Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/121205

Algorithm Comparison for Student-Supervisor Matching in Supervisorship System Development: K-Means vs. One-to-Many Gale-Shapley

Azamat Serek[1,*] *and Meirambek Zhaparov*[2]

[1]Department of Information Systems, SDU University, Kaskelen, Kazakhstan
[2]ICT Faculty, Paragon International University, Phnom Penh, Cambodia

Received: 2 Sep. 2023, Revised: 20 Nov. 2023, Accepted: 22 Nov. 2023
Published online: 1 Dec. 2023

Abstract: This paper presents a comparative analysis of the K-Means and Gale-Shapley algorithms for matching students to supervisors. The two algorithms are evaluated based on their performance in terms of preference satisfaction, balance of workload, time complexity, space complexity, and maximum and minimum workloads. The experimental results show that the Gale-Shapley algorithm outperforms the K-Means algorithm on all criteria, as shown in Table 4. Specifically, the Gale-Shapley algorithm achieves a preference satisfaction score of 0.74 and a balance of workload score of 0.5, compared to 0.34 and 0.2 for the K-Means algorithm. Additionally, the Gale-Shapley algorithm has a time complexity of $O(\text{num\_students} \times \text{num\_supervisors})$ and a space complexity of $O(\text{num\_students} + \text{num\_supervisors})$, which is comparable to the K-Means algorithm. Finally, the Gale-Shapley algorithm has a maximum workload of 6 and a minimum workload of 3, compared to 15 and 3 for the K-Means algorithm. Based on the experimental results, the paper concludes that the Gale-Shapley algorithm is the superior algorithm for matching students to supervisors. It achieves a higher level of preference satisfaction and balance of workload than the K-Means algorithm, and it is still relatively efficient to run. The paper also discusses the advantages and disadvantages of both algorithms and provides recommendations for future work.

Keywords: matching algorithm, students and supervisors, intersection of computer science and education

1 Introduction

Scientific supervisors are university employees whose job it is to oversee their students' work as they implement their theses or dissertations. It is common knowledge that a graduate program's effectiveness greatly depends on the supervisor-student connection. A poor establishment of their relationship generally leads to the graduation's total failure [1]. Effective supervision-student contact is essential for the graduation project to be completed successfully [2].

Supervisor who is also called an academic advisor or faculty supervisor, is a member of the faculty who helps students with their academic and research activities. Depending on the academic program and study level, undergraduate, master's, or doctorate, a supervisor's responsibilities may vary. Good results provided by Victor Sanchez-Anguix, Rithin Chalumuri, Reyhan Aydogan, Vicente Julian [3] states that personalized learning can be achieved by matching supervisors with students according to their preferences. Student needs and goals can be met with a customised academic experience designed by a supervisor who has similar research interests, pedagogical approaches, or communication styles. Motivating and engaging students also can be organized by a good supervisor-student match. Students are more likely to be driven to succeed in their coursework and research projects and to contribute to a more lively academic community when they have a solid connection with their supervisors, by Glenice Ives and Glenn Rowley [4]. Students feel more confident asking for help and voicing concerns when there is a positive working relationship based on shared expectations and values. Supervisors offer students educational advice and encouragement all through their time at university. A suitable supervisor is aware of a student's objectives, areas of strength, and fields for improvement [5]. They can then provide guidance to promote the best possible academic growth.

The relationship of scientific supervisors with students is important for several reasons:

* Corresponding author e-mail: azamat.serek@sdu.edu.kz

1. Mentoring and guidance: Throughout the research process, scientific supervisors act as mentors to their students, offering advice and assistance. They give students advise on career advancement, feedback on their work, and assistance in developing research skills.
2. Professional networking: Students looking for possibilities to work with other researchers or find jobs after graduation may find it helpful if scientific supervisors have developed networks within their field of study.
3. Opportunities for research: Students can engage in ongoing research projects or partnerships led by scientific supervisors, which offer invaluable research experience and expose them to the most recent advancements in their field.
4. Personal development: Students can gain valuable life skills from scientific supervisors, including communication, problem-solving, and critical thinking. During the study process, they can also offer encouragement and emotional support.
5. Academic achievement: A key element influencing a student's academic performance is their interaction with their scientific supervisors. The successful completion of research projects, publications, and presentations—all significant turning points in an academic career—can result from strong mentoring and coaching.

