Comparing academic performance across course topics: a pilot study

Laura Fink
Faculty of entrepreneurship
GEA College
Ljubljana, Slovenia
laura.fink@gea-college.si

Bojan Cestnik
Faculty of entrepreneurship
GEA College
Ljubljana, Slovenia
bojan.cestnik@gea-college.si
bojan.cestnik@temida.si

ABSTRACT

In this paper, we examine the academic performance of students in different courses to determine whether good performance in one course is related to good performance in other courses. Although certain predictive models emphasize the importance of course content for learning success, there are few studies that address how student performance in different courses is related to similar course topics, learning goals, competences, and skills. By creating a preliminary framework that examines how academic performance is related to different course topics, we attempt to make a first step further towards addressing the research gap regarding the interrelatedness of student achievement not only in different course topics but in different competence areas. We examined a set of student grades from eleven different courses at the faculty from areas such as entrepreneurship, management and leadership, business informatics, mathematics, economics, marketing and market analysis, innovation and creativity, English, finance and accounting, business law, and human resource management. We show that students with more exam retakes on average reached a lower grade rank than the students who only registered for the exam once. We used linear regression to show the significance of the relationships between student performance in the Informatics course compared to their achievement in other courses. With a correlation matrix coefficient, we measured the strength of reciprocal interrelatedness between the grade ranks students attained in each of the eleven courses. The results of this preliminary study indicate a possible stronger association between academic achievement in courses that have similarities in terms of content or focus, such as business administration and entrepreneurship (correlation coefficient of 0.58). Further studies with detailed comparison of course-specific competences are needed for accepting the finding that interrelatedness between achievements in courses from similar versus different disciplines is stronger. The preliminary model could further be improved by a broader range of courses, input explanatory student factors, and application of advanced analytical techniques.

KEYWORDS

course-specific competence, learning analytics, prediction models, student academic achievement, student academic performance

1 INTRODUCTION AND RELATED WORK

Are students who achieve excellent results in one course more likely to achieve outstanding results in another? And vice versa? This is the key question that triggered this preliminary research into how students perform in different courses. As a result, we are examining the relationships between students' academic achievement in different courses from two different undergraduate study programs at the faculty of entrepreneurship. Despite some prediction models suggesting the course content as one of the input explanatory variables, there is a significant lack of detailed research on the relationship between students' achievement in different courses from the same, similar, or entirely different discipline. Therefore, this preliminary study aims to develop a pilot model to investigate the relationship between academic success in various courses and their respective course topics. The findings of this study could be further reinforced and interpreted by the comparison of course-specific competences and learning objectives.

Apart from the prediction model of academic achievement, interrelatedness between academic achievement in different courses compared to the course main topics, competences, and learning objectives is still largely missing. Since previous studies that would previously investigate the interrelatedness of student achievement in different courses are, to the best of our knowledge, entirely nonexistent, let us draw attention to the research studies that are related to the field and that have led to this investigation. OECD [12, 13], for example, show a positive association between literacy and numeracy skills. Moreover, in our previous research [2], we showed that students who achieved better academic achievement in word skills were on average also more likely to achieve better achievement in excel skills. However, this does not imply that it is sufficient to develop only some of these skills, such as solely including word skills, excel skills, literacy skills, or just numeracy skills in the curriculum. Furthermore, this does not imply that students in a real situation cannot achieve much better results in a certain type of skill compared to another. Additionally, Fink and Vadnjal [3] conducted a pilot study that compared the development of

* Comparing academic performance across course topics: a pilot study.
† Fink, L., Cestnik, B.

Information Society 2023, 7–11 October 2024, Ljubljana, Slovenia

https://doi.org/10.70314/is.2024.cog.10

generic and course-specific competences during a higher education course.

