https://doi.org/10.31891/csit-2025-1-10
UDC 37.012:311.2:519.223.5

Kateryna BEREZKA, Oksana BASHUTSKA
West Ukrainian National University
Nataliya NAVOLSKA
WSB Merito University in Wrocław
Vasil MELNYCHENKO
West Ukrainian National University

ANALYSIS OF STUDENT PERFORMANCE BASED ON CANONICAL CORRELATION ANALYSIS

The article examines the application of Canonical Correlation Analysis (CCA) to investigate the relationships between student performance outcomes across different groups of disciplines. The disciplines were categorized into the following groups: mathematics, programming and algorithms, systems design, networks and distributed systems, applied software and technologies, and economic and managerial disciplines. The study aims to identify dependencies between these discipline groups that influence overall academic performance.

The analysis revealed that discrete mathematics plays a key role in shaping programming skills, with performance in mathematical disciplines significantly correlating with outcomes in other fields. Both strong and weak correlations were identified between specific discipline groups. The use of CCA provided deeper insights into the relationships between subjects, offering new opportunities for optimizing the educational process.

The findings of the article have both theoretical and practical significance, contributing to the improvement of educational approaches and methods for assessing academic performance.

Future directions is the application of the findings to enhance academic performance prediction models. This research provides a foundational basis for the use of advanced statistical methods to uncover significant dependencies in education, paving the way for improvements in learning outcomes and educational strategies. The study demonstrates the potential of Canonical Correlation Analysis (CCA) as a reliable method for exploring relationships between disciplines and students' academic performance, providing valuable insights for the development of educational strategies and improvement of curricula, which will contribute to enhancing learning effectiveness and optimizing resources.

Keywords: canonical correlation analysis, discipline groups, canonical correlation, canonical model, student performance.

Катерина БЕРЕЗЬКА, Оксана БАШУЦЬКА
Західноукраїнський національний університет
Наталія НАВОЛЬСЬКА
Університет ВШБ Меріто в Вроцлаві
Василь МЕЛЬНИЧЕНКО
Західноукраїнський національний університет

АНАЛІЗ УСПІШНОСТІ СТУДЕНТІВ НА ОСНОВІ КАНОНІЧНОГО КОРЕЛЯЦІЙНОГО АНАЛІЗУ

Сучасна освіта стикається з низкою викликів, які вимагають комплексного підходу до їх вирішення . У статті розглядається застосування канонічного кореляційного аналізу (CCA) для дослідження взаємозв'язків між результатами успішності студентів у різних групах дисциплін. Було здійснено поділ дисциплін на групи: математичні, програмування та алгоритми, системне проектування, мережі та розподілені системи, прикладне програмне забезпечення та технології, економічні та управлінські дисципліни. Метою дослідження є встановлення залежностей між групами дисциплін, які впливають на загальну академічну успішність. На основі аналізу було встановлено, що дискретна математика відіграє ключову роль у формуванні навичок програмування, успішність у математичних дисциплінах суттєво корелює з результатами в інших галузях. Встановлені тісні і слабкі зв'язки між деякими групами дисциплін. Використання CCA дозволило глибше зрозуміти зв'язок між предметами, що відкриває нові можливості для оптимізації навчального процесу.

Результати статті мають як теоретичне, так і практичне значення, сприяючи вдосконаленню навчальних підходів та методів оцінювання академічної успішності.

Майбутніми напрямками є застосування отриманих результатів для вдосконалення моделей прогнозування академічної успішності. Це дослідження забезпечує фундаментальну основу для використання передових статистичних методів для виявлення значущих залежностей в освіті, прокладаючи шлях до покращення результатів навчання та освітніх стратегій. Дослідження демонструє потенціал канонічного кореляційного аналізу як надійного методу дослідження взаємозв'язків між дисциплінами та академічною успішністю студентів, надаючи цінну інформацію для розробки освітніх стратегій та вдосконалення навчальних програм, що сприятиме підвищенню ефективності навчання та оптимізації ресурсів.

Ключові слова: канонічний кореляційний аналіз, групи дисциплін, канонічна кореляція, канонічна модель, успішність студентів.

Introduction

Modern education faces a range of challenges that require a comprehensive approach to address them [1]. One of the primary tasks is adapting educational programs to rapid technological advancements and the evolving demands of employers to prepare students for the challenges of the modern labor market.

The challenges of contemporary education can be effectively addressed through the application of mathematical modeling for analyzing and forecasting educational processes [2].