The impact of the student-supervisor relationship on final grades, student happiness, and the supervisor's perceived contribution to the student's learning was examined in this study [6]. The results of the experimental investigations demonstrated that supervisors ought to have close ties to one another and that control levels ought to be precisely calibrated.

According to this research [7], supervisor-student interaction mediates the positive moderating effect of team member support on the connection between abusive supervision and psychological capital. Abuseful supervision is found to have a detrimental impact on psychological capital.

In general, both parties' success depends on the interaction between scientific supervisors and students. In addition to offering chances for academic achievement, networking, and personal and professional development, it also promotes a sense of belonging and teamwork within the academic area. A student is inspired to persevere in completing their graduation work when a supervisor shares expertise with them.

Many universities still assign students to supervisors by hand, which can lead to biases and blunders because it doesn't take into account the psychological fit between the supervisors and the students. While efforts to match students with supervisors do exist, most of them do not take into account a wide range of measures to confirm the success of matching, particularly in academic settings. Using the k-means and one-to-many Gale-Shapley algorithms, a set of matching methods that can be used to match students to supervisors are multidimensionally analyzed in this work. The following measures are used to compare the algorithms that pair students with supervisors:

1. **Workload balance**: This measure can be used to assess how supervisors' workloads are distributed. It is computed as the supervisors' minimum to maximum workload ratio. The workload balance score will be higher if the algorithm divides the workload equitably.
2. **Preference satisfaction**: This indicator demonstrates how happy students and supervisors are with their matches. It is calculated as the average absolute difference between the supervisor's and the student's preferences for each matched student-supervisor pair. A lower preference satisfaction score indicates that the algorithm is matching students and supervisors based on comparable preferences.
3. **Maximum workload**: This metric measures the maximum number of pupils that can be assigned to a supervisor. It is useful for identifying supervisors who may require assistance because of an excessive workload.
4. **Minimum workload**: The bare minimum of students a supervisor can have is measured by this metric. Finding supervisors who may not be completely occupied and who may have more students assigned to them is useful.

2 Related work

The objective of matching is to pair two objects in a way that is appropriate and widely accepted. These limitations and guidelines should be considered by the matching algorithm when building a match that produces stable matched pairings. The context of the items involved in the matching process may be governed by a set of rules and preferences.

Matching is essential in many facets of life, from daily operations to market activity. It can be used to match a set of entities while accounting for their diverse tastes and membership in distinct sets, pair store products with potential customers, and more [8,20,11,10,9]. One-to-one, many-to-one, and many-to-many matching tasks are examples of abstract matching tasks.

2.1 One-to-one matching

One-to-one matching has been the subject of extensive research [13, 14, 15, 16]. One-to-one matching, as the name suggests, is the process of matching every element in one set to exactly one element in the other. Stated otherwise, a matching occurs when there is a single, exclusive relationship between the elements of the two sets.

2.2 Many-to-one matching

Numerous components from one collection can be paired with just one from another using a type of matching known as many-to-one matching [17, 18, 19]. To put it another way, it's a matching where there isn't necessarily a single, exclusive correlation between the elements of the two sets.

2.3 Many-to-many matching.

Many-to-many matching is a type of matching where numerous elements in one set can be linked with numerous elements in another set in a process known as "many-to-many matching." Put otherwise, it is a matching when the components of the two sets do not always have a unique, exclusive correspondence.

2.4 Matching students to supervisors

The problem of matching students to supervisors is considered to be many-to-one matching, because one student can have only one supervisor, but one supervisor can have many students. There are a lot of work dedicated on matching students to supervisors.

Assigning research tasks to graduate students involves several crucial procedures, one of which is placing them with supervisors. Carefully connecting students with supervisors can ensure a positive experience for both sides and boost the effectiveness of research initiatives. In recent years, there has been an increased focus on developing efficient tools and algorithms for matching students with supervisors.

