Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

A PROPOSED MODEL FOR THE SELECTION OF WORKERS ON CROWDSOURCING PLATFORMS UTILIZING NESTED CRITERIA MATCHING

AHMAD CUCUS *
Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan, Malaysia.
*Corresponding Author E-mail: ahmad.cucus@ubl.ac.id

AL-FAHIM MUBARAK ALI
Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan, Malaysia.

ZAFRIL RIZAL M. AZMI
Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan, Malaysia.

Abstract

The objective of this study is to examine the concept of crowdsourcing and the corresponding method for employee selection. In recent years, there has been a growing trend towards the utilization of crowdsourcing, wherein both businesses and individuals harness the power of collective cooperation to provide solutions, ideas, or contributions across several domains, including but not limited to product development and scientific research. This study examines the concept of crowdsourcing as a means of gathering occupations or tasks completed by individuals with diverse qualifications. In this study, we examine the worker selection mechanism as explored in prior research and put forth a novel worker selection model incorporating a profile matching algorithm. The subsequent phase involves enhancing the profile matching algorithm to accommodate nested criteria for matching worker requirements. Upon concluding the investigation, a comparative analysis was conducted to assess the outcomes of matching workers' criteria using both single criterion and nested criteria. Additionally, the proposed formula was implemented to evaluate the case involving nested criteria. The findings reveal substantial disparities between the suggested workforces, particularly in terms of their composition. The proposed workforce with nested criteria exhibits a reduced and more precise numerical representation compared to the proposed workforce with a single criterion. This demonstrates that the utilization of the proposed selection model offers a viable solution to the challenge of identifying workers with layered criteria.

Keywords: Crowdsourcing, Nested Citeria, Profile Matching, Match Algorithm, big data, data collections, worker selections, worker validations.

1 INTRODUCTION

The concept of crowdsourcing has gained significant popularity in the realm of online platforms in recent years. The practice of obtaining a desired service or information by soliciting contributions from a large number of persons has garnered significant public interest, sometimes referred to as crowdsourcing. [1]. The concept pertains to leveraging the collective experience and knowledge of a networked community to successfully accomplish a task. Crowdsourcing has demonstrated its applicability across diverse businesses, encompassing domains such as graphic design, research, and social advocacy.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

Presently, crowdsourcing has emerged as a very efficient approach for expeditiously and effectively accomplishing substantial initiatives. Within the realm of business, the utilization of crowdsourcing has the potential to elicit innovative solutions from a broader community, include clients in the process of developing new products, and even facilitate the recruitment of temporary labour for specialized jobs [2].

The utilization of crowdsourcing has brought about a significant transformation in the way tasks are accomplished. However, it is important to note that this approach often proves to be unsuccessful, resulting in the expenditure of valuable resources such as time and money for practitioners [3]. The utilization of crowdsourcing and human computation facilitates the engagement of a substantial cohort of collaborators in tasks that are conventionally performed by extensively skilled annotators. By consolidating the collective responses of these contributors, a more comprehensive dataset is generated, which accommodates for the presence of ambiguity.

The process of worker selection in crowdsourcing can provide challenges due to the inherent variances in worker skills. In the context of crowdsourcing, the allocation of tasks is undertaken with the objective of breaking them down into smaller units, so enabling numerous individuals within an online community to contribute towards their completion.[4] Selecting the most competent individuals for a certain activity might be a considerable challenge when faced with the decision of assigning people to do these duties. The potential cause of this issue could be attributed to the presence of skill gaps among the employees, which could result in a decline in the overall quality of the project's ultimate deliverable. [5]. In order to ensure the efficient completion of duties, it is imperative to select the appropriate personnel. The process of staff selection holds significant importance. Workers are often selected for available job positions based on their dependability, cost-effectiveness, physical proximity, and other relevant factors.[6]

Selection models can serve as a method for identifying appropriate personnel through the utilization of crowdsourced matching algorithms. The matching algorithm identifies workers who exhibit the highest degree of compatibility with the jobs or projects assigned by the assignor. The method of crowd selection is implemented with the purpose of identifying appropriate crowd workers for a certain task. Without the appropriate selection of crowd workers, the process of crowdsourcing becomes futile.[7].

In the context of assigning tasks, individuals responsible for delegating assignments have the ability to employ a process known as selection. This process involves carefully choosing workers who possess specific attributes that align with the requirements of the tasks at hand. These attributes may include factors such as skill level, prior experience, and other relevant credentials. Subsequently, the assignor has the capability to employ a matching algorithm in order to ascertain the worker who possesses the most suitable qualifications for the given assignment. Through the utilization of a selection process that employs a matching algorithm, the assignor can effectively ascertain the most suitable worker for a certain assignment. This approach serves to enhance the overall quality of the project's outcome by ensuring that the chosen worker possesses the requisite skills and qualifications.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

Normally, the process of worker selection in crowdsourcing employs either single-row criteria or multiple criteria. For instance, a common approach involves seeking workers who possess specific qualifications, such as being master graduates, having over 5 years of relevant experience, and having completed a minimum of 20 hours of training. Although opting for a worker based solely on one criterion may seem more straightforward, it is important to consider the potential drawbacks and limitations associated with this approach. There are several challenges that may arise when relying solely on a single criterion.

Limited Assessment: Focusing just on a single element may lead to an inadequate appraisal of an employee's overall suitability for the position. By exclusively prioritizing a single element, the candidate exposes themselves to the possibility of disregarding other vital abilities, competencies, or traits that are needed for achieving success in the given position.

A restricted perspective: Depending solely on a single criterion restricts the range of possibilities and fails to consider the full range of skills and qualities that a worker can contribute. The aforementioned statement fails to acknowledge the intricate nature of individuals and the vast array of valuable contributions they are capable of making. The absence of equilibrium can be observed in situations when a diverse range of abilities and attributes are required for different positions and tasks. Exclusively depending on a single criterion may result in an imbalanced workforce comprising individuals who possess expertise in certain domains but lack proficiency in others, hence compromising the attainment of a comprehensive and high-performing workforce.

