Journal of Computer Science and Technology Studies (JCSTS))
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

A Counseling System of Multiple Intelligence Theory Combined With kNN Classification Algorithm

Hieu Le Ngoc[1] and Thanh Luong Van[2]
[12]*Department of Information Technology, Ho Chi Minh City Open University, Ho Chi Minh City, Vietnam*
✉ **Corresponding Author**: Thanh Luong Van, **E-mail**: hieu.ln@ou.edu.vn, 1751010132thanh@ou.edu.vn

ARTICLE INFORMATION

Received: July 08, 2021
Accepted: August 20, 2021
Volume: 3
Issue: 2
DOI: 10.32996/jcsts.2021.3.2.2

KEYWORDS

Multiple intelligence, career path counseling, machine learning, k nearest neighbor

ABSTRACT

Choosing the right career is always a big issue, an important concern for everyone. To have a job, which is suitable for you, firstly you must look at yourself, called the self, and you should be aware of what the self is then you can promote the strength of your own self and avoid your weakness. To help discover more about yourself, during researching and studying, we come up with the idea that we would propose a career counseling system based on Howard Gardner's theory. The system uses the theory of multiple intelligences (Abenti & Daradoumis, 2020) which is combined with the K-nearest neighbors (KNN) (Tang, Ying; Tang, Ying; Hare, Ryan; Wang, Fei-Yue;, 2020) algorithm to assist people and to give out a suitable suggestion about career path for them. We use the results of the eight intelligences retrieved from the KNN classification algorithm to give users the consulting for their career paths. This system is built with a dataset based on 56 multiple-choice questions. These include 48 multiple choice questions based on Howard Gardner's theory of multiple intelligences (Bravo, Leonardo Emiro Contreras; Molano, Jose Ignacio Rodriguez; Trujillo, Edwin Rivas, 2020), (businessballs, 2017) and 8 multiple choice questions which are the labels of the classifier. We divided the dataset into 8 subsets corresponding with 8 Intelligences defined by Howard Gardner with the collected dataset. In each subset, we build the KNN classifier model using KNN classification algorithm. This processing of 8 subsets come out with the results accuracy for the 8 Intelligences: linguistic intelligence (80.95%), logical-mathematical intelligence (82.14%), musical intelligence (96.43%), bodily-kinesthetic intelligence (82.14%), spatial-visual intelligence (82.14%), interpersonal intelligence (89.29%), intrapersonal intelligence (88.1%), existential intelligence (78.57%). With the outcome of 8 models, we have tested with 5 students and compared them to their actual intelligences. The comparison results tell us about the valuable potential in career path of the proposed counselling system, the advantages of this combination between Multiple Intelligence and KNN classifier.

1. Introduction

In today's society, the development of socio-economic and science and technology have brought many changes to people's lives, which also gives us too many new job opportunities and challenges. With the variety of economic sectors, we have too many kinds of jobs and many working fields. So, this leads to a big difficulty for us to choose a right career that can fit us the best. Choosing a career that best fits yourself means that this job matches your ability, and you feel happy to work with it. Therefore, *"how to choose a matching job"* has become one of the hot topics for human beings in contemporary society.

Choosing the right career has never been an easy decision, especially in the evolution 5.0 of technology. At the present time, people have the freedom to choose their favourite professional job in accordance with their abilities and interests. But there are too many majors and works for us to choose (Azmir, Johari, & Mahmud, 2017). It is not simple that we just can do the job by

understanding its own strengths and weaknesses. We also need to get involved in it for a certain time which is enough to understand it. That is why a lot of career path guidance and counselling organizations are born. These activities are helping people to get the right, advice when choosing a career, deciding a future path. In fact, there are quite a few cases of choosing the wrong career path.

With the outstanding development of modern industry, the application of new modern technologies like Machine Learning is an inevitable trend of today's society (Shafiee, Nor Safika Mohd; Mutalib, Sofianita;, 2020). There are many good algorithms in machine learning that help combine with multiple intelligence theories (Nasri, Rahimi, Nasri, & Talib, 2021) to determine people's intelligences. In detail, they are their strength and weakness. This recognition of people's intelligences can tell us what careers can fit them the most. The aim of this paper is to provide a counselling system that can help people identify their intelligences then they can make the right decisions to choose their own jobs.

Through the process of studying machine learning algorithms to approach this study goal, the KNN algorithm is quite suitable for the topic. With the potential of KNN (Awoyelu, Iyabo; Awoyelu, Tolulope; M, Awoyelu T.;, 2020), (Assegie, 2021) and the current situation of career guidance, we realize that we can build a solution to make choosing a career easier. To solve the above problem, we propose a career guidance system based on the theory of multiple intelligences. We apply a machine learning algorithm, specifically the KNN classification algorithm, to classify the intelligences' level then help conclude the jobs they can do.

This article is organized into 6 sections. Section 1 is the introduction of this paper. In the second section, we would discuss some related works. Section 3 is the methodology of this study, they are about the theory of multiple intelligences (MI), career guidance, machine learning and k nearest neighbors (KNN); And then, we propose a classification model based on the k nearest neighbor (KNN) classification algorithm, that can help to identify people's intelligences. Section 4 would introduce to data sets that we collected and how we process these data. Section 5 is the experimental results. Finally, it is Section 6 which is the conclusion of the topic.

2. Literature Review

In 2019, DeLauney and Bryan (DeLauney & Bryan, 2019) published the study "*Multiple Intelligence*" which introduced the theory of multiple intelligences. The meaning of multiple intelligences can be misunderstood that people with different skin-colors will have different intelligences together. This article is an approach to people with disabilities via the view of multiple intelligences. For example, people with disabilities often have low linguistic intelligence, but visual-spatial intelligence can be much higher than normal, so we can provide appropriate treatment and access to them. On the other hand, this article also shows the application of multiple intelligences theory in the classroom and the positive benefits that it has brought during class time. Applying this theory helps increasing independence and self-direction in learning. The application of this theory in teaching has brought these advantages: teachers can present lectures in many different approaches; students approach the topic to give multidimensional and diverse opinions and judgments on the issue; students can increase self-confidence, independence, and self-study. Besides, the limitations also showed the difficulties in teaching the same subject for different intelligences, difficulties in evaluating students based on intelligence, difficulties in teaching students with different intelligences, etc. To teach and follow students with superior intelligence, we lack human and material resources, leading to difficulties in teaching.

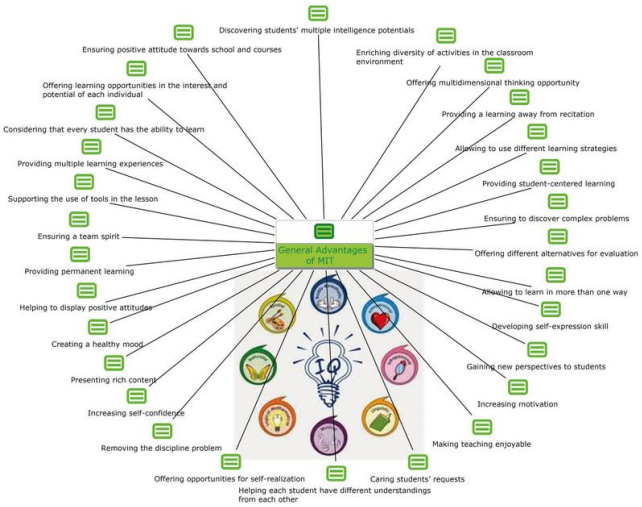

Fig 1. Advantages of MI (DeLauney & Bryan, 2019)

Fig 1 shows us the advantages that MI brings to us in education. Through this study, we can see the advantages of multiple intelligences theory. This theory can help us understand more about people and their own intelligences. We can also see the difficulties that need to be overcome when applying the theory of multiple intelligences in practice.

In 2016, Jasjit S. Sodhi, Maitreye Dutta, and Naveen Aggarwal (Sodhi, Jasjit S.; Dutta, Maitreyee; Aggarwal, Naveen;, 2016) published the study "*Efficacy of Artificial Neural Network based Decision Support System for Career Counseling*", researched on building a career guidance system based on artificial neural network (ANN) algorithms. This study was conducted with 100 students of 10 grades. The study consisted of 6 phases. Phase 1 is to collect data to be used for the training of ANN. Phase 2 is about transforming the collected data into neural network inputs and normalization of data. Phase 3 is the construction of training dataset. Phase 4 is about Select, Train, and Test Network. Phase 5 is to deploy developed network application. Phase 6 is the analysis and interpretation of results. With each output node in the ANN algorithm corresponding to each industry among these industries: engineering, science, office, government, and stakeholders in society. The classification results were obtained from the above techniques with an accuracy of 91% in all test cases. Based on the techniques presented in this paper, we can refer to the classification procedures of this paper and the data transformation and survey methods.

In 2018, Sabriye and Ayten Cokcaliskan (Sener & Çokçaliskan, 2018) launched the study "*An Investigation between Multiple Intelligences and Learning Styles*". In this study, the authors use multiple intelligences theory to help students find out their strengths and weaknesses, thereby measuring their learning style. The study used quantitative methods to conduct data analysis.

Variables	Categories	N	%
Gender	1. Female	45	51.1
	2. Male	43	48.9
Grade	1. Fifth Grade	16	18.18
	2. Sixth Grade	27	30.68
	3. Seventh Grade	28	31.81
	4. Eighth Grade	17	19.31
	Total	88	100.0

Fig 2. Gender data (Sener & Çokçaliskan, 2018)

This study used a dataset with 45 female (51.1%), 43 male (48.9%). The research subjects are students in the 5th to 8th grade age range. In which, the 7th graders accounted for the highest percentage of 31.81% (27 people), the 5th graders accounted for the least with 18.18 % (16 people). Through research, the authors have concluded that students have all intelligences as well as all learning styles, but learning styles are mainly tactile and auditory with the three highest scores, belongs to three intelligences: linguistic intelligence, spatial-visual intelligence, and bodily-kinesthetic intelligence. The goal of the study is to find out their strengths and strengths so that they can overcome their weaknesses and develop their strengths. From there, it helps to have clear and specific orientations, which is also very useful in choosing the right career. This publication shows the usefulness of multiple intelligences theory and its application in career path orientation. However, this paper stops at statistics and data analysis, without applying machine learning and data processing techniques. Therefore, there are many applications of the theory of multiple intelligences that we can develop with AI and machine learning techniques.

