 https://doi.org/10.24191/jeesr.v21i1.012

Job classification: A Review on Data, Features, and Methods

Muhammad Haziq Haikal Hisham, Mohd Azri Abdul Aziz*, and Ahmad Asari Sulaiman

***Abstract*— In this review paper, job classification is viewed as a process to classify or to recommend jobs to the group job candidates according to the criteria set. Job classification has wide view of its definition, in this review paper, we are focusing on the job classification in classifying certain jobs area into different categories according to the skills required. Job recommender area also had been explored as it has the same application as the job classification is this study. The purpose of this review is to study on the data, features, and methods used in classifying jobs. Type of data used such as nominal and numerical, features used such like CGPA, demographic factors, and different methods used to classify the data like job recommender systems and classifiers had been figured out in this review paper. Considering the recent works in this area, several recommendations for future works are presented to further improve the performance of job classification especially for graduates.**

***Index Terms*— Classification methods, Data, Features, Job classification, Job recommender system**

I. INTRODUCTION

JOB classification serves a big area to be studied for. Hence, it has attracted many researchers to further study about it. There are studies with several specific purposes conducted under this research area, among them are to predict the best fit candidate for a job [1], to recommend jobs to the candidates [2], to predict graduates' job placement according to their study's performance [3], etc. To achieve their own potential, it is important for the employers to have people with the right competences who can fit their culture [4]. Today's organizations are looking for the employees which not just has the basic academic knowledge but also has the ability to link their skills set and the needs of the respective job [5]. Thus, the hiring process play an important role in recruiting the right employees for the suitable job positions. Not only the candidates need to be classified according to their skills and background course, but the jobs need to be classified accordingly too.

Normally, hiring process can be done through the classical method that includes the resume or curriculum vitae (CV) review and the interview session. K.Appadoo et al [1] had highlighted in their research about the recruitment process that has been changing with the increasing usage of the internet for the past few years. Job applications also have moved online with emails and direct applications on the various recruitment platforms. The authors also stated that this situation had caused a phenomenon of "Information Overload" in the decision support environment that leads to poor decision making. To take an advantage on this situation, researchers should grab the opportunity to explore deeper in this field to utilize the overloaded data wisely instead of wasting it as unexplored data.

From a certain point of view, the term 'job classification' can also be interpreted as 'job recommender'. Where the main idea is to classify certain group of jobs' candidates according to the feature requirements and recommend jobs that match their features. The recommender systems are being used to figure out the interested items for a certain user by employing a variety of information resources that is related to the users and items [6]. Fortunately, it was found that recruiting the appropriate person is a challenge faced by most companies, as well as the unavailability of certain candidates in some skill areas has long been identified as a major obstacle to companies success [7]. Thus, by furthering the study in this area, it will contribute towards the idea to solve the major problem faced by the companies by recommending the most suitable candidates for the job advertised.

In this paper, we reviewed previous and recent techniques explored by researchers through several angles such as the data, features, and the methods used in the job classification. Thus, several recommendations are discussed for the future used and improvement in the job classification area. Main application of this review paper is to get the idea on how to classify and recommend jobs to the groups of graduates. Thus, the 'job classification' and 'job recommender' keywords have been explored since both keywords can be applied in implementing the idea of the system. This sub-section will include methods that were used by researchers in classifying the data for the job classification and job recommendation study. 'Job recommender' keyword has been explored to get the information about the job recommender systems. While 'job classification' keyword has been explored to get to know more about various classifiers used in the job classification study.

This rest of this paper is organised as follows. Section II is a discussion and review on the data used in the classification.

This manuscript is submitted on 22nd December 2021 and accepted on 28th March 2022. M.A.A.Aziz and A.A.Sulaiman are with the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor (azriaziz@uitm.edu.my).

Muhammad Haziq Haikal Hisham was with the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA. (email: 2020402226@isiswa.uitm.edu.my).

*Corresponding author
Email address: azriaziz@uitm.edu.my

Section III discusses on features, section IV explains about the various methods and section V is for the conclusion of this review paper.

II. DATA

Data used in the job classification/recommender study in various classification fields is reviewed as below. Since the 'job recommender' research studies have the same concept as the 'job classification' from a certain point of view and has wider source, this review paper includes the review on 'job recommender' papers that were conducted in various field of study by previous researchers.

Job recommender system can be used in various field: a) E – recruitments/websites [7]–[10], to recruit manpower/talent for an organization, including to increase selling rates customer services and to reduce customers search time. b) Social networks [11]–[15], beBee, Facebook, and LinkedIn are used to get information and interest of users on social networks for advertisements/recommend jobs. c) Students and graduates [16]–[18], to recommend jobs to students, graduates and former graduates by using their historical information during university years. d) Application systems (framework) [19]–[23], JRDP, FoDRA, PrivateJobMatch, Skills2Job, and IHR+ are among frameworks that were used in job recommendation process with various purposes where JRDP is a job recommender system based on ontology for disable people, FoDRA is a content-based job recommender with a more flexible way, PrivateJobMatch is a privacy-oriented deferred multi-match recommender system for stable employment, Skills2Job is a recommender system that encodes job offer embeddings on graph databases (student abstract), and IHR+ used a mobile reciprocal job recommender system to extend the traditional recommendation algorithm in a people-to-people recommender system where both of the subjects and objects of the system are people, etc.

As for data used, Table I shows that the researchers used a wide data range, from 5 until 7000 candidates contributed. Al-Otaibi et al. [7] said that for the sake of simplicity, their research study only used data that were obtained from 5 different candidates CVs which contain the information of job title, job description, qualifications, and skills of the candidates. Also, one of the reasons why the number of data used in the job recommender study is high because each of the individual data is considered to be unique since every human being has different background history and personality. Hence, a single person cannot be chosen several times unlike book or movie [7]. Thus, by having bigger number of data can possibly lead to more unique outcome of the study. It could be said that in the job recommender studies, data used are in big value as it can probably give a clear and better view of the patterns of the results compared to the small volume of data studies although the accuracy of the studies are the same. This will enable the researchers to see analytics in a the results easier as it will help them to discover patterns, comprehend information, and form an opinion [24].

By referring to the information illustrated in Table 1, the number of data used in the job recommender studies are in a big scale from 5 to more than 200,000 amount of data that focus on e-recruitment [7]–[10], social networks [11]–[15], students/graduates [16]–[18], and embedded recommender area [19]–[23]. Normally, the job recommender studies only involved with three different sides of stakeholders which are candidates, employers/companies, and jobs. Where in several studies, the candidates could be different according to the objective of the studies. For instance, in [16]–[18], the candidates were focussed on the students/graduates and in [19], the study focussed on the disable people as candidates.