Unfairness and prejudice may arise in the selection process when it is exclusively reliant on a single criterion, leading to unintended discriminatory outcomes. There is a potential for inadvertent preferential treatment or discrimination against specific individuals or collectives, unrelated to their work performance. Requesters may encounter missed opportunities when they fail to consider a broader variety of characteristics, resulting in the loss of applicants that possess exceptional talents, experiences, or potential. These people have the ability to contribute to the organization in unforeseen ways.

In order to tackle these issues, it is recommended to take into account a blend of criteria that encompass the essential qualifications, abilities, and traits necessary for the particular position. The utilization of a holistic approach, which integrates several criteria or nested criteria, facilitates a more exhaustive and equitable assessment process. Consequently, this strategy enhances the quality of recruiting decisions and augments the likelihood of identifying the most appropriate candidate for the position.

When making a selection of a worker, the requester has the option to employ nested criteria in order to assess their appropriateness for a specific function or assignment. Nested criteria refer to the practice of dividing the evaluation process into multiple hierarchical levels or stages, with each level or stage specifically targeting distinct facets of the worker's credentials. Through the utilization of layered criteria, the requester is able to systematically examine workers across several dimensions, hence facilitating a

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

comprehensive evaluation process and enhancing the likelihood of identifying the best appropriate candidate for the given position.

2 RELATED RESEARCH

2.1 Worker Quality On Crowdsourcing

The concept of crowdsourcing was first proposed by Jeff Howe in a seminal piece published in Wired magazine in June 2006. According to Howe, crowdsourcing is a method employed to effectively manage and engage with transient workforce [8]. The present discourse centers on research endeavors pertaining to the processing of large-scale datasets, sometimes referred to as big data. Data management is considered to be a fundamental component of the fourth industrial revolution. The processes of data consolidation and integration are crucial in facilitating the efficient functioning of companies and aiding in informed decision-making. The term "Big Data" encompasses datasets that provide significant challenges for standard processing methods and databases due to their immense size, diverse nature, and rapid velocity [9].

The objective of this study was to collect data through the utilization of the crowdsourcing methodology. The utilization of human cognitive abilities, commonly referred to as crowdsourcing or human computation, is employed as a paradigm to address challenges that are currently beyond the capabilities of computers. The field of computer-human interaction is experiencing significant growth. This study diverges from existing Artificial Intelligence (AI) algorithms by prioritizing the utilization of human cognition to tackle computational obstacles.[10]

The inclusion of quality control in crowdsourced data management is crucial due to the potential for the generation of inaccurate and irrelevant data. The intentions of people engaging in data collection operations remain uncertain to data collectors. The presence of undisclosed intentions on the part of the worker may impede the effectiveness of the crowdsourcing process, as indicated by the provided statistics. Data collectors exhibit a diverse spectrum of knowledge and skill levels. Workers who possess an adequate level of competence are incapable of performing some tasks. The data presented in the accompanying chart indicates that the process of quality control necessitates greater exertion in collecting data that contains mistakes and noise, hence resulting in data of superior quality.[11]

2.2 Worker Selection On Crowdsourcing

Profile matching is a technique that can be employed to obtain worker recommendations that align with the specific task needs. The job-matching process commonly involves the analysis of a candidate's résumé, which is then compared to the available possibilities listed. [12]. The objective of profile matching is to evaluate the degree of compatibility between two profiles. The approach employed in this scholarly paper is founded on the principles of matching theory, wherein filters are utilized to represent profiles within lattices, and matching values are employed to assess the degree of compatibility between the profiles. A greater matching value corresponds to a higher degree of compatibility.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

[13].

The process of profile matching involves evaluating a candidate's similarity to a known example, prioritizing this resemblance above commonly utilized identity measures. During the process of profile matching, an individual's capacity to fulfill job skills may be assessed. In this context, it often pertains to the capacity for discernment; a smaller "gap," also known as the "gap," signifies a greater number of substantial opportunities for prospective candidates. Profile matching is a decision-making approach that acknowledges the necessity for employees to fulfill a range of anticipated variables.[14]. In order to determine the disparity between the employee profile and the standard profile, the subsequent formula is utilized:

$$Gap\ (WVal, SVal) = SVal - WVal \qquad (1)$$

The term "SVal" is used to represent the standard value of task criteria, whereas "WVal" is used to denote the value of worker criteria. The framework of Profile Matching will be utilized to calculate the similarity scores between the two profiles. The similarity scores of both profiles are calculated based on the shared attributes extracted from the profiles. The similarity scores obtained are further adjusted to enhance the accuracy of the results by considering the relative importance of each feature. The new similarity value will frequently increase or decrease depending on the relative importance of each feature.[14].

3 DATA AND METHOD

3.1 Data Set

The researchers utilized Kaggle datasets as the basis for conducting model testing. Kaggle provides a cloud-based data science platform in addition to its extensive datasets. [15]. A number of fields are included in the dataset, including enrollee id, city, city development index, gender, relevant experience, enrolled university, education level, major discipline, experience, company size, company type, last new job, training hours. but for this research we only use 5 fields such as registered id, gender, education level, experience, training hours, dataset contains of 2230 rows, the table will be a test table for selecting worker criteria with nested criteria, The graph in figure 1 below is a dataset profile that will become test data

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

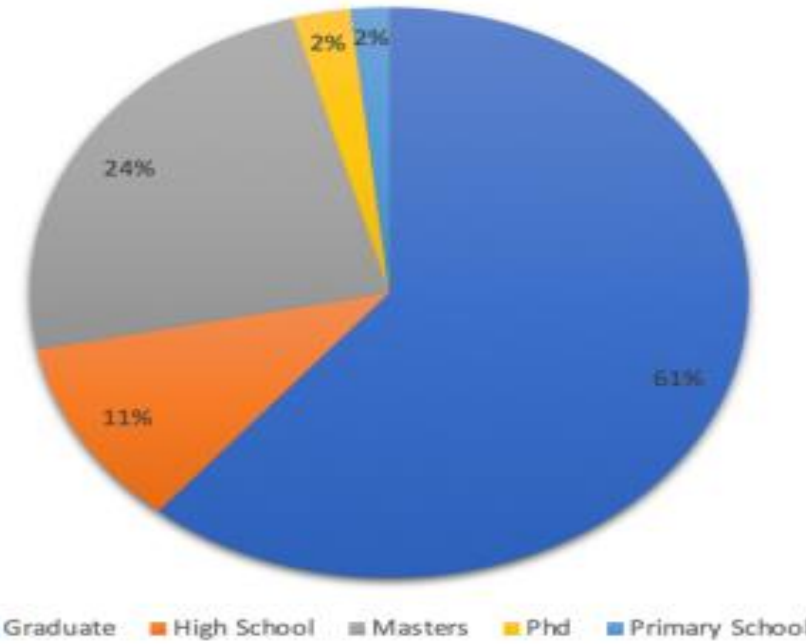

Fig 1: Data size based on education level

Figure number 1 above shows the size of the data in the dataset consisting of education level, and figure number 2 below shows chart of experience data sizes, the highest number at the education level is Graduate at 61%, and Experience for 21 years with a total of 11%. for large data based on training hours of more than 300 types (labels)