This study proposes the use of Canonical Correlation Analysis (CCA) as a research method [3]. CCA is a powerful tool for examining the relationships between two sets of variables and has broad applicability in the educational context. It can be utilized to analyze the interdependence of student performance across various disciplines with factors influencing the educational process. Additionally, CCA can evaluate the effectiveness of different teaching methods by comparing their impact on performance indicators. All these aspects contribute to informed decision-making for improving educational programs and policies.

Object of the Study: The process of analyzing student performance.

Subject of the Study: Algorithms for evaluating the dependencies of student performance across different discipline groups based on Canonical Correlation Analysis (CCA).

Research Methods: The study is based on the application of Canonical Correlation Analysis.

Purpose of the Study: To investigate the dependencies between disciplines for assessing student performance using Canonical Correlation Analysis.

Analysis of Publications Related to Research Methods

Canonical analysis is widely applied in various fields, including education, psychology, economics, biology, and security, to study the relationships between two sets of variables and uncover multidimensional dependencies. In [4], the authors applied canonical analysis in the construction industry to identify factors that have a greater or lesser impact on the annual economic performance of construction enterprises.

In article [5], canonical analysis was used to determine key relationships between the leading sustainability rankings of countries and the most common socio-economic indicators worldwide, resulting in the development of an optimal canonical model.

Shafto M.G., Degani A., & Kirlik A. [6] employed Canonical Correlation Analysis (CCA) to study the interrelations between the operational environment and pilots' actions and reactions during flights. This method revealed patterns and operational dependencies between independent variables (such as environmental conditions) and dependent variables (pilots' decisions and control settings) in the context of automated flight control systems. This approach is instrumental in identifying critical multidimensional dependencies, enhancing operational safety, and improving training methods.

Another study [7] utilized canonical analysis to identify the key factors influencing the performance of agricultural enterprises at various stages of their life cycle, particularly during the stagnation phase. This enabled the identification of negative impacts from internal and external environments and the development of measures to mitigate them at early stages.

In [8], canonical analysis was applied to uncover the relationships between cybersecurity indicators and countries' capabilities to counter financial crimes. The method identified a strong correlation between these groups of indicators, establishing a causal relationship where the level of financial monitoring serves as a cause and the level of cybersecurity as a consequence.

Article [9] proposed an adaptive method of canonical correlation analysis tailored for working with microbiome data and other high-dimensional datasets. This method considers the compositional nature of microbiome data and integrates structural information, such as taxonomic groups among bacterial taxa. The results demonstrate improved selectivity of relevant taxa and increased analytical accuracy compared to existing approaches.

In [10], the use of canonical analysis for studying relationships between different types of variable sets, particularly in ecological research, was reviewed. The study highlighted the possibilities and limitations of canonical analysis as a tool for simplifying the correlation structure between variables. Examples from plant ecology were provided to illustrate its practical contributions to achieving ecological goals.

In [11], sparse canonical correlation analysis (SCCA) was proposed to investigate relationships between single nucleotide polymorphisms and cortical thickness in Alzheimer's patients. This approach helps identify genetic markers that potentially influence neurodegeneration and cognitive function.

The results of canonical analysis in [12] reveal a strong positive relationship between internal competition, trade openness, and entrepreneurial culture, with a Pearson correlation coefficient of 0.857. Regression analysis showed that factors related to internal competition have a significantly greater impact on shaping entrepreneurial culture than those related to trade openness.

Finally, in [13], canonical analysis was used to evaluate relationships between key determinants of economic security and globally recognized indicators of economic and national security in the context of sustainable development. This study identified primary dependencies between indicators such as the Human Development Index (HDI) and the Global Peace Index (GPI), along with other macroeconomic metrics outlined in the research.

Presentation of the Main Material

Canonical Correlation Analysis (CCA) is a statistical method that allows for the investigation of the relationship between two sets of variables. The core idea is to simplify the correlation structure between these sets of variables into the simplest possible form by constructing pairs of canonical variables [3].

Let us introduce the notation:

X – the matrix of independent variables (dimension $n \times p$), where n is the number of observations, and p is the number of variables in the set.

Y – the matrix of dependent variables (dimension $n \times q$), where q is the number of variables in the set.

u=**Xa** – the canonical variable for the **X** group, which is a linear combination of variables from the **X** set.

v=**Yb** – the canonical variable for the **Y** group, which is a linear combination of variables from the **Y** set.

Canonical coefficients **a** and **b** – the weight vectors of the canonical variables that need to be determined.

Algorithm of the Method:

1. Computation of Covariance Matrices:

Covariance matrix for **X**: $\mathbf{S}_{XX} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X}$.

