1. **Career Recommendation System based on Student Preferences**

* One of the most common types of career recommendation systems is based on clustering algorithms [11]. Clustering algorithms group individuals based on their preferences, skills, and interests, and then recommend careers that are most closely aligned with their cluster
* Another common type of career recommendation system is based on decision trees. Decision trees use a set of rules to recommend careers based on an individual's input data. Other types of career recommendation systems include neural networks, support vector machines, and random forests. These systems are typically more complex and require larger amounts of data to generate accurate recommendations
* The paper provides various methods used for job recommendations include knowledge- based, collaborative, hybrid, and content-based filtering [12]. competitions, validations along with them regarding the reciprocal and time-based nature of Job referrals, Ethical aspects in job recommender system and JRS at scale: notes from LinkedIn [4] are also considered as a part of methodology. []🡪 Yang, Shuo, et al. "Combining content-based and collaborative filtering for job recommendation system: A cost-sensitive Statistical Relational Learning approach." Knowledge-Based Systems 136 (2017): 37-45.
* The **first phase** involves data extraction and feature selection. The dataset used for this phase is collected and consists of 6901 rows and 20 columns.
* **Column names:** 'Logical quotient rating', 'hackathons', 'coding skills rating', 'public speaking points', 'self-learning capability?', 'Extra-courses did', 'certifications', 'workshops', 'reading and writing skills', 'memory capability score', 'Interested subjects', 'interested career area', 'Type of company want to settle in?', 'Taken inputs from seniors or elders', 'Interested Type of Books', 'Management or Technical', 'hard/smart worker', 'worked in teams ever?', 'Introvert', 'Suggested Job Role'
* Decision Tree and Random Forest models have the highest accuracy score of 87%, while the SVM model has an accuracy score of 83%. A binary classification model's performance is measured using measures such as precision, recall, and F1-score.
* It is crucial to remember, however, that the accuracy of these models is determined by the quality and quantity of data utilised to train them. Therefore, it is necessary to collect reliable and diverse data to train the models and continually update them to ensure their accuracy. Overall, career recommendation based on student preferences using machine learning is a promising area of research that has the potential to help students make informed career choices and lead fulfilling lives.

2. **Career Recommendation for College Students Based on Deep Learning and Machine Learning. (China’s higher education) – 2022.**

Traditional recommendation algorithms mainly include three types: collaborative filtering recommendation algorithm, content-based recommendation algorithm, and hybrid recommendation algorithm. collaborative filtering recommendation algorithm can use existing historical activity interaction information and a series of user comparison information with the same behaviour to make collaborative similarity matching recommendation without relying on user or item content information [11]. content-based recommendation algorithm is an algorithm scheme that uses the user’s profile or project description information to recommend an algorithm scheme [12], which combines the attributes and characteristics of the project by means of data mining or information retrieval to build the user’s profile model. hybrid recommendation algorithm aims to obtain better prediction or recommendation performance by combining the content-based recommendation algorithm and collaborative filtering recommendation algorithm, taking the advantages of the two. Traditional recommendation algorithms need to perform a lot of experiments on the selection of combination methods and the sequence of combinations to find a better combination method, and the weight distribution of the recommendation results obtained by different methods also needs to be experimentally tested and analysed in order to make up for this.

* *User-Based Collaborative Filtering Recommendation Algorithm.*  user-based collaborative filtering recommendationalgorithm mainly implements similar recommendations for users to find other users with similar behaviours or constructs a user’s own preference model and performs a predictive score based on their own preference model to complete personalized recommendation. core of the user-based collaborative filtering recommendation algorithm is similarity calculation. main function of similarity calculation is to measure the similarity *u,v* between user *u* and user *v*.
* *Content-Based Recommendation Algorithm*. content-based recommendation algorithm is to use thecontent that users are interested in to calculate the similarity to achieve relevant recommendations. By means of data mining or information retrieval, a data file model belonging to the user is constructed by combining the attributes and characteristics of the project. recommendation algorithm adopted by this algorithm is to use the content that the user is interested in to calculate the similarity to realize the recommendation. algorithm uses *x* to represent the similarity evaluation of item *s* recommended to user *c*, where *x* is based on *s*, indicating that the recommendation of item *c* is obtained by quantitative calculation of other items with similar meta-information to item *s*.
* *Experimental Data.* . experimental data set used in this article is collected from the employment guidance platform for college students. . data set used in the experiment mainly contains information about college students themselves (gender, grades, interests, and so on), job requirements information, and user-post interaction behaviour information.
* A hybrid convolutional neural network is proposed, which uses convolution operation to learn high-level features to achieve personalized employment recommendation; (2) the training optimization strategy of the hybrid convolutional neural network is studied, aiming at the activation function, pooling processing, and loss function, and the feasibility of the optimization method is verified through simulation experiments;