As previously mentioned, one could apply findings about the relationships between various course topics, contents, competences, or objectives to the development of predictive models and predictive algorithms. Prediction models are often used to find out ahead of time which students are likely to drop out or fail external exams. The schools aim to take intervention measures and steps to stop the bad predictions from coming true, and the success rate can be raised [16, 9]. There are a lot of different input explanatory variables that can have a big effect on how well and especially how accurately prediction models perform [16, 4, 7, 8, 9, 11, 1]. Some prediction models include related concepts to the course content [7], such as the course's learning objective, the course's main competences, the course topic [4], or even course preparedness [10].

In addition to the combination of input explanatory variables, various statistical methodologies and techniques, as well as different types of measurements and academic achievements, significantly influence the prediction model's power and performance. While adding additional or all relevant factors to the model does not always improve its performance, the right combination of input explanatory variables significantly influences the accuracy and other model performance measures [4]. In the end, the right combination of input variables largely determines the model's explanatory power, the accuracy of its predictions, and other performance measures [6, 15].

Different prediction models are based on different methodologies and include different input and output variables. Francis and Babu [4], for example, compare prediction models of student achievement that include the topic of the course as the input explanatory variable. They demonstrated that the course topic, along with many other factors, is one of the explanatory variables for academic achievement. However, their model found that academic factors, including the course topic, were less accurate in predicting students' academic achievement than demographic factors, behavior factors, and other factors such as absence days, parental satisfaction, and school survey responses. They developed, compared, and assessed performance measures of several prediction models. The model that included academic factors, behavior, and additional input explanatory variables showed the greatest improvement in accuracy. On the other hand, adding demographic factors on top of that reduced the accuracy of academic achievement prediction. Clearly, the addition of additional input variables, for example, the topic of the course, in different models contributes differently to improving prediction accuracy and other performance measures, depending on other input variables in the model.

2 HYPOTHESES

In this preliminary study, we aim to build a preliminary pilot research model on which we will test the interrelatedness between academic achievement in different courses. We suggest the following hypotheses:

H1. Mostly there are reciprocal relationships between a student's performance in one course and their performance in another.

H2. As the number of exam retakes increases, the student's grade rank decreases.

3 DATASET AND METHODOLOGY

The dataset collection and preparation included several phases. First, we have collected students' grades for different courses at the higher education institution. The initial dataset included the grades of 223 students for 67 different courses.

In the second phase, we have refined and further prepared the dataset. Based on some simple data exploration and visualization techniques, such as plotting the missing values, plotting the distribution of the number of grades available per course, and plotting the distribution of the number of exam retakes per course, we have decided to eliminate the data of courses with less than 60 students' grades per course. With that, we narrowed further analysis to the following eleven selected courses: Business Economics, Informatics, Management and Leadership, Marketing and Market Analysis, Entrepreneurship, Business English, Accounting, Creativity and Innovation, Business Mathematics and Statistics, Business Law, and Human Resource Management.

Table 1: Number of grades per one and two courses

Course (Short name)	Nr. of students/grades	
	per course	for chosen course and Informatics (%)
Informatics for entrepreneurs (P07_IE)	160	
Business economics (P05_BE)	139	120 (86 %)
Business law for entrepreneurs (P06_BLE)	79	71 (90 %)
Human resource management (P09_HRM)	66	58 (88%)
Management and leadership (P15_ML)	156	142 (91%)
Marketing and market analysis (P16_MMA)	90	73 (81%)
Entrepreneurship (P27_ENT)	191	148 (77%)
Business English (P28_BE)	152	133 (88%)
Accounting for entrepreneurs (P33_AE)	143	131 (92%)
Creativity and innovation in entrepreneurship (P36_CIE)	173	142 (82%)
Business mathematics and statistics (P39_BMS)	139	119 (86%)

We then compared the number of students' grades available per one course with the number of grades available per two courses (the selected course and the informatics) and calculated the share of students that also took the exam in informatics compared to the number of students who took the exam in the selected course only. Students' grades include both those that indicate a student has passed the course and those that indicate a student has not. Since we included only one grade per student in further analysis, the number of students' grades reflects the number of students who took the exam in each course. Not all students took exams in all eleven subjects. The reasons for this are varied, including the fact that the courses are from two different programs.