Covariance matrix for **Y**: $\mathbf{S}_{YY} = \frac{1}{n-1}\mathbf{Y}^T\mathbf{Y}$.

Cross-covariance matrix between **X** and **Y**: $\mathbf{S}_{XY} = \frac{1}{n-1}\mathbf{X}^T\mathbf{Y}$.

2. Finding such vectors **a** and **b** that maximize the correlation between **u** and **v**:

$$\rho = \frac{\mathbf{a}^T\mathbf{S}_{XY}\mathbf{b}}{\sqrt{\left(\mathbf{a}^T\mathbf{S}_{XX}\mathbf{a}\right)\left(\mathbf{b}^T\mathbf{S}_{YY}\mathbf{b}\right)}}.$$

3. Finding the eigenvalues of vectors **a** and **b**. To find the optimal **a** and **b**, the following equations are solved:

$$\left(\mathbf{S}_{XX}^{-1}\mathbf{S}_{XY}\mathbf{S}_{YY}^{-1}\mathbf{S}_{YX}\right)\mathbf{a} = \boldsymbol{\lambda}\mathbf{a},$$

$$\left(\mathbf{S}_{YY}^{-1}\mathbf{S}_{YX}\mathbf{S}_{XX}^{-1}\mathbf{S}_{XY}\right)\mathbf{b} = \boldsymbol{\lambda}\mathbf{b},$$

where $\boldsymbol{\lambda}$ represents the squared canonical correlations ρ^2.

4. Computation of Canonical Variable:

The canonical variables are calculated as:

$$\mathbf{u} = \mathbf{Xa},\ \mathbf{v} = \mathbf{Yb}.$$

5. Significance testing: The significance of canonical correlations is verified using the Wilks' Lambda test or other statistical tests.

Experimental Research

The research was conducted based on the average grades from 18 academic disciplines during 2015–2024 for undergraduate students majoring in "Computer Engineering" at the Faculty of Computer Information Technologies of the West Ukrainian National University (Tables 1–2).

Table 1

Average Grades of Disciplines by Groups: Mathematical, Programming and Algorithms, Applied Software, and Technologies

Year	Average Grades of Disciplines in Group 1				Average Grades of Disciplines in Group 2			Average Grades of Disciplines in Group 5	
	HM	DM	PTMS	P	FP	SP	NP	ASCSN	CL
2015	68,78	73,25	70	76,68	70,1	70,4	68	74,42	77,76
2016	73	72,4	71	78,41	68,57	74,33	67,9	76	74,83
2017	72,2	75,38	75,3	64,84	76,56	75,68	70,2	72,1	79
2018	75,6	74	71,8	69,95	76,38	71,58	70,52	71	70,33
2019	61,88	74,53	73,52	71,6	75,01	74,57	70,55	73,33	72,43
2020	73,32	77,5	71,74	74,39	76,98	70,27	73,29	73,24	74,5
2021	73,87	75,82	71,05	81,07	81,61	72,89	79,82	78,4	79,5
2022	74,18	76,52	72,96	77,88	78,01	75,5	76,1	78,13	75,95
2023	67,95	66,69	65,8	70,16	72,61	73,89	73,47	71,42	61,17
2024	62,45	63,91	63,94	63,46	73,59	65,89	66,47	60,8	61,78

To determine models for studying the relationships between performance in various groups of disciplines, the disciplines were categorized based on their orientation and the nature of learning outcomes into the following groups: mathematical disciplines, programming and algorithms, system design, networks and distributed systems, applied software and technologies, and economic and managerial disciplines.

The group of mathematical disciplines (Group 1) included:

– Higher Mathematics (HM);

– Discrete Mathematics (DM);

– Probability Theory and Mathematical Statistics (PTMS);

– Physics (P).
The group of programming and algorithms (Group 2) included:
– Fundamentals of Programming (FP);
– System Programming (SP);
– Network Programming (NP).
The group of system design (Group 3) included:
– Computer-Aided Design Systems in Computer Systems and Networks (CADCSN);
– Design of Computer Systems Based on Graphics Processors (DCSGP);
– Technologies for Designing Computer Systems (TDCS);
– Computer Circuitry (CC).
The group of networks and distributed systems (Group 4) included:
– Computer Networks (CN);
– Design and Administration of Computer Networks (DACN);
– Systems for Processing Distributed Databases (SPDD).
The group of applied software and technologies (Group 5) included:
– Applied Software for Computer Systems and Networks (ASCSN);
– Computer Logic (CL).
The group of economic and managerial disciplines (Group 6) included:
– Project Economics in Computer Engineering (PECE);
– Qualification Work (QW).