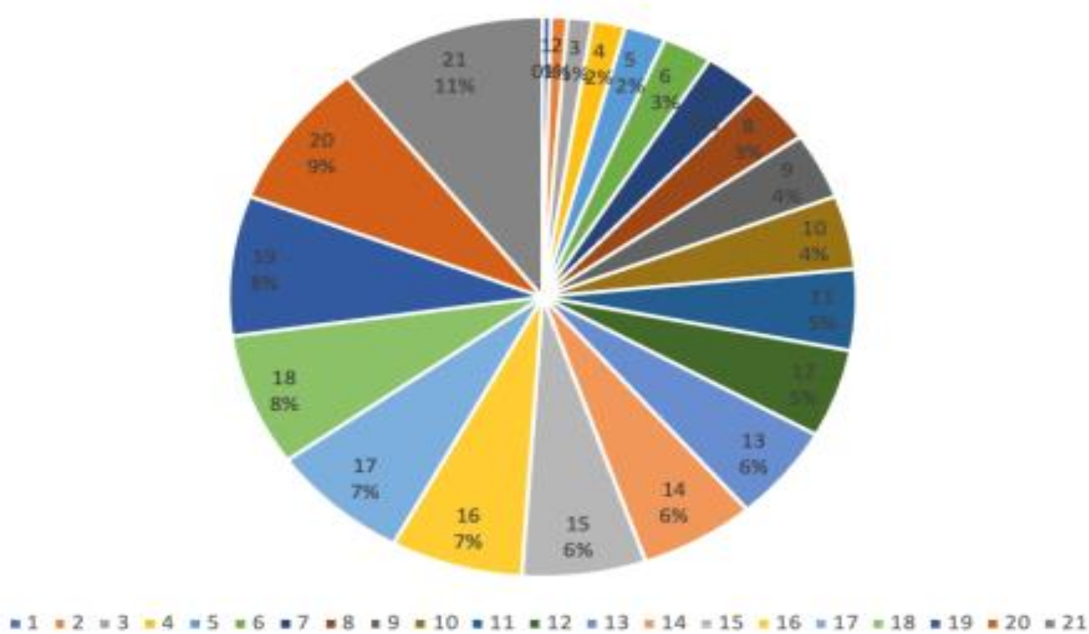

Fig 2: Data size based on Experience

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

3.2 Proposed Model

The paradigm of worker selections on the proposed crowdsourcing platform is depicted in Figure 4, as presented in this research. The initial stage of the crowdsourcing activity involves the implementation of a selection process for the workers involved. The framework facilitates the identification and selection of individuals who possess the requisite skills and abilities to effectively carry out the assigned activities. When individuals engage in crowdsourcing endeavors, a computational algorithm is employed to selectively screen and categorize workers based on their suitability for specific tasks or activities.

The stages in this proposed framework are Worker Profile Data Collections [16], Task Requirement [17], Worker Selections [12], Proposed Selected Worker[18], Task Asignment [19]. The worker selection process in this model encompasses two types of criteria: single row criteria and nested criteria. Single row criteria consist of a single row containing multiple criteria, where worker selection is solely based on one independent variable followed by several dependent variables. On the other hand, nested criteria involve multiple dependent variables followed by dependent variables.