The courses with the highest student grades include Entrepreneurship (191), followed by Creativity/Innovation (173) and Informatics (160) (Table 1). More than 77% of students (148) who took the entrepreneurship exam also took the informatics exam. Similarly, 82% of students who took the exam in creativity/innovation (142), and 91% of those who took the exam in management/leadership (142) also took the exam in informatics. In other words, students who attended the exam in the Informatics course often also attended the exam in the Entrepreneurship, Management/Leadership, and Creativity/Innovation courses, as shown in Table 1.

Then, we continued by calculating the average grade ranks and number of exam retakes (table 2) for each of the courses included in the analysis. The grade ranks range from 0 to 10, where 0 represents not attending, 1 to 5 represents failed, 6 satisfactory, 7 average, 8 good, 9 very good, and 10 excellent. As shown in Table 2, the students on average achieved better grades in the HRM course than in other courses. In comparison, the students on average achieved the lowest average grade rank in the Mathematics/Statistics and Economics course compared to other courses. The students also, on average, most commonly retook the exam in these two courses. Additionally, we found that 38 students attended at least one exam deadline for each of the eleven courses.

Table 2: Average grade rank and number of exam retakes

Course (Short name)	Avg. grade	Avg. nr. of retakes
Informatics for entrepreneurs (P07_IE)	7.7	1.2
Business economics (P05_BE)	6.6	1.8
Business law for entrepreneurs (P06_BLE)	7.1	1.3
Human resource management (P09_HRM)	8.0	1.1
Management and leadership (P15_ML)	7.3	1.2
Marketing and market analysis (P16_MMA)	7.5	1.3
Entrepreneurship (P27_ENT)	7.4	1.3
Business English (P28_BE)	7.3	1.2
Accounting for entrepreneurs (P33_AE)	7.0	1.3
Creativity and innovation in entrepreneurship (P36_CIE)	7.2	1.3
Business mathematics and statistics (P39_BMS)	6.4	2.0

In the third phase, we continued with the data exploration and visualization. We plotted the distribution of the number of grades achieved per grade rank for each of the selected eleven courses. An example of the distribution for the Entrepreneurship course is provided in Figure 1.

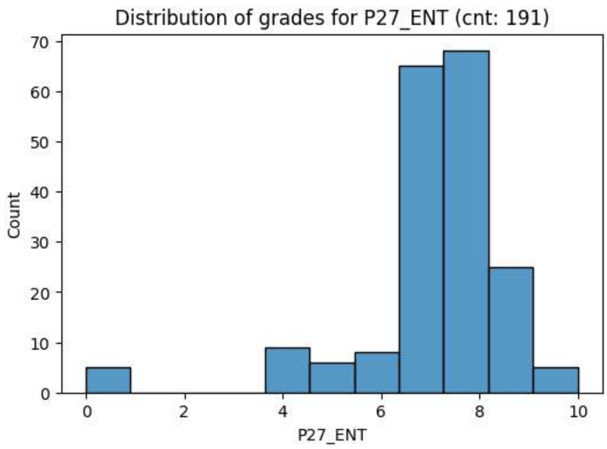

Figure 1: Distribution of number of grades per each grade rank for Entrepreneurship course example

In the next phase, we focused our investigation on how the grades that students achieved in the Informatics course behave compared to the grades they achieved in the remaining ten courses. We performed and visualized ten linear regression models describing the relationships between grade ranks students achieved in the informatics course and the grade ranks students achieved in other selected courses. When performing the linear regression, we included the grade rank achieved in informatics as an independent variable. Though we are aware that linear regression models assume the influence of the independent variable on the dependent variable and not the reciprocal relationships per se, we decided to mention this limitation and work further with the results obtained from the regression analyses in this preliminary pilot study.