Table 2

Average Grades of Disciplines in the Groups: System Design, Networks and Distributed Systems, Economic and Managerial Disciplines

Year	Average Grades of Disciplines in Group3				Average Grades of Disciplines in Group 4			Average Grades of Disciplines in Group 6	
	CADCSN	DCSGP	TDCS	CC	CN	DACN	SPDD	PECE	QW
2015	71,5	71,88	73,63	68,79	72,83	76,07	76,76	77,69	79,5
2016	60	72,83	73,3	75,17	70,47	74	75,2	67	82,18
2017	78,5	71,74	76,6	75,34	76,17	75,63	74,97	77,89	75
2018	71,97	72,47	75,55	72,93	72,01	72,93	72,09	67,86	79,68
2019	72,73	74,7	77,94	74,31	70,41	72,53	72	69,88	79,72
2020	74,09	73,62	76,24	72,7	75,83	72,82	68,84	71,23	78,7
2021	72,55	75,58	74,38	74,74	73,7	73,25	64,23	65,31	78,74
2022	83,24	69,45	69,38	67,77	75,91	76,18	69,23	69	82,67
2023	70,24	73,96	73	64,67	70,46	70,86	65,29	71	80,2
2024	60,81	75,58	76,06	67,87	65,67	71,12	61,65	67,85	79,88

Models for Studying Relationships Between Performance in Different Subject Groups

The relationships between average grades in different subject groups were studied. We will examine in greater detail how performance in mathematical disciplines influences performance in other subject groups. The research was conducted using the STATISTICA statistical analysis software package.

Figure 1 shows the results of canonical correlation analysis (CCA) for the datasets: Programming and Algorithms as the first (left) dataset and Mathematical Disciplines as the second (right) dataset.

N=10	Canonical Analysis Summary Canonical R: ,99611 Chi?(12)=31,206 p=,00184	
	Left Set	Right Set
No. of variables	3	4
Variance extracted	100,000%	88,0223%
Total redundancy	51,0450%	48,2991%
Variables: 1	FP	HM
2	SP	DM
3	NP	PTMS
4		P

Fig. 1. Overall Results of Canonical Analysis

The data analysis (Fig. 1) shows that the canonical correlation coefficient R equals 0.99611, indicating a very strong relationship between programming disciplines and mathematical disciplines. The results of the canonical analysis are statistically significant ($p=0.00184<0.01$) at a significance level of 0.01. The left group explains 100% of its own variance, while the right group explains 88.0223%. This indicates that the set of programming variables is fully interrelated, whereas the mathematical set exhibits some level of residual variation.

The overall redundancy indicates that the left group has 51.0450% redundancy, meaning half of its variance is explained by the right group. For the right group, this measure is 48.2991%, suggesting that nearly half of the variance in mathematical disciplines is explained by programming disciplines.

This demonstrates that the constructed canonical model has sufficiently high accuracy, implying that students' performance in programming disciplines largely depends on their results in mathematical disciplines.

The following characteristic roots were obtained (Fig. 2).

	Eigenvalues (Лист1 in dla_stat_		
Root	**Root 1**	Root 2	Root 3
Value	0,992228	0,600522	0,372722

Fig. 2. Table of Characteristic Roots

Their significance was tested using $\chi 2$ (Fig. 3).

	Chi-Square Tests with Successive Roots Removed (Лист					
Root Removed	**Canonicl R**	Canonicl R-sqr.	Chi-sqr.	df	p	Lambda Prime
0	0,996107	0,992228	31,20620	12	0,001839	0,001947
1	0,774934	0,600522	6,91981	6	0,328356	0,250584
2	0,610510	0,372722	2,33183	2	0,311651	0,627278

Fig. 3. Significance Testing of Characteristic Roots

It is evident that the first root is statistically significant ($p<0.01$).

Let us analyze the first pair of canonical variables, which have the strongest correlation: $R=0.996107$. Consider the factor loadings of the left set of variables (Fig. 4). System Programming has the highest factor coefficient (0.603053), indicating its largest contribution to this canonical root. The variables FP (0.032684) and NP (-0.238365) have significantly smaller coefficients, indicating a weak and negative contribution to the first root. Thus, among the programming disciplines, System Programming is the most important factor in explaining the relationship with mathematical disciplines.

