 ISSN 2579-7298

Analysis of Graduate Success Patterns Based on Association Rule Mining to Increase the Achievement of the Performance Index of Higher Education Graduates

Wasilah [a,1*], Halimah [a,2]

[a] Darmadajaya Institute of Informatics and Business, Jl. ZA Pagar Alam No. 93 Labuhan Ratu,Lampung 35142, Indonesi
[1] wasilah@darmajaya.ac.id , [2] halimahyunus@darmajaya.ac.id
* corresponding author

ARTICLE INFO

Article history
Received 09 May 2023
Revised 18 July 2023
Accepted 03 Sep 2023

Keywords
Primary performance index of graduates,
apriori algorithm,
Algoritm,
looking for job via online,
Work Placement

ABSTRACT

The main performance index (KPI) of graduates is an important factor for universities. However, there is still very little research that discusses the achievement of IKU graduates. If this is managed well, universities can take more effective action to improve the quality of education and prepare graduates to be successful in their careers. This research will utilize data mining to photograph the achievements of implementing one of the main performance indices in higher education. The research stages are initial study, Knowledge Discovery Database (KDD), data mining process, interpretation and evaluation. The KDD stage includes data collection, data pre-processing consisting of data selection, data integration, data cleaning, and data transformations. Data processing was carried out using a priori algorithms and Rapidminner tools. Processing using a minimum support value of 10% and a confidence value of 60% produces 36 rules. There are 6 rules that show students who have a high GPA will get a job quickly and with a salary greater than 1.2 UMR. A maximum of 20% of graduates meet this rule with a confidence value of 86%. There are 7 rules indicating which studentsLooking for jobs online will get a job quickly even though they do not show compliance with a salary greater than 1.2 UMR. A maximum of 10% of graduates meet this rule with a confidence value of 76%. This gives the image that a high GPA does not guarantee eligibility for a job. This condition requires an analysis of the quality of learning.Looking for jobs onlinehas an impact on the speed of getting a job even though it has no impact on salary eligibility. Work placement and obtaining information from faculty career does not appear at the 10% support value. Work placement and obtaining information from faculty career appears at a support value of 3%. This shows that work placement and obtaining information from faculty careers are still not functioning enough to produce graduates who are successful in getting work for < 6 months and a salary > 1.2 minimum wage.

1. Introduction

Decision support systems (DSS) have been widely used to assist decision making in higher education. DSS can assist in analysis, performance measurement, monitoring and business management for institutions, both in administrative and academic work[1] [2][3][4]. Investments in DSS in Higher Education Institutions aim to improve decision-making processes. DSS has been adopted and implemented to support higher education business processes. DSS can help institutions make better and more effective decisions. Therefore, investment in DSS is considered important to increase the efficiency and effectiveness of decision-making processes in higher education[5].

Implementation of DSS in higher education includes Measuring students' skills, Students' enrollment and admission, Program design for accreditation and improved university ranking, etc. Decision making methods widely used in higher education include Analytical Hierarchy Process (AHP), decision tree classification algorithms, and various frameworks such as DDDM, ANP/SNA, synchronous Delphi, DEA, and reference text methods.[6][1][3][7].

Decision quality can be improved by creating new data and information that can be applied to the decision domain and facilitate understanding and learning for decision makers[8]. Various studies have been carried out to improve the quality of information to support decision making. Improving the quality of information in decision making can be done by involving the use of data mining. Several studies state that decision support systems utilizing data mining can be used to make the right decisions[9]. The involvement of data mining in decision making in organizations becomes more effective and scientific by modeling, classifying, grouping and finding correlations between data[10].

The main performance index (KPI) of graduates is an important factor for universities [11]-[13]. However, there is still very little research that discusses the achievement of IKU graduates. A well-managed main performance index can be a reference for determining more effective actions. Higher education can improve the quality of education and prepare graduates to be successful in their careers. It is hoped that the results of this analysis can become a reference for universities to improve the main performance index scores of their graduates. This will have an impact on increasing accreditation.

The principles in designing Main Performance Indicators for Higher Education that need to be considered include: increasing the relevance of higher education to the needs of industry, the business world and the world of work. Apart from that, universities also need to prioritize targets for adaptation to change. In this regard, the study program must involve partners from industry, the business world or the world of work. Eight Main Higher Education Performance Indicators as indicators of changes that will have the greatest impact on the quality of graduates, the quality of lecturers, and the quality of the curriculum.