In 2019, Tehseen Mehraj and Asifa Mehraj Baba (Mehraj & Baba, 2019) published a study on "*Scrutinizing Artificial Intelligence based Career Guidance and Counseling Systems: and Appraisal*", the authors highlight the applications of artificial intelligence in education such as supporting the assessment of academic performance, supporting assessment, grades, ... Thereby we can see the importance of artificial intelligence in education. In addition, the authors also briefly presented the laws of artificial intelligence. However, this study only mentioned AI as their advantages but did not go into specific applications. This study serves as a reference list of applications and artificial intelligence, contributing to the theoretical support of this paper.

In 2020, the author Ashok Panigrahi and Vijay Joshi (Panigrahi & Joshi, 2020) have studied "*Use of artificial intelligence in education*", to help apply AI technology in education to improve learning quality and efficiency. In this study, the authors mentioned the limitations of traditional educational methods and why it is necessary to apply AI technology in education. In addition, the authors have come up with a solution to bring AI technology into education, that is, using AI, reaching children first by giving lectures on AI for students to study and research.

In 2020, I. Ketut Sudarsana (Sudarsana, Surp, Sena, & Setiawan, 2020) and colleagues published the study "*Technology-Based Hinduism Learning and Multiple Intelligences*", in this study the authors applied Howard Gardner's theory of multiple intelligences combined with the classroom reversal method to help Studying as well as studying about Hinduism becomes easier, more effective, and enjoyable. Applying the theory of multiple intelligences in teaching has made teaching livelier and more interesting. At the same time, the application of this theory makes it easier to teach different intelligence objects. Specifically, it is possible to apply music, add visual images, and add group activities to make the lectures no longer monotonic.

In 2020, Nor Safika Mohd Shafiee and Sofianita Mutalib (Shafiee, Nor Safika Mohd; Mutalib, Sofianita;, 2020) have studied about *"Predicting mental health problems of higher education students using machine learning"*, this study aimed to detect mental health markers of high school students using machine learning to solve problems. This study used classification algorithms to classify the causes of mental health problems. Through this study, it is possible to see the application of machine learning in education. On the other hand, it is also better to understand the classification algorithms such as KNN, SVM presented in the research. This paper outlines the reasons for the impact on student mental health and a review of previous studies on student mental health described in Fig 3. This paper shows the high applicability of classification machine learning algorithms in education and career guidance. This is a solid premise that helped us carry out this study. This article stops at the evaluation level and gives some comment, not going into any specific model.

Author	Techniques	Variables	Tool	Significances
[31]	Random forest, NB, SVM, KNN	Perceived Stress Scale (PSS) questionnaire	Not stated	The accuracy of random forest 83.33%, NB 71.42%, SVM 85.71% and KNN 55.55%
[51]	DT, SVM, ANN, BN	Outgoing activity, toileting activity, sleeping activity, disease, mental status (GDS)	Not stated	Normal: DT 95.1%, SVM 75.6%, ANN 96.7%, BN 92.7% Mild depression: DT 94.3%, SVM 75.6%, ANN 96.2%, BN 91.3% Severe depression: DT 99.4%, SVM 99.5%, ANN & BN 100%
[52]	KNN, SVM BN	Depression level, sex, grade, major, technical, novel, amusing, psychological,	SPSS	The accuracy of KNN 76.6%, SVM 82%, NB 64.2%
[53]	Random forest, random tree, MLP, SVM	Age, MMSE score, neurological condition, depression (GDS), MoCA test	WEKA	The accuracy of random forest 95.45%, random tree 93.93%, MLP 92.42%, SVM 92.42%
[54]	XGBoost	socio-demographic information, earthquake-related experience, sleep, mood, somatic symptoms and everyday functioning were assessed, CRIES	Python	The combination of earthquake experience, everyday functioning, somatic symptoms and sleeping correctly predicted 683 out of 802 cases of probable PTSD, translating to a classical accuracy of 74.476% (85.156% sensitivity and 60.366% specificity) and an area under the curve of 0.80.
[55]	SVM, Linear Discriminant analysis, ensemble, KNN, DT	Mean heart rate, Standard deviation of heart rate, Mean EDA, Standard deviation of EDA, Minimal EDA, Maximal EDA, Root mean square EDA, Mean ECG, Standard deviation of ECG, Energy of ECG, Absolute value of ECG, Root mean square ECG	MATLAB	The accuracy of SVM 91%, Linear Discriminant Analysis 90.7%, ensemble 89.5%, KNN 87.6%, DT 86.1%

Fig 3. Summary of Previous Research on Data Mining Techniques used in Mental Health Problems (Shafiee, Nor Safika Mohd; Mutalib, Sofianita;, 2020)

In 2020, Abhishek S. Rao and associates (Rao, Abhishek S.; Bola Sunil Kamath, Ramya R; Chowdhury, Shreya; Pattan, Shreya A; Kundar, Raveena Krishna;, 2020) published the study "*Use of Artificial Neural Network in Developing a Personality Prediction Model for Career Guidance: A Boon for Career Counselors*", the article uses artificial neural network (ANN) algorithms to build a career counseling system. The study used a 36-question questionnaire to collect data. Use chaos matrix and K-Folds Cross-Validation to evaluate the performance of the model. About the Extroverted and Introverted category, there are 194 of the 200 test samples which are correctly classified with an accuracy of 97%. For Sensation and Intuition, 191 of the 200 test samples are completely classified with an accuracy of 95.5%. For Feeling and Thinking, 199 of the 200 test samples are correctly classified and had an accuracy of 99.5%. For Assessment and Perception, 195 of the 200 tested samples are properly classified and had an accuracy of 97.5%. This paper also reinforces artificial intelligence, classification techniques and deep learning and demonstrates the usefulness of machine learning and AI in career counseling.

	MBTI Categories			
Evaluation matrices	**E vs I**	**S vs N**	**F vs T**	**J vs P**
Sensitivity	96.87	97.29	100	100
Specificity	97.11	93.25	99	92.18
Precision	96.87	94.73	99	96.45
Accuracy	97	95.5	99.5	97.5

Fig 4 Performance Metrics of the Proposed Classifier (Rao, Abhishek S.; Bola Sunil Kamath, Ramya R; Chowdhury, Shreya; Pattan, Shreya A; Kundar, Raveena Krishna;, 2020)

Since the early 2010s, there have been many publications related to the study of classification algorithms. In 2014, the group of authors included Debo Cheng, Shichao Zhang, Zhenyun Deng, Yonghua Zhu, and Ming Zong (Cheng Debo, Zhang Shichao, Deng Zhenyun, Zhu Yonghua, Zong Ming, 2014) studied about *"K nearest neighbor (KNN) Algorithm with Data-Driven k Value"* and proposed a new k Nearest Neighbor (K nearest neighbor (KNN)) algorithm based on sparse learning, to overcome the limitations of the previous K nearest neighbor (KNN) algorithm, such as a fixed k value for each test sample and ignoring the correlation of samples. Specifically, the paper will reconstruct the test samples by training the samples to learn the optimal k value for each test sample, then use the KNN algorithm with the learned k value to conduct all the tests. All types of tasks, such as classification, regression, and missing value entry, will be the same. The rationale of the proposed method is that different test samples should be assigned different k values in the K nearest neighbor (KNN) algorithm and learning the optimal k value for each test sample must consider the correlation of the data. Therefore, during reconstructing, the proposed method is designed to achieve the minimum reconstruction error through the least-squares loss function and uses the standard l1 regularization term to generate the sparseness. The element sparsity in the reproducibility coefficient, i.e., the sparsity, occurs in the element of the coefficient matrix. For efficiency, Location Preservation Forecasting is used to keep local structures of the data.

Table 1. Comparison of classification accuracy

Dataset	*k*NN	L-*k*NN	S-*k*NN
Cleveland	0.7048 ± 0.0029	0.7619 ± 0.0030	**0.8048 ± 0.0038**
Ionnosophere	0.6743 ± 0.0031	0.8286 ± 0.0025	**0.8571 ± 0.0025**
Heart	0.5381 ± 0.0031	0.6741 ± 0.0033	**0.7889 ± 0.0043**
Seeds	0.8048 ± 0.0033	0.8714 ± 0.0015	**0.9238 ± 0.0011**

Table 2. Comparison of RMSE

Dataset	*k*NN	L-*k*NN	S-*k*NN
Mpg	3.8080 ± 0.4417	3.5909 ± 0.3662	**3.4693 ± 0.3158**
Triazines	0.1477 ± 0.0017	0.1357 ± 0.0016	**0.1242 ± 0.0013**
Concreteslup	0.0176 ± 0.000036	0.0151 ± 0.000033	**0.0139 ± 0.0000339**
Bodyfat	0.000021± 3.7921e-09	0.000014 ± 3.3355e-09	**0.000013 ± 3.8892e-09**

Fig 5. kNN Results of Debo Cheng et al. (Cheng Debo, Zhang Shichao, Deng Zhenyun, Zhu Yonghua, Zong Ming, 2014)

The proposed S-kNN algorithm performs better than other algorithms, such as L-kNN and kNN. More specifically, the proposed S-kNN algorithm improves on average by 5.97% and 16.32% compared with the L-kNN algorithm and the kNN algorithm in classification accuracy, respectively. The S-kNN algorithm performs better than the L-kNN algorithm because the proposed S-kNN method uses the LPP regularization term to preserve the local structure of the data. In particular, the proposed method improved 11.48% compared to L-kNN, on the Heart dataset. Both S-kNN and L-kNN are better than kNN algorithm. Finally, the experimental results on the real dataset show that the proposed K nearest neighbor (KNN) algorithm is better than the modern algorithms in terms of different learning tasks such as classification, regression, and missing input value. This study outlines some ways to process and select the optimal k value, thereby also presents knowledge about KNN algorithms, thereby can help us realize the potential of KNN and its application into classification problems in the real world.