There are two attribute types that were observed in the studies, which are nominal variable that requires no ranking sequence and numerical variable that is indicated as integer numbers. From four focus area stated in the Table 1, it can be observed that in e-recruitment, students/graduates, and embedded recommender focus area, nominal variable was mostly used as variables. While the job recommender in social network studies used numerical variable the most compared to the other three focus area. This is because the experiments involved with the number of user entries on the social networks as the variables [13].

TABLE I
DATA USED IN PREVIOUS STUDIES IN JOB RECOMMENDER SYSTEM

Focus area	Data used	Attribute Name	Attribute Type	Ref
E-recruitment	1 job description and 5 candidates CV	Candidates' profile (personal information about employee)	Nominal	[7]
		Candidates' profile (professional positions held)	Nominal	
		Candidates' profile (educational experiences)	Nominal	
		Collaboration measures (job profile ratings from candidates)	Numerical	
		Job's profile	Nominal/ Numerical	
	7000 candidates, 400 employers, and 8000 jobs	Candidate	Nominal	[9]
		Employer	Nominal	
		Job	Nominal	
	25,000 pieces of data from 250 companies	Job seekers' portfolio	Nominal	[10]
		Corporate recruitment	Nominal	
		Job search history	Nominal	
		User feedback	Numerical	
Social network	203,990 user entries	Facebook/LinkedIn users (candidate, review, validation, all)	Nominal	[12]
		Job categorization	Nominal	
	15,625 positive examples	Candidate dataset	Numerical	[13]
	3,394 entries	Random dataset	Numerical	
	6,650 entries	Feedback dataset	Numerical	
	25,669 entries	All dataset	Numerical	

	6 collected datasets	Candidate, Feedback, Random, Review, Validation, and All	Nominal	[14] [15]
Students/ Graduates	5000 graduates (estimated 4100 former graduates), 620 employers, and more than 11,000 jobs	Graduate, Employer, and Job	Nominal	[16]
	658 Master graduates and their 3830 ratings to 856 employers that associate	Students, employers, rating, and acceptance	Nominal/ Numerical	[17]
	103 job offers, 72 students grading data point, and 2 courses	Students, job descriptions, and course descriptions	Nominal/ Numerical	[18]
Embedded recommend er	3 cases with jobs and top candidates	Case 1: database developer job, top candidates match for all four aspects	Nominal	[25]
		Case 2: licensed practical nurse job, top candidates meet the LPN license criteria and other required certificates	Nominal	
		Case 3: Regional sales representative, candidates can be selected as long as meet the criteria set	Nominal	
	Job posting during 13 weeks in 2012	Job posting (job title, requirements, description, closing date, posting date, etc)	Nominal	[26]
		Job seekers (education level and major, location, job history, management experience, etc)	Nominal	
		User job application history	Nominal	
	10 people with different disabilities	Type of disabilities	Nominal	[19]

III. FEATURES

Many different features/criteria that are concerned by employers/recruiters when hiring an employee were discussed by researchers in their study. Some of them are candidates' demographic factors, academic performance, non-cognitive skills, communication skills, etc.

CGPA (Cumulative Grade Point Average) is one of many features that is concerned by employers when hiring, especially for fresh graduates [27][28]. It is normally will be considered by the employers during the first phase of the hiring process which is during the resume or curriculum vitae (CV) review session [29]. Especially with a higher number of candidates situation, this feature is useful as it can be used by employers to short list the suitable candidates for the job advertised [2][3][27]. Candidates with higher CGPA tend to be hire easier than others as some of the employers state 'the higher, the better' in term of CGPA of candidates, as it is linearly related to the candidates' educational performance [30]. Many researchers had used CGPA in their studies as it is closely related and most likely can describe a quality or which group of the students they belong to, either low, average, or excellent students [31]–[40]. However in [41], it is stated that CGPA is a worthless criteria for hiring and it predicts nothing. This is because, most of the time in the working environment, soft skills are being used instead theoretical practice learned during university years. Thus, it is too general to conclude that a candidate with higher CGPA can fit to a certain job position as CGPA only reflect about candidates' education performance.

Strong soft skills or non-cognitive skills level is a feature studied in the job classification too. Some of the examples of non-cognitive skills are perseverance, conscientiousness, motivation, sociability [30], leadership [42]–[45], etc. These skills will help an employee to advance once they are part of the company [44]. Researchers in [30] have found that both cognitive (grade performances) and noncognitive skills are important selection criteria, but employers give high priority to noncognitive skills (when selecting candidates among students with intermediate-level school degree). This is supported by senior vice president of people operations for Google [41], he had stated that proportion of people without any college education at Google has increased over time, as high as 14% at some team. This event highlighted that the noncognitive skills is looked up more than cognitive skills.

Out of many noncognitive skills, leadership skill is highlighted by several researchers [42]–[45], where it is important for candidates to have it especially for ones in engineering field. Leadership skill is a general view about skills in a person such like initiative/confidence, interpersonal interactions, communication, engagement, and teamwork. According to [42], leadership skill is a skill where ones can interact, engage, lead, and influence other people effectively toward accomplishing the shared final goal and contributing positively to diverse communities. Specifically focusing in engineering field, leadership is an essential skill for engineering graduates to develop and proficient according to academia and industry [43]. This is also supported in a study that was conducted in India [44], the author had found that leadership skill is among the skills that is required by the employers in engineering field. By referring to this statement, it shows that the leadership skill affects the employability rate of an engineering graduate.

Ones with strong communication skills have higher employability rate compared to others as the research finding in [46] stated that the excellent communication skills are among the top three requirement required by Malaysian employers along with good attitude toward professionalism, and excellent teamwork ability. Commonly, communication skill term has the same meaning as effective communication, where it highlights ones' interpersonal skills with colleagues and oral/written communication skills with clients [42]. By referring to a study in [5], communication skills has a significant impact on employability skills and plays a major role in determining the career-ability of a candidate. Other than that, in [44], it is mentioned that communication skill is concerned by employers as it can shows ones' ability to convey idea clearly with confidence either verbally or written, ability to present confidently in front of the audience, and ability to practice active listening skills and reply.

Furthermore, having an internship experience can affect the graduate's employability and considered as an affecting factor/feature in the job classification too according to some researchers [47][45]. Internship gives a working experience for the graduates to know about the difference between theoretical practice and practical practice of the course, so that they can be prepared for their future work. In [47], internship is the most effective way for the employers/recruiters to recruit an employee. The author also stated that the internship can determine the students' commitment level and preference for the real-world experience. It was also concluded by researchers in [33] that internship is one of the attributes that contained the most information that affects students' employability rate. Besides, in [48], it states that by providing internship, students will be exposed and developed the skill-sets that industries looking for . Hence, by concerning this feature, it shows that employers tend to hire candidates that are well prepared and have the exposure for the working environment.