This research will utilize data mining to photograph the achievements of implementing one of the main performance indices in higher education. Research related to the use of data mining in analyzing KPI compliance is still very rare. This is important to do because success in fulfilling the IKU can illustrate the success of higher education institutions in adapting to change. The aim of the research is to use data mining, especially the a priori algorithm, to find out patterns that support graduates' success in getting jobs. The novelty that will be produced in this research is that the analysis is carried out based on the main performance index reference for the independent campus program. Apart from that, the attributes used are a combination of academic and non-academic attributes.

The research results can help provide information on the current dominant patterns in producing graduates who are successful in getting decent jobs according to the main performance index (KPI). This information can help university management in making related decisions.

2. Method

2.1 Association Rules Algorithm

Association is one method of data mining. In DM, association rule mining is important [14]. This method looks for similar patterns in data that often appear. The association rule algorithm is an algorithm that finds attributes that "go together". The association rules algorithm departs from the pattern "If antecedent, then consequent," along with measurements of support (coverage) and confidence (accuracy) associated with the rules. Association rule algorithms include: Apriori algorithm, FP-Growth algorithm, GRI algorithm [15].

An association rule is said to be interesting if the support value is greater than the minimum support value and the confidence value is greater than the minimum confidence value. Association rule mining is the process of finding patterns, correlations, associations, or causal structures that usually occur in different data in various types of databases, including relational data, transactional data, and various other forms of data storage [16][17] . Usually this method is used to find specific

consumer behavior from a certain group. The most basic algorithm of Association Rule Mining is the Apriori algorithm[14]. The a priori algorithm is one of the algorithms that is widely applied in the Association rule mining algorithm. The main process carried out in the a priori algorithm is to obtain the highest frequency pattern. The principle of this algorithm is merging and pruning. High frequency patterns are element patterns in the database that have a support value exceeding a certain limit (minimum support). The Apriori algorithm aims to identify itemsets that frequently appear, so that it can be used to find important patterns or association rules in transactional data. In higher education, the implementation of the Apriori algorithm includes: identifying suitable teaching methods for Outcome-Based Education (OBE)[18], the use of cellphones in learning success[19], enhance college enrollment management[20].

2.2 Research Stages

The stages in this research are the initial study, the Knowledge Discovery Database stage which includes data collection, data pre-processing (data selection, data integration, data cleaning, and data transformations), data mining process using a priori algorithm methods, and interpretation and evaluation. The stages of this research are depicted in a flowchart as in Figure 1.

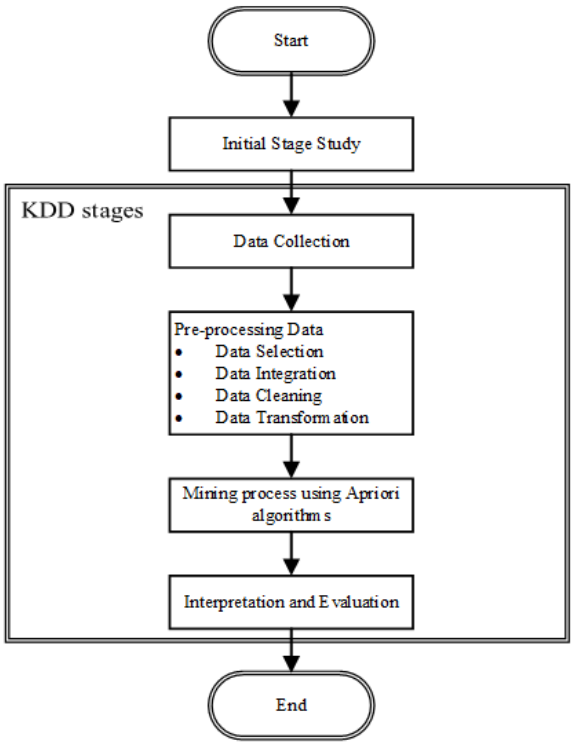

Fig. 1. Research Stages

1) Preliminary Study

Based on the phenomena that occur, an analysis is carried out to understand the scope, details of the problem and problem solving issues.

2) Data Collection

Data collection was carried out from various related sources. The data source must be determined precisely because it will affect the quality of the pattern that will be produced.