	Factor Structure, left set (Лист1 i		
Variable	**Root 1**	Root 2	Root 3
FP	0,032684	-0,342685	0,938882
SP	0,603053	0,696331	0,389167
NP	-0,238365	0,255057	0,937085

Fig. 4. Factor Structure of the Left Set of Variables

The first canonical root explains only 14.05% of the total variation of the variables in the left set (Fig. 5). The redundancy of Root 1 is 13.94%, indicating that only a small portion of the variation in programming variables is explained by the variables in mathematical disciplines.

	Variance Extracted (Proportions), left set			
Factor	**Variance extractd**	Reddncy.		
Root 1	0,140519	0,139427		
Root 2	0,222455	0,133589		
Root 3	0,637026	0,237434		

Fig. 5. Proportions of Explained Variance for the Left Set of Variables

Let us consider the factor loadings of the right set of variables (Fig. 6). Discrete Mathematics has the highest factor coefficient (0.711733), indicating its largest contribution to the first canonical root. The variables PTMS (0.334895) and Physics (-0.357688) have moderate contributions, with PTMS being positive and Physics negative. Higher Mathematics shows a weak connection. Thus, among the mathematical disciplines, Discrete Mathematics is the most important factor in explaining the relationship with programming disciplines.

Variable	Factor Structure, right set (Лист		
	Root 1	Root 2	Root 3
HM	0,038938	0,284932	0,726547
PTMS	0,334895	0,175317	0,908855
DM	0,711733	0,175743	0,662515
P	-0,357688	0,736657	0,542613

Fig. 6. Factor Structure of the Right Set of Variables

The first canonical root explains only 18.7% of the total variation of the variables in the right set (Fig. 7). The redundancy of Root 1 is 18.56%, indicating that only a small portion of the variation in mathematical discipline variables is explained by the programming variables.

Variable	Variance Extracted (Proportions), right set	
	Variance extractd	Reddncy.
Root 1	0,187044	0,185590
Root 2	0,171368	0,102910
Root 3	0,521811	0,194491

Fig. 7. Proportions of Explained Variance for the Right Set of Variables

To calculate the values of canonical variables, we use the canonical weights of the left and right sets (Fig. 8). Canonical variables represent latent indicators:

$$v = 1,017FP + 1,031SP - 1,448NP\,,$$
$$u = 0,048HM - 1,379PTMS + 1,971DM - 0,160P$$

are statistically significant at the 0.01 level and reliable. The standardized variable *v* can be considered as a quantitative assessment of programming and algorithm disciplines, while *u* represents a quantitative assessment of mathematical disciplines.

Variable	Canonical Weights, left set (Ли		
	Root 1	Root 2	Root 3
FP	1,01745	-1,20983	0,588099
SP	1,03093	0,46788	0,134886
NP	-1,44754	1,01784	0,421895

a)

Variable	Canonical Weights, right set (Л		
	Root 1	Root 2	Root 3
HM	0,04769	0,20787	0,19107
PTMS	-1,37881	-2,29358	1,85226
DM	1,97057	1,77157	-1,01856
P	-0,16042	1,40029	-0,27173

b)

Fig. 8. Canonical Weights: a) Left Set; b) Right Set of Variables

A further review of models determining the relationship between performance in different subject groups will be examined less thoroughly. The summary of the CCA results for various sets of disciplines with the group of mathematical disciplines is presented in Table 3.

Table 3

Overall Results of Canonical Analysis of Mathematical Disciplines with Groups 3–6

Indicator	**Canonical Analysis Summary of Mathematical Disciplines with Groups of Disciplines**							
	System Design		**Networks and Distributed Systems**		**Applied Software and Technologies**		**Economic and Managerial Disciplines**	
	Canonical R=0,997 Chi²=34, 565 p=0,00457		Canonical R=0,917 Chi²=12,256 p=0,42541		Canonical R=0,976 Chi²=18,837 p=0,01578		Canonical R=0,712 Chi²=4,1701 p=0,84145	
Variance extracted	100%	100%	100%	89,22%	100%	80,6862%	100%	28,1586%
Total redundancy	48,7521%	77,0083%	61,4954%	54,5611%	85,9098%	63,6098%	36,8768%	9,41997%
Variables: 1	CADCSN	HM	CN	HM	ASCSN	HM	PECE	HM
2	DCSGP	DM	DACN	DM	CL	DM	QW	DM
3	TDCS	PTMS	SPDD	PTMS		PTMS		PTMS
4	CC	P		P		P		P

As seen in Table 3, a strong relationship between mathematical disciplines is observed with the groups of System Design and Applied Software and Technologies. The results of the canonical analysis are statistically significant ($p<0.05$) at a significance level of 0.05. A weaker relationship is noted with the group of Networks and Distributed Systems (R=0.917, p=0.42541) and an even weaker one (R=0.712, p=0.84145) with the group of Economic and Managerial Disciplines, which are studied in the 4th year. The latter models are statistically insignificant at the 0.05 significance level.