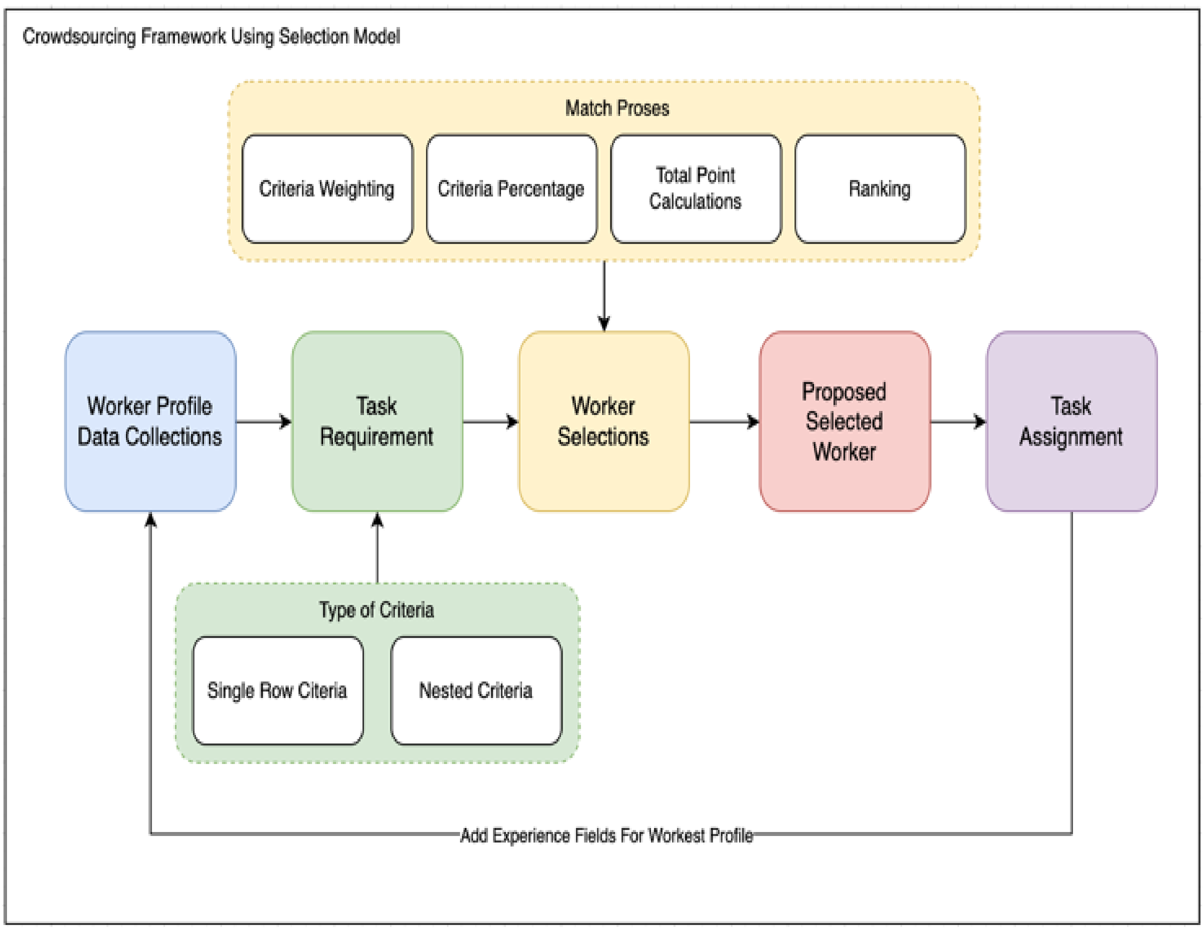

Fig 3: Crowdsourcing Framework Using Selecitons Model

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

Figure 3 illustrates the worker selection process, which involves the determination of criterion weights. In the provided case example, the key variable is identified as education level, followed by the dependent variables of experience and training hours. The subsequent sections outline the stages inside the selection model.

First is Criteria Weighting: In order to ascertain the relative importance of the criteria, the researcher puts forth the subsequent formula.

$$WGap = \sum_{i=0}^{n} \frac{Wval_i - Sval_i}{Max(Wval_i - Sval_i, Wval_n - Sval_n)} x100 \qquad (2)$$

Where,

W = Total Worker Weight Value

Wval = Value of Worker Criteria

Sval = Value of Standard Criteria

Max() = The maximum gap value of the criteria that all workers have in the dataset

The allocation of points will be based on the identification of key criteria and secondary criteria. The main criterion will account for 60% of the total points, while the secondary criteria will account for the remaining 40%.

Second step is Total Point Calculcations, calculate the total points by adding up the primary criteria and secondary criteria, by using the formula (3) below

$$TW = (PWval\ x\ 60\%) + (SWval\ x\ 40\%) \ (3)$$

Where,

TW = Total Point of Worker

PWval = Total Worker's Primary Criteria Points

SWval = Total Worker's Secondary Criteria Points

Third step is Ranking, apply weights to several criteria and compute a weighted score for each value. Calculate the cumulative scores by assigning appropriate weights to each value, and thereafter arrange the values in a ranked order based on their cumulative scores. the next step is proposes selected workers, we divide the selected workers into four parts, high recommended, medium recommended, low recommended, unrecommended.

3.3 Algorithm

Our approach is to construct a selections model by using profile match algorithms. We put some changes to the profile matching algorithm to be able to solve matching cases with nested criteria. Profile Matching Algorithm Using Nested Criteria

1. Determine the dataset to be used.
2. determine the key criteria and dependent criteria.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

3. dependent criteria are variables that are associated with the key criteria
4. Specify the data to be matched, and store it in an array or table.
5. Main variable match between dataset and data profile.
6. calculate gap companion criteria according to primaries matching criteria between dataset and data profile using formula
7. calculate the value of the total gap
8. Rank the result

In the previous algorithm, various values are defined, including Worker List which represents the variable value obtained from the dataset of workers, and Worker Looking which denotes the standard criteria value for identifying workers. Subsequently, ascertain the weight value associated with the connection between the Worker List and Worker Standard Criteria.

The subsequent procedure involves determining the weight and worker point values using the iteration function. Once the worker point values have been obtained, they will be ranked based on their position, with a smaller value compared to the preceding Euclidean value computation process.

4 IMPLEMENTATIONS

to test the selection model with nested criteria, we perform a worker search input with its type, single row criteria and nested criteria,

4.1 Single Row

To solve single row questions, we use a profile matching algorithm with the following steps:

First, Define the attributes: Determine the attributes that are relevant to workers profiles.

1. Education Level
2. Experience
3. Training Level

Second, Gather profile data: Collect the profile data for each user or entity in your system. Get Data Kaggle dataset

Table 1: Dataset of Worker from Kaggle

No	EId	gender	EL	Exp	TH
1	32403	Male	Graduate	9	21
2	9858	Female	Graduate	5	98
3	31806	Male	High School	1	15
4	27385	Male	Masters	11	39
5	27724	Male	Graduate	25	72

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

Third, gap calculations typically refer to the process of identifying and measuring the difference or discrepancy between two values or sets of values. The term "gap" implies a shortfall or distance that needs to be bridged or addressed [20]

Table 2: Require Worker Criteria

No	EL	Exp	TH
1	Graduate	5	100

The data shown in Table 2 illustrates the specific factors that the requester will consider during the evaluation process. Next, we will examine the discrepancy between the requirements possessed by the workers listed in table 1 and the necessary criteria data outlined in table 2. The process of determining the gap value is conducted through the utilization of Formula 1. The values presented in Table 3 are derived from the computation of the gap value, which signifies the disparity between the data presented in Table 1 and Table 2.

Table 3: Require Worker Criteria

E Id	gender	EL Gap	Exp Gap	TH Gap
32403	Male	Graduate	-4	79
9858	Female	Graduate	0	2
31806	Male	High School	4	85
27385	Male	Masters	-6	61
27724	Male	Graduate	-20	28

Forth is an assign weights, assign weights to each attribute based on their relative importance in determining the matching score. In table number 3 above,

Fifth is education Level Weighting, Researchers provide significance to every degree of education. A weight of 100 is assigned to the Group Education level. The weight assigned to those holding a Master's degree is 80, whereas individuals engaged in research at the SMA level are assigned a weight of 50. The Graduate group is assigned a weight of 100 due to their precise alignment with the job prerequisites.