Demographic attributes such like age and gender [49] are among the hiring factors too. Researchers in [50] had concluded that due to the factors like short-term and long-term goal issues, company's profit, and behaviours of the workers, employers prefer to hire young candidates compared to the old ones. However, several researchers had supported that age is not a factor for a declining performance [51], while others claim that performance may even increase when ones are older. Thus, it is too general to conclude that certain group of age is favourable to be hired by employers, it is all depend on the employers themselves.

Other than that, gender identity as a feature can be considered as balance because researchers in [27] said that women earn more benefits when the employers hiring are based on the academic qualification due to women's high academic success across educational level compared to men. Contrarily in [49], men would be easier to get hired because men are expected to perform well in the working environment since the employers think men are more capable compared to women.

In addition, nowadays, job hiring often can be done with the involvement of the computerized technology too, for instance, the job portals will filter the candidates according to the criteria set and the shortlisted candidates will get the opportunity for to go through the interview session with the employers/recruiters. Normally, job portals consist of two profiles, which are job profile and candidate profile [11][30]–[33]. Where, candidates' profile is to indicate the students' real-life profile and the job profile is to show the details about the job advertised. With the existing of this two profiles, candidates and jobs can be classified accordingly to the requirement/criteria set by the employers/recruiters.

For candidates' profile, candidates' features are related to details that represent themselves and information about them such as resume, most recent job title, skills, location, gender, education level, education history, working experience, type of job needed, amount of salary needed, and job history [25], [26], [52]–[54]. Most of the time, features like job title, skills, location, gender, education level, education history, work experience, and type of job are categorised as nominal [25], [26], [52]–[54], where it has no ranking sequence. To be added, researchers in [55] found that information like type of working, time arrangement, languages, study, work regime, competencies, and work experience are less important information after getting a conclusion from job seekers and job mediators preference. But according to the researchers in [11][56] has found that candidates' features like main education, required skills, desired skills, experience, language, salary range, and geographic area are important and were used in the study because these features can be defined in metric values and can be calculated to relate the candidates' profile with the jobs' profile for recommendation.

Job profile consists of information and details about the job advertised like job title, description, requirement, skills required, location, salary offers, applicants' sex preference, range of age of the applicants, gender [11][13][17][30]–[32]. Research in [55] has found that distance to a job place and the job title are ranked in the first place as the most important information that job candidates considered about, as it is important for them to know whether they can easily reach the location of the job. The authors also stated that the job postal code as an important parameter since some cities are much easier to reach than others. In [57], the authors emphasized the important of the job title and its description as for instance, the job title and description were used by the researchers in their own specific module to predict/recommend the most related job according to the job candidates' profile. This is supported by researchers in [58], where it was stated that the job title and its description are the main attributes that hold meaningful data for the recommendation process. Furthermore, in [25], not only job title and description, the researchers had used job title and skill as it is stated that traditionally, these two features have been focuses in job classification and recommendation task. These two features are important semantic entities as they are (semi-) structured fields and consists of much information in the job-related documents. Throughout the study, three different case studies had been conducted by using three different job titles and skills as one of the criteria and had come out with impressive quality of classification.

Table 2 shows the summary of features discussed above. To be conclude, it is observed that CGPA is the most used and one of the important features that were used in the job classification and job recommender area by many researchers.

IV. METHODS

Various methods were used in previous studies by different researchers in classifying jobs. Classifiers and specific systems are among the methods used by them. The details are explained as below:

TABLE II
THE SUMMARY OF FEATURES/CRITERIA HIGHLITED BY PREVIOUS RESERACHERS

Features/criteria highlighted	**References**
CGPA (Cumulative Grade Point Average)	[27][28][29][31]–[37][38]–[40]
Leadership skill	[42]–[45]
Non-cognitive/Soft skill	[30][42]–[45]
Communication skill	[5][44][46]
Internship experience	[33][45][47]
Age	[49][50][51]
Gender	[27][49]
Candidate's profile on job portals	[25][26][52]–[54]
Job profile on job portals	[11][13][17][30]–[32]

A. Job Recommender Systems

Job recommender systems have gained much attention in both industries and academia since it can obtain the appropriate item like candidates and jobs from the online recruiting information based on the preference of users such as recruiters and applicants [53]. Job recommender systems are frequently used in online recruitment platform such as in CareerBuilder [25][26], Facebook, LinkedIn [15], etc. It is mainly to be used for recommending jobs for the job seekers or the users of the job recommender system [11][14][30]–[32][34] by analysing and determine the similarity of the jobs and candidates content [56]. Each data in the system is important as it can indicates valuable note for the system such as the users' information is to represent the users' profile, jobs' information is to show the jobs' advertised, and the jobs applied by users reflect the preferences of the job applicant [53].

Nowadays, many job recommendation researches had been done by various researchers. One of them is to study the performance of the job recommendation by using an embedding-based approach [9][11][13][14][30]–[34][52]. Content-based recommender system has the advantages in reducing and generalizing the cold-start problem [25]. The recommendation process matches up the users' attributes from the user profiles against the set of properties of item content that the users has liked in the past [20]. This embedding technique allows an easy multi-feature convolution to achieve reliable and efficient item retrieval [25]. Then, a case-based job recommender system is a system that is capable of deciding, on the fundamental of a series of metrics and same case stored in the system, whether the system is likely to be recommended to a user [56]. Referring to the job recommendation study in [26], an experiment about the performance of content-based recommender, case-based recommender, and the hybrid approach of a combination of content features with well-structured features was conducted. The experiment used several feature-based item representations, together with various feature weighting schemes. The results show that the content-based and case-based recommender were outperformed by the hybrid approach in the recommendation process, especially when using the document frequency (DF) weighting scheme. Where DF is used to notice a feature that is shared by many cases is more likely to be important from a job requirements perspective.

Besides, there are also a profile-based job recommender system [52][53][57][58]. The job seekers' profile in this system will be filtered based on their skill set for the job recommendation process [52]. This profile-based system can solve one of the job recommenders' problems where the job applicants do not update the user profile in a timely manner. Where the system will automatically update the user profile dynamically based on the jobs applied and behaviours of the job applicants [53]. A profile-based recommender will ease candidate to find jobs as candidates will be recommended with jobs based on the matched skills of the candidates and jobs [58]. There are several ways to improve the quality of the profile-based job recommender based on previous studies. In [58], one of the techniques that were used in the profile-based job recommender system is text clustering method. Where there are two text clustering methods were used in this study which are Word2vec and K-means clustering techniques. Purposedly, these two techniques were used to group a given text document set about jobs into clusters to match it with the candidate's information for the job recommendation process. While in [57], user's career progression model was combined with profile-based job recommender system to improve the suggestion of suitable jobs to the users. This improvised system consists of several individual modules such as title transition module (TT), title description module (TDM), and Description-Description module (DDM) that was successfully proven had improved the quality of the job recommendation.