3. Methodology

Choosing the right methods and algorithms for the data set is very important and it is necessary to build an efficient, accurate classifier model. This section describes the classification method and algorithm applied to build the classifier model for the career path counseling system.

3.1 Theoretical basis

3.1.1 Multiple intelligence theory

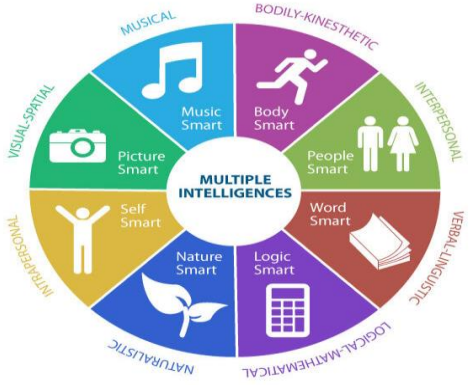

Fig 6. Multiple intelligence (DeLauney & Bryan, 2019)

Multiple intelligences (MI) theory is a theory about human intelligences. Theory of multiple intelligences (Ahmad & Dzulkarnain, 2020), (Alez-Treviño, Valencia-Hernández, & Jesús Arrona-Palacios, 2020) published by Dr. Howard Gardner in his book in 1993 whose name is "Frames of Mind". Accordingly, he observed that (Sudarsana, Surp, Sena, & Setiawan, 2020), (Yavich & Rotnitsky, 2020), (Abenti, Heather Freya, 2020), (Aubrey L. C. Statti;Kelly M. Torres, 2020) intelligence is the potential to solve problems or to create something that is recognized in one or more cultures. He also stated that humans always have possessed intelligences and that no two people have the same development of intelligences. According to him, intelligence (businessballs, 2017) is not just one type but also a set of intelligences. According to his research, there are 9 types of intelligence including (Oasis, 2020): linguistic intelligence, mathematical logic intelligence, musical intelligence, visual intelligence, motor intelligence, and intelligence. Introspective intelligence, natural intelligence, existential intelligence, interpersonal intelligence. However, in this article, we only use 8 types of intelligence: linguistic intelligence, logical-mathematical intelligence, musical intelligence, spatial-visual intelligence, existential intelligence, intrapersonal intelligence, interpersonal intelligence.

The theory of multiple intelligences has many applications in education (Alez-Treviño, Valencia-Hernández, & Jesús Arrona-Palacios, 2020), (Ahmad & Dzulkarnain, 2020) as well as a career choice (Atela, Richard Juma; Othuon, Prof. Lucas; Agak, Prof. John .O., 2019), (Azmir, Johari, & Mahmud, 2017). The advantage of this theory is that a person's intelligence is assessed not by a specific number but by dividing the intelligence into different parts. Thereby, based on the characteristics of each person, different rates of intelligence will have different development directions.

3.1.2 Career guidance and career counseling

Vocational guidance (Biswas, Al Amin; Majumder, Anup; Mia, Md. Jueal; Basri, Rabeya; Zulfiker, Md. Sabab;, 2021), (Kiselev, et al., 2020), (T.V, 2020) are activities to support all individuals to choose and develop professional expertise that best fits their individual abilities, and at the same time, it can satisfy human resource needs for all career fields (labor market) at the local and national levels. For a long time, we still think that career guidance is just choosing a career that we love, choosing a university that is right for us. Reality shows (DIỆP, 2021), (Kiselev, Pavel; Feshchenko, Artem; Matsuta, Valeriia; Bogdanovskaya, Irina;, 2020) that students often choose the career path according to feelings, due to temporary feelings, or follow the trends of the society, because there is no suitable orientation.

However, this is only the tip of the iceberg, one of many career-oriented activities. Vocational terms (tạo, 2020), (Lester & Hofmann, 2020) (career counseling) if we understand correctly, it is a combination of many areas such as career assessment (career assessment), career management (career management), career development (career development) ... Career counseling (Atela, Othuon, & Agak, Relationship Between Types of Intelligence And Career Choice Among Undergraduate Students Of Maseno University, Kenya, 2019), (Lee, Lai, & Wahid, 2017), (Dr. Gufran Ahmad Ansari, 2017) is an activity that helps you to recognize and understand yourself as well as the career path in certain social circumstances, thereby making career choices, choosing personal development paths, and getting suitable careers. Career counseling activities include two purposes: short-term and long-term. In the short term, career counseling will help you find the right direction when you come to the point where you need to decide to choose a career. In the long run, career counseling will support you to overcome difficulties in the process of finding a job, a working environment, and building a career path, providing the necessary knowledge and skills for you that you can self-direct and make the right decisions. In recent years in Vietnam (tạo, 2020), career counseling has received more attention than before. However, in this area that is still poor and do not have the conditions to develop the infrastructure as well as educational facilities. This also has not received the attention of many people and it is still quite superficial.

3.1.3 Machine Learning and KNN

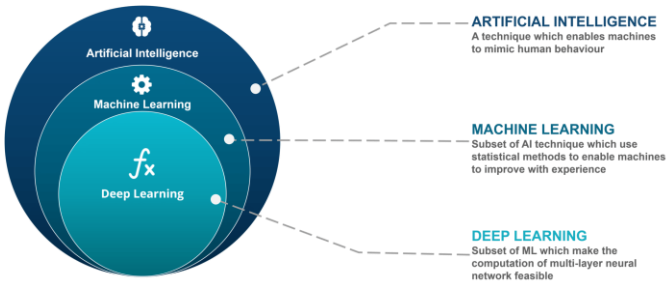

Fig 7. Machine learning (Abenti, Heather Freya, 2020)

Machine learning is a branch of AI (Kolachalama, Vijaya B.; Garg, Priya S.;, 2018), (Iatrellis, Savvas, Kameas, & Fitsilis, 2020), (Haji, 2020) that focuses primarily on building learned models based on a set of data and improving accuracy over time without having to rely on programming to do so. K nearest neighbors (KNN) (Cheng Debo, Zhang Shichao, Deng Zhenyun, Zhu Yonghua, Zong Ming, 2014), (Zahour, Benlahmar, Eddaouim, & Hourrane, 2020), (Nasution, Fadliansyah; Zamzami, Elviawaty Muiza;, 2019) is a

non-parametric classification algorithm commonly used in data mining, statistical probabilities and is considered one of the simplest algorithms in ML. Machine Learning algorithms are appreciated for performing classification effectively. With the initial idea of choosing one k nearest neighbors (Li, Bin; Wan, Sailuo; Xia, Hua; Qian, Fengshou;, 2020) for each test sample, the next step is to use the trained k nearest neighbors to predict for a new test sample or, in other words, determine a new class using the k nearest neighbors that have been trained. More specifically, to implement the KNN algorithm (Taha, 2017), it is necessary to gather training data, the distance measure (the way is calculated using the Euclidean formula, Manhattan ...), and the k value. After calculating the distance measured from the training objects, the KNN (Taha, 2017), (Hu, Juan; Peng, Hong; Wang, Jun; Yu, Wenping;, 2020) selects the k nearest neighbors to the object and uses these k neighbors to determine the class for that object.

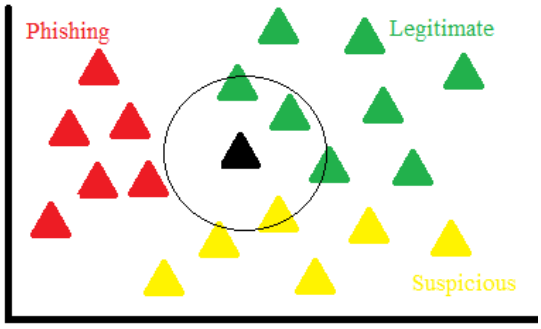

Fig 8. KNN (Assegie, 2021)

Fig 8 shows that there are 3 classes of networking attacks, including Phishing, Legitimate, Suspicious. The problem is that there is an object (black triangle). How can we determine whether this object belongs to? In machine learning, there are many classification algorithms (sklearn, 2020), but the simplest and most suitable classification algorithm for this study is the KNN classification algorithm. With KNN (Yulianto, Lili Dwi; Triayudi, Agung; Sholihati, Ira Diana;, 2020) In a Classification problem, the label of a new data point (or the result of a test question) is inferred directly from the K nearest data points in the training set. The label of the test data can be decided by primary voting between the nearest points, or it can be inferred by assigning different weights to each nearest point and then deducing it off the label. The Regression problem in KNN, (Shichao Zhang, Senior Member, IEEE, Xuelong Li, Fellow, IEEE, Ming Zong, Xiaofeng Zhu, Ruili Wang, 2018) the output of a data point will equal to the output of the nearest known data point itself (in the case of K = 1), or the weighted average of the outputs of the nearest points, or the relation system based on the distance to the nearest points. In summary, KNN is an algorithm to find the output of a new data point by relying only on the information of K data points in its nearest training set (K-neighbors), regardless of some data.

With the study of the theoretical bases, we find that building a career counseling system in education is extremely necessary. With the advantage of the multiple intelligence theory, the strong development of classification algorithms in machine learning, KNN is one of the potential algorithms to support the development of this application. Therefore, this paper uses KNN classifier with classifier models to build a career recommendation system applying multiple intelligences theory. This is a powerful combination of technology and education theory.
KNN with Sklearn

3.1.4 Sklearn

Scikit-Learn (Paper, 2019) is a Python library that provides simple and efficient tools for the implementation supervised and unsupervised machine learning algorithms. The library is accessible for everyone because it is open source and can be used for commercial purposes. Scikit-Learn is built on NumPy, SciPy and matplolib library, which means it is reliable, powerful and the core of Python language. Scikit-Learn focuses on data modeling rather than loading, data cleaning, munging, or manipulation.
Sklearn Classification (sklearn, 2020)

SKLearn classification is for identifying which category an object belongs to. It has many applications such as spam detection, image recognition etc. Most common algorithms of SKLearn are SVM, nearest neighbors, random forest, and more.
KNeighborsClassifier

Neighborhood-based classifiers (sklearn, 2020), (Nugroho, Aldi; Riady, Osvaldo Richie; Calvin, Alexander; Suhartono, Derwin, 2020) do not attempt to build a common internal model, but it only stores the versions of the training data. Neighborhood classification is calculated based on the votes of the nearest points.

