation with respect to users’ diversified interest.

Most of these systems failed in taking a holistic approach of deciding the optimal explicit and implicit attributes of the students. Secondly, the concept of social media mining for the most suitable and relevant information with respect to the personality traits and area of interest of the student is not found included.

3. THE PROPOSED FRAMEWORK AND METHODOLOGY

The proposed framework deals with processing of query from the user and then predicting the right career option to the person. In order to predict and guide the perfect career path authors have used number of refinement techniques. The first technique we used to extract the information is MBTI technique known as Myer Briggs Type Indicator. MBTI classifies the personality of individual in four categories.

Category I : Extraversion (E) or Introversion (I)

Category II : Sensing(S) or Intuition (N)

Category III : Feeling (F) or Thinking (T)

Category IV : Perceiving (P) or Judging (J)

By combination of all the above personality traits we can decide on the individual personality which an individual has and his inclination towards the liking of the individual goals.

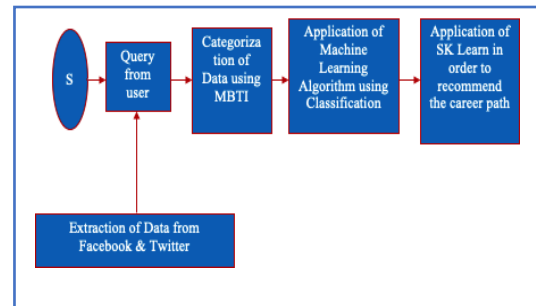


Fig -1: Recommender System Based on Facebook and Twitter Comments

Recommendation system has an everlasting impact on our daily life be it small issues like waking up in the morning and deciding the priority of activities to do. There has been tremendous growth in the field of recommender system with the development of Artificial Intelligence (AI) systems into our daily life. Social Media is also playing a significant role as people especially Young Generations like to post about their liking/disliking on twitter as well as Facebook, these two provide an open forum for posting ideas/opportunities. Based on one’s posting one can trace out his category in terms of MBTI format and then fed it into classification algorithm to deduct the necessary deductions.

3.1 Data Collection

In order to illustrate the working of Data collection we have taken snapshot of Facebook and Twitter posting by and individual. The data we get from Facebook and Twitter needs to be scrubbed i.e. we refine the data set before making it workable. Seeing the post, the person responded.

Table -1: Posts and its associated responses.

Serial	Type of Posts	Response	Comment keyword
1.	Engineering Innovation	Likes	Excellent, Good, Knowledgeable, More Info Desired
2.	Engineering Forums	Likes	Gives Comments of Discussion
3.	Medical Forums	Likes	Interacts by asking Questions/Suggestions
4.	Movie Forums	Likes/dislikes	Romance, Adults, Comedy
5.	Administrative Services Forum	Likes/Dislikes	Preparation, Roles, Responsibility
6.	Arts/Painting	Likes	Interacts by asking the ideas behind it
7.	Armed Forces Forum	Likes	Give Inspirational comments
8.	Social Service Forums	Likes	Encouraging Comments/Suggestion
9.	Academic Forums	Likes	Interact by asking questions
10	News	Like/dislike	Politics, Legal, Technical, Entertainment, Social issue.

Armed Forces, Academicians, Politician and Arts (performing and nonperforming).

3.3 Encoding

We have used one hot encoding method, in which text-based values are transformed into binary form represented by '1' or '0' for 'True' or 'False' respectively. The following table indicate the observation of the individual interest and its encoding.

Coding formula used is: Engaged with the Forum by Liking, Commenting or Sharing = 1
Not Engaged with the Forum = 0

Table -2: Encoding of Individual Interest

Sno	Name	Forum/Page/Domain Engaged by the Individual						Inference	
		Arts	Admin	Forces	Academician	Doctor	Engineer		Politician
1.	Ram	0	1	0	0	0	0	1	Admin & Politics
2.	Sam	1	0	0	0	0	0	0	Arts
3.	Ravi	0	0	1	0	0	0	0	Armed Forces
4.	Santa	0	0	0	0	0	0	1	Politics
5.	Rita	0	0	0	1	0	0	0	Academician
6.	Pillai	0	0	0	0	1	0	1	Doctor & Politician
7.	Radha	0	1	0	1	0	0	0	Admin & Academic
8.	Tom	0	0	1	0	1	0	0	Armed Force & Doctor
9.	Abdul	0	1	0	0	0	1	0	Admin & Engineering
10.	Max	0	0	0	1	0	0	0	Academician

The comment and post by the individual in Social Media reflects his personality. It has been extensively studied by various people. Thus, we can relate the comment and post with the relative MBTI personality traits. The screenshot shown below is from the Kaggle repository.