Table 4: Ducation Level Weighting

No	Education Level	Weight
1	Graduate	100
2	Masters	80
3	High School	50

Sixth is Experience Weighting, In Table number 5 below, the researcher tries to divide the weights into five parts, the weights are 10, 30, 50, 80, and 100. experience (years) of all workers in the dataset, then the researcher calculates the intervals for each class to get the appropriate weight desired, the values obtained are as in Table number 5

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

Table 5: Experience Weighting

No	Minimum Value	Maximum Value	Weight
1	-20	-15	10
2	-15	-10	30
3	-10	-5	50
4	-5	0	80
5	0	5	100

Seventh is Training Hours Weighting, In the fifth table presented, the researcher endeavors to partition the weight into five distinct categories, specifically denoted as 10, 30, 50, 80, and 100. The researcher partitions the weight by considering the minimum difference between the training hours requested by the worker and the training hours recorded in the dataset. Subsequently, the researcher computes the interval for each class in order to determine the desired weight. The resulting values are presented in Table 6.

Table 6: Training Hours Weighting

No	Minimum Value	Maximum Value	Weight
1	-239	-171	10
2	-171	-103	30
3	-103	-35	50
4	-35	33	80
5	33	101	100

Eight is The similarity scores are categorized into four groups: high recommended, medium recommended, low recommended, and unrecommended. The range of similarity scores is from 0 to 100. The researcher then assigns weights to each group based on class intervals, as presented in Table 7.

Table 7: Worker Type Based on Total Point (Total Scores)

Minimum Value	Maximum Value	Worker Type
0	25	Un Recommended
26	50	Low Recommended
51	75	Medium Recommended
76	100	High Recommended

Ninth is Combine attribute scores: Calculate a weighted average or sum of the similarity scores for all attributes to obtain an overall matching score for the profiles. Multiply each similarity score by its corresponding attribute weight and aggregate them accordingly. And assigning the type of worker according to the total points earned by the worker, table number 8 displays the value of the primary factor (experience) of 60 percent of the weighted value, and the secondary factor (Training Hours) of 40 percent, we also display the number of points earned by the worker and the type employee recommendation to the applicant.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

Table 8: Final Result For Single Row Criteria

No	EId	PF	SF	Total Point	Type of Worker Recommendation
1	32403	60	36	96	High
2	9858	60	36	96	High
3	31806	30	40	70	Medium
4	27385	48	30	78	High
5	27724	60	18	78	High

Tenth is Display the matches: Exhibit the profiles that have been successfully matched as the output of the algorithm. The top matches are presented to the user or requester for additional analysis. Our recommendations or tailored suggestions, categorized as High Recommendation and Medium Recommendation, are provided to facilitate the completion of the work.

4.2 Nested Criteria

Tests were conducted using the dataset provided in table 1, employing layered criterion matching, Sample case

“We are currently seeking individuals to collaborate on research assignments that meet the following criteria: I possess a Master's degree in education, accompanied by five years of professional experience and a training term of ten hours. The individual possesses a Bachelor's degree together with a decade of professional experience, supplemented by a training period of at least 20 hours. The individual possesses a wealth of experience spanning more than two decades, along with a comprehensive training period totaling 100 hours “. In accordance with the aforementioned criteria, a table or array is generated, containing the fields as presented in Table 9.

Table 9: Require Worker Criteria

No	Education Level	Experience	Training Hours
1	Graduate	5	100
2	High School	10	150
3	Masters	2	80

To carry out the test, we perform calculations to determine the weight of workers in the dataset with the following criteria:

1. Enrollee_id : 67762
2. Education Level : Graduate
3. Experience : 10
4. Training Hours : 120

The next step is to calculate the GAP and determine points based on the criteria in Table 9. Max value of Experience where education level is graduate follow the formula

$$Maxexp = Max(Wexpval_i - Sexpval_i, Wexpval_n - Sexpval_n \quad (4)$$

Where:

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

Maxexp = Maximum Value of Experience where education level "Graduate"

Wexpval = Value of Workers Experience (years)

Sexpval = Value of Standard Experience from requester (years)

From the iteration process and calculations performed, the Max Exp value is 20. Max value of Training hours education level is graduate:

$$Maxth = Max(Wthval_i - Sthval_i, Wthval_n - Sthval_n) \quad (5)$$

Where:

Maxth = Maximum Value of Training hours where education level "Graduate"

Wthval = Value of Workers Training hours (years)

Sthval = Value of Standard Training hours from requester (years)

From the iteration process and calculations performed, the Maxexp value is 232. To find out the Gap Experience from the problems above, we try to apply formula number (2) as follows

$$WGapExp = \frac{10 - 5}{20} x\, 100$$

$$WGapExp = 25$$

To find out the Gap Jam Training from the problems above, we try to implement formula number (2) as follows

$$WGapTh = \frac{120 - 100}{232} x\, 100$$

$$WGapTh = 9$$

To get total points, the main criterion is a lot of experience 60%, dan Training Hours 40%

Worker Point = (25*60%) + (9*40%)

Worker Point = 16

The last stage involves ascertaining the interpretation of the data set by dividing it with reference to table 6. The subsequent procedure involves the computation of the GAP and the subsequent determination of the points assigned to each worker within the dataset. These values are presented in Table 9, wherein each worker is assigned a Gap Experience Weight and a Training Hours Gap Weight. Additionally, the table displays the total points data and the worker recommendations that we offer to the applicant. Specifically, we prioritize workers who have been classified as having high or medium recommendation types.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

5 RESULT AND DISCUSSION

This section will examine the output results produced by the two methods that were applied to the dataset. The comparison will focus on the utilization of the Profile matching algorithm on datasets with single row criteria matching types, as well as the application of enhanced profile matching algorithms on datasets with nested criteria matching types. The data is presented in the form of graphs and diagrams, facilitating the visualization of the ultimate outcomes of the two ways.

5.1 Interpretation of Data Derived by Profile Matching Algorithm

Figure 4 presents the output findings categorized by suggestion type. The results indicate that workers with high recommendation types exhibit the highest frequency, while workers with medium recommendation types are observed to have a comparatively lower frequency. Moreover, the data group depicted in Figure 4 below represents the outcome of aligning the worker requirements mandated by the requester, specifically for categories of employees characterized by low recommended and unrecommended kinds, and possessing a limited workforce or even without any workers with unrecommended types.