We performed additional analysis based on the correlation matrix between the grade ranks students achieved in each of the eleven selected courses. We draw a correlation matrix with significant ($p < 0.05$) correlations among regression coefficients between the eleven selected courses to determine which of the eleven courses is related to another one. We then examined the strength and significance of the reciprocal relationship that the correlation matrix coefficient measures.

Next, we further compared the characteristics of our data for the selected eleven courses with the characteristics of the data for all the courses. We used data visualization techniques such as plotting to compare the distribution of the average grade of eleven selected courses and all the courses. The distribution of average grade for the selected eleven courses seems fairly similar to the distribution of average grade for all the courses, as shown in Figure 2.

Although we cannot claim that an analysis of the entire data set would yield similar results to the analysis of the selected eleven courses based solely on the similar distribution of average grades, we cannot completely rule out the possibility that on average, somewhat similar associations would emerge among the other courses.

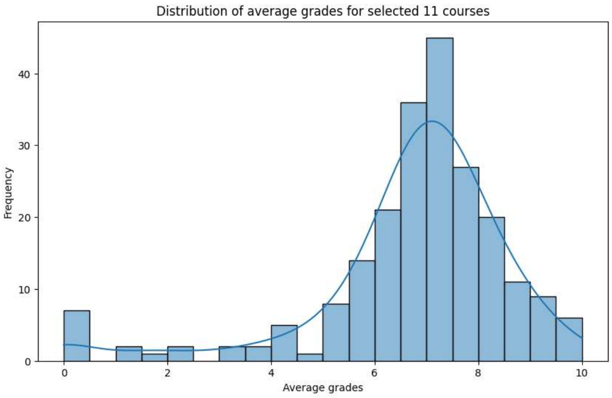

Figure 2: Distribution of number of grades per each grade rank per course example

In the final phase of this preliminary pilot research, we performed the regression analysis between the grade rank that the student achieved and the number of exam retakes of the student.

4 RESEARCH FINDINGS

The eleven selected courses served as the basis for the analysis, which focused on identifying reciprocal relationships between the grade ranks of one course and those of another course. We selected the courses based on the number of grades available after exploring, visualizing, cleaning, and refining the data. In addition to linear regression models, we calculated the correlation matrix's correlations (Figure 3) that investigate the reciprocal association between the grade ranks students achieved in one course compared to another.

Overall, we investigated 55 reciprocal relationships (H1) between the grade ranks of one course with the grade ranks of another course. Among these, forty-six correlations are significant ($p < 0.05$) compared to nine correlations that are not significant. Among the significant correlations, nine correlation coefficients exhibit a moderate relationship (between 0.5 and 0.7) between the grade ranks of one course and the grade ranks of another course. Twenty-seven correlations show a weak correlation (between 0.3 and 0.5), while ten correlations show a negligible or low correlation (below 0.3) between the grade ranks of two selected courses.

The strength of significant coefficients varies from 0.19 all the way up to 0.58. Many of the coefficients that we would otherwise have placed in the group of weak correlations are very close to 0.5, which indicates moderate correlation. There are also many coefficients that we have otherwise placed in the group of negligible correlations close to the value of 0.3, which indicates a weak correlation.

Based on these results, we can accept hypothesis 1 that foresees the existence of reciprocal relationships between a student's performance in one course and their performance in another in most cases when comparing different pairs of courses.

Let's examine the concrete correlations between the grades of the two courses. There is a moderate correlation between Economics and Mathematics/Statistics (0.60), followed by Economics and Entrepreneurship (0.58), Informatics and Mathematics/Statistics (0.57), Marketing/Market Analysis and Law (0.57), Marketing/Market Analysis and Entrepreneurship (0.55), Economics and Accounting (0.54), Marketing/Market Analysis and Creativity/Innovation (0.52), and Economics and Informatics (0.51), and Informatics and Marketing/Market analysis (0.50).