(<https://www.kaggle.com/datasnaek/mbti-type>).

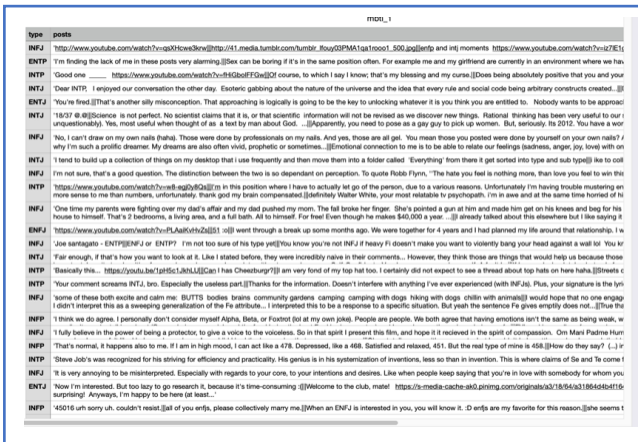


Fig -2: MBTI Personality Type and Comment/Post

3.2 Feature Selection

To identify the most relevant features to our objective is very important. We are preserving only the features that correlate strongly with the output. We have also restricted our stream into seven professions only to avoid greater application of time and effort. The seven professional streams selected are Doctor, Engineering, Administrative Services,

3.4 MBTI Personality Traits through Questioners

The questioners of MYERS-BRIGGS TYPE INDICATOR by Katharine C. Briggs & Isabel Briggs Myers downloaded from <https://www.16personalities.com> can be used to find out the individual personality of the students out of the 16 possible traits.

3.5 MBTI Personality Traits and Career Options

The type of personality traits and the corresponding suitable career option already available in literature. Number of studies and books are written on these aspects. The following table provide an illustration

based on the book, "Do What You Are", written by Paul Tieger et.al [26]

Table -3: MBTI Personality Traits and Career Matrix

SNO	Type	MBTI Code	Career	Career Code
1.	ISTJ	16	Armed Force	1
2.	ISFJ	15	Politician	2
3.	ISFP	14	Arts	3
4.	ISTP	13	Engineer	4
5.	ESFP	12	Therapist	5
6.	ESTP	11	IAS	6
7.	ENFP	10	Entrepreneur	7
8.	INFJ	9	Social Worker	8
9.	INFP	8	Professor	9
10.	ENFJ	7	Consultant	10
11.	INTP	6	Research	11
12.	ENTJ	5	Doctor	12
13.	ENTP	4	Strategic Planner	13
14.	INTJ	3	Scientist	14
15.	ESFJ	2	Legal Assistant	15
16.	ESTJ	1	Lawyer	16

3.6 Decision Matrix

The three inputs which are obtained from different methods are used as input to the decision matrix. The inputs are, Social Media mining based individual interest, individual personality traits predicted through Social Media and the MBTI questioners-based personality traits. These are utilized to give the final recommendation of the career option. The random forest algorithm is used to obtain the result. An illustration is represented diagrammatically below:

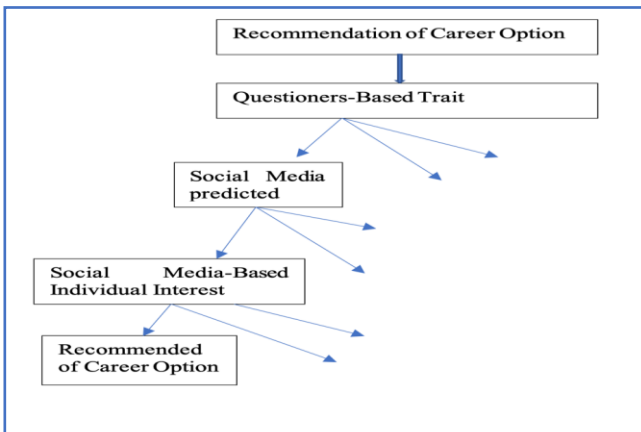


Fig -3: Career Recommendation Decision Tree

4. RESULT AND DISCUSSION

The following table [Table:4] represent the recommender and actual career of ten individual under study. The Serial numbers is taken as the code of the profession. Using ordinary least square regression, we found out the regression line. Our dependent variable was Career, we create a variable y as Career. The following are the variables:

$$y = \text{data}[\text{'Career'}]$$

$$x1 = \text{data}[\text{'MBTI'}]$$

Table -4: Recommended vs Actual

SNO	Name	MBTI Trait	MBTI Code	Recommended Career		Actual Career	
				Category	Code	Category	Code
1.	Ram	ISTJ	16	FORCES	1	DOCTOR	12
2.	Sam	ESTJ	1	LAWER	16	LAWER	16
3.	Ravi	ENTJ	5	DOCTOR	12	POLITICIAN	2
4.	Santa	INTP	6	RESEARCH	11	PROFESSOR	9
5.	Rita	INFJ	9	SOCIAL WORK	8	SCIENTIST	14
6.	Pillai	ISFP	14	ARTS	3	ARTS	3
7.	Radha	ISFJ	15	POLITICIAN	2	POLITICIAN	2
8.	Tom	INTJ	3	SCIENTIST	14	SCIENTIST	14
9.	Abdul	ISTP	13	ENGINEER	4	ENGINEER	4
10.	Max	ESTP	11	IAS	6	IAS	6

Based on the input MBTI traits we are predicting the Career options. Our interest is only in the MBTI and the recommended career. Our model is given as

$$\hat{y} = 17.0 - 1.0 * x1 \text{ OR Career} = 17.0 - 1.0 * \text{MBTI}$$

Table -5: OLS Regression Results

Dep. Variable:	CAREER	R-squared:	1.000
Model:	OLS	Adj. R-squared:	1.000
Method:	Least Squares	F-statistic:	Inf
Date:	Tue, 16 Feb	Prob (F-	0.00

	2021	statistic):				
Time:	19:41:58	Log-Likelihood:		Inf		
No. Observations:	10	AIC:		-inf		
Df Residuals:	8	BIC:		-inf		
Df Model:	1					
Covariance Type:	Nonrobust					
	Coef	std err	T	P> t 	[0.025	0.975
Const	17.0000	0	Inf	0.000	17.000	17.000
MBTI	-1.0000	0	-inf	0.000	-1.000	-1.000
Omnibus:	2122.865		Durbin-Watson:		Nan	
Prob(Omnibus):	0.000		Jarque-Bera (JB):		3.750	
Skew:	0.000		Prob(JB):		0.153	
Kurtosis:	0.000		Cond. No.		22.4	

The Scatter plot and regression line of the prediction is given below in [Figure: 4].

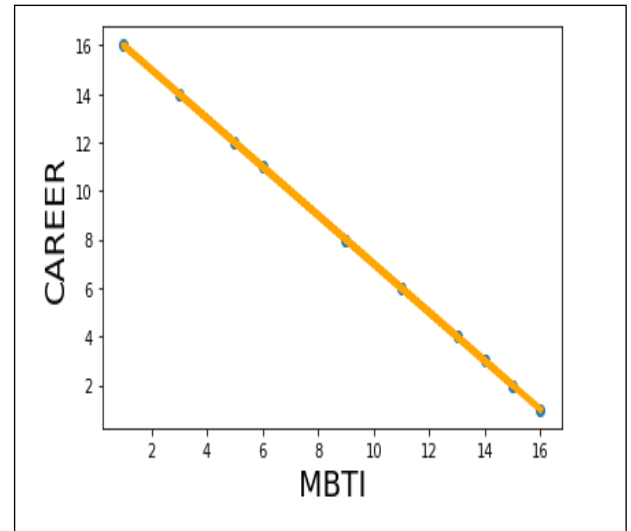


Fig -4: Scatter Plot and the regression Line

The plot above [Figure: 4] shows the scatter plot of x_1 vs y , which is given by the input dataset. The x_1 are the encoded values of MBTI Traits as determined by the refinement of the traditional MBTI traits determination method through social media factor. The y values are the encoded values of the recommended career based on MBTI and interest of the individual found out through mining of social media. The regression line is plotted between predicted encoded values of Career and the MBTI.