The issue of assigning supervisors to final-year students was examined by Ismail et al. The university's students may have limited their selection pool by taking into account only the lecturers they knew, as they were unaware of the types of faculty members they had and their interests. Based on the preferences of both the students and the supervisors, the authors developed a recommendation system that pairs final-year project students with possible supervisors. In order to put the system into place, first student data indicating how interested they were in the project names of the supervisors' prior work was gathered. 51 data records in total were gathered. After that, the recommendation engine received the data, used the Euclidean distance function to compute similarities, and produced potential supervisors to students [8].

Kawagoe T. and Matsubae looked into the problem of pairing supervisors with students. In Japanese universities, supervisors are chosen by students, and each has a minimum quota. This could lead to issues with conflicts of interest between students. The authors built a student-supervisor matching engine and employed a deferred acceptance (DA) technique to address this problem. In order to complete the matching process, students first had interviews with supervisors to determine their preferred hierarchy. Following this procedure, supervisors were given the list of students who had ranked their supervisors. Based on these rankings, supervisors created a priority order by ranking the students themselves [20].

Considerable academic research has been sparked by the search for the best way to match academic supervisors with students. Unique research by Maedeh Mosharraf and Fattaneh Taghiyareh, [21] established the foundation for algorithmic methods in student-supervisor matching. The effectiveness of machine learning and data analysis is explored in recent works by Areeba Ahmed, Saif ur Rehman [22] and Yuanyuan Fan, Ana Evangelista and Hadi Harb [23], who show how algorithms analyse large data sets to determine the best pairings based on interpersonal dynamics, academic accomplishments, and research interests. The modern environment places a strong emphasis on the value of comprehensive matching standards. Christine Phund, Angela Byars-Winston, Janet Branchaw, Sylvia Hurtado, Kevin Eagan [24] support a methodology that takes into account mentoring philosophies, communication styles, and shared expectations in addition to academic metrics. This all-encompassing viewpoint strives for compatibility outside of conventional bounds while acknowledging the complex nature of academic relationships. As the academic environment changes, new research emphasises the necessity of dynamic systems that can adjust to new requirements. According to B. Chachuat, B. Srinivasan, D. Bonvin, [25] real-time adaptation is necessary to meet the changing goals and priorities of both managers and students. Yamada, S., Cappadocia, M. C., and Pepler, D. [26] examine the moral issues raised by student-supervisor matching programs. As these systems use more data, privacy, biases, and fairness issues start to

Table 1: Comparison with other research work

Paper	Methodology	Key Findings
[30]	semi-structured interviews as a method of gathering data. In total, 14 supervisor interviews and 23 supervisor invitations from 16 different institutions were held.	Discovered significant discrepancies between the demands of the real students and the supervisors' view of their roles.
[8]	Utilizing a recommendation engine, students and supervisors can match and use the Euclidean distance technique by determining similarity based on the answers to a questionnaire.	An easy-to-implement method that can help students choose their supervisors and lower the likelihood of errors throughout the assignment process.
[31]	gathering information on research topic choices from supervisors and students, then using a genetic algorithm to find matches in this area	introducing a genetic algorithm that can find supervisors for students.
[32]	evaluating student proposals using information retrieval, then suggesting supervisors in light of the findings.	Supervisors are grouped according to the research they have done. 53.09% accuracy was the average. Retrieving information from supervisors' profiles by looking only at the top ten terms can be useful.
[33]	A model of students' interests is created via reinforcement learning, in which students can modify terms from faculty members' papers to describe what interests them.	The authors addressed situations in which students find it difficult to formulate precise questions about their research interests or in which there is a dearth of information about possible supervisors.
[20]	Two-sided matching is used, and type-specific minimum and maximum quotas are taken into account. There were 67 supervisors and 254 pupils in all.	Justified jealousy amongst students of the same "type" is eliminated by the system.
This study	gathering information about how each party views the relationships between students and supervisors, creating preferences based on this information, and then algorithms to carry out matching based on the metrics.	The experimental results show that the Gale-Shapley algorithm outperforms the K-Means algorithm on all criteria, as shown in Table 4. Specifically, the Gale-Shapley algorithm achieves a preference satisfaction score of 0.74 and a balance of workload score of 0.5, compared to 0.34 and 0.2 for the K-Means algorithm.

surface. The study makes the case for a careful examination of ethical issues and suggests solutions for problems arising from algorithmic decision-making. A recommendation system is presented by Putney, M. W., Worthington, E. L., & McCullough [27] to pair students with possible supervisors according to their interests. Amos Tversky and Eldar Shafir [28] use a deferred acceptance mechanism to handle conflicts of interest in student-supervisor matching. Horst Bunke [29] reviews matching theory and algorithms, investigate applications, and discuss big data and stability concerns in matching algorithms. Together, these studies advance our knowledge of and ability to enhance the student-supervisor matching procedure.