3) Data Pre-processing

In data mining, the pre-processing stage is a process that has a very important role. Pre-processing has several stages, namely:

a) Selecting

This research uses college graduates X in 2021 and 2022. The data consists of academic data and study tracers.

b) Integration

The data integration process is the process of combining various sources. In this research, the merging was carried out on data based on graduation year, namely 2021 and 2022. Next, the merging was based on academic data sources and study tracer data.

c) Cleaning

This data cleaning stage is carried out with the aim of eliminating incomplete, irrelevant, redundant and empty data. At this stage data cleaning is also carried out by deletion. In this research, deletion was carried out on empty data, as well as data that had inappropriate values. Inappropriate values include: salary and waiting time with a minus value, salary with an undefined UMR value, and salary with a value of 0. This process causes the data to be reduced from 678 to 368.

d) Transformation

After the data has been cleaned, the next thing to do is transform the data into tabular format which contains Boolean format in Microsoft Excel. This needs to be done because the rapidminner tool can read Boolean format. In the process of transforming data into data in tabular format, this is done by changing the existing data with the values 1 and 0.

4) Mining Process Using the Apriori Algorithm

At this stage the mining process is carried out by applying an a priori algorithm. This algorithm is used to analyze the association rules of several attributes. Attributes are used as parameters in forming data association rules. The attributes used are: faculty. GPA category, work placement, looking for jobs online, obtain information from faculty career center, waiting period [21][22][23] and salary category[24].

5) Interpretation and Evaluation

After analysis, the results were obtained in the form of rules that show association relationships. The rules are then interpreted. Rules are interpreted based on the resulting patterns. The results of this interpretation become the basis for developing recommendations

3. Results and Discussion

3.1 Collecting and Cleansing

The data used in the research are study tracer data and academic data for graduates in 2021 and 2022 at private university This process resulted in a reduction in data from 678 to 368.

3.2 Selection and Transform

The data processed consists of academic data and non-academic data. Academic data consists of Faculty and GPA. Meanwhile, tracer data consists of: work placement, looking for jobs online, obtaining information from faculty career center, waiting period and salary category. Next, the transformation process is carried out into tabular format.

The tabular in this study contains several items as follows:

ATTRIBUTE	TABULAR 1	TABULAR 0
Faculty	FEB	ILKOM
IPK Category	>=3.0	3>
Work Placement	Yes	No
Looking For Job Via Online	Yes	No
Obtain Information From Faculty Career Center	Yes	No
Wating Period	< 6 Mont	>= 6 Mont
Salary Category	Greater Than 1.2 UMR	Smaller Than 1.2 UMR

3.3 Mining Process Using the Apriori Algorithm

Data processing uses a priori algorithms and Rapidminner tools. The minimum support value used is 10% and the confidence value is 60%. The formation of the a priori algorithm process in the rapidminner application is carried out through the following stages: import data, withdraw data to the work page, enter the items "numerical to binominal, fp growth, create association rules, determine the support value and confidence value. The resulting algorithm logic scheme is:

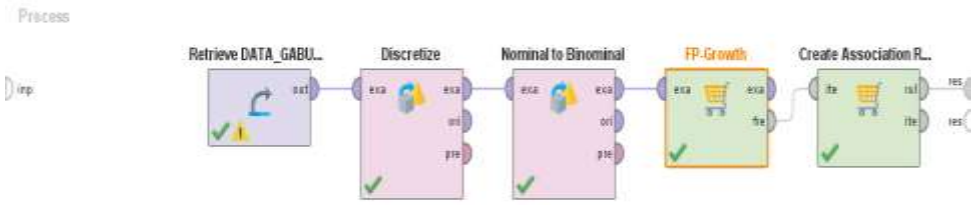

Fig. 2. Apriori Algorithm Logic Scheme

Based on the results of the calculation process using the a priori algorithm, 36 rules were produced. The resulting rules are displayed in table 1 below:

Table 1. Association Rules for University Graduates

No	Premises	Conclusion	Support	Confident
28	Ilkom, Salary Category	Waiting Period	0.101	0.907
37	Ilkom, Salary Category	IPK Category	0.106	0.953
13	Ilkom, Looking For Job Via Online	Waiting Period	0.109	0.75
12	Feb, Salary Category	IPK Category, Waiting Period	0.114	0.746
23	IPK Category, FEB, Salary Category	Waiting Period	0.114	0.83
33	Waiting Period, FEB, Salary Category	IPK Category	0.114	0.936
22	Feb, Salary Category	Waiting Period	0.122	0.797
26	Ilkom, Looking For Job Via Online	IPK Category	0.13	0.893
27	Feb, Salary Category	IPK Category	0.138	0.898
8	Feb, Looking For Job Via Online	Waiting Period	0.166	0.736
9	Feb, Looking For Job Via Online	IPK Category, Waiting Period	0.166	0.736
15	IPK Category, FEB, Looking For Job Via Online	Waiting Period	0.166	0.762
41	Waiting Period, FEB, Looking For Job Via Online	IPK Category	0.166	1
21	Salary Category	IPK Category, Waiting Period	0.208	0.792
25	IPK Category, Salary Category	Waiting Period	0.208	0.86
34	Waiting Period, Salary Category	IPK Category	0.208	0.941
40	Feb, Looking For Job Via Online	IPK Category	0.218	0.966
24	Salary Category	Waiting Period	0.221	0.842
31	Salary Category	IPK Category	0.242	0.921
6	Looking For Job Via Online	IPK Category, Waiting Period	0.265	0.708
14	IPK Category, Looking For Job Via Online	Waiting Period	0.265	0.761
36	Waiting Period, Looking For Job Via Online	IPK Category	0.265	0.953
11	Looking For Job Via Online	Waiting Period	0.278	0.743
7	Ilkom	IPK Category, Waiting Period	0.296	0.713
19	IPK Category, Science and Technology	Waiting Period	0.296	0.776
30	Waiting Period, Ilkom	IPK Category	0.296	0.919
18	Ilkom	Waiting Period	0.322	0.775
32	Looking For Job Via Online	IPK Category	0.348	0.931
29	Ilkom	IPK Category	0.382	0.919
10	FEB	IPK Category, Waiting Period	0.436	0.74
17	IPK Category, FEB	Waiting Period	0.436	0.774
39	Waiting Period, FEB	IPK Category	0.436	0.966
16	FEB	Waiting Period	0.452	0.767
38	FEB	IPK Category	0.564	0.956
20	IPK Category	Waiting Period	0.727	0.776
35	Waiting Period	IPK Category	0.727	0.943

Table 1 shows that there are 6 rules indicating that students who have a high GPA will get a job quickly and with a minimum salary of 1.2 UMR. The 6 rules are 12,23,33,21,25,34. A maximum of 20% of graduates meet this rule with a confidence value of 86%. In graphic form it is shown in Figure 2 below:

Fig. 3. Graph of the 10 lowest Association Rules

The graph above shows that the rules related to GPA category, salary category and waiting period appear at support values of 10% -20%. This percentage is included in the low category.This gives the image that a high GPA does not guarantee eligibility for a job. This condition requires an analysis of the quality of learning.

Looking for jobs online has an impact on waiting for jobs as shown in rules 13, 8, 9, 15, 6, 14 and 11.These rules show students whoLooking for jobs onlinewill get a job quickly even if they do not meet the minimum salary of 1.2 UMR. A maximum of 10% of graduates meet this rule with a confidence value of 76%. In graphic form it is shown in Figure 3 below:

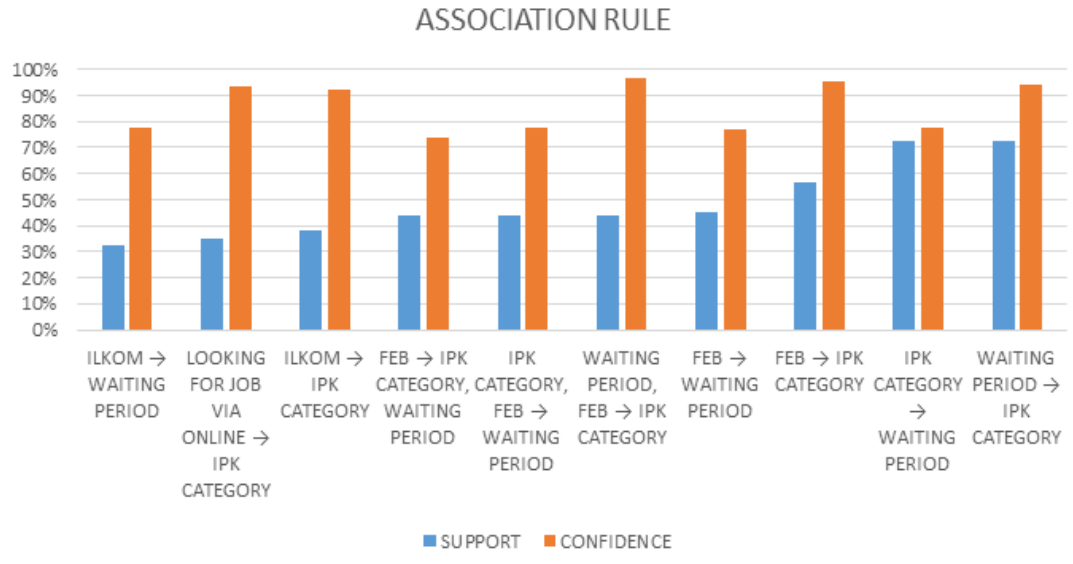

Fig. 4. Graph of the 10 highest Association Rules

The graph above shows that rules related to waiting periods and looking for jobs via online appear at support values of 30% -70%.This gives an image thatLooking for jobs online has quite a high impact on the speed of getting a job. However, this was not accompanied byminimum salary of 1.2 UMR.Work placement and obtaining information from faculty career does not appear at the 10%

support value. Work placement and obtaining information from faculty career appears at a support value of 3%. This shows that work placement and obtaining information from faculty career is still not very effective in producing graduates who are successful in getting jobs for < 6 months and a salary > 1.2 minimum wage.The results of data processing show that the GPA of graduates above 3.0 reaches 95%. Very few students graduate with a GPA lower than 3. Several studies state that a good GPA will have an impact on the ease of getting a job. The resulting pattern does not describe these ideal conditions. Based on these findings, several recommendations for improvement can be produced aimed at higher education management.

4. Conclusion

Based on the results of the process of forming graduate success patterns which was carried out using the a priori algorithm and Rapiminner tools, the results show as follows:

1. Processing using a minimum support value of 10% and a confidence value of 60% produces 36 rules.
2. There are 6 rules that show students who have a high GPA will get a job quickly and with a salary greater than 1.2 UMR. A maximum of 20% of graduates meet this rule with a confidence value of 86%.
3. There are 7 rules that show that students who are looking for a job online will get a job quickly even though they do not show compliance with a salary greater than 1.2 UMR. A maximum of 10% of graduates meet this rule with a confidence value of 76%.
4. Work placement and obtaining information from faculty career does not appear at the 10% support value.
5. Work placement and obtaining information from faculty career appears at a support value of 3%. The results above show that work placement and obtaining information from faculty careers are still not very functional in producing graduates who are successful in getting work for < 6 months and a salary > 1.2 minimum wage.

The resulting recommendations include the need to increase the role of work placement and obtain information from faculty careers.Besides thatSpecial attention also needs to be paid to increasing graduate competency as a companion to the GPA score. It is necessary to ensure that graduates have accompanying competencies in accordance with the field of work that is the opportunity. This fulfillment can help increase the number of graduates who have a high GPA to get a job within 6 months and a salary of more than 1.2 UMR. This is one of the efforts to increase the success of the main performance index of graduates. The success of the main performance index of graduates is one of the factors that has an impact on increasing higher education accreditation.

References

[1] F. Ş. YÜKSEL, A. N. KAYADELEN, and F. ANTMEN, “A systematic literature review on multi-criteria decision making in higher education,” *Int. J. Assess. Tools Educ.*, vol. 10, no. 1, pp. 12–28, 2023, doi: 10.21449/ijate.1104005.

[2] S. Doltsinis, P. Ferreira, M. M. Mabkhot, and N. Lohse, “A Decision Support System for rapid ramp-up of industry 4.0 enabled production systems,” *Comput. Ind.*, vol. 116, p. 103190, 2020, doi: 10.1016/j.compind.2020.103190.

[3] B. D. Nikola Kadoić, N. Begicevic, “A new method for strategic decision-making in higher education,” *Cent. Eur. J. Oper. Res.*, vol. 26, pp. 611–628, 2018, doi: 10.1007/s10100-017-0497-4.

[4] L. S. Ling and S. Krishnasamy, “Information Technology Capability (ITC) Framework to Improve Learning Experience and Academic Achievement of Mathematics in Malaysia,” *Electron. J. e-Learning*, vol. 21, no. 1, pp. 36–51, 2023, doi: 10.34190/ejel.21.1.2169.