5. CONCLUSION

Choosing the right career is very important for a successful life. Most of the youngsters suffer with the dilemma of choosing the right career. It is more so because of the fast-changing world with multiple options available. Finding the option most appropriate with the type of personality trait always give and advantage. Moreover, the interest of the individual capture indirectly through Social Media usage habit of the Individual will add on to the suitability of the recommended career. In this paper we have proposed a career recommender system

In [Table : 5], the Adjusted R-squares indicates the fitting of the model. R-square values range from 0 to 1. A higher value generally indicates a better fit, assuming certain conditions are met. In our case we have a value of 1.000. The constant coefficient is 17.0000. MBTI coefficient represents the change in the output Y due to a change of one unit in the MBTI (everything else held constant). std err reflects the level of accuracy of the coefficients. The lower it is, the higher is the level of accuracy. Here it is 0. P >|t| is our p-value. A p-value of less than 0.05 is considered to be statistically significant.

based on MBTI personality traits of the individual and the individual interest. The personality trait is predicted through questioner as done by the Myers-Briggs, supplemented by Social Media based Personality Traits prediction. The interest of the individual is obtained through the Social Media usage habit of the individual.

REFERENCES

- [1] Anusha, R., Asoka, S., Chathra, H., Maheshika, D., Narmada, W., & Savindhi, S. (2006). Artificial Intelligence Approach to Effective Career Guidance. Sri Lanka Association for Artificial Intelligence (SLAAI). Proceeding of the third annual sessions, Colombo, pp. 32 - 42.
- [2] Faliagka, Evanthia & Ramantas, Konstantinos & Tsakalidis, Athanasios & Viennas, Emmanouil & Kafenza, Eleanna & Tzimas, Giannis. (2011). An Integrated e-Recruitment System for CV Ranking based on AHP.. 147-150. Research Gate.
- [3] Saraswathi, S., Reddy, M. H. K; Kumar, S. U; Suraj, M. and Shafi, S. K. (2014). 'Design of an Online Expert System for Career Guidance', International Journal of Research in Engineering and Technology, Vol. 3, Issue 7, pp. 314 - 319.
- [4] Razak, T. R; Hashim, M. A; Noor, N. M; Halim, I. H. A, and Shamsul, N. F. F. (2014). 'Career Path Recommendation System for UiTM Perlis Students Using Fuzzy Logic', 2014 5th International Conference on Intelligent and Advanced Systems.
- [5] O. A, Abisoye, I. Alabi, S. O, Ganiyu, O. Abisoye, and J. Omokore(2015), "A Web-Based Career Guidance Information System for Pre-Tertiary Institutions Students in Nigeria", the International Journal of Scientific Research in Science, Engineering and Technology, Vol. 1, No. 3, 2015.
- [6] Raj Kishor Bisht (2015)"A Fuzzy Based Conceptual Framework For Career Counselling", Advanced Computational Intelligence: An International Journal (ACIJ), Vol.2, No.4, October 2015.
- [7] A. Iwayemi, B. F. Oladejo, D. S. Adeleke(2016) "Career Advisor Expert System Based on Myers Briggs Personality Assessment" Department of Computer Science, University of Ibadan, Ibadan, Nigeria.
- [8] Verma, Prabal & Sood, Sandeep (2017). "Student career path recommendation in engineering stream based on three-dimensional model". Computer Applications in Engineering Education. 10.1002/cae.21822.
- [9] Ghalia, Khaled and Maryam, Juma A. (2017). "A Rule-Based System for Advising Undergraduate Students", Journal of Theoretical and Applied Information Technology, Vol. 95, No. 11, pp. 2453 - 2465.
- [10] Peker, M; Guruler, H; Sen, B; and Istanbulu, A. (2017). "A New Fuzzy Logic Based Career Guidance System: WEB-CGS", Technical Gazette 24, Vol. 6, pp. 1863-1868.
- [11] Gorad, N; Zalte, I; Nandi, A. and Nayak, D. (2017). "Career Counselling using Data Mining", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 4.
- [12] H. Takci, K. Gürkahraman and A. F. Yelkuvan, "Measurement of the appropriateness in career selection of the high school students by using data mining algorithms: A case study," (2017) Federated Conference on Computer Science and Information Systems, Prague, 2017, pp. 113-117, doi: 10.15439/2017F283.
- [13] Gupta M V, Patil P, Deshpande S, Arisetty S and Asthana S (2017) " Fuzzy expert system for career counselling", International Journal of Recent Innovations : Trends Computer Communication. 5(12) pp 239- 43.
- [14] Razak, Tajul Rosli & Abd. Halim, Iman Hazwam & Jamaluddin, Muhammad Nabil Fikri & Ismail, Mohammad Hafiz & Fauzi, Shukor. (2019). "An Exploratory Study of Hierarchical Fuzzy Systems Approach in A Recommendation System", Jurnal Intelek. 14. 10.24191/ji.v14i2.233.
- [15] Natividad, M & Gerardo, Bobby & Medina, Rujii. (2019). "A fuzzy-based career recommender system for senior high school students in K to 12 education ". IOP Conference Series: Materials Science and Engineering. 482. 012025. 10.1088/1757-899X/482/1/012025.
- [16] Faraaz Ansari, Shaikh Saad, Shaikh Shareem (2019) "Career Counselling Chatbot" International Research Journal of Engineering and Technology (IRJET), e-ISSN: 2395-0056 Volume: 06 Issue: 04.
- [17] Tehseen Mehraj, Asifa Mehraj Baba (2008), "Scrutinising Artificial Intelligence based Career Guidance and Counselling Systems".
- [18] Bikse, Veronika & Libkowska, Una & Rivza, Peteris & Rivza, Baiba. (2013). "Using AHP for the Evaluation of the Development of Career Education in Latvia". 10.13033/isahp.y2013.089.
- [19] Ademiluyi, Oluwadamilola, Ukaoha, Kingsley and Otokiti, Kareem (2020), "A Neuro Fuzzy-based Guidance and Counselling System for Students", Afr. J. MIS, Vol. 2, Issue 1, pp. 39 - 54.
- [20] Hosam F. El-Sofany, Samir A. El-Seoud.(2019) "A Cloud-based Educational and Career Guidance Model using Fuzzy Logic Concepts". The 2nd. International Conference on Network Technology (ICNT 2019), and 8th International Conference on Software and Information