The summary of the CCA results for various sets of disciplines with the group of programming disciplines is presented in Table 4.

Table 4

Overall Results of Canonical Analysis of Programming and Algorithms Disciplines with Groups 3–6

Indicator	**Canonical Analysis Summary of Programming and Algorithms Disciplines with Groups of Disciplines**							
	System Design		**Networks and Distributed Systems**		**Applied Software and Technologies**		**Economic and Managerial Disciplines**	
	Canonical R=0,978 Chi^2=23,419, p=0,02441		Canonical R=0,908 Chi^2=12,280 p=0,19803		Canonical R=0,954 Chi^2=15,814 p=0,01481		Canonical R=0,707 Chi^2=5,3636 p=0,49811	
Variance extracted	68,8247%	100%	100%	100%	100%	60,2144%	100%	51,5000%
Total redundancy	39,9538%	75,1087%	41,5756%	59,0987%	48,9096%	35,5008%	26,3223%	19,2363%
Variables: 1	CADCSN	FP	CN	FP	ASCSN	FP	PECE	FP
2	DCSGP	SP	DACN	SP	CL	SP	QW	SP
3	TDCS	NP	SPDD	NP		NP		NP
4	CC							

As seen in Table 4, a strong relationship between the disciplines of Programming and Algorithms is observed with the groups of System Design and Applied Software and Technologies. The results of the canonical analysis are statistically significant ($p<0.05$) at a significance level of 0.05. A weaker and statistically insignificant relationship at the 0.05 significance level is noted with the group of Networks and Distributed Systems (R=0.908, p=0.19803) and an even weaker one (R=0.707, p=0.49811) with the group of Economic and Managerial Disciplines.

The summary of the CCA results for various sets of disciplines with the group of System Design disciplines is presented in Table 5.

Table 5

Overall Results of Canonical Analysis of System Design Disciplines with Groups 4–6

Indicator	**Canonical Analysis Summary of System Design Disciplines with Groups of Disciplines**					
	Networks and Distributed Systems		**Applied Software and Technologies**		**Economic and Managerial Disciplines**	
	Canonical R=0,932 Chi^2=17,558, p=0,12986		Canonical R=0,907 Chi^2=12,775 p=0,11988		Canonical R=0,962 Chi^2=17,823 p=0,02343	
Variance extracted	100%	82,4299%	100%	52,1690%	100%	81,5893%
Total redundancy	77,7908%	49,4197%	77,3767%	32,7603%	57,5705%	34,0825%
Variables: 1	CN	CADCSN	ASCSN	CADCSN	PECE	CADCSN
2	DACN	DCSGP	CL	DCSGP	QW	DCSGP
3	SPDD	TDCS		TDCS		TDCS
4		CC		CC		CC

As seen in Table 5, a strong relationship between System Design disciplines is observed with the group of Economic and Managerial Disciplines. The results of the canonical analysis are statistically significant ($p<0.05$) at a significance level of 0.05. A weaker relationship is noted with the group of Networks and Distributed Systems (R=0.932, p=0.12986) and Applied Software and Technologies (R=0.907, p=0.11988). These latter models are statistically insignificant at the 0.05 significance level.

The summary of the CCA results for the group of Networks and Distributed Systems with the groups of Applied Software and Technologies, Economic and Managerial Disciplines, as well as the Applied Software and Technologies group with the group of Economic and Managerial Disciplines, is presented in Table 6.

As seen in Table 6, a weak relationship is observed with the group of Applied Software and Technologies disciplines (R=0.861, p=0.19232) and Economic and Managerial Disciplines (R=0.655, p=0.62011). A weak relationship is also noted between the group of Applied Software and Technologies and the group of Economic and Managerial Disciplines (R=0.633, p=0.489588). The canonical models are statistically insignificant at the 0.05 significance level.