K-NN classification has 2 phases: the first one is to identify the nearest neighbor which the element belongs, and the second phase is to determine the category of that neighbor. We assume that we have a training dataset D generated from (xi) i∈ [1, n], the training samples (with n = | D |). The examples are described by a set of features F and any numeric features have been normalised to the range [0,1]. Each training example is labelled with a class label yj ∈ Y . Our objective is to classify an unknown example q. For each xi∈D we can calculate the distance between q and xi as follows:

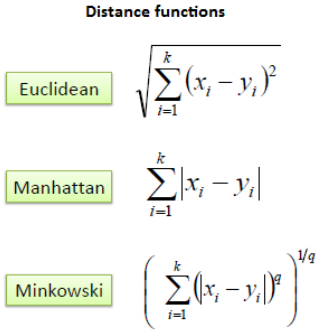

Fig 9. Formulas of distance calculation (sklearn, 2020)

KNeighborsClassifier (sklearn, 2020), (Cunningham P, Delany SJ, 2020), (Sadali, Muhamad; Putra, Yupi Kuspandi; Putra, Yupi Kuspandi;, 2019) using uniform weighting: that is, the value assigned to the query score is calculated from the majority of votes of the nearest points. In some cases, it is advisable to consider the neighbors for high efficiency. This can be done via the weights keyword. The default value (Geler, Zoltan; Kurbalija, Vladimir; Ivanovi´c, Mirjana; Radovanovi´c, Milo˘s;, 2020), weights = 'uniform', assigns a uniform weight to each neighbor, with weights = 'distance' assigning a weight proportional to the inverse of the distance.

3.2 Career path counseling system based on Multi Intelligence Theory and KNN

To build a career path counselling system, we use a K nearest neighbor (KNN) classification model to classify people's intelligences based on the theory Multiple intelligences. With the proposed system, we would have the 03 following steps for counselling:

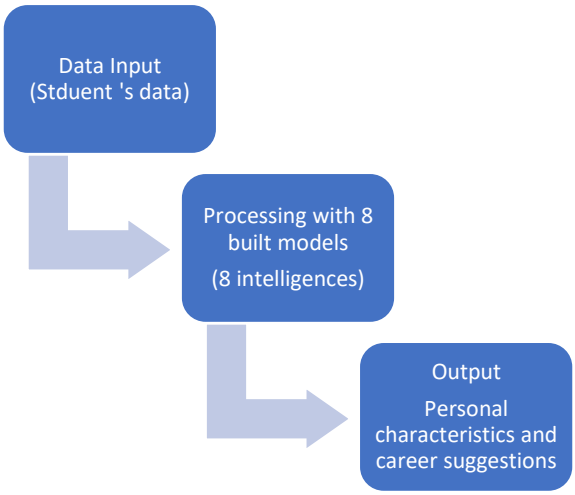

Fig 10. The processing flow of Counselling system

Step 1: From the input data which is collected from student or whoever wants to get counsel about career path, we handle the data as the following sequences:
Splitting data: the collected data is split into 8 parts corresponding to 8 intelligences.
Weighting the fields / attributes of all 8 parts of data (Give weight to questions of the survey)
Calculating the *Total* attribute, which is the total score of the weighted questions for each part (each intelligence)

Step 2: After handling the input data, we proceed and input the data into 8 classifier models built from the K nearest neighbor (KNN) algorithm. The input data is 8 pieces, respectively with 8 intelligences.

Step 3: The models' outcomes will classify the user's information, and we can get the level of each intelligence. Based on these obtained results and Multiple Intelligence Theory, we will make career suggestions which are related to the strength and weakness.

Building classification models for career path counseling system: To have a good counselling system, we should have good classifier models to classify people's intelligences. We would propose the models building as the followings:

- *Collecting data:* The dataset for training to build the classifier models is the survey result of 56 multiple-choice questions. They are 6 questions of linguistic intelligence, 6 questions of logical-mathematical intelligence, 6 questions of musical intelligence, 6 questions of bodily-kinesthetic, 6 questions of spatial - visual, 6 questions of interpersonal intelligence, 6 questions of intrapersonal intelligence, 6 questions of existential intelligence, and 8 questions for self-assessment their own 6 intelligences, respectively. With this data, we separate into 8 parts corresponding to 8 intelligences, each part includes 6 normal attributes and 1 label attribute.
- *Processing data: train to build model:* The classifier models are the core of the career counseling system based on the theory of multiple intelligences. The process of building these models is described as following diagram:
-

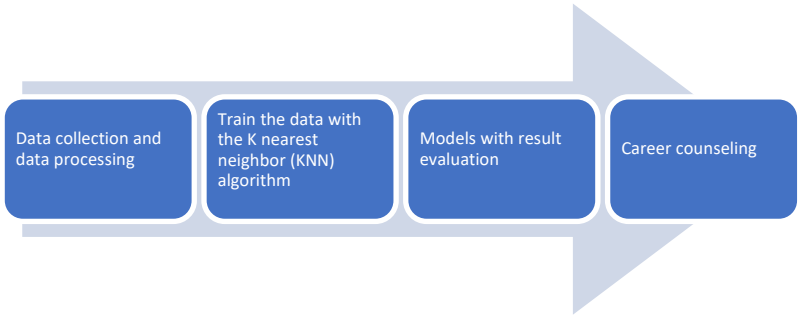

Fig 11. Process of building KNN Classifier Models

Step 1: Collect and process data

Firstly, people participate in the survey with their personal information (name, gender, and age) and intelligences' information as below:

6 questions about linguistic intelligence
6 questions about logical-mathematical intelligence
6 questions about musical intelligence
6 questions about bodily-kinesthetic intelligence
6 questions about spatial-visual intelligence
6 questions about interpersonal intelligence
6 questions about intrapersonal intelligence
6 questions about existential intelligence
8 questions that show the level of intelligences (from 1 to 10)

After having a complete data set, we synthesize the data obtained from the above survey, and then we proceed to convert the data into the corresponding score attributes (numeric fields) for each question. Then divide the data set into 8 sub-set, with each set, we have 6 questions (normal attributes) of the intelligence and 1 question (label attribute) of self-assessment about the level of the respective intelligence. We give weight for 6 questions and calculate their points. We add a new attribute *Total*, the total score of the 6 attributes (except the label attribute). After processing the dataset, we have 8 subsets with 8 intelligences, and each one has 7 normal attributes and 1 label attribute. All attributes are in numeric type.

Step 2: Build the models with K nearest neighbors (KNN) algorithm: With the 8 sub-datasets above, we proceed to run K nearest neighbors (KNN) for each subset and make assessment of each result.

Step 3: Classification results: With the results from the K nearest neighbor (KNN) algorithm for each part, we evaluate the precision and loss for each part.

Step 4: Career consulting: When the expected accuracy is obtained (we expect the rate is around 75% and above), we proceed with the user's input data with these built models. Based on the previous study (Atela, Richard Juma; Othuon, Prof. Lucas; Agak, Prof. John .O., 2019), we combine the outcomes with th and propose some suitable occupations / jobs / careers for users.

3.2 Data and Data Processing

3.2.1 About the collected dataset of intelligences

The dataset consists of 56 data fields corresponding with the answers of 56 questions. The dataset is collected based on a survey of 419 people. These 56 questions are 48 questions from the Multiple Intelligences Theory (businessballs, 2017) and 8 questions for individuals to self-assess their 8 bits of intelligence. This dataset was collected mainly in Ho Chi Minh City and the surrounding area. The survey is subjected with the age from 25 to 45 years old.

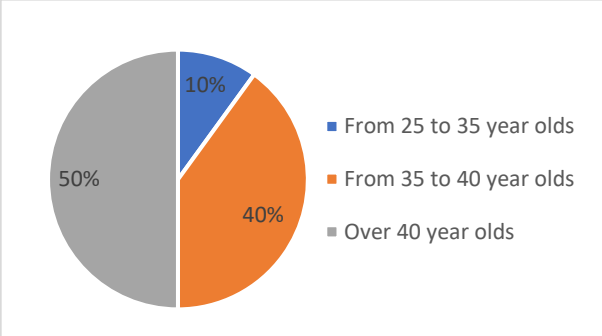

Fig 12. Age of people who joined the survey for data collection

In Figure 12, we see that people from 25 to 35 years old account for 10% (41 people), people from 35 to 40 years old account for 40% (168 people), people over 40 years old account for 50% (210 people). With this objective, the dataset is more trust with people's self-assessment because they are aware of their strengths and weakness at a late age. This dataset stores the score results of the survey questions. Here in table 1, are the data from question 1 to question 6 of the dataset, it is also the first intelligence among 8 intelligences. There are 6 questions of linguistic intelligence; the last question (the seventh question) is self-assessment about the level of linguistic intelligence.

Table 1 Questions of Linguistic Intelligence

Question	**Data Type**	**Scale**
1 - At school one of your favourite subjects is / was Literature?	Integer	Dichotomous scale
2 - When talking to someone, you tend to listen to the words they use not just what they mean.	Integer	Dichotomous scale
3 - You can easily learn a new language.	Integer	Semantic differential scale (wiki, Semantic differential, 2017)
4 - You prefer to use flexible words like using metaphors, idioms, transforming sentence structure or an inversion in a sentence	Integer	Dichotomous scale
5 - You have good rhetoric and reasoning skills.	Integer	Rating scale (wiki, Rating Scale, 2017)
6 - You find it easy to make up stories	Integer	Rating scale (wiki, Rating Scale, 2017)
7 – How do you point your level of Linguistic Intelligence?	Integer	Rating scale (wiki, Rating Scale, 2017)

For the rest of 7 intelligences, the data is also with the same described and constructed according to Table 1. So on, we collected the dataset of 56 attributes.

3.2.2 Data processing With the surveyed data, we need to process the data for better accuracy with KNN classification algorithm.