Tables 1 shows comparison with other research work.

2.5 Knowledge Gap

Based on the literature review, it can be observed that none of the found work performed multidimensional analysis of matching students to supervisors specific to the context of university settings in terms of psychological perceptions of students and supervisors regarding their workflows.

3 Methods and materials

3.1 Dataset

The purpose of this dataset is to investigate views of what constitutes a healthy student-supervisor dynamic. A questionnaire that included both general and particular inquiries about the respondent and their opinions of the student-supervisor dynamic was used to gather the results. Whether the individual is a student or supervisor, as well as the university, are among the basic inquiries.

The particular questions are given as statements and pertain to opinions about what constitutes a healthy student-supervisor relationship. On a scale of 1 to 5, respondents were asked to indicate how much they agreed or disagreed with these statements. The statements outline the supervisor's duties, which include choosing a research topic, choosing a suitable theoretical framework and methodology, creating a suitable program and schedule for the student's research and study, upholding a strictly professional relationship, insisting on frequent meetings, making sure the student is working consistently and on task, demanding to see all drafts of the work, and, if needed, helping with the thesis writing and making sure the presentation is faultless. It is thought of these variables as discrete quantitative variables. The questions for defining perceptions were used from [38].

A total of 130 records were gathered, of which 20 came from supervisors and 110 from undergraduate students. Just 45% of supervisors only hold an MSc degree, compared to 55% who hold a PhD.

3.2 Algorithms

3.2.1 KMeans

K-means clustering is a type of unsupervised learning algorithm that groups data points into a predefined number of clusters. The algorithm works by iteratively assigning each data point to the cluster with the nearest mean [34]. The K-means algorithm can be summarized as follows: [35]

- Initialize the cluster centers. This can be done randomly or by using a heuristic method, such as choosing the k most distant data points as the initial cluster centers.
- Assign each data point to the cluster with the nearest mean.
- Update the cluster means by averaging the values of the data points assigned to each cluster.
- Repeat steps 2 and 3 until convergence.

The K-means algorithm is guaranteed to converge to a local minimum of the sum of squared errors (SSE) objective function. The SSE objective function is defined as the sum of the squared distances between each data point and its assigned cluster center [36]. The K-means algorithm is typically terminated when the cluster assignments no longer change or when a predefined number of iterations have been completed [37].

3.2.2 One-to-many Gale Shapley

The Gale-Shapley algorithm is an algorithm for solving the stable marriage problem. One-to-many Gale Shapley is a one-to-many matching algorithm, which means that each person is matched to at most one other person, but one person can be matched to multiple other people. The Gale-Shapley algorithm works by having each person propose to their most preferred person who has not yet rejected them. If the proposal is accepted, the two people are engaged. If the proposal is rejected, the proposer proposes to their next most preferred person who has not yet rejected them. This process continues until all people are engaged or have proposed to everyone they prefer [**?**]. The Gale-Shapley algorithm is guaranteed to find a stable matching, if one exists. A stable matching is a matching such that there are no two pairs of people who would both prefer to be matched to each other than to their current partners The Gale-Shapley algorithm has been applied to a variety of real-world problems, including school choice, kidney exchange, and resident assignment.

The one-to-many Gale-Shapley algorithm is a generalization of the Gale-Shapley algorithm that allows one person to be matched to multiple other people. This is in contrast to the original Gale-Shapley algorithm, which is a one-to-one matching algorithm, meaning that each person is matched to exactly one other person. The one-to-many Gale-Shapley algorithm works as follows:

- Each person proposes to their most preferred person who has not yet rejected them.
- If a person is proposed to by multiple people, they choose to accept the proposal from the person they prefer the most.
- If a proposal is rejected, the proposer proposes to their next most preferred person who has not yet rejected them.
- This process continues until all people are engaged or have proposed to everyone they prefer.