Engineering (ICSIE 2019). Cairo, Egypt on 4-6 May 2019. <https://doi.org/10.1145/3328833.3328846>.

- [21] Comendador, Benilda & Becbec, Wilmarie & Guzman, John. (2020). "Implementation of Fuzzy Logic Technique in a Decision Support Tool: Basis for Choosing Appropriate Career Path", *International Journal of Machine Learning and Computing*. 10. 339-345. [10.18178/ijmlc.2020.10.2.940](https://doi.org/10.18178/ijmlc.2020.10.2.940).
- [22] Ankush Daharwal, Prof. Sandeep Gore, Aishwarya Bhagwat, Shraddha Dethe, Sunny Chavan, (2020) "Career Guidance System using Machine Learning For Engineering Students (CS/IT)" June 2020, *International Research Journal of Engineering and Technology (IRJET)*, e-ISSN: 2395-0056, Volume: 07 Issue: 06.
- [23] Bhuvaneshwaran.B, Akshaya.N , Akshaya.S ,Devesh.M , Bhuvaneshwaran.B et al (2020), *International Journal of Computer Science and Mobile Computing*, Vol.9 Issue.3, March- 2020, pg.177-181.
- [24] Awoyelu, Iyabo & O., Oguntoyinbo & Awoyelu, Tolulope. (2020). "Fuzzy K-Nearest Neighbour Model for Choice of Career Path for Upper Basic School Students" , *International Journal of Education and Management Engineering*. 10. 18-32. [10.5815/ijeme.2020.04.03](https://doi.org/10.5815/ijeme.2020.04.03).
- [25] Manoharan, Saravanapriya & Senthilkumar, Radha. (2020). "An Intelligent Fuzzy Rule-Based Personalized News Recommendation Using Social Media Mining", *Computational Intelligence and Neuroscience*. 2020. 1-10. [10.1155/2020/3791541](https://doi.org/10.1155/2020/3791541).
- [26] Paul D.Tieger, Barbara Barron, and Kelly Tieger (2014) Fifth Edition, "Do What You Are", published by Little Brown Company.