1) Split the dataset into 8 parts

We proceed to split the data set into 8 subsets corresponding to 8 intelligences. The components of each subset are described as following:

Table 2 Component Data Fields of an Intelligence

Fields	Q1	Q2	Q3	Q4	Q5	Q6	Total	Label
Weighted	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Sum	No	No	No	No	No	No	Yes	No

- From Q1 to Q6: These are question 1 to question 6, they are about the intelligence.
- Total: the field sums up the scores of 6 questions (after giving weight for each question). We add this Total field to increase the accuracy and the fitness of the model.
- Label: the label attribute that is self-assessed by survey participants through self-assessment score.

2) Give weights to data

To highlight some important ideas and contents that strongly related to the intelligence, we should give weight to the respective questions to emphasize the importance by multiple the score of that question with the weight. The weighting of these questions is based on some research (businessballs, 2017), (Ahmad & Dzulkarnain, 2020). One of the previous studies, the author (Geler, Zoltan; Kurbalija, Vladimir; Ivanovi´c, Mirjana; Radovanovi´c, Milo˘s;, 2020)presents the characteristics of each intelligence through which we use it to conduct the analysis of questions:

- With Linguistic Intelligence, we multiply the scoring values of questions 2 and 6 by 10.
- With Logical-mathematical Intelligence, we multiply the scoring values of questions 1, 2, 4 and 5 by 10.
- With Musical Intelligence, we proceed to multiply the scoring value by 10 for question number 1. For question number 3, if the value of question number 1 is not 0, then multiply the value of question number 3 by 10 otherwise keep the same value.
- With Physical Intelligence, we proceed to multiply the scoring value by 5 for question number 1, and multiply the scoring value of question number 3, 4 by 2.
- With Image-spatial Intelligence, we proceed to multiply the scoring value of questions 5, 6 by 10.
- With Personal Interactive Intelligence, we proceed to multiply the scoring value of questions 1, 3, 6 by 10.
- With Inner Intelligence, we multiply the value of question 4 by 10.

After giving weight for all the necessary attributes, we generate the Total attribute for the respective Intelligence with each Intelligence. So totally we get 64 attributes for the dataset.

3) Normalize data with StandardScaler

We use *StandardScaler*, which is built in Python language, to normalize these attributes of the dataset. This will set the mean to 0 and the standard deviation to 1.

4. Results and Discussion

To conduct this Research, we use *Google Colaboratory* with Python language for experimental development. We also use the K nearest neighbor (KNN) algorithm provided by *SKlearn library* of Python to approach this study. Below are the results of 8 models corresponding to 8 intelligences. We would present in detail the experimental results of 8 intelligences.

4.1 Result of models

4.1.1 Models of 8 Intelligences

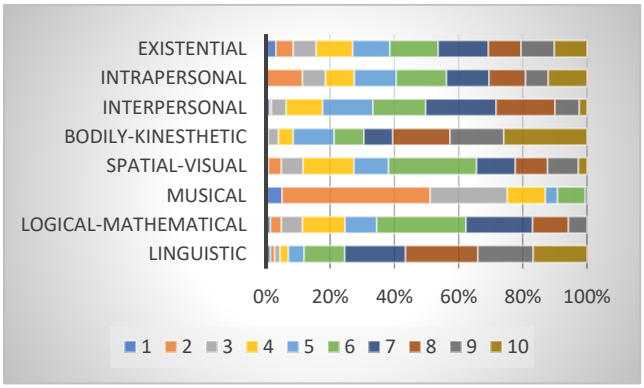

Fig 13. Label Distribution of 8 Intelligences

Figure 13 shows the distribution of each intelligence, the percentage of each label in each intelligence. With this figure, we can see that all the distributions of labels (from 1 to 10) are quite reasonable. This means the dataset collected is useful and fitting with Multiple Intelligence Theory.
We would like to present the methodology of building KNN model for the first Intelligence, Linguistic Intelligence. We followed these steps as below:

Step 1: With the dataset of Linguistic Intelligence, we split it into 2 subsets training set and testing set with the ratio of 80/20, respectively. We use the library *sklearn.model_selection* in Python with its method *train_test_split*:

```
from sklearn.model_selection import train_test_split
Linguistic_train, Linguistic_test, LB_Linguistic_train,
LB_Linguistic_test = train_test_split(X, Y, test_size = 0.20)
```

Fig 14. Splitting the dataset of Linguistic Intelligence

As *fig 14*, we use the method *train_test_split* to split the data:

- *Linguistic_train, Linguistic_test* are training set and testing set of the Linguistic model.
- *LB_Linguistic_train, LB_Linguistic_test* are the label of the training set and testing set.
- *X* is the data of 6 linguistic questions after weighting and the total of these 6 values.
- *Y* is the label of the data; it is the self-assessment value for Linguistic Intelligence.
- *Test_size = 0.20* means that we divide the data into training set and testing set, in which the testing set is 20% of overall.

Step 2: To avoid the overfitting and underfitting of the dataset, we use the method *StandardScaler*.

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
scaler.fit(Linguistic_train)
Linguistic_train = scaler.transform(Linguistic_train)
Linguistic_test = scaler.transform(Linguistic_test)
```

Fig 15. Handling data with StandardScaler

Step 3: We proceed to build the KNN model for linguistic Intelligence with the method *KNeighborsClassifier* in the library class *sklearn.neighbors*.

```
from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors = 10,
                                  algorithm='ball_tree',
                                  weights='distance')
classifier.fit(Linguistic_train,
               LB_Linguistic_train)
```

Fig 16. Building KNN Model for Linguistic Intelligence

We use the *KneighborsClassifier* in the *SKLearn* with

- *n_neighbors = 10:* quantity of neighbors which is used for *kneighbors queries* (sklearn, 2020).
- *Algorithm = 'ball_tree'*: we use the *ball_tree* algorithm to get better performance for the model.
- *Weights= 'distance':* we give weight depending on the distance.

After all, we calculate the fitness of the model.

Step 4: Chúng tôi tiến hành đưa ra dự đoán về y_prediect (nhãn dự đoán của tập dữ liệu) như sau.

```
Linguistic_pred = classifier.predict(Linguistic_test)
```

Fig 17. Predicting label set

Step 5: We use *sklearn.metrics* to calculate the accuracy and loss of the built model.

```
result = confusion_matrix(LB_Linguistic_test, Linguistic_pred)
print("Confusion Matrix:")
print(result)
```

Fig 18. Calculate the confusion Matrix