4 Results

The results in Table 2 show that the K-Means algorithm achieves a moderate level of preference satisfaction (0.34) and balance of workload (0.2). This suggests that the algorithm is able to find a reasonable match between students and supervisors, even though it does not guarantee to find the optimal match.

The maximum workload of 15 and minimum workload of 3 suggest that the algorithm is able to distribute the workload among supervisors relatively evenly. However, it is important to note that these are just the maximum and minimum workloads, and the actual workload distribution may vary.

Overall, the results in Table 2 suggest that the K-Means algorithm is a viable approach for matching students to supervisors. It achieves a moderate level of preference satisfaction and balance of workload, and it is relatively efficient to run.

Table 2: Summary of Metrics in K-Means

Metric	Value
Preference Satisfaction	0.34
Balance of Workload	0.2
Time Complexity	$O(\text{num\_students} \times \text{num\_supervisors})$
Space Complexity	$O(\text{num\_students} + \text{num\_supervisors})$
Maximum Workload	15
Minimum Workload	3

The results in Table 3 show that the Gale-Shapley algorithm achieves a high level of preference satisfaction (0.74) and balance of workload (0.5). This suggests that the algorithm is able to find a very good match between students and supervisors.

The maximum workload of 6 and minimum workload of 3 suggest that the algorithm is able to distribute the workload among supervisors very evenly. This is a significant improvement over the K-Means algorithm, which had a maximum workload of 15 and a minimum workload of 3.

Overall, the results in Table 3 suggest that the Gale-Shapley algorithm is a superior approach for matching students to supervisors than the K-Means algorithm. It achieves a higher level of preference satisfaction and balance of workload, and it is still relatively efficient to run.

Table 3: Summary of Metrics of One-to-Many Gale Shapley

Metric	Value
Preference Satisfaction	0.74
Balance of Workload	0.5
Time Complexity	$O(\text{num\_students} \times \text{num\_supervisors})$
Space Complexity	$O(\text{num\_students} + \text{num\_supervisors})$
Maximum Workload	6
Minimum Workload	3

Table 4 shows comparison of these algorithms in terms of the metrics.

Table 4: Experimental results for matching students to supervisors

Criterion	K-Means	Gale-Shapley
Preference satisfaction	0.34	0.74
Balance of workload	0.2	0.5
Time complexity	$O(\text{num\_students} \times \text{num\_supervisors})$	$O(\text{num\_students} \times \text{num\_supervisors})$
Space complexity	$O(\text{num\_students} + \text{num\_supervisors})$	$O(\text{num\_students} + \text{num\_supervisors})$
Maximum workload	15	6
Minimum workload	3	3

5 Discussion

Based on the results of the experiments, the Gale-Shapley algorithm is the superior algorithm for matching students to supervisors. It achieves a higher level of preference satisfaction and balance of workload than the K-Means algorithm. Additionally, the Gale-Shapley algorithm is guaranteed to find a stable matching, if one exists.

However, the Gale-Shapley algorithm is more complex to implement than the K-Means algorithm and can be less efficient to run for large datasets.

If the highest level of preference satisfaction and balance of workload is the most important criterion, then the Gale-Shapley algorithm is the recommended algorithm. However, if simplicity of implementation and efficiency are more important criteria, then the K-Means algorithm may be a better choice.

K-Means Algorithm.

- Pros:
  - Relatively simple to implement
  - Relatively efficient to run
- Cons:
  - Moderate level of preference satisfaction and balance of workload
  - Can be sensitive to the initialization of the cluster centers

One-to-Many Gale-Shapley.

- Pros:
  - High level of preference satisfaction and balance of workload
  - Guaranteed to find a stable matching, if one exists
- Cons:
  - More complex to implement than K-Means
  - Can be less efficient to run than K-Means for large datasets

One area of future work is to develop more efficient implementations of the Gale-Shapley algorithm. This would make the algorithm more practical for use with large datasets.

Another area of future work is to develop hybrid algorithms that combine the advantages of the K-Means and Gale-Shapley algorithms. For example, it may be possible to develop an algorithm that uses the K-Means algorithm to initialize the Gale-Shapley algorithm. This could improve the efficiency of the Gale-Shapley algorithm without sacrificing its high level of performance.