Other than that, there is a rating prediction-based job recommender system was proposed to help the college students in securing their first job [17]. This system was mainly studied the job recommendation from student's perspective. It will generate a list of potential employers as a recommendation for the students according to the feedback and rating given by the graduates who have already received job offers from employers.

Lastly, there is the use of the taxonomy-based job recommender systems that has been applied on Facebook and LinkedIn profiles in [13][15]. This system uses the O*NET-SOC taxonomy that is developed mainly to recommend jobs to the users according to the users' attributes on the social network platforms like Facebook and LinkedIn. It contains a vector model with two similarity functions based on the AND and OR fuzzy logic's operators. The results obtained from [15] proved that taxonomy-based vector model improved the performance of the job recommender systems and reduced the difference of quality between LinkedIn and Facebook data in the task of the job recommendation.

To be concluded, each of the recommender system has its own way of working that is different from each other. For instances, content-based job recommender system recommend jobs to the users by matching up the users' attributes from the user profiles against the properties of item content that the users has liked from the [25], case-based job recommender is processed by a series of metrics that analyse the similarity between users and job offers that will lead to the best proposed individual solutions [26][56], profile-based job recommender system can recommend jobs by exploiting users' interest and job, then proposing a personalized recommendations of users and jobs [52][53][57][58], rating prediction-based job recommender system will recommend jobs according to the feedback received about the employers whether it recommended or not [17], and taxonomy-based job recommender system that recommends jobs according to the users' activities on the social media platform [13][15].

B. Classifiers

1. Naïve Bayes (NB)

Generally, Naïve Bayes is a method that is based on the theory of probability [59]. It is the simplest probabilistic classifier in both training and classifying stage [52][53]. It tends to work well on the text classification and normally takes less time to train when compared to other classification model like Support Vector Machine (SVM) [62]. This method consists of a set of supervised learning algorithms that are based on Bayes' Theorem with an assumption of independence among predictors. In other words, each features that exist in this algorithm is not related to other features, they are independent from each other [54][55]. Naïve Bayes classifier can calculate the most possible output based on the input given. This method allows each of the features to contribute towards the output fairly and independently from other features [60]. This classifier can be used in both continuous and categorical variables [63].

There are several studies in the job classification and recommender area that used NB as the classifier [65][66][67][68]. For instance, to simplify and solve the problem of the overflowing information about attributes and criteria of a job position on job search engine websites, a collaboration of NB classification and web scraping technique was proposed in [65]. The results show that NB classification algorithm gives consistent accuracy above 70% (the average is 71.87%) in the testing of five-time classification on eight categories. In another job recommendation study [66], NB classifier was used to give recommendation to the users about the most suitable job for them. The study had been concluded with two situations, for Division Ratio 0.2 the average accuracy is 91.33% and for Division Ratio 0.25, the average accuracy is 92.74%. Referring to [68], NB was used to recommend job position for the users according to users' resume. The classifier takes information as academic performance, education, professional experience, established publications, exhibits and projects, grants, etc. Besides, NB classifier also been used in job seeker profile classification based on their social media account (Twitter) data [67]. NB classifier algorithm was used and collaborated with W-IDF (Weighted-Inverse Document Frequency) weighting to classify the personality of the recruits.

However, according to several studies [67][69]–[72], there are some weaknesses detected in NB classifier that leads to a poor classification performance. In [67], the experiment conducted by using NB as the classifier algorithm for job seeker profile classification for 120 personal Twitter data had given a poor classification accuracy with only 36.67%. Several factors might affect the classifier performance. As NB classifier is based on the theory of probability and assumption [59][69], it was found that NB gives a poor assumption that is rarely true in reality which will harm its performance especially when dealing with complex attribute dependencies [69][72]. Besides, overfitting and underflow are two major limitations in NB which cause the classifier to not perform well during the experiment when dealing with high dimensional data like gene expression data [70].

2. Decision Tree (DT)

Decision tree (DT) is the most commonly used classification algorithm due to its ease of execution and easier to understand compared to other classifiers [73]. DT is a flowchart like tree structure. It consists of three typical nodes which are decision nodes (symbolised by square), chance nodes (denoted by circle), and end nodes (signifies by triangles) [74]. Normally, there are two types of DT: classification tree and regression tree. Where classification tree is used to predict which data belongs to which labels and the regression tree is generally used where the class label is real number [74]. Several techniques are used in DT such like J48, NBTree, Reptree, CART, ID3, C4.5, C5.0, CHID, CHAID, and Hunt's algorithm [73]–[77]. Note that C4.5 is an improvement of ID3 algorithm. According to previous experiments [73][75], C4.5 and J48 (Java version of C4.5) gives the best performance accuracy compared to other models when dealing with small datasets. Besides, CART model can be considered to be used instead of other models because CART can handle outliers well, which ID3 and C4.5 performance are affected by outliers [78].

Furthermore, in [78], the job recommendation system is quite unique where the job seekers need to write an essay for the system to classify the suitability of the applicants' character with the company's work culture. Hybrid DT, the combination of DT (CART model) and Particle Swarm Optimization (PSO) classifier gives better results than ordinary DT classifier in the job recommendation based on the essay submitted. DT classifier also being used [79] to make job recommendations based on candidate's profile and the preserving candidate's job behaviour or preferences. It can be said that DT in this study is being used as a content-based match up. Since video resume or video CV has become one of the requirement for certain recruiters, researchers in [80] had conducted a study about using DT as one of the algorithms used for the job candidate screening from video CVs. The DT model is said to ease the interpretability and implementation process.

On the other hand, the disadvantage of serial DT such as ID3, C.45 and CART is that the algorithm gives low classification accuracy when the training data is large [73]. Furthermore, there are several more limitations of DT algorithm that is summarized in [76]: 1) DT has fragmentation problem if the data is gradually separated into smaller segments. 2) Replication problem can be detected when sub-trees are replicated in DT. 3) Problem in partitioning continuous data such as age. 4) Repetition problem, happens when the features are repeatedly tested along a path in a DT. 5) ID3 and C4.5 tend to take multi valued attributes due to the formula of the entropy and information gain that leads to difficulty in the making of the root of the tree. 6) ID3 model does not backtrack in searching which causes the algorithm easily converged local optimal answer but not global optimal answer. 7) DT does not provide incremental learning because it requires abandoning the original DT. 8) Issue in handling of range inputs because of the limitation of the membership grade and entropy calculation method. 9) DT does not express the output easily in mathematical relations like XOR, Parity Check, and multiplexer problem. 10) Overfitting DT that does not generalize the data well.

3. Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is an arithmetical model that is inspired by the organization and functional feature of biological neural networks [81]. It is a powerful and complex modelling tool for modelling nonlinear functions that often involved in the real world systems [82]. ANN is a biological motivated intelligent method that is generally made of several highly interrelated nodes or processing element units that functionality is loosely based on animal neuron [83]. It consists of three layers of arranged nodes: input layer, hidden layer, and output layer. General rule of an ANN is that it is an adaptive system that adjusts its structure depending on the internal and external information that involved in the process throughout the learning process [63][65]. To be added, ANN is an algorithm that does not make assumptions about the nature of data distribution, and the function is learned from training samples [85].

In a job recommendation study [86], it was found that Neural Network give better results in F1 score compared to Decision Tree, Naïve Bayes, and Support Vector Machine (SVM). The results show that NN best learning rate is 0.001, the best number of iterations for one hidden layer is 50 and 30 for two hidden layers, and it was proven that the number of hidden layers together with the items/neurons in hidden layers do not affect the result.

Same as other classifiers, there are some limitations in ANN. Research in [82] had found that ANN is limitedly used in educational research area due to some reasons like the difficulty for an ANN model to provide the black-box nature (suitable explanation), the time needed for the neural network training, and proneness to over-fitting. Besides, researchers in [87][88] had concluded some other limitations of ANN algorithm: 1) ANN is not a daily life general purpose solver. 2) No structured methodology available in ANN. 3) No single normalized paradigm of ANN development. 4) Predictable output quality of ANN. 5) Many ANN systems do not provide detail and explanation of the way to solve the problems. 6) Black box Nature. 7) Bigger computational burden. 8) Empirical nature of model development.

4. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) is named as one of the ten most popular algorithms [89] and the most popular supervised learning method [90]. It is known for its simple, easy, and example-based algorithm. KNN classifier it to classify unlabelled observations by labelling them according to the class of the most similar labelled example [91]. KNN is one of the many algorithms that is often used in classification research. Mainly, this algorithm is to classify the objects according to the features and training data which has the nearest distance to a new object and it is based on the Euclidean equation formula [89]. It is a nonparametric method that does not depends on building model at the training phase, and a classification rule that is based on a given similarity function between the instance to be classified and the training instances [90]. KNN is being used in the classification area due to several convincing factors like its simpleness, good in handling noise, and uses a huge computerized [89].

In a job recommendation study called "Resume Classification and Matching" [92], KNN algorithm was used as one of the CV recommender models. It is to determine whether the candidates' resume or CV are the nearest to the job description or not. Firstly, an open-source library called "gensim" was used to get job description and CVs on a similar scale for the generation of the summary of the provided text in the provided word limit. Once the summary of the job description and CVs was generated, then KNN was applied to find which CVs that are closely matching with the provided job description for the recommendation process.

5. K-means clustering

Clustering is one of the operations used in unsupervised information retrieval, automatic topic extraction, text mining, feature selection, pattern recognition, and document organization, machine learning, data mining, image analysis, and bioinformatics, and it is a process of separating the knowledge into groups of related items. [93]–[99]. K-means clustering algorithm is a popular and the oldest partitional method [99]. It is numerical, iterative, non-deterministic, unsupervised [97], theoretically proven that k-means algorithm has been showing to perform as well as or better than some other clustering methods, and has an appealing computational efficiency [98]. K-means clustering algorithm is also well known for its efficiency in clustering big data sets [100]–[102], a widely used clustering method for its fast convergence speed and simplest principle [101][94], and the algorithm's results depend extremely on the initial cluster centres [103].

Generally, about the job recommender system by using K-means clustering algorithm, the job offers with similar features are grouped into clusters and the job seekers with the related features will be recommended the jobs for them according to the job features match. Referring to [58], k-means clustering is one of the selected clustering algorithms in this job recommendation study due to its variant algorithm (Spherical k-means) that considers all vectors as normalized and using cosine dissimilarity based on the angle between the vectors. The researchers had used k-means clustering algorithm in the text clustering method to group the given text document about the job offers into clusters in a way that the documents in a same cluster are more similar between each other. As a result, job offers are grouped into job clusters based on the common features/criteria and it is matched to the job seekers according to their interactions.

V. CONCLUSION

In this paper, a review on job classification/recommender focusing on its data, features used, and methods to classify the data. The review covers on the overall data used by researchers in the study, various preference of features used according to idea of the studies, and the variation of methods used in classifying the data of the experiment. The keyword-related papers were collected from digital library and reviewed. The most relevant ones have been included in this review paper.

It was highlighted that the most common feature and criteria in all research is CGPA, either from ‘job classification’ or ‘job recommender’ keyword. It is because CGPA can give a brief representation of a graduate, especially academically. Followed by jobs’ profile on the job portals, candidates’ profile on the job portals, non-cognitive skills, and others. Oddly, gender identity also identified as a feature/criterion that affect ones’ employability status. Other than that, the recommender systems are the commonly used method in today’s applications to recommend/classify jobs compared to classifiers.

To date, most of the work in the area of ‘job classification’ and ‘job recommender’ are not related to the classification of graduates of final year students. The number of this area of study can be said second to none. Specifically in Malaysia, with the unemployability rate for the fresh graduates in 2019 reached up to 170,300 graduates which is 5.5% more than previous year which is 161,300 graduates [104]. It is a good opportunity to contribute and further study to improve the employability rate of the fresh graduates by recommending jobs from employers that suit them the most.

VI. REFERENCES

[1] K. Appadoo, S. Muhammad Bilaal, and M.-D. Zahra, “JobFit : Job Recommendation using Machine Learning and Recommendation Engine,” *IEEE Asia-Pacific Conf. Comput. Sci. Data Eng.*, pp. 1–6, 2020.

[2] G. Özcan and S. G. Öguducu, “Applying different classification techniques in reciprocal job recommender system for considering job candidate preferences,” *2016 11th Int. Conf. Internet Technol. Secur. Trans. ICITST 2016*, pp. 235–240, 2017.

[3] S. Venkatachalam, “Data Mining Classification and analytical model of prediction for Job Placements using Fuzzy Logic,” pp. 765–770, 2021.

[4] R. M. Lima, D. Mesquita, C. Rocha, and M. Rabelo, “Defining the Industrial and Engineering Management Professional Profile: A longitudinal study based on job advertisements,” *Producao*, vol. 27, no. Specialissue, pp. 1–15, 2017.

[5] V. Ajit and P. B. Deshmukh, “Factors Impacting Employability Skills of Engineers,” *Int. J. Sci. Res.*, vol. 2, no. 4, pp. 30–32, 2013.

[6] S. T. Al-Otaibi, “A survey of job recommender systems,” *Int. J. Phys. Sci.*, vol. 7, no. 29, pp. 5127–5142, 2012.