```
result1 = classification_report(LB_Linguistic_test, Linguistic_pred)
print("Classification Report:",)
print (result1)
result2 = accuracy_score(LB_Linguistic_test,Linguistic_pred)
print("Accuracy:",result2)
print("Loss:", hamming_loss(LB_Linguistic_test,Linguistic_pred))
```

Fig 19. Accuracy and Loss of the Model

We use the method *accuracy_score* to calculate the accuracy of the model. To get the loss, we use the method *haming_loss* for calculating the loss result.

The above steps are a sequence of building Intelligence Model for classifier. We would apply the same sequential process for build the rest Models. They are 7 models (Existential Model, Intrapersonal Model, Interpersonal Model, Bodily-Kinesthetic Model, Spatial Model, Musical Model, and Logical-Mathematical Model).

In *fig 20* is the result of ground truth and predict labels for Linguistic Intelligence Model. In this chart, the orange dots represent ground truth and the blue dots are predict labels. In this chart, the horizontal dimension is the score of the observed elements (the score of who has joined the survey), the vertical dimension is the label of Linguistic Intelligence which is 1 to 10. We would do the same for other Intelligences. The results are shown in *fig 21 to fig 27*.

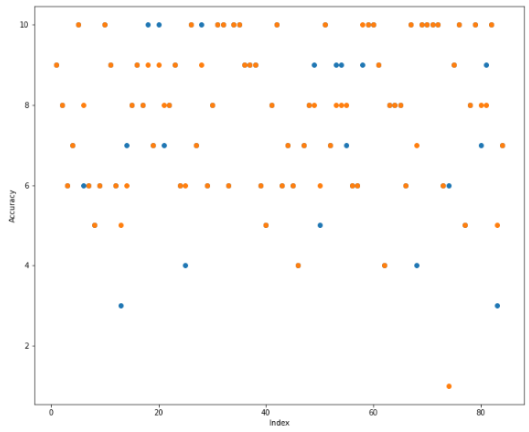

Fig 20. Ground truth and predict labels of Linguistic Intelligence

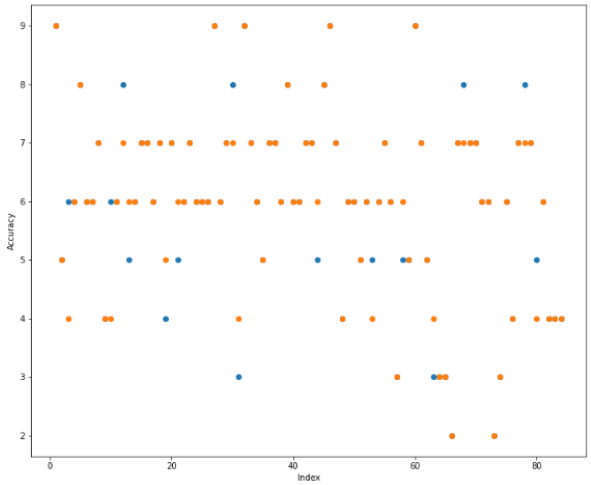

Fig 21. Ground truth and predict labels of Logical-Mathematical intelligence

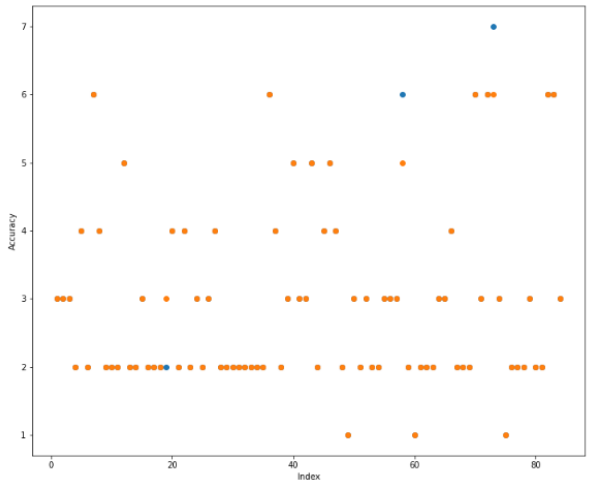

Fig 22. Ground truth and predict labels of Musical intelligence

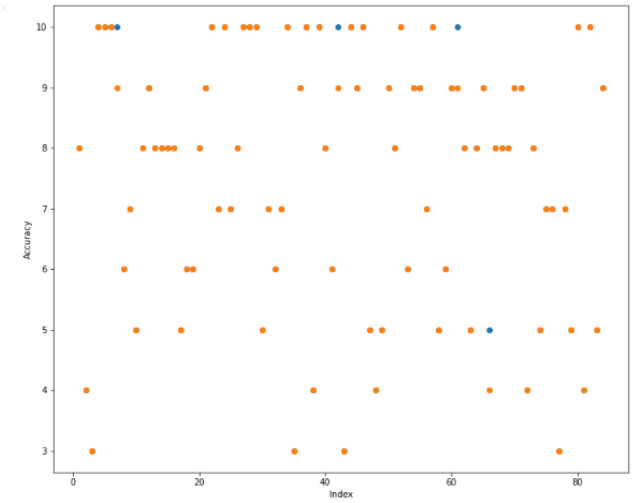

Fig 23. Ground truth and predict labels of Bodily-Kinesthetical intelligence

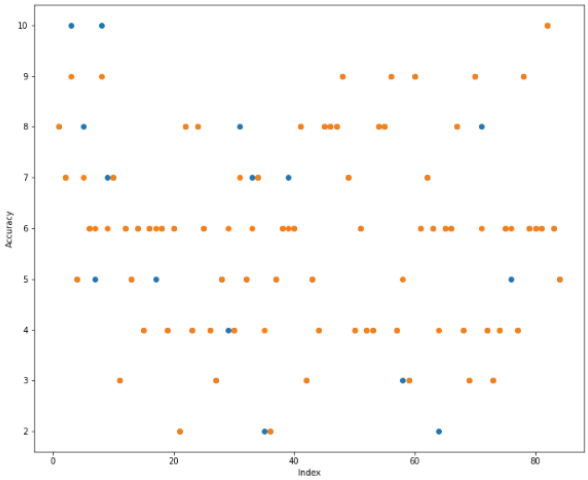

Fig 24. Ground truth and predict labels of Spatial-Visual intelligence

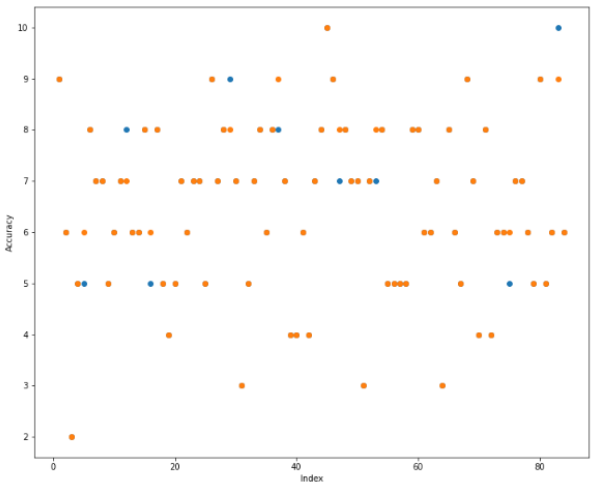

Fig 25. Ground truth and predict labels of Interpersonal intelligence

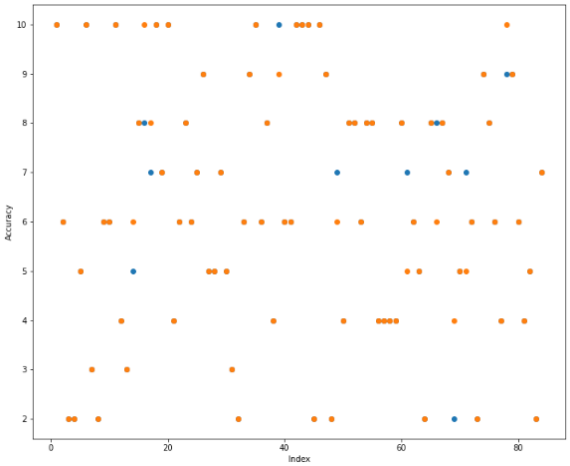

Fig 26. Ground truth and predict labels of Intrapersonal intelligence

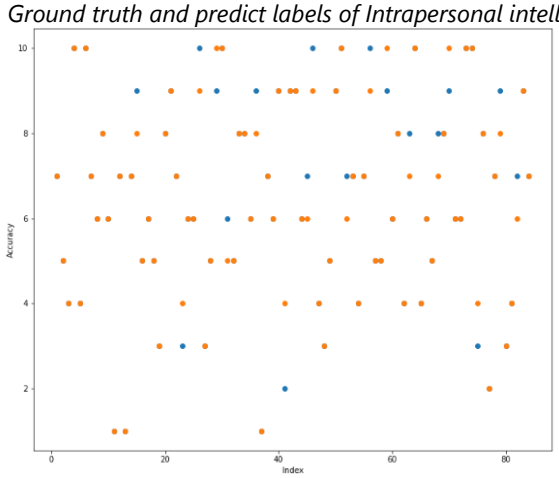

Fig 27. Ground truth and predict labels of Existential intelligence

From the figure 20 to figure 27, we can see that the ground truth and prediction labels are almost the same. The distributions are quite stable. We can see that ground truth dots and the predicted labels dots are displayed in the same area. Only a few dots of the predicted label deviate from the correct label of the data set. To have a better understanding about the models, we check the accuracy of this first model. We proceed to calculate the algorithm's accuracy through the method *sklearn.metrics.accuracy_score* and the loss using the loss function *sklearn.metrics.hamming_loss* to calculate the wrong labeled values. We get the results for 8 intelligences as following:

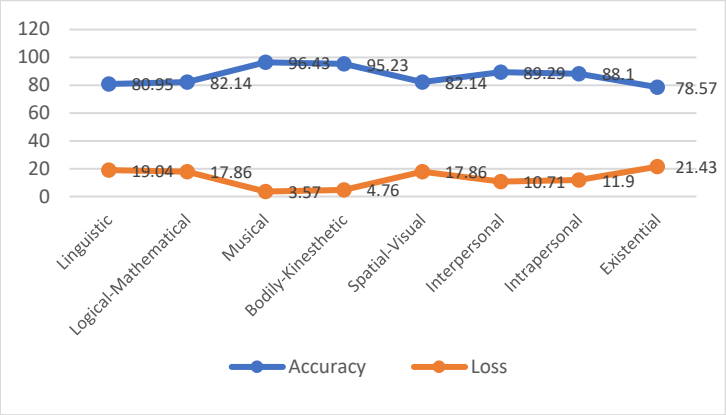

Fig 28. Accuracy and Loss of 8 Models

From the chart above, we can see that the accuracy of all the models is quite high, and it has good reliability to apply to the career counseling system. After getting the 8 models, we then proceed with the data of 5 students as input to run the 8 models for counselling information. The results of the 5 students are as follows:

Result from the 1st student:

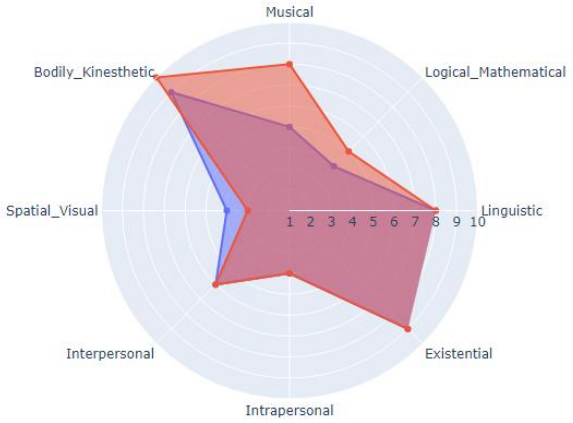

Fig 29. Intelligences Distribution of the 1st student

The fig 29 is the chart which shows the level of 8 intelligences with levels predicted by the system (the blue polygon and its area), and the red part is from the student. The red area is the student's self-assessment of his 8 intelligences. From the above chart we can draw the following conclusions for the career path counselling:

- The linguistic intelligence, interpersonal intelligence, intrapersonal intelligence, and existential intelligence have a match between the results self-assessed by the student and the results classified by the proposed system.
- The spatial-visual intelligence, bodily kinesthetic intelligence, and logic-mathematical intelligence have the slightly differences between the student self-assessment and the results given by the system, the different distance is 1.
- The musical intelligence has the much difference between the user's self-assessment and the result given by the system, the distance is 3.

According to the chart above, we can see that physical intelligence, linguistic intelligence and existential intelligence have the highest scores. Based on the research of Richard Juma Atela (Atela, Richard Juma; Othuon, Prof. Lucas; Agak, Prof. John .O., 2019), we can see this student should have the following careers:

- *Physical intelligence*: Athletes, choreographers, dancers, music directors, firefighters, artisans.
- *Linguistic intelligence*: Poets, journalists, public relations, writers, teachers, politicians, lawyers, interpreters.
- *Existential intelligence*: Preacher, philosopher, star reader, astrologer

Result from the 2nd Student:

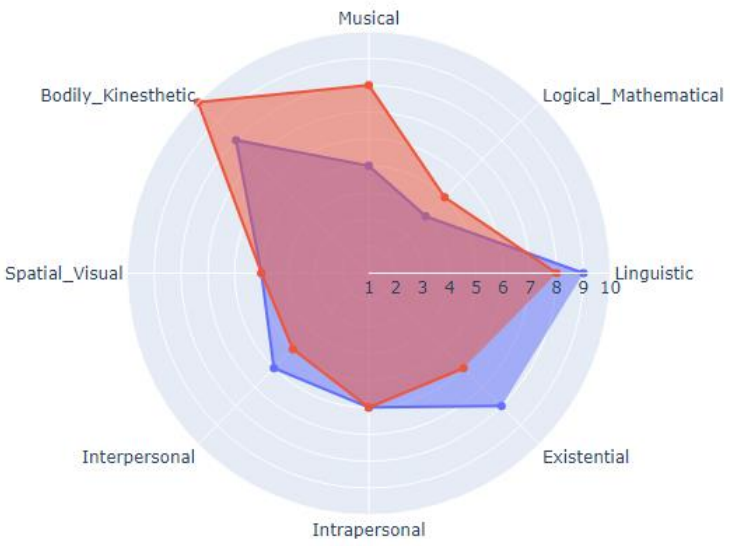

Fig 30. Intelligences Distribution of the 2nd Student

The result chart of the 2nd student shows the level of 8 intelligences, we can draw the following conclusions:

- The outcome of intrapersonal intelligence and spatial-visual intelligence from both results are the same.
- There are slightly differences between the student self-assessment and the results given by the system about interpersonal intelligence and logical mathematical intelligence. The distance is 1.
- The bodily-Kinesthetic intelligence has the difference between the student self-assessment and the results given by the system. The distance is 2.
- Musical intelligence has much difference between the student self-assessment and the results given by the system. The distance is 3.

With this 2nd student, we can see that he is strongest at the linguistic intelligence. Careers (Atela, Richard Juma; Othuon, Prof. Lucas; Agak, Prof. John .O., 2019) which are suitable for him:

- *Linguistic intelligence*: Poet, journalist, public relations, writer, teacher, politician, lawyer, interpreter

Result from the 3rd Student:

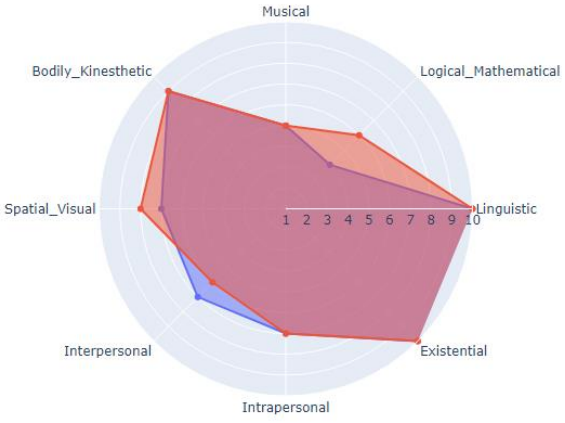

Fig 31. Intelligences Distribution of 3rd Student

Based on the result chart of the 3rd student, we can draw the following conclusions:

- About the Bodily-Kinesthetic intelligence, Intrapersonal intelligence, Existential intelligence, Linguistic intelligence, and Musical intelligence, there are a match between the self-assessed results and the results classified by the system.
- With the Spatial-Visual intelligence and Interpersonal intelligence, they have slightly difference, and the distance is 1.
- The Logical-Mathematical intelligence has the difference. The distance is 2.

The chart also tells us that Linguistic intelligence, Bodily-Kinesthetic intelligence, and Existential intelligence have the highest score. Careers (Atela, Richard Juma; Othuon, Prof. Lucas; Agak, Prof. John .O., 2019) are suitable for this student:

- *Linguistic intelligence:* Poet, journalist, public relations, writer, teacher, politician, lawyer, interpreter.
- *Bodily-Kinesthetic intelligence:* Athlete, choreographer, dancer, music director, firefighters, artisans, P.E tutors
- *Existentialist intelligence:* Preacher, philosopher, star reader, astrologer

Result from the 4th student:

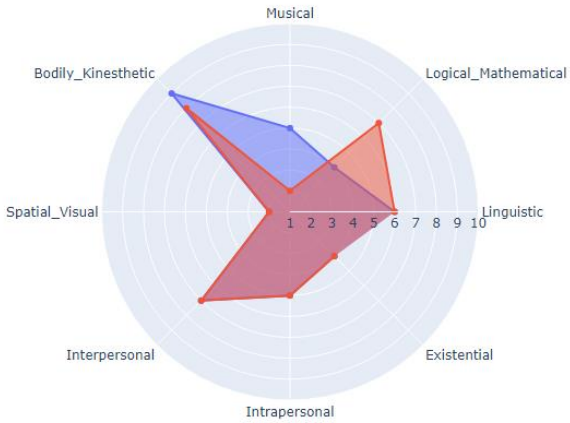