6 Conclusion

In conclusion, the Gale-Shapley algorithm is the superior algorithm for matching students to supervisors. It achieves a significantly higher level of preference satisfaction and balance of workload than the K-Means algorithm, as evidenced by the experimental results. Additionally, the Gale-Shapley algorithm is guaranteed to find a stable matching, if one exists. This is a critical property, as it ensures that no student or supervisor is left in a situation where they are matched to someone they prefer less than another available match.

However, it is important to note that the Gale-Shapley algorithm is more complex to implement than the K-Means algorithm and can be less efficient to run for large datasets. This is because the Gale-Shapley algorithm requires each student to propose to their most preferred supervisor, and then to iterate through this process until all students are matched. In contrast, the K-Means algorithm simply groups students together based on their preferences, without the need for any iteration.

Overall, the Gale-Shapley algorithm is the recommended algorithm for matching students to supervisors, particularly when a high level of preference satisfaction and balance of workload is essential. However, if simplicity of implementation and efficiency are more important criteria, then the K-Means algorithm may be a better choice.

References

[1] E., D. A growing phobia. *Nature*. (2017)
[2] Yarwood-Ross L, H. As others see us: what PhD students say about supervisors. *Nurse Researcher*. pp. 38-43 (2014)
[3] Sanchez-Anguix, V., Chalumuri, R., Aydoğan, R. & Julian, V. A near Pareto optimal approach to student–supervisor allocation with two sided preferences and workload balance. *Applied Soft Computing*. **76** pp. 1-15 (2019)