3. **Mining Campus Big Data: Prediction of Career Choice Using Interpretable Machine Learning Method. – China (2022)**

* In this study, we used eXtreme Gradient Boosting (XGBoost), a machine learning (ML) technique, to predict the career choice of college students using a real-world
* dataset collected in a specific college. Specifically, the data include information on the education and career choice of 18,000 graduates during their college years
* In addition, SHAP (Shapley Additive exPlanation) was employed to interpret the results and analyse the importance of individual features.
* Finally, we secured a data set containing 10,292 students and 20 features, as shown

**4. A fuzzy-based career recommender system for senior high school students in K to 12 educations. 🡪 Conference (2019) (Philippines)**

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.

* 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 as it can easily cope with classification tasks in feature spaces of large dimensionality as found out by in their paper.
* 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 counselling (27) (28), as well as to analyse students’ academic performances (29).
* 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.
* with a total of 716 students.
* 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

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.

**5. PCRS: Personalized Career-Path Recommender System for Engineering Students (2020) (Palestine)**

* Provide guidance and help high school students choose engineering discipline.
* The design of PCRS is based on fuzzy intelligence of N-layered architecture and uses students' academic performance, personality type, and extra-curricular skills.
* The association between personality type and engineering discipline was built using a sample of 1250 engineering students enrolled in seven engineering disciplines.

The recommender system exploits the personal an educational information retrieved from the target users, and it is implemented based on four main phases, as shown in

1) Obtaining student's personal information including

gender, high school grades in STEM courses, and a list

of extra-curricular interests.

2) Determining student's personality type based on a

self-administered personality test.

3) Processing input data to construct a personal and academic profile for each student.

4) Building a fuzzy recommender system to provide students with personalized and user-speci\_c ranking of engineering disciplines.

The objective of PCRS is to mimic the logic of high-school advisors and provide guidance and recommendations to students after analysing their academic performance in addition to their personal. PCRS starts by collecting data which is categorized as follows; personal information, academic performance, and personality type.

INPUT DATA PROCESSING

The data processing phase is necessary for PCRS to convert the personal input data into numerical values that can be used as input variables for the membership functions of the fuzzy-logic system. Although qualitative inputs are acceptable for fuzzy logic systems, the quantifying process of inputs is necessary to calculate a common score to be used for all engineering disciplines. As a result, PCRS's design can be easily extended into other career specialities by following the same proposed processing phase. Another importance of the processing phase is to have a common membership function for fuzzy-logic system to produce a different recommendation for each engineering discipline, and no need to have a membership function per discipline.

IMPLEMENTATION OF THE FUZZY-LOGIC SYSTEM

Fuzzy system consists of three parts; 1) a fuzzi\_er which converts the input to a linguistic variable using the membership functions such as triangular, sigmoid, trapezoidal, or gaussian, 2) an inference engine which uses If-Then rules to convert the fuzzy input to the fuzzy output. And 3) a defuzzi\_er which converts the fuzzy output of the inference engine to output using membership functions similar to the ones used by the fuzzi\_er. The choice of fuzzy logic intelligence to implement PCRS comes from the main vision of the work which is to mimic the work of a school advisor who can provide a personalized and suitable recommendation to students based on their academic and personal pro\_les. Hence, the fuzzy-logic is used to provide a sorted list of recommended choices of future career disciplines among which the student can choose from.

The rules were derived after consulting with educational experts to determine the suitability degree of an engineering discipline based on persona and academic scores for each student.

In order for PCRS to provide an engineering discipline recommendation, the following steps will be repeated for each engineering discipline

1) The academic and personal scores of students are processed based on the engineering discipline.

2) The processed scores are entered as input data to corresponding membership functions.

3) a suitability rate is generated by the fuzzy logic based on input data.

4) Steps (1-3) are repeated for all engineering disciplines.

5) The suitability rates of all disciplines are sorted in an ascending order and presented to the student as the output of PCRS.