Fig 32. Intelligences Distribution of 4th student

As the 1st student, we can draw the following conclusions:

- With the Spatial-Visual intelligence, linguistic intelligence, existential intelligence, interpersonal intelligence, intrapersonal intelligence, there are a match between the student self-assessment and the results given by the system.
- About the Bodily-kinesthetic intelligence, this student has slightly difference between the student self-assessment and the results given by the system, the distance is 1.
- Logical-mathematical intelligence and musical intelligence have the difference between the student self-assessment and the results given by the system, the distance is 3.

The chart above shows that this student is strongest about Bodily/ Kinesthetic intelligence with the highest score. Careers (Atela, Richard Juma; Othuon, Prof. Lucas; Agak, Prof. John .O., 2019) are suitable for this student:

- *Bodily-Kinesthetic:* Athlete, choreographer, dancer, music director, firefighters, artisans, P.E tutors

Result from the 5th Student

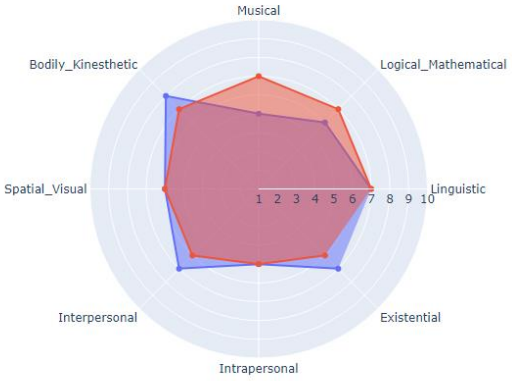

Fig 33. Intelligences Distribution 5th student

As the 1st student, we can draw the following conclusions:

- With Intrapersonal intelligence, Linguistic intelligence, Spatial-Visual intelligence, there are a match between the student self-assessment and the results given by the system.
- About Bodily-Kinesthetic intelligence, Interpersonal intelligence, Existential intelligence, and Logical-Mathematical intelligence, this student has a slight difference the between self-assessment and the results given by the system, the distance is 1.
- Musical intelligence has the difference between the student self-assessment and the results given by the system, the distance 2.

Through the chart above, we can see that Bodily-Kinesthetic intelligence has the highest score. Based on research by Richard Juma Atela (Atela, Richard Juma; Othuon, Prof. Lucas; Agak, Prof. John .O., 2019) the Careers which is suitable for the 5th student are:

- *Bodily/ Kinesthetic:* Athlete, choreographer, dancer, music director, firefighters, artisans, P.E tutors

5. Conclusion

In this paper, we have proposed a career recommendation system that combines the theory of multiple intelligences and the K nearest neighbor (KNN) classification algorithm. We have collected a dataset of 419 people joining the survey, with the age of 25 years and older, as the training dataset of the predictive model. After the experimental work of 38 test records, the program outputs the results of the level of each intelligence with the following accuracy respectively: linguistic intelligence (80.95%), logical-mathematical intelligence (82.14%), musical intelligence (96.43%), kinesthetic intelligence (95.23%), visual-spatial intelligence (82.14%), interpersonal intelligence (89.29%), intrapersonal intelligence (88.1%), existential intelligence (78.57%). Thereby, we rely on the ratio of these intelligences to generate the suitable careers (Abenti, Heather Freya, 2020). This study also shows the potential of AI, especially machine learning techniques in career path counseling, as well as the core value of multiple intelligences theory. To improve the quality of career counseling, this study should improve the accuracy of the training dataset, or se other classification methods in machine learning, or even deep learning to increase the accuracy of the predictive model. In this paper, there is some limitation about the dataset and its size. We can enhance and improve the effectiveness of this study by upgrading and focusing on taking dataset and qualifying it with professional scholar in Multiple Intelligences. These datasets can be figured and adapted with variant and diversity regions around the world. This is a good fundamental for a useful application for people who want to understand more about self-career in the future. With this achievement, we believe that AI and machine learning will well support human beings with the best and most suitable career orientation in the future. This research received no external funding

Funding: This research received no external funding.

Acknowledgements: We are so grateful that Ho Chi Minh City Open University has enabled us to carry out this study. Besides, we would also like to express our sincere thanks to the people who have joined the surveyed and supported us to contribute the dataset of this research

Conflicts of Interest:

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

[1] Abenti, H. F., & Daradoumis, T. (2020). Exploring Multiple Intelligence Theory Prospects as a Vehicle for Discovering the Relationship of Neuroeducation with Imaginative/Waldorf Pedagogy: A Systematic Literature Review. *Education Sciences, 10*(11).

[2] Abenti, Heather Freya. (2020). How do I teach you? An examination of multiple intelligences How do I teach you? An examination of multiple intelligences. *Language & Communication, 73*, 29-33.

[3] Ahmad, N. A., & Dzulkarnain, S. N. (2020). Utilization of Gardner‟s Multiple Intelligence Theory for School Counselling System with Usability Testing. *International Journal of Recent Technology and Engineering (IJRTE), 8*(6), 2253-2260.

[4] Alez-Treviño, I. &.-R., Valencia-Hernández, G., & Jesús Arrona-Palacios, A. (2020). Assessment of multiple intelligences in elementary school students in Mexico: An exploratory study. *Heliyon, 6*(4).

[5] Assegie, T. A. (2021). K-Nearest Neighbor Based URL Identification Model for Phishing Attack Detection. *10.35940/ijainn.B1019.041221*.

[6] Atela, R. J., Othuon, P. L., & Agak, P. J. (2019). Relationship Between Types of Intelligence And Career Choice Among Undergraduate Students Of Maseno University, Kenya. *Journal of Education and Practice, 10*, 87-110.

[7] Atela, R. J., Othuon, P. L., & Agak, P. J. (2019). Relationship Between Types of Intelligence And Career Choice Among Undergraduate Students Of Maseno University, Kenya. *Journal of Education and Practice, 10*, 87-110.

[8] Aubrey L. C. Statti;Kelly M. Torres. (2020). Multiple Intelligence Theory in the Digital Age of Learning. 18.

[9] Awoyelu, Iyabo; Awoyelu, Tolulope; M, Awoyelu T.;. (2020). Fuzzy K-Nearest Neighbour Model for Choice of Career Path for Upper Basic School Students. *Education and Management Engineering, 4*, 18-32.

[10] Azmir, A. M., Johari, K. S., & Mahmud, I. (2017). Relationship Between Multiple Intelligence and Career. *Proceedings of the 1st Progress in Social Science, Humanities and Education Research Symposium (PSSHERS 2019), 464*, 671-677.

[11] Biswas, Al Amin; Majumder, Anup; Mia, Md. Jueal; Basri, Rabeya; Zulfiker, Md. Sabab;. (2021). Career Prediction with Analysis of Influential Factors Using Data Mining in the Context of Bangladesh. *Proceedings of International Conference on Trends in Computational and Cognitive Engineering, Advances in Intelligent Systems and Computing* .

[12] Bravo, Leonardo Emiro Contreras; Molano, Jose Ignacio Rodriguez; Trujillo, Edwin Rivas. (2020). Exploration of a System to Determine the Academic Performance of Engineering Students Through Machine Learning. *International Journal of Advanced Science and Technology, 29*(7), 11894 - 11905.

[13] businessballs. (2017). *Multiple Intelligences Test - based on Howard Gardner's MI Model*. (businessballs) Retrieved from businessballs: businessballs.com