[7] S. T. Al-Otaibi and M. Ykhlef, “Job Recommendation Systems for Enhancing E-recruitment Process,” *Proc. Int. Conf. Inf. Knowl. Eng.*, pp. 1–7, 2012.

[8] A. Mogenet, T. A. N. Pham, M. Kazama, and J. Kong, “Predicting online performance of job recommender systems with ofline evaluation,” *RecSys 2019 - 13th ACM Conf. Recomm. Syst.*, pp. 477–480, 2019.

[9] Y. Lu, S. El Helou, and D. Gillet, “A recommender system for job seeking and recruiting website,” *WWW 2013 Companion - Proc. 22nd Int. Conf. World Wide Web*, pp. 963–966, 2013.

[10] C. Lun, C. Ting, and Y. Lu, “A hybrid - feedback recommender system for employment websites,” *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2020.

[11] A. Rivas, P. Chamoso, A. González-Briones, R. Casado-Vara, and J. M. Corchado, “Hybrid job offer recommender system in a social network,” *Expert Syst.*, vol. 36, no. 4, pp. 1–13, 2019.

[12] E. Malherbe, “Field Selection for Job Categorization and Recommendation to Social Network Users,” no. Asonam, 2014.

[13] M. Diaby, E. Viennet, and T. Launay, “Toward the next generation of recruitment tools: An online social network-based job recommender system,” *Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2013*, pp. 821–828, 2013.

[14] M. Diaby, E. Viennet, and T. Launay, “Exploration of methodologies to improve job recommender systems on social networks,” *Soc. Netw. Anal. Min.*, vol. 4, no. 1, pp. 1–17, 2014.

[15] M. Diaby and E. Viennet, “Taxonomy-based job recommender systems on Facebook and LinkedIn profiles,” *Proc. - Int. Conf. Res. Challenges Inf. Sci.*, 2014.

[16] Y. Ding, Y. Zhang, L. Li, W. Xu, and H. Wang, “A reciprocal recommender system for graduates’ recruitment,” *Proc. - 2016 8th Int. Conf. Inf. Technol. Med. Educ. ITME 2016*, pp. 394–398, 2017.

[17] R. Liu, Y. Ouyang, W. Rong, X. Song, C. Tang, and Z. Xiong, “Rating prediction based job recommendation service for college students,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9790, pp. 453–467, 2016.

[18] A. Jacobsen and G. Spanakis, “It’s a match! Reciprocal recommender system for graduating students and jobs,” *EDM 2019 - Proc. 12th Int. Conf. Educ. Data Min.*, no. July, pp. 580–583, 2019.

[19] S. Shishehchi and S. Y. Banihashem, “JRDP: A job recommender system based on ontology for disabled people,” *Int. J. Technol. Hum. Interact.*, vol. 15, no. 1, pp. 85–99, 2019.

[20] N. D. Almalis, G. A. Tsihrintzis, N. Karagiannis, and A. D. Strati, “FoDRA - A new content-based job recommendation algorithm for job seeking and recruiting,” *IISA 2015 - 6th Int. Conf. Information, Intell. Syst. Appl.*, 2016.

[21] A. Saini, F. Rusu, and A. Johnston, “PrivateJobMatch: A privacy-oriented deferred multi-match recommender system for stable employment,” *RecSys 2019 - 13th ACM Conf. Recomm. Syst.*, pp. 87–95, 2019.

[22] A. Giabelli, L. Malandri, F. Mercorio, M. Mezzanzanica, and A. Seveso, “Skills2Job: A recommender system that encodes job offer embeddings on graph databases,” *Appl. Soft Comput.*, vol. 101, 2021.

[23] H. Wenxing, C. Yiwei, Q. Jianwei, and H. Yin, “IHR+: A mobile reciprocal job recommender system,” *10th Int. Conf. Comput. Sci. Educ. ICCSE 2015*, no. Iccse, pp. 492–495, 2015.

[24] A. E. Shadare, S. M. Musa, and C. Akujuobi, “Data visualization,” no. December, 2016.

[25] J. Zhao *et al.*, *Embedding-based Recommender System for Job to Candidate Matching on Scale*, vol. 1, no. 1. Association for Computing Machinery, 2021.

[26] X. Guo, H. Jerbi, and M. P. O. Mahony, “An Analysis Framework for Content-based Job Recommendation,” 2014.

[27] N. Quadlin, “The Mark of a Woman’s Record: Gender and Academic Performance in Hiring,” *Am. Sociol. Rev.*, vol. 83, no. 2, pp. 331–360, 2018.

[28] L. A. Rivera, “Hiring as Cultural Matching: The Case of Elite Professional Service Firms,” *Am. Sociol. Rev.*, vol. 77, no. 6, pp. 999–1022, 2012.

[29] M. Humburg and R. van der Velden, “Skills and the graduate recruitment process: Evidence from two discrete choice experiments,” *Econ. Educ. Rev.*, vol. 49, pp. 24–41, 2015.