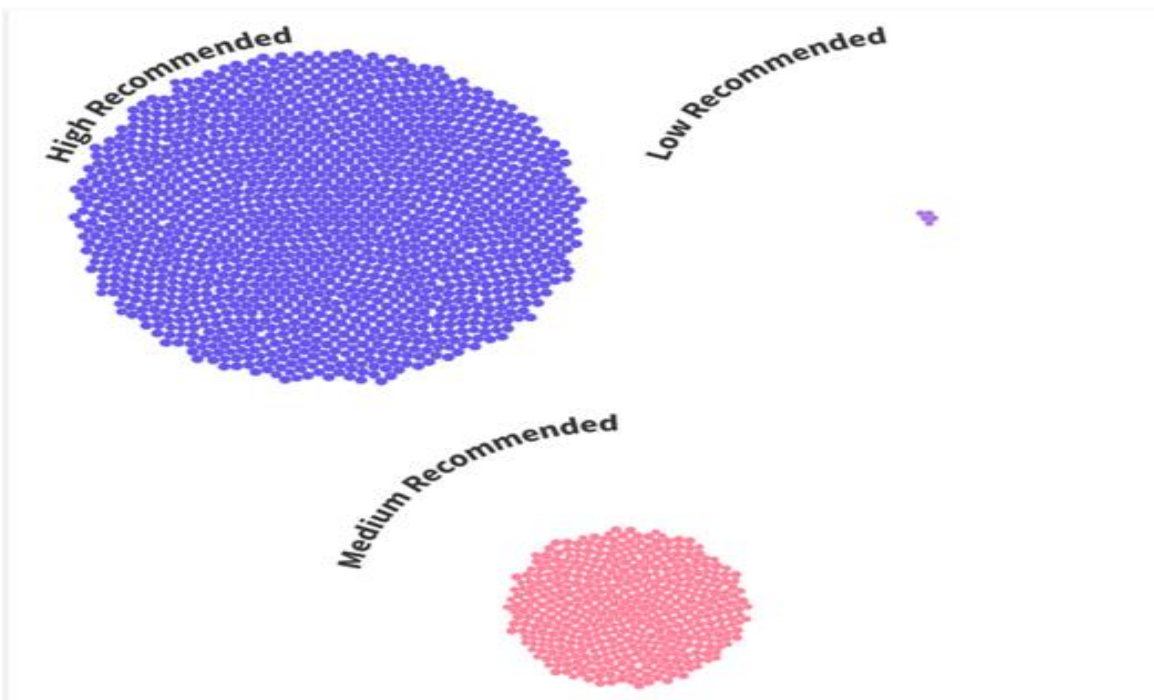

Fig 4: Worker group based on recommendation type using profile matching

The data depicted in Figure 5 illustrate the outcomes of the individuals' exposure to the suggested categories of workers. The majority, comprising 76%, consisted of workers who met the high recommendation criterion. This was followed by workers who were recommended at a medium level, accounting for 24% of the sample. Lastly, there were workers who fell into the low recommendation category. The quantity is approximately 0.1%.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

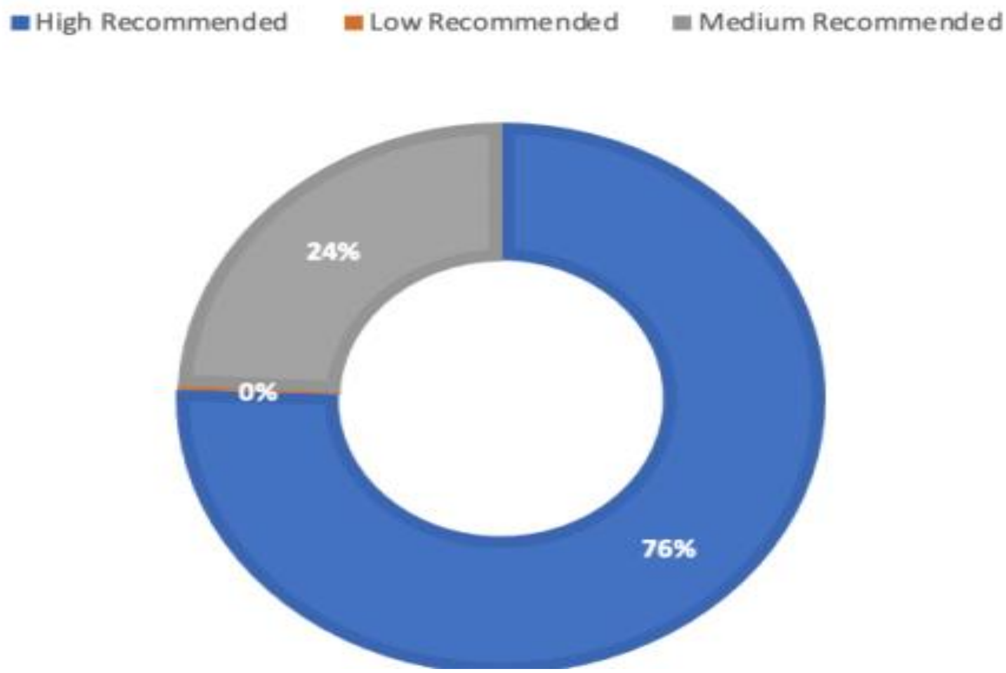

Fig 5: Percentage Of Worker Based On Recommendation Type.

Based on the findings depicted in figures 3 and 4, it is evident that employing profile matching based on single type criteria yields a substantial number of worker recommendations. This abundance of recommendations has the potential to adversely impact the quality of work performed by the workers. As the number of recommended workers increases, the likelihood of individuals lacking a suitable match for the job also escalates. This observation underscores the fact that the utilization of the profile matching algorithm does not yield highly specific outcomes in terms of identifying workers who align closely with the specified criteria.