[14] Cheng Debo, Zhang Shichao, Deng Zhenyun, Zhu Yonghua, Zong Ming. (2014). kNN Algorithm with Data-Driven k Value. *International Conference on Advanced Data Mining and Applications*. Guilin.

[15] Cunningham P, Delany SJ. (2020). k-Nearest neighbour classifiers: : 2nd Edition. *Mult Classif Syst*, pp. 34:1–17.

[16] DeLauney, & Bryan. (2019). Multiple Intelligence.

[17] DIỆP, N. (2021, 4 24). *baotuoitre*. (baotuoitre) Retrieved 5 17, 2021, from https://tuoitre.vn/lam-the-nao-de-khong-chon-nham-nghe-khong-ngoi-nham-truong-20210424092401746.htm

[18] Dr. Gufran Ahmad Ansari. (2017). Career Guidance through Multilevel Expert System Using Data Mining Technique. *Information Technology and Computer Science*.

[19] Geler, Zoltan; Kurbalija, Vladimir; Ivanovi´c, Mirjana; Radovanovi´c, Milo˘s;. (2020). Weighted kNN and constrained elastic distances for time-series classification. *Expert Systems With Applications, 162*.

[20] Haji, E. e. (2020). Proposal of a Digital Ecosystem Based on Big Data and Artificial Intelligence to Support Educational and Vocational Guidance. *International Journal of Modern Education and Computer Science*.

[21] Hu, Juan; Peng, Hong; Wang, Jun; Yu, Wenping;. (2020). kNN-P: A kNN classifier optimized by P systems. *Theoretical Computer Science, 817*, 55-65.

[22] Iatrellis, O., Savvas, I. K., Kameas, A., & Fitsilis, P. (2020). Integrated learning pathways in higher education: A framework enhanced with machine learning and semantics. *Education and Information Technologies*.

[23] Kiselev, P., Kiselev, B., Matsuta, V., Feshchenko, A., Bogdanovskaya, I., & Kosheleva, A. (2020). Career guidance based on machine learning: social networks in professional identity construction. *Procedia Computer Science*, 158-163.

[24] Kiselev, Pavel; Feshchenko, Artem; Matsuta, Valeriia; Bogdanovskaya, Irina;. (2020). CAREER ROBOT FOR MANAGING COLLEGE MAJOR CHOICE. *Proceedings of INTED2020*. Valencia, Spain.

[25] Kolachalama, Vijaya B.; Garg, Priya S.;. (2018). Machine learning and medical education. *nature partner journals*.

[26] Lee, M. F., Lai, C. S., & Wahid, W. A. (2017). Relationship between Multiple Intelligence and Career Interest among the Relationship between Multiple Intelligence and Career Interest among the. *IEEE Xplore*.

[27] Lester, S., & Hofmann, J. (2020). Some pedagogical observations on using augmented reality in a vocational practicum. *British Educational Research Association*.

[28] Li, Bin; Wan, Sailuo; Xia, Hua; Qian, Fengshou;. (2020). The Research for Recommendation System Based on Improved KNN Algorithm. *International Conference on Advances in Electrical Engineering and Computer Applications*.

[29] Mehraj, T., & Baba, A. (2019). Scrutinising Artificial Intelligence based Career Guidance and Counselling Scrutinising Artificial Intelligence based Career Guidance and Counselling. *International Journal of Interdisciplinary Research and Innovations, 7*(1), 402-411.

[30] Nasri, N., Rahimi, N. M., Nasri, N. M., & Talib, M. A. (2021). A Comparison Study between Universal Design for Learning-Multiple Intelligence (UDL-MI) Oriented STEM Program and Traditional STEM Program for Inclusive Education. *sustainnability, 13*(2), 554.

[31] Nasution, Fadliansyah; Zamzami, Elviawaty Muiza;. (2019). Prediction of Vocational Students Behaviour using The kNearest Neighbor Algorithm. *4th International Conference on Computing and Applied*. Medan, Indonesia.

[32] Nugroho, Aldi; Riady, Osvaldo Richie; Calvin, Alexander; Suhartono, Derwin. (2020). Identification of Student Academic Performance using the KNN Algorithm. *Engineering, MAthematics and Computer Science, 2*(3), 115-122.

[33] Oasis, M. (2020, 3 13). *MI Oasis*. (MI Oasis) Retrieved 5 17, 2021, from https://www.multipleintelligencesoasis.org/blog/2020/3/13/7-things-you-should-know-about-multiple-intelligences

[34] Panigrahi, A., & Joshi, V. (2020). USE OF ARTIFICIAL INTELLIGENCE IN EDUCATION. *SSRN Electronic Journal, 55*(5), 64-67.

[35] Paper, D. (2019). *Introduction to Scikit-Learn*. Berkeley: Apress.

[36] Rao, Abhishek S.; Bola Sunil Kamath, Ramya R; Chowdhury, Shreya; Pattan, Shreya A; Kundar, Raveena Krishna;. (2020). Use of Artificial Neural Network in Developing a Personality Prediction Model for Career Guidance: A Boon for Career Counselors. *International Journal of Control and Automation, 13*(4), 391 - 400.

[37] Sadali, Muhamad; Putra, Yupi Kuspandi; Putra, Yupi Kuspandi;. (2019). Evaluation Of Lecturer Education And Teaching Performance Through E-Monevin Using K-Nearest Neighbor (K-NN) Algorithm. *The 5th Hamzanwadi International Conference of Technology and Education*.

[38] Sener, S., & Çokçaliskan, A. (2018). An Investigation between Multiple Intelligences and Learning Styles. *ERIC*.

[39] Shafiee, Nor Safika Mohd; Mutalib, Sofianita;. (2020). Prediction of Mental Health Problems among Higher Education Student Using Machine Learning. *Education and Management Engineering*.

[40] Shichao Zhang, Senior Member, IEEE, Xuelong Li, Fellow, IEEE, Ming Zong, Xiaofeng Zhu, Ruili Wang. (2018). Efficient kNN Classification With Different Numbers of Nearest Neighbors. *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*. Guangxi.

[41] sklearn. (2020). *scikit-learn*. (scikit-learn) Retrieved 4 15, 2021, from https://scikit-learn.org/stable/modules/neighbors.html#nearest-neighbors-classification

[42] Sodhi, Jasjit S.; Dutta, Maitreyee; Aggarwal, Naveen;. (2016). Efficacy of Artificial Neural Network based Decision Support System for Career Counseling. *Indian Journal of Science and Technology*.

[43] Sudarsana, I. K., Surp, N. K., Sena, I. G., & Setiawan, I. G. (2020). Technology Based Hinduism Learning And Multiple Intelligences. *European Union Digital Library*.

[44] T.V, A. O. (2020). The Use of Digital Technologies in the Activities of Teachers in the Vocational Guidance of Students with Mental Retardation. *Digitalization of Education: History, Trends and Prospects*.

[45] Taha, A. A. (2017). Phishing Websites Classification using Hybrid SVM and KNN Approach. *International Journal of Advanced Computer Science and Applications*.

[46] Tang, Ying; Tang, Ying; Hare, Ryan; Wang, Fei-Yue;. (2020). A Personalized Learning System for Parallel Intelligent Education. *IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, 7*(1).

[47] tạo, b. g. (2020, 10 8). *bộ giáo dục và đào tạo*. (bộ giáo dục và đào tạo) Retrieved 5 17, 2021, from https://moet.gov.vn/bovoinguoidan/Pages/hoi-dap.aspx?ItemID=7269
[48] wiki. (2017). *Rating Scale*. (wikipedia) Retrieved 4 22, 2021, from https://en.wikipedia.org/wiki/Rating_scale
[49] wiki. (2017). *Semantic differential*. (wikipedia) Retrieved 4 21, 2021, from https://en.wikipedia.org/wiki/Semantic_differential
[50] Yavich, R., & Rotnitsky, I. (2020). Multiple Intelligences and Success in School Studies. *International Journal of Higher Education, 9*, 107-117.
[51] Yulianto, Lili Dwi; Triayudi, Agung; Sholihati, Ira Diana;. (2020). Implementation Educational Data Mining For Analysis of Student Performance Prediction with Comparison of K-Nearest Neighbor Data Mining Method and Decision Tree C4.5. *Jurnal Mantik, 4*(1), 441-451.
[52] Zahour, O., Benlahmar, E. H., Eddaouim, A., & Hourrane, O. (2020). A Comparative Study of Machine Learning Methods for Automatic Classification of Academic and Vocational. *A Comparative Study of Machine Learning Methods for Automatic Classification of Academi, 14*(8).