**6. Personalized Career-Path Recommendation Model for Information Technology Students in Indonesia – 2024**

* career-path recommendation model (CPRM) to provide guidance and help college students choose information technology jobs.
* The association between personality type and college students was established using samples of 104 computer science students enrolled in private universities in Indonesia.
* CPRM has been validated by Information Technology (IT) professionals and three psychologists in Indonesia through focus group discussions. The evaluation results showed that more than 83% of respondents were satisfied with the recommendation model.
* The model considers educational data mining grounded theory (EDM-GT) data integration and hierarchically related concepts
* educational data mining-grounded theory (EDM-GT), which combines technical skills, especially in the IT field, and validates the model through the involvement of experts in the fields of IT and psychology to measure a person's interest in a job

The main contributions of this research as follow:

1) Construct a new framework by integrating educational and psychological fields using educational data mining and grounded theory (EDM-GT) techniques to personalize students' information.

2) The framework uses a personalized recommender system to map the profiles of the job and subjects according to the personality type of the student.

3) Extracting and integrating data from various sources and mapping them is used to obtain a comprehensive knowledge of the recommended job and subject.

4) To validate the proposed model, we involved expert judgment, as well as IT professionals, and psychologists.

Our framework consists of six stages as follows:

1) Extract all the information retrieval from various sources to construct the model.

2) Construct the IT jobs’ profiles by extracting all the important information regarding job features and storing it in the database.

3) Construct the subject profile through subject structure documents.

4) Build a student profile using a personality type test. Four attributes have been identified to profile the student into the model and user preference for the recommended subjects.

5) Build dynamic personalized CPRM mapping to link the student profile with job and subject profile.

6) Use an EDM-GT technique to obtain recommender system using a personalized Naive Bayes (p-NB) algorithm.

Suggestion of recommended jobs and subjects for students are provided in the last step.

**7. Student’s Career Interest Prediction using Machine Learning -2022**

* We have used advanced machine learning algorithms like SVM, Random Forest decision tree, XG boost for predictions. After training and testing the data with these algorithms we take the most accurate results into consideration for further processing
* This paper has been proposed as an improvised conditional network-based deep support vector machine (ICGAN- DSVM) algorithm.ICGAN focuses on addressing the issue of low data volume that is using fewer datasets by mimicking new training datasets whereas DSVM extends SVM from shallow learning to deep learning.

Nunsina, Tulus, Zakarias Situmorang,” Analysis Optimization K-Nearest Neighbor Algorithm with Certainty Factor in Determining Student Career”, 2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology

8. **Student Career Recommendation System Using Content-Based Filtering Method. (2022) - Malaysia**

* The goal of this project is to create a system that only recommends careers for students studying computer science. Web scraping was used to extract the career information from the JobStreet website. The recommendation is made using a content-based filtering technique that compares one thing to another based on the user’s preferences.
* To create an efficient recommendation engine, recommendation systems use various forms of filtering, such as collaborative [6], content-based [7], and hybrid [8]. If people purchase an item from the website, they will be shown additional items based on the quality and specifications of the item that they purchased.
* *Collaborative Filtering:* Collaborative filtering is a recommendation system approach that focused on the assumption that [6] an item will be recommended [9] similarity among a group of users or items to predict missing ratings then make appropriate recommendations. According to Geetha [8], collaborative filtering is based on user similarity rather than item similarity, thus it generates more powerful recommendations.
* *Content-based Filtering:* Content-based filtering is based on an item’s description [12] and a preference profile of the user [5]. From all of the information provided by the users, these systems will create a list of item profiles for them [9]. In simple word, this method will compare previously rated items and suggests the best match [8]. In Content-based filtering, there are three process steps [5]. First is the content analyzer. To move item representation from the original data space to the target one, feature extraction methods are used to evaluate data items. The second is profile learner, This model creates a user profile by gathering data on user preferences and attempting to generalize it [13]. The third, filtering component, this approach uses user data profile to discover matching items related to the item list, but in new item and show this new item to the user [14].
* *Hybrid filtering:* This approach combines collaborative filtering techniques with content-based filtering techniques

Three steps to apply the content-based filtering to recommend a list of careers which are data review using Python, implement TFIDF content-based filtering method and vector space model and embed python code to visualize the obtained results.

* *The implementation of TF\*IDF:* The number of times a word appears in a document is known as TF (term frequency). Meanwhile IDF is the inverse document frequency (IDF) of a word that will measure of its importance in the whole corpus. The TF\*IDF technique is used to weigh a term in a document and give its relevance depending on how many times it appears in the document. The phrase is rarer and more essential if
* the TF\*IDF score (weight) is greater, and vice versa. *Vector space model:* Using vector space model, each product is kept as a vector of its characteristics in an ndimensional space, and the angles between the vectors are computed to indicate their similarity. Following that, user profile vectors are built based on the career ID that the user has chosen, and the similarity between an item and a user is calculated in the same way. Cosine similarity is a measure of how similar two vectors are. It can be used on objects in a dataset to show their closeness to one another using keywords.