Based on moderate correlations, we can speculate that Economics, Informatics, Mathematics, and Accounting courses are closely connected, partially because students have to use the numeracy skills in these courses. Therefore, future research could address the question of whether the syllabus of these courses reflects shared similar competences. Although rare, previous research [12, 13] on the interrelatedness of competences has shown that people who are more proficient in literacy skills are usually also more proficient in numeracy skills, and vice versa, additional inquiry into the similar and different competences that are developed within these courses would provide an important insight and more thoroughly address the gap in the literature on the interrelatedness between competences that is largely still missing. Therefore, we also need to investigate further to what degree are the numeracy skills included in the syllabus of the courses that otherwise aim at developing soft, social, and other professional skills, such as the Entrepreneurship and Marketing/Market Analysis courses. Additionally, based on the analysis, it would also make sense to check whether the Marketing/Market Analysis and Creativity/Innovation courses foster the development of related skills. In general, we can speculate that the type of competences is important for student achievement, but to confirm this, we would have to perform the qualitative analysis of the similar competences and learning objectives in the future.

The majority of the relationships are statistically significant but weak. Weak correlation exists between Entrepreneurship and Mathematics/Statistics (0.49), Entrepreneurship and Informatics (0.47), Economics and Management/Leadership (0.47), Entrepreneurship and Law (0.45), English and Entrepreneurship (0.45), Economics and English (0.43), Creativity/Innovation and Entrepreneurship (0.42), HRM and Law (0.41), Accounting and HRM (0.41), Informatics and Accounting (0.41), Management/Leadership and Accounting (0.40), English and Mathematics/Statistics (0.38), Mathematics/Statistics and Law (0.38), Marketing/Market Analysis and Law (0.37), English and Accounting (0.37), Management/Leadership and English (0.36), Entrepreneurship and HRM (0.36), Law and Economics (0.34), Management/Leadership and Entrepreneurship (0.34), Entrepreneurship and Accounting (0.33), Accounting and Mathematics/Statistics (0.33), Management/Leadership and Marketing/Market Analysis (0.31), Marketing/Market Analysis and Economics (0.31), Management/Leadership and Law (0.31), Informatics and Law (0.30), English and HRM (0.30), Creativity/Innovation and Accounting (0.30).

A low correlation below 0.3 exists between Accounting and Law (0.29), Marketing/Market Analysis and English (0.29), Marketing/Market Analysis and Accounting (0.28), Economics and Creativity/Innovation (0.27), Informatics and Management/Leadership (0.26), Management/Leadership and Creativity/Innovation (0.25), Management/Leadership and Mathematics/Statistics (0.25), English and Informatics (0.22), English and Creativity/Innovation (0.20), and Creativity/Innovation and Mathematics/Statistics (0.19).

Based on the weak and low correlations, we can speculate that these correlations exist in courses that do not necessarily share that much of similar competences and learning objectives as those that exhibit moderate correlations. As mentioned previously, further qualitative research is required to substantiate these assumptions.

Since not all students took the exams in all the courses, the number of observed instances for each pair of two courses ranges from 46 to 166, depending on the particular pair of courses. Not only that, we found that the low number of observed instances importantly contributed to the statistical insignificance of some correlation coefficients. The relationships calculated based on a low number of observed instances, such as, for example, between Mathematics/Statistics and Law (46 observed instances), Marketing/Market Analysis and Mathematics/Statistics (46), HRM and Economics (47), Informatics and HRM (58), English and Law (58), Management/Leadership and HRM (59), and Creativity/Innovation and HRM (59 observed instances), exhibit insignificant relationships. The exemptions are the insignificant relationship between the Informatics and Creativity/Innovation course with 142 observed instances and between the Creativity/Innovation and Law course with 74 observed instances. We therefore speculate that the Informatics and Law courses strive to develop different competences than the Creativity/Innovation course. This preliminary analysis is a useful basis for further research and analysis.