5.2 Data interpretation from the enhanced profile matching for nested criteria algorithm.

Figure 6 displays the output results categorized by recommendation type. The results indicate that workers with a high recommendation type exhibit a smaller population compared to other types. The largest group consists of workers with an unrecommended type. Workers with a low recommended type have a slightly larger population than those with a medium recommended type. In this model, we propose that workers with a high recommended type and medium recommended type be assigned to carry out the requested task.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

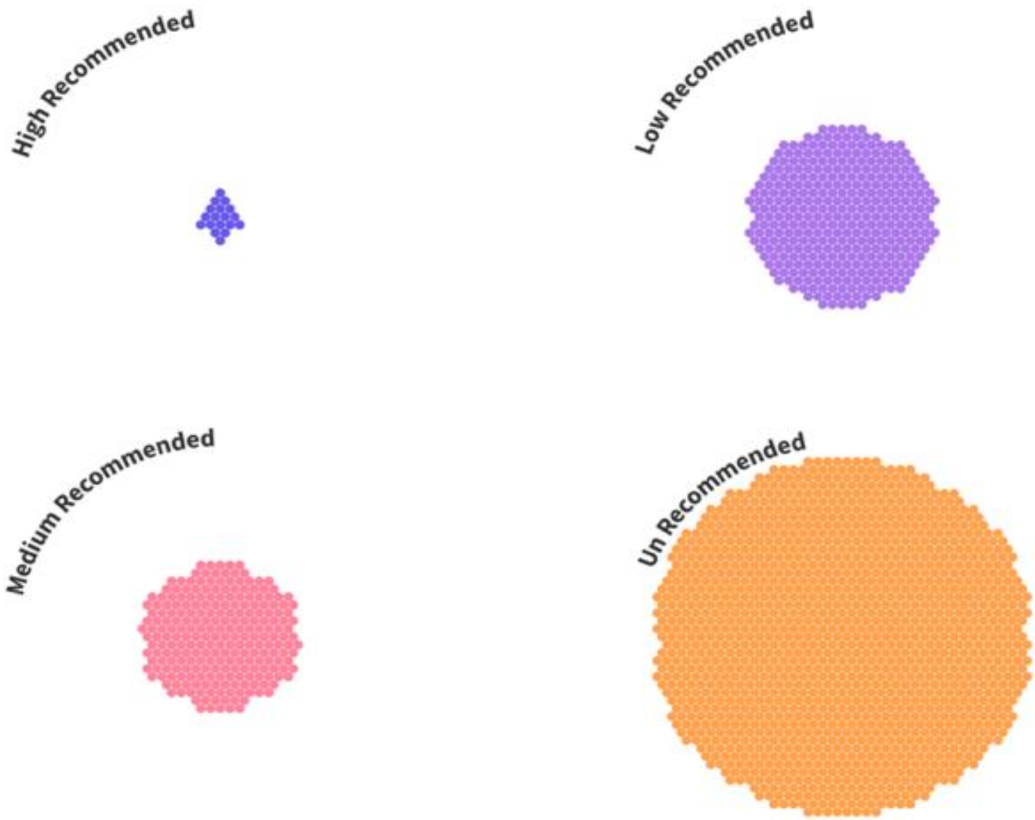

Fig 6: Worker group based on recommendation type using enhanced profile matching for nested criteria

Figure 6 illustrates the distribution of data for different categories of workers. Highly recommended workers constitute 1% of the entire worker dataset, moderately recommended workers make up 12%, workers with a low recommendation comprise 17%, and the remaining 70% are classified as unrecommended. It is worth noting that workers in the unrecommended category are assigned weight values ranging from zero to 25 points.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

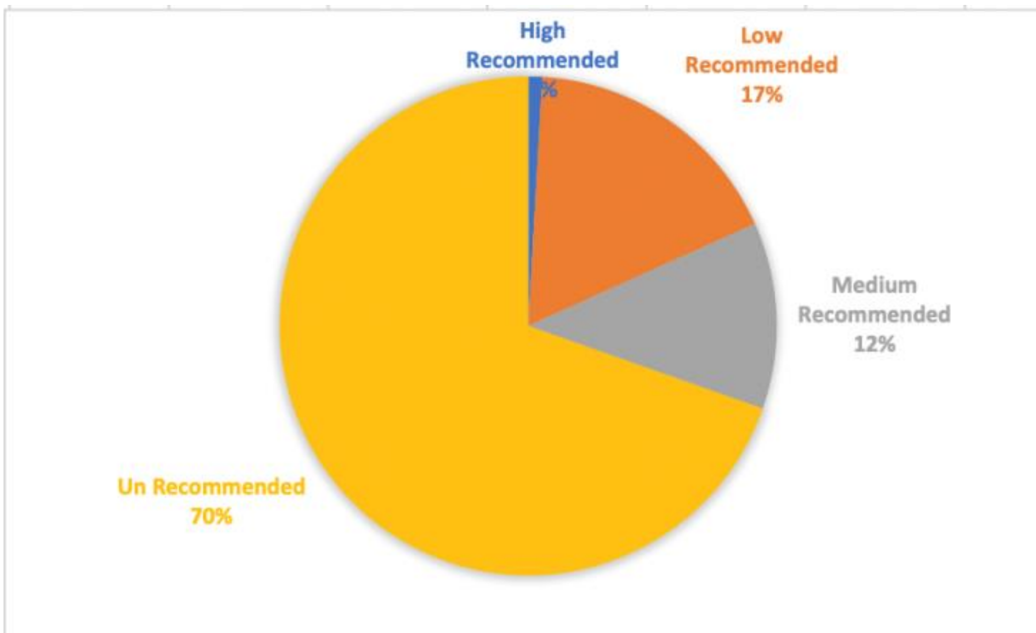

Fig 7: Percentage of worker based on recommendation type using nested criteria

Figures 6 and 7 depict the outcomes of worker suggestions employing nested criteria, which are accompanied by an improved profile matching algorithm. The two presented graphs demonstrate that the implementation of the suggested nested criteria and algorithm results in a reduced number of recommended personnel. The worker table consists of two categories: high recommended workers, comprising 1% of the total, and medium recommended workers, including 12% of the total. Combined, these two categories account for 13% of the complete worker table. It is observed that the likelihood of matching job criteria is greater inside this subset.