Table 6

Overall Results of Canonical Analysis of Networks and Distributed Systems Disciplines with Groups 5–6 and Applied Software and Technologies with Group 6

Indicator	Canonical Analysis Summary of Networks and Distributed Systems Disciplines with Groups of Disciplines				Canonical Analysis Summary of Applied Software and Technologies Disciplines with Groups of Disciplines	
	Applied Software and Technologies		Economic and Managerial Disciplines		Economic and Managerial Disciplines	
	Canonical R=0,861 Chi^2=8,6817, p=0,19232		Canonical R=0,655 Chi^2=4,4194 p=0,62011		Canonical R=0,633 Chi^2=3,4237 p=0,48958	
Variance extracted	100%	84,0103%	100%	79,1981%	100%	100%
Total redundancy	62,3010%	53,0684%	25,7960%	28,5160%	28,9039%	6,67451%
Variables: 1	ASCSN	CN	PECE	CN	PECE	ASCSN
2	CL	DACN	QW	DACN	QW	CL
3		SPDD		SPDD		

Model for Assessing the Impact of Individual Subjects on Overall Performance

To determine which subjects have the greatest impact on the overall average grade, we conducted an analysis of the dependency between performance in individual disciplines and overall performance, calculated as the arithmetic mean of all disciplines in Tables 1-2 for each year. Table 7 presents the results of the CCA for only the significant disciplines at the 0.05 significance level.

Table 7

Results of the canonical analysis of the relationship between academic performance in individual disciplines and overall academic performance

Indicator	Disciplines							
	DM	PTMS	SP	CADCSN	CN	DACN	ASCSN	CL
Canonical R	0,9556	0,883	0,6584	0,7142	0,9023	0,7185	0,8689	0,9022
Chi^2	18,344	11,375	4,262	5,351	12,622	5,447	10,547	12,617
p	0,00002	0,00075	0,03899	0,02072	0,00038	0,01961	0,00117	0,00038
Variance extracted, %	100	100	100	100	100	100	100	100
Total, % redundancy	91,335	78,054	43,347	57,006	81,418	51,629	75,494	81,404

According to Table 7, the greatest impact on the overall average grade from the group of mathematical disciplines is observed in Discrete Mathematics (R=0.9556, p=0.00002), from the group of Programming and Algorithms – System Programming (R=0.6584, p=0.03899), from the group of System Design – Computer-Aided Design Systems in Computer Systems and Networks (R=0.7142, p=0.02072), from the group of Networks and Distributed Systems – Computer Networks (R=0.9023, p=0.00038), and from the group of Applied Software and Technologies – Computer Logic (R=0.9022, p=0.00038).

The high value of % redundancy (91.335) indicates that the results of the overall average grade from all disciplines largely depend on Discrete Mathematics. This confirms the strong relationship between a student's overall performance (average grade) and the results from Discrete Mathematics.

Conclusions

The study demonstrates the potential of Canonical Correlation Analysis (CCA) as a reliable method for exploring relationships between variables in educational contexts. By analyzing student performance in various groups of disciplines, the research reveals significant correlations that can serve as guidelines for the development of future educational strategies and curricula. Specifically, the following points should be noted:

1. The strong canonical correlation between mathematical disciplines and programming suggests that success in fundamental mathematical courses significantly influences student performance in programming.
2. A close relationship between the disciplines of Programming and Algorithms is observed with the groups of System Design and Applied Software and Technologies.
3. A strong relationship between the disciplines of System Design is observed with the group of Economic and Managerial Disciplines, which includes the qualification work.
4. The results highlight the importance of Discrete Mathematics as a critical factor influencing overall academic performance.

The use of CCA provides educators and administrators with a deeper understanding of the interdependent nature of disciplines, allowing data-driven decision-making.

Practical Implications. The conclusions drawn from this study can contribute to targeted interventions and resource allocation to improve student performance in specific disciplines.

Future Research Directions. The application of the findings to enhance academic performance prediction models.

This research provides a foundational basis for the use of advanced statistical methods to uncover significant dependencies in education, paving the way for improvements in learning outcomes and educational strategies.

The study demonstrates the potential of Canonical Correlation Analysis (CCA) as a reliable method for exploring relationships between disciplines and students' academic performance, providing valuable insights for the development of educational strategies and improvement of curricula, which will contribute to enhancing learning effectiveness and optimizing resources.