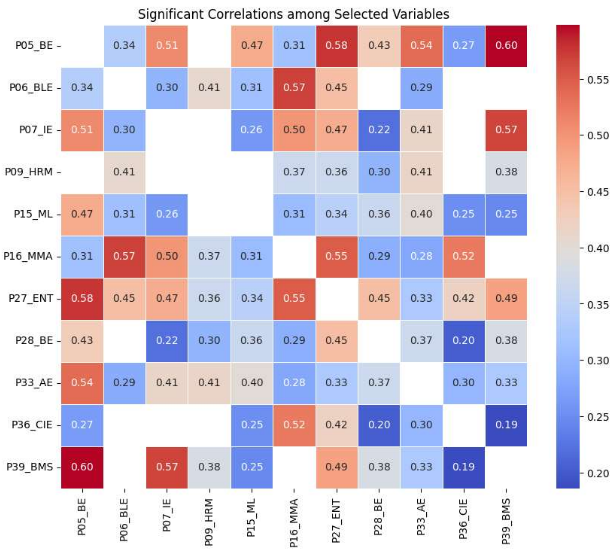

Figure 3: Correlation matrix with significant correlations

The results of the regression analysis between the grade rank that the student achieved and the number of exam retakes that the student took (H2), that are shown in Figure 4, suggest that the students with more exam retakes on average reached a lower grade rank than the students who only registered for the exam once (β = -0.47, p = 0.00). The full equation is displayed below (equation 1). Based on these results, we accept hypothesis 2 that as the number of exam retakes increases, the student's grade rank decreases.

$$Grade\ rank = 8.25 - 0.47 * Nr.of\ exam\ retakes + residuals \quad (1)$$

Figure 4 also shows that most students register for the exam once, fewer students register for the exam twice, and even fewer students register for the exam a third time.

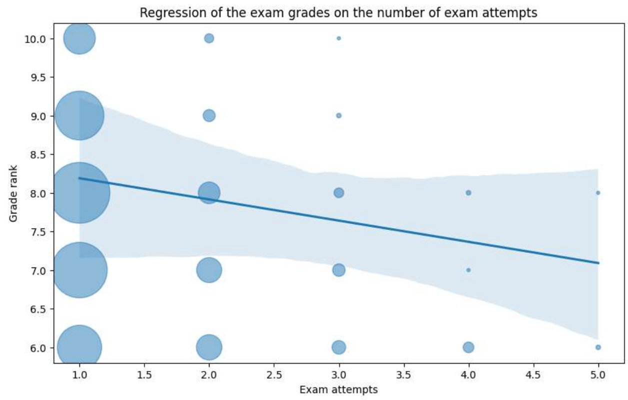

Figure 4: Regression analysis between number of exam retakes and average grade achieved

5 CONCLUSION

The purpose of this study was to lay the groundwork for further research regarding the correlation between different course-specific competences and to present initial findings regarding the correlation between students' academic achievement in different courses. In this paper, we aim to enhance our comprehension of the intricate relationship between competences, an area that remains largely unexplored. The preliminary analysis revealed the existence of interrelatedness among grades students achieve in different courses, and showed that a student's academic performance in one course influences their performance in another. Based on the analysis we accept hypothesis 1 that foresees the existence of reciprocal relationships between a student's performance in one course and their performance in another in most cases when comparing different pairs of courses. We also show that students with more exam retakes on average reached a lower grade rank than the students who only registered for the exam once. With that, we accept hypothesis 2 that as the number of exam retakes increases, the student's grade rank decreases.

To determine whether there is a stronger correlation between academic achievements in courses from the same or similar discipline than in courses from completely different disciplines, further research is required to explore how much the interrelatedness between courses depends on the competences, learning goals, and discipline of the course.