6 CONCLUSION

Based on the conducted tests, there are two notable distinctions between the outcomes obtained from matching data based on single criteria and matching data based on nested criteria. Particularly, for workers with high recommendations, there are significant variations. When employing a single criterion, the number of workers categorized as having a high recommendation demonstrates a substantial increase of 70%. Conversely, when utilizing nested criteria, this figure is merely 1%. This discrepancy can be attributed to disparities in the calculation mechanism of the GAP (Generalized Average Precision) and the profile matching algorithm. The profile matching algorithm assigns weights based on user input, resulting in a potentially different weighting mechanism for each user. In contrast, the proposed weighting mechanism algorithm automatically determines weights using a predefined formula. Consequently, the weighting values remain fixed and are not subject to user interference. This feature facilitates the user's ability to perform GAP calculations and determine weighting, as the proposed formula streamlines the process.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

References

[1] T. A. Gautre and T. H. Khan, "An analysis of question answering system for education empowered by crowdsourcing," *Proc. 2nd Int. Conf. Inven. Syst. Control. ICISC 2018*, no. Icisc, pp. 955–958, 2018, doi: 10.1109/ICISC.2018.8398942.

[2] S. S. Bhatti, X. Gao, and G. Chen, "The Journal of Systems and Software General framework , opportunities and challenges for crowdsourcing techniques : A Comprehensive survey," vol. 167, 2020.

[3] J. Chamberlain, U. Kruschwitz, and M. Poesio, "Optimising crowdsourcing efficiency: Amplifying human computation with validation," *it-Information Technol.*, 2018, doi: 10.1515/itit-2017-0020.

[4] J. Chen, R. Mo, S. Yu, D. Chen, J. Chu, and J. Gong, "The optimized selection strategy of crowdsourcing members in cloud-based design and manufacturing platform," *Adv. Mech. Eng.*, vol. 12, no. 2, pp. 1–16, 2020, doi: 10.1177/1687814020903853.

[5] Z. Ibrahim, H. Aris, and A. Mansur, "Automated Validation of Crowdsourced Data," *2018 IEEE 16th Student Conf. Res. Dev. SCOReD 2018*, pp. 1–6, 2018, doi: 10.1109/SCORED.2018.8711108.

[6] A. Li, W. Jiang, X. Li, X. Chen, and G. Wang, "Independent worker selection in crowdsourcing," *Proc. - 2020 IEEE 19th Int. Conf. Trust. Secur. Priv. Comput. Commun. Trust. 2020*, pp. 1173–1178, 2020, doi: 10.1109/TrustCom50675.2020.00157.

[7] R. Shang *et al.*, "Selection of Crowd in Crowdsourcing for Smart Intelligent Applications: A Systematic Mapping Study," *Sci. Program.*, vol. 2021, pp. 8–10, 2021, doi: 10.1155/2021/9368128.

[8] B. J. Howe, "The Rise of Crowdsourcing," no. 14, pp. 1–5, 2006.

[9] A. Juneja and N. N. Das, "Big Data Quality Framework: Pre-Processing Data in Weather Monitoring Application," *Proc. Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. Trends, Prespectives Prospect. Com. 2019*, pp. 559–563, 2019, doi: 10.1109/COMITCon.2019.8862267.

[10] E. Law and L. Von Ahn, "Human computation," *Synth. Lect. Artif. Intell. Mach. Learn.*, vol. 13, pp. 1–120, 2011, doi: 10.2200/S00371ED1V01Y201107AIM013.

[11] G. Li, J. Wang, Y. Zheng, and M. J. Franklin, "Crowdsourced Data Management: A Survey," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 9, pp. 2296–2319, 2016, doi: 10.1109/TKDE.2016.2535242.

[12] A. Cucus, L. B. Aji, A. F. Bin Mubarak Ali, A. Aminuddin, and L. D. Farida, "Selection of Prospective Workers Using Profile Matching Algorithm on Crowdsourcing Platform," *ICOIACT 2022 - 5th Int. Conf. Inf. Commun. Technol. A New W. to Make AI Useful Everyone New Norm. Era, Proceeding*, pp. 122–126, 2022, doi: 10.1109/ICOIACT55506.2022.9972155.

[13] J. Martinez-Gil, A. L. Paoletti, G. Rácz, A. Sali, and K. D. Schewe, "Accurate and efficient profile matching in knowledge bases," *Data Knowl. Eng.*, vol. 117, no. November 2017, pp. 195–215, 2018, doi: 10.1016/j.datak.2018.07.010.

[14] L. G. Rodriguez and E. P. Chavez, "Feature selection for job matching application using profile matching model," *2019 IEEE 4th Int. Conf. Comput. Commun. Syst. ICCCS 2019*, pp. 263–266, 2019, doi: 10.1109/CCOMS.2019.8821682.

[15] L. Quaranta, F. Calefato, and F. Lanubile, "KGTorrent: A dataset of python jupyter notebooks from kaggle," *Proc. - 2021 IEEE/ACM 18th Int. Conf. Min. Softw. Repos. MSR 2021*, pp. 550–554, 2021, doi: 10.1109/MSR52588.2021.00072.

[16] J. Y. Zhu, B. Tang, and V. O. K. Li, "A five-layer architecture for big data processing and analytics A five-layer architecture for big data processing and analytics 39," 2019. [Online]. Available: https://github.com/amplab-extras/.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 01:2024
DOI: 10.5281/zenodo.10612592

[17] L. Cui, X. Zhao, L. Liu, H. Yu, and Y. Miao, “Complex crowdsourcing task allocation strategies employing supervised and reinforcement learning,” *Int. J. Crowd Sci.*, vol. 1, no. 2, pp. 146–160, 2017, doi: 10.1108/IJCS-08-2017-0011.

[18] J. Wang, Y. Yang, S. Wang, C. Chen, and ..., “Context-aware Personalized Crowdtesting Task Recommendation,” *IEEE Trans. …*, 2021, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9435104/.

[19] M. Li, Y. Zheng, X. Jin, and C. Guo, “Task Assignment for Simple Tasks with Small Budget in Mobile Crowdsourcing,” *Proc. - 14th Int. Conf. Mob. Ad-Hoc Sens. Networks, MSN 2018*, pp. 68–73, 2018, doi: 10.1109/MSN.2018.00018.

[20] C. G. Harris, “Finding the Best Job Applicants for a Job Posting: A Comparison of Human Resources Search Strategies,” *IEEE Int. Conf. Data Min. Work. ICDMW*, vol. 2017-Novem, pp. 189–194, 2017, doi: 10.1109/ICDMW.2017.31.