[30] P. Protsch and H. Solga, “How Employers Use Signals of Cognitive and Noncognitive Skills at Labour Market Entry: Insights from Field Experiments,” *Eur. Sociol. Rev.*, vol. 31, no. 5, pp. 521–532, 2014.
[31] K. Bunkar, U. K. Singh, B. Pandya, and R. Bunkar, “Data mining: Prediction for performance improvement of graduate students using classification,” *IFIP Int. Conf. Wirel. Opt. Commun. Networks, WOCN*, pp. 3–7, 2012.
[32] R. Asif, A. Merceron, S. A. Ali, and N. G. Haider, “Analyzing undergraduate students’ performance using educational data mining,” *Comput. Educ.*, vol. 113, pp. 177–194, 2017.
[33] Z. Othman, S. W. Shan, I. Yusoff, and C. P. Kee, “Classification techniques for predicting graduate employability,” *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 4–2, pp. 1712–1720, 2018.
[34] D. M. Costello-Nichitas, “from the SAGE Social Science Collections . All Rights,” *Hisp. J. Behav. Sci.*, vol. 9, no. 2, pp. 183–205, 1987.
[35] F. Ahmad, N. H. Ismail, and A. A. Aziz, “The Prediction of Students ’ Academic Performance Using Classification Data Mining Techniques,” vol. 9, no. 129, pp. 6415–6426, 2015.
[36] N. Buniyamin, U. Mat, and P. M. Arshad, “Educational Data Mining for Prediction and Classification of Engineering Students Achievement,” pp. 49–53, 2015.
[37] Harwati, A. P. Alfiani, and F. A. Wulandari, “Mapping Student’s Performance Based on Data Mining Approach (A Case Study),” *Agric. Agric. Sci. Procedia*, vol. 3, pp. 173–177, 2015.
[38] E. Osmanbegovic, “Data Mining Approach for Predicting Student Performance,” 2012.
[39] I. Singh, A. S. Sabitha, and A. Bansal, “Student performance analysis using clustering algorithm,” *Proc. 2016 6th Int. Conf. - Cloud Syst. Big Data Eng. Conflu. 2016*, pp. 294–299, 2016.
[40] N. Elassy, “Quality Assurance in Education Article information :,” 2015.
[41] T. L. Friedman, “How to Get a Job at Google,” *New York Times*, no. 1, p. 4, 2014.
[42] M. Abdulwahed, W. Balid, M. O. Hasna, and S. Pokharel, “Skills of engineers in knowledge based economies: A comprehensive literature review, and model development,” *Proc. 2013 IEEE Int. Conf. Teaching, Assess. Learn. Eng. TALE 2013*, pp. 759–765, 2013.
[43] B. L. Hartmann and C. T. Jahren, “Leadership: Industry needs for entry-level engineering positions,” *IEEE Eng. Manag. Rev.*, vol. 44, no. 3, pp. 76–85, 2016.
[44] A. Prof and D. Mishra, “Engineering Employability Skills Required By Employers in India,” pp. 961–964, 2016.
[45] D. Won, G. Bravo, and C. Lee, “Careers in collegiate athletic administration: Hiring criteria and skills needed for success,” *Manag. Leis.*, vol. 18, no. 1, pp. 71–91, 2013.
[46] M. Y. Yuzainee, A. Zaharim, and M. Z. Omar, “Employability skills for an entry-level engineer as seen by Malaysian employers,” *2011 IEEE Glob. Eng. Educ. Conf. EDUCON 2011*, pp. 80–85, 2011.
[47] K. Peltonen, “Education + Training Article information : learning perspective,” *Educ. + Train.*, vol. 57, no. 5, 2015.
[48] J. Moody and C. Bolt-lee, “Showcasing Expanding,” pp. 21–33.
[49] S. T. Fiske, A. J. C. Cuddy, P. Glick, and J. Xu, “A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition,” *J. Pers. Soc. Psychol.*, vol. 82, no. 6, pp. 878–902, 2002.
[50] D. Abrams, H. J. Swift, and L. Drury, “Old and Unemployable? How Age-Based Stereotypes Affect Willingness to Hire Job Candidates,” *J. Soc. Issues*, vol. 72, no. 1, pp. 105–121, 2016.
[51] C. Göbel and T. Zwick, “Age and Productivity: Sector Differences,” *Economist*, vol. 160, no. 1, pp. 35–57, 2012.
[52] D. Bahl, V. Desai, S. Vibhandik, and I. Fatma Vijay Kolekar, “A Profile Based Job Recommender System,” *Int. Res. J. Eng. Technol.*, vol. 11, pp. 2395–56, 2016.
[53] W. Hong, S. Zheng, and H. Wang, “Dynamic user profile-based job recommender system,” *Proc. 8th Int. Conf. Comput. Sci. Educ. ICCSE 2013*, no. Iccse, pp. 1499–1503, 2013.
[54] N. D. Almalis, G. A. Tsihrintzis, and E. Kyritsis, “A constraint-based job recommender system integrating FoDRA,” *Int. J. Comput. Intell. Stud.*, vol. 7, no. 2, p. 103, 2018.
[55] F. Gutiérrez, S. Charleer, R. De Croon, N. N. Htun, G. Goetschalckx, and K. Verbert, “Explaining and exploring job recommendations,” pp. 60–68, 2019.
[56] A. González-Briones, A. Rivas, P. Chamoso, R. Casado-Vara, and J. M. Corchado, “Case-Based Reasoning and Agent Based Job Offer Recommender System,” *Adv. Intell. Syst. Comput.*, vol. 771, pp. 21–33, 2019.
[57] B. Heap, A. Krzywicki, W. Wobcke, M. Bain, and P. Compton, “Combining Career Progression and Profile Matching in a Job Recommender System,” pp. 396–408, 2014.
[58] D. Mhamdi, R. Moulouki, M. Y. El Ghoumari, M. Azzouazi, and L. Moussaid, “Job recommendation based on job profile clustering and job seeker behavior,” *Procedia Comput. Sci.*, vol. 175, pp. 695–699, 2020.
[59] W. F. W. Yaacob, S. A. M. Nasir, W. F. W. Yaacob, and N. M. Sobri, “Supervised data mining approach for predicting student performance,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 16, no. 3, pp. 1584–1592, 2019.
[60] S. L. Ting, W. H. Ip, and A. H. C. Tsang, “Is Naïve bayes a good classifier for document classification?,” *Int. J. Softw. Eng. its Appl.*, vol. 5, no. 3, pp. 37–46, 2011.
[61] S. Taheri and M. Mammadov, “Learning the naive bayes classifier with optimization models,” *Int. J. Appl. Math. Comput. Sci.*, vol. 23, no. 4, pp. 787–795, 2013.
[62] V. Narayanan, I. Arora, and A. Bhatia, “Fast and Accurate Sentiment Classification Using an Enhanced Naive Bayes Model,” pp. 194–201, 2013.
[63] K. Yadav and R. Thareja, “Comparing the Performance of Naive Bayes And Decision Tree Classification Using R,” *Int. J. Intell. Syst. Appl.*, vol. 11, no. 12, pp. 11–19, 2019.
[64] I. A. A. Amra, “Naïve Bayesian,” pp. 909–913, 2017.
[65] C. Slamet, R. Andrian, D. S. Maylawati, Suhendar, W. Darmalaksana, and M. A. Ramdhani, “Web Scraping and Naïve Bayes Classification for Job Search Engine,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 288, no. 1, 2018.
[66] S. Choudhary, S. Koul, S. Mishra, A. Thakur, and R. Jain, “Collaborative job prediction based on Naïve Bayes Classifier using python platform,” *2016 Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solut. CSITSS 2016*, pp. 302–306, 2016.
[67] A. D. Hartanto, E. Utami, S. Adi, and H. S. Hudnanto, “Job seeker profile classification of twitter data using the naïve bayes classifier algorithm based on the DISC method,” *2019 4th Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2019*, pp. 533–536, 2019.
[68] H. Chaudhari, N. Yadav, and Y. Shukla, “A predictive analysis on job recruitment,” pp. 6 (5 .)-6 (5 .), 2018.

[69] E. P. F. Lee *et al.*, "An ab initio study of RbO, CsO and FrO (X2∑+; A2∏) and their cations (X3∑-; A3∏)," *Phys. Chem. Chem. Phys.*, vol. 3, no. 22, pp. 4863–4869, 2001.

[70] B. Chandra and M. Gupta, "Robust approach for estimating probabilities in Naïve-Bayes classifier for gene expression data," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 1293–1298, 2011.

[71] B. Stewart, "Improving performance of naive bayes classifier by including hidden variables," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1415, pp. 272–280, 1998.