Since this is a preliminary pilot analysis, we considered additional opportunities to improve our research in the future. We could enhance this study by utilizing additional methods, such as cluster analysis, network analysis, and structural equation modeling, along with techniques used to make predictions like data mining [11, 1], neural networks [5], or decision trees [14]. Furthermore, we could enrich the model by increasing the number of observations and the number of courses included in the analysis.

To capture more subtleties in these relationships, we could potentially build and test the model's performance with additional input variables such as information about general knowledge, broad outlook, ambitions, and psychological characteristics. Additionally, we could compare different cohorts of students, the semester, the study program and track, forms of study (full-time, part-time), types of study programs (undergraduate, postgraduate, higher vocational program), and study modes (classroom, blended, online).

REFERENCES

[1] Balqis Albreiki, Nazar Zaki and Hany Alashwal.2021. A systematic literature review of student'performance prediction using machine learning techniques. *Education Sciences, 11*(9), 552.

[2] Laura Fink and Bojan Cestnik. 2023. Reinforcing digital skills: a quantitative analysis of digital skills development and its impact on higher educational outcomes. In Szymanska, K. (ed.), Janczewski, R. A. (ed.). *Problemy i wyzwania ekonomii i zarządzania w XXI wieku : wybrane aspekty*. Poznań: Grupa Wydawnicza FNCE. P. 263-274.

[3] Laura Fink and Jaka Vadnjal. 2024. Generic and course-specific competences in comparison. In Fošner, A. (ed.). *ABSRC Ljubljana: conference proceedings,* GEA College - Faculty of Entrepreneurship, p. 10.

[4] Bindhia K. Francis and Suvanam Sasidhar Babu. 2019. Predicting academic performance of students using a hybrid data mining approach. *Journal of medical systems*, 43(6), 162.

[5] Leon Gerritsen. 2017. *Predicting student performance with Neural Networks.* Master Thesis. Tilburg University, Netherlands.

[6] Ramin Ghorbani and Rouzbeh Ghousi. 2020. Comparing different resampling methods in predicting students' performance using machine learning techniques. *IEEE access, 8*, 67899-67911.

[7] Judith M. Harackiewicz, Kenneth E. Barron, John M. Tauer and Andrew J. Elliot. 2002. Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of educational psychology, 94*(3), 562.

[8] Shaobo Huang and Ning Fang 2013. Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education*, 61, 133-145. https://www.sciencedirect.com/science/article/abs/pii/S0360131512002102

[9] Kitsadaporn Jantakun, Thiti Jantakun, and Thada Jantakoon. 2022. The architecture of system for predicting student performance based on data science approaches (SPPS-DSA architecture). *International Journal of Information and Education Technology, 12*(8), 778-785.

[10] René F. Kizilcec, and Hansol Lee. 2022. Algorithmic fairness in education. In Holmes, W., Porayska-Pomsta, K. (Eds.), *The ethics of artificial intelligence in education*, Routledge. p. 174-202.

[11] Abdallah Namoun and Abdullah Alshanqiti. 2020. Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences, 11*(1), 237.

[12] OECD. 2019A. *Skills Matter: Additional Results from the Survey of Adult Skills,* OECD Skills Studies, OECD Publishing, Paris, https://doi.org/10.1787/1f029d8f-en.

[13] OECD. 2019B. *The Survey of Adult Skills: Reader's Companion,* Third Edition, OECD Skills Studies, OECD Publishing, Paris, https://doi.org/10.1787/f70238c7-en

[14] Harikumar Pallathadka, Alex Wenda, Edwin Ramirez- Asís, Maximiliano Asís-López, Judith Flores-Albornoz and Khongdet Phasinam. 2023. Classification and prediction of student performance data using various machine learning algorithms. *Materials today: proceedings,* 80, 3782-3785.

[15] David M. W. Powers. 2011. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation". *Journal of Machine Learning Technologies, 2*(1): 37–63.

[16] Mustafa Yağcı. 2022. Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments, 9*(1), 11.