[4] Ives, G. & Rowley, G. Supervisor selection or allocation and continuity of supervision: Ph. D. students' progress and outcomes. *Studies In Higher Education*. **30**, 535-555 (2005)
[5] Westhuizen, P. & De Wet, J. The training needs of supervisors of postgraduate students in the social sciences and humanities: research in higher education. *South African Journal Of Higher Education*. **16**, 185-195 (2002)
[6] Kleijn, R., Mainhard, M., Meijer, P., Pilot, A. & Brekelmans, M. Master's thesis supervision: Relations between perceptions of the supervisor–student relationship, final grade, perceived supervisor contribution to learning and student satisfaction. *Studies In Higher Education*. **37**, 925-939 (2012)
[7] Liao, Z. & Liu, Y. Abusive supervision and psychological capital: a mediated moderation model of team member support and supervisor-student exchange. *Frontiers Of Business Research In China*. **9**, 576 (2015)
[8] Ismail, M., Razak, T., Hashim, M. & Ibrahim, A. A Simple Recommender Engine for Matching Final-Year Project Student with Supervisor. *ArXiv Preprint ArXiv:1908.03475*. (2019)
[9] Akbarpour, M., Li, S., & Gharan, S. *Thickness and Information in Dynamic Matching Markets*. In *Journal of Political Economy*, **128**, 783-815 (2020).
[10] Wang, L., Li, Y., Huang, J., & Lazebnik, S. *Learning Two-Branch Neural Networks for Image-Text Matching Tasks*. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**, 394-407 (2018).
[11] Ren, J., Xia, F., Chen, X., Liu, J., Hou, M., Shehzad, A., Sultanova, N., & Kong, X. *Matching Algorithms: Fundamentals, Applications and Challenges*. In *IEEE Transactions on Emerging Topics in Computational Intelligence*, **5**, 332-350 (2021).
[12] Kawagoe, T., & Matsubae, T. *Matching with Minimal Quota: Case Study of a Student-Supervisor Assignment in a Japanese University*. Available at SSRN 3429626 (2020).
[13] Roth, A. & Sotomayor, M. *Two-sided matching*. In *Handbook Of Game Theory With Economic Applications*, Vol. 1, pp. 485-541 (1992).
[14] Becker, G. *A Theory of Marriage: Part I*. In *Journal Of Political Economy*, **81**, 813-846 (1973).
[15] Bergstrom, T. & Bagnoli, M. *Courtship as a Waiting Game*. In *Journal Of Political Economy*, **101**, 185-202 (1993).
[16] Gale, D. & Shapley, L. College admissions and the stability of marriage. *The American Mathematical Monthly*. **120**, 386-391 (2013)
[17] Che, Y., Kim, J. & Kojima, F. Stable matching in large economies. *Econometrica*. **87**, 65-110 (2019)
[18] Ismaili, A., Hamada, N., Zhang, Y., Suzuki, T. & Yokoo, M. Weighted matching markets with budget constraints. *Journal Of Artificial Intelligence Research*. **65** pp. 393-421 (2019)
[19] Kelso Jr, A. & Crawford, V. Job matching, coalition formation, and gross substitutes. *Econometrica: Journal Of The Econometric Society*. pp. 1483-1504 (1982)
[20] Kawagoe, T. & Matsubae, T. Matching with Minimal Quota: Case Study of a Student-Supervisor Assignment in a Japanese University. (2020), Available at SSRN 3429626
[21] Mosharraf, M. & Taghiyareh, F. An evolutionary-based educational expert system to maximize student-supervisor compatibility. *2012 6th IEEE International Conference On E-Learning In Industrial Electronics (ICELIE)*. pp. 84-89 (2012)
[22] Ahmed, A. & Rehman, S. A Machine Learning Approach for Advisors to Discover in Higher Education. (2022)
[23] Fan, Y., Evangelista, A. & Harb, H. An automated thesis supervisor allocation process using machine learning. *Global Journal Of Engineering Education*. **23**, 20-30 (2021)
[24] Pfund, C., Byars-Winston, A., Branchaw, J., Hurtado, S. & Eagan, K. Defining attributes and metrics of effective research mentoring relationships. *AIDS And Behavior*. **20** pp. 238-248 (2016)
[25] Chachuat, B., Srinivasan, B. & Bonvin, D. Adaptation strategies for real-time optimization. *Computers & Chemical Engineering*. **33**, 1557-1567 (2009)
[26] Yamada, S., Cappadocia, M. & Pepler, D. Workplace bullying in Canadian graduate psychology programs: Student perspectives of student–supervisor relationships.. *Training And Education In Professional Psychology*. **8**, 58 (2014)
[27] Putney, M., Worthington, E. & McCullough, M. Effects of supervisor and supervisee theoretical orientation and supervisor-supervisee matching on interns' perceptions of supervision.. *Journal Of Counseling Psychology*. **39**, 258 (1992)
[28] Tversky, A. & Shafir, E. Choice under conflict: The dynamics of deferred decision. *Psychological Science*. **3**, 358-361 (1992)
[29] Bunke, H. Graph matching: Theoretical foundations, algorithms, and applications. *Proc. Vision Interface*. **2000** pp. 82-88 (2000)
[30] Orellana, M. & Others Improving doctoral success by matching PhD students with supervisors. *International Journal Of Doctoral Studies*. **11** pp. 87 (2016)
[31] Sanchez-Anguix, V., Chalumuri, R. & Julian, V. A multi-objective evolutionary proposal for matching students to supervisors. *Distributed Computing And Artificial Intelligence, 15th International Conference*. pp. 94-102 (2019)
[32] Wijanto, M., Rachmadiany, R. & Karnalim, O. Thesis Supervisor Recommendation with Representative Content and Information Retrieval. *Journal Of Information Systems Engineering And Business Intelligence*. **6**, 143-150 (2020)
[33] Gao, Y., Ilves, K. & Głowacka, D. Officehours: A system for student supervisor matching through reinforcement learning. *Proceedings Of The 20th International Conference On Intelligent User Interfaces Companion*. pp. 29-32 (2015,3)
[34] MacQueen, J. Some methods for classification and analysis of multivariate observations. *Proceedings Of The Fifth Berkeley Symposium On Mathematical Statistics And Probability*. **1** pp. 281-297 (1967)
[35] Hartigan, J. & Wong, M. Algorithm AS 136: A k-means clustering algorithm. *Applied Statistics*. **28**, 100-108 (1979)
[36] Lloyd, S. Least squares quantization in PCM. *IEEE Transactions On Information Theory*. **28**, 129-137 (1982)
[37] Arthur, D. & Vassilvitskii, S. k-means++: The advantages of careful seeding. (2007)

[38] Supervisor & Research Student Expectation Questionnaire, *Online*, Available:
https://www.notredame.edu.au/__data/assets/pdf_file/0019/13078/Supervisor-Research-Student-Expectation-Questionnaire.pdf, [Accessed: 23-Apr-2022].