[72] L. Jiang, H. Zhang, and Z. Cai, "A novel bayes model: Hidden naive bayes," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 10, pp. 1361–1371, 2009.

[73] A. Priyam, R. Gupta, A. Rathee, and S. Srivastava, "Comparative Analysis of Decision Tree Classification Algorithms," pp. 334–337, 2013.

[74] M. Batra and R. Agrawal, "Comparative analysis of decision tree algorithms," *Adv. Intell. Syst. Comput.*, vol. 652, pp. 31–36, 2018.

[75] M. Pandey and V. Kumar Sharma, "A Decision Tree Algorithm Pertaining to the Student Performance Analysis and Prediction," *Int. J. Comput. Appl.*, vol. 61, no. 13, pp. 1–5, 2013.

[76] B. R. Patel and K. K. Rana, "A Survey on Decision Tree Algorithm For Classification," *Ijedr*, vol. 2, no. 1, pp. 1–5, 2014.

[77] R. Hasan, S. Palaniappan, A. R. A. Raziff, S. Mahmood, and K. U. Sarker, "Student Academic Performance Prediction by using Decision Tree Algorithm," *2018 4th Int. Conf. Comput. Inf. Sci. Revolutionising Digit. Landsc. Sustain. Smart Soc. ICCOINS 2018 - Proc.*, pp. 1–5, 2018.

[78] W. K. Jati and L. Kemas Muslim, "Optimization of Decision Tree Algorithm in Text Classification of Job Applicants Using Particle Swarm Optimization," *2020 3rd Int. Conf. Inf. Commun. Technol. ICOIACT 2020*, pp. 201–205, 2020.

[79] A. Gupta and D. Garg, "Applying data mining techniques in job recommender system for considering candidate job preferences," *Proc. 2014 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2014*, pp. 1458–1465, 2014.

[80] H. Kaya, F. Gurpinar, and A. A. Salah, "Multi-modal Score Fusion and Decision Trees for Explainable Automatic Job Candidate Screening from Video CVs," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2017-July, pp. 1651–1659, 2017.

[81] F. Ahmad, S. Z. Yahya, Z. Saad, and A. R. Ahmad, "Tajweed Classification Using Artificial Neural Network," *2018 Int. Conf. Smart Appl. Commun. Networking, SmartNets 2018*, vol. 2, no. 11, pp. 8–14, 2018.

[82] E. T. L. L. Sun and Q. Yang, "Modelling , prediction and classification of student academic performance using artificial neural networks," *SN Appl. Sci.*, no. July, 2019.

[83] P. M. Arsad, N. Buniyamin, and J. L. A. Manan, "A neural network students' performance prediction model (NNSPPM)," *2013 IEEE Int. Conf. Smart Instrumentation, Meas. Appl. ICSIMA 2013*, no. July 2006, pp. 26–27, 2013.

[84] S. A. Naser, I. Zaqout, M. A. Ghosh, R. Atallah, and E. Alajrami, "Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology," *Int. J. Hybrid Inf. Technol.*, vol. 8, no. 2, pp. 221–228, 2015.

[85] Y. Shao and R. S. Lunetta, "Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points," *ISPRS J. Photogramm. Remote Sens.*, vol. 70, pp. 78–87, 2012.

[86] Y. Cai, R. Lin, and Y. Kang, "A Personalized Company Recommender System for Job Seekers," pp. 1–6, 2015.

[87] M. Mishra and M. Srivastava, "e e /, e," pp. 29–31, 2014.

[88] V. Sharma, S. Rai, and A. Dev, "A Comprehensive Study of Artificial Neural Networks," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 2, no. 10, pp. 278–284, 2012.

[89] Okfalisa, I. Gazalba, Mustakim, and N. G. I. Reza, "Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification," *Proc. - 2017 2nd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2017*, vol. 2018-Janua, pp. 294–298, 2018.

[90] J. Derrac, F. Chiclana, S. García, and F. Herrera, "Evolutionary fuzzy k-nearest neighbors algorithm using interval-valued fuzzy sets," *Inf. Sci. (Ny).*, vol. 329, pp. 144–163, 2016.

[91] Z. Zhang, "Introduction to machine learning: K-nearest neighbors," *Ann. Transl. Med.*, vol. 4, no. 11, pp. 1–7, 2016.

[92] P. K. Roy, S. S. Chowdhary, and R. Bhatia, "A Machine Learning approach for automation of Resume Recommendation system," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 2318–2327, 2020.

[93] M. G. H. Omran, A. P. Engelbrecht, and A. Salman, "An overview of clustering methods," *Intell. Data Anal.*, vol. 11, no. 6, pp. 583–605, 2007.

[94] R. Chouhan and A. Purohit, "An approach for document clustering using PSO and K-means algorithm," *Proc. 2nd Int. Conf. Inven. Syst. Control. ICISC 2018*, no. Icisc, pp. 1380–1384, 2018.

[95] A. J. M. Rani and L. Parthipan, "Clustering analysis by improved particle swarm optimization and K-means algorithm," *IET Conf. Publ.*, vol. 2012, no. 624 CP, pp. 83–88, 2012.

[96] A. Ahmadyfard and H. Modares, "Combining PSO and k-means to enhance data clustering," *2008 Int. Symp. Telecommun. IST 2008*, pp. 688–691, 2008.

[97] N. A. M. Isa, S. A. Salamah, and U. K. Ngah, "Adaptive fuzzy moving K-means clustering algorithm for image segmentation," *IEEE Trans. Consum. Electron.*, vol. 55, no. 4, pp. 2145–2153, 2009.

[98] H. Xiong, J. Wu, and J. Chen, "K-means clustering versus validation measures: A data-distribution perspective," *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 39, no. 2, pp. 318–331, 2009.

[99] K. P. Sinaga and M. S. Yang, "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80716–80727, 2020.

[100] I. Khan, Z. Luo, J. Z. Huang, and W. Shahzad, "Variable weighting in fuzzy k-Means clustering to determine the number of clusters," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 9, pp. 1838–1853, 2020.

[101] T. S. Xu, H. D. Chiang, G. Y. Liu, and C. W. Tan, "Hierarchical K-means Method for Clustering Large-Scale Advanced Metering Infrastructure Data," *IEEE Trans. Power Deliv.*, vol. 32, no. 2, pp. 609–616, 2017.

[102] S. Mjahed, S. El Hadaj, K. Bouzaachane, and S. Raghay, "Improved PSO based K-Means clustering applied to fault signals diagnosis," *2018 Int. Conf. Control. Autom. Diagnosis, ICCAD 2018*, 2018.

[103] Z. Fan, "Clustering of College Students Based on Improved K-means Algorithm," 2016.

[104] D. F. Kamarudin, "Pandemik tingkat pengangguran graduan," Malaysia, 2021.