References

1. Rebukha L Z. et al. Innovative learning technologies in the context of modernization of modern education. Ternopil: WUNU, 2022. 143 p. http://dspace.wunu.edu.ua/handle/316497/48105
2. Prus A. Mathematical modelling in the eduation system and tasks of mathematical modelling. Chapter in book: “Main trends and traditions in mathematics teaching”. Zhytomyr, PP «Ruta», 2024. Pp. 227-258.
3. Johnson, R. A., & Wichern, D. W. Applied Multivariate Statistical Analysis (6th ed.). (2007). Pearson. 773p.
4. Gumenna-Derij M., Khorunzhak N., Poprozman N., Berezka K., Kruchak L. Modeling, accounting and control of formation and use of resources (on the example of the construction industry). Independent Journal of Management & Production (Special Edition ISE, S&P). 2022. Vol. 13, No. 3. P. 123-144. DOI: https://doi.org/10.14807/ijmp.v13i3.1901.
5. Kovalchuk O.Ya., Perih V.M. Multiple analysis of global sustainable development. Scientific Bulletin of KSU. Series «Economic Sciences». 2018. Vol 1, No 29. Pp. 40-46.
6. Shafto M. G., Degani A., Kirlik A. Canonical Correlation Analysis of Data on Human-Automation Interaction. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 1997. 41(1). P. 62-65. DOI:https://doi.org/10.1177/107118139704100116.
7. Volkova N. A. Koliada A. L. Analytical methods of search of efficiency reserves of agricultural enterprises functioning. Ekonomichnyy analiz. 2017. 27(2). Pp. 171-177.
8. Kuzmenko O., Yarovenko H., Radko V. Preliminary analysis of the convergence process of cyber security systems and financial monitoring of countries. Economy and Society. 2021. №32. DOI: https://doi.org/10.32782/2524-0072/2021-32-37.
9. Deng L., Tang Y., Zhang X., Chen J. Structure-adaptive canonical correlation analysis for microbiome multi-omics data. Front. Genet. 2024. 15:1489694. DOI: https://doi.org/10.3389/fgene.2024.1489694.
10. Gittins R. Canonical analysis: a review with applications in ecology. Berlin; New York: Springer-Verlag, 1985. 351 p. DOI: https://doi.org/10.1007/978-3-642-69878-1.
11. Kim B.-H., Seo S. W., Park Y. H., Kim J., Kim H. J., Jang H., Yun J., Kim M., Kim J. P. Clinical application of sparse canonical correlation analysis to detect genetic associations with cortical thickness in Alzheimer’s disease. Front. Neurosci. 2024. 18:1428900. DOI: https://doi.org/10.3389/fnins.2024.1428900.
12. Milan K., Živković J. Domestic Competition, Trade Openness and Entrepreneurial Culture: Canonical Correlation Analysis. South East European Journal of Economics and Business. 2024. Vol. 19, no. 1, Sciendo. Pp. 18-31. DOI: https://doi.org/10.2478/jeb-2024-0002.
13. Kovalchuk O., Berezka K., Masonkova M., Chudyk N., Ukhach V., Pilyukov Y. Statistical Modeling of Determinants Influencing Economic Security in the Context of Sustainable Development and National Security. Proceedings of International Conference on Applied Innovation in IT (Koethen, GERMANY, 7 March 2024), Vol. 12, Is. 1 Pp. 197-203. DOI: https://doi.org/10.25673/115700.

Kateryna Berezka **Катерина Березька**	PhD, Associate Professor, Associate Professor Department of Applied Mathematics West Ukrainian National University https://orcid.org/0000-0002-9632-4004 e-mail: k.berezka@wunu.edu.ua	Кандидат технічних наук, доцент, доцент кафедри прикладної математики, Західноукраїнський національний університет
Oksana Bashutska **Оксана Башуцька**	PhD, Associate Professor, Associate Professor Department of Economic Cybernetics and Informatics West Ukrainian National University https://orcid.org/0000-0002-2445-896X e-mail: oksana.bashutska@wunu.edu.ua	Кандидат економічних наук, доцент, доцент кафедри економічної кібернетики та інформатики, Західноукраїнський національний університет
Nataliya Navolska **Наталія Навольська**	PhD, Associate Professor, Associate Professor Uniwersytety WSB Merito, Wroclaw, Poland https://orcid.org/0000-0003-1375-5990 e-mail: navol.natalya@gmail.com	Кандидат економічних наук, доцент, доцент ВШБ Університет Меріто, Вроцлав, Польща
Vasil Melnychenko **Василь Мельниченко**	senior lecturer Department of Economic Expertise and Land Management, Director of the Research and Educational Center for Automated Support of the Educational Process and Work with the Unified State Electronic Database on Education (USEDE) West Ukrainian National University https://orcid.org/0000-0001-8384-0934 e-mail: v.melnychenko@wunu.edu.ua	старший викладач кафедри економічної експертизи та землевпорядкування, директор ННЦ автоматизованого супроводу освітнього процесу та роботи з ЄДЕБО, Західноукраїнський національний університет