[5] A. H. Mohammed, A. H. Mousa, N. M. Almeyali, and I. S. Nasir, “M2cim-dss: A model for measuring continuance intention in decision support systems,” *Indones. J. Electr. Eng. Informatics*, vol. 9, no. 3, pp. 756–765, 2021, doi: 10.52549/.v9i3.3032.

[6] Z. J. S. Akanmu, "A conceptual framework for students' data-driven decision making process in higher education institutions," *Adv. Sci. Lett.*, vol. 21, pp. 2256–2260, 2015, doi: 10.1166/asl.2015.6275.
[7] U. A. Urinov, "Formation Of Students 'Practical Skills In Technical Higher Education," *Am. J. Soc. Sci. Educ. Innov.*, vol. 02, no. 10–52, pp. 313–320, 2020, doi: : https://doi.org/10.37547.
[8] N. M. S. DD. Steiger, N. M. Steiger. Steiger, "Decision Support as Knowledge Creation: An Information System Design Theory," *Hawaii Int. Conf. Syst. Sci.*, 2007, doi: 10.1109/HICSS.2007.162.
[9] A. R. Onik, "Predictive Analysis and Warehousing of Web Log Data," *Int. J. Comput. Appl.*, 2017, doi: 10.5120/IJCA2017914257.
[10] N. W. Yuan Guo, "The internet of things-based decision support system for information processing in intelligent manufacturing using data mining technology," 2020.
[11] E. C. Dina Yuniar, Heti Mulyati, "Faktor-faktor yang mempengaruhi penyelesaian masa studi program pascasarjana di Institut Pertanian Bogor," *J. Akuntabilitas Manaj. Pendidik.*, vol. 7, no. 2, 2019, doi: 10.21831/AMP.V7I2.25084.
[12] S. Jian-ron, "Performance-Based Assessment in Evaluating Student Learning Outcomes," *Educ. Comput. Sci.*, 2006.
[13] D. DIKTI, *Buku Panduan IKU Perguruan Tinggi Negeri*. 2021.
[14] D. Liu, "Construction of Higher Education Management and Student Achievement Evaluation Mechanism Based on Apriori Algorithm," *Mob. Inf. Syst.*, 2022, doi: 10.1155/2022/5375825.
[15] Z. Bin, "On the Development of Association Rules," *J. Jilin Agric. Sci. Technol. Coll.*, 2008.
[16] S. Sivanantham, V. Mohanraj, Y. Suresh, and J. Senthilkumar, "Association Rule Mining Frequent-Pattern-Based Intrusion Detection in Network," *Comput. Syst. Sci. Eng.*, vol. 44, no. 2, pp. 1617–1631, 2023, doi: 10.32604/csse.2023.025893.
[17] D. Yan, "The Correlation Analysis of University Courses Based on Association Rule," *Comput. Sci. Educ.*, 2010.
[18] H. F. Yanna Qiao, "Study on OBE Teaching Concept in the Context of Deep Learning for the Construction of University Mathematics Microcourses," *Comput. Intell. Neurosci.*, 2022.
[19] P. Gao, "Research on Analysis of Students Using Mobile Phones in Ideological and Political Classrooms by Apriori Algorithm of Association Rules," *Procedia Comput. Sci.*, vol. 208, pp. 12–17, 2022, doi: 10.1016/j.procs.2022.10.003.
[20] J. W. Hequn Wu, Mende Enhe, "Research on Application of Data Mining Based on Improved APRIORI Algorithm in Enrollment Management in Colleges and Universities," *Lect. Notes Electr. Eng.*, 2018, doi: DOI:10.1007/978-981-13-3648-5_136.
[21] A. D. Nicole Gürtzgen, "Do Digital Information Technologies Help Unemployed Job Seekers Find a Job? Evidence from the Broadband Internet Expansion in Germany," *Eur. Econ.*, 2018, doi: 10.2139/ssrn.3194424.
[22] X. L. Dan Zhou, Sibo Yang, "Internet Use and Job Satisfaction: Evidence from China," *Int. J. Environ. Res. Public Health*, 2022, doi: DOI:10.3390/ijerph191912157.
[23] R. U. Manuel Denzer, Thorsten Schank, "Does the Internet Increase the Job Finding Rate? Evidence from a Period of Internet Expansion," *Soc. Sci. Res. Netw.*, 2018, doi: 10.2139/ssrn.3249878.
[24] C. V.-B. F. Castellacci, "Internet use and job satisfaction," *Comput. Human Behav.*, 2019, doi: DOI:10.1016/j.chb.2018.09.001.