www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

Student elective course selection patterns and satisfaction determinants identified through educational data mining

Serhiy O. Semerikov[1,2]✉, Olha V. Bondarenko[1,3], Pavlo P. Nechypurenko[1,3], Tetiana A. Vakaliuk[2,4] & Iryna S. Mintii[5,6]

Digital transformation in higher education is reshaping how institutions design and deliver their curricula, with a growing emphasis on student agency and personalized learning paths. This study employs educational data mining techniques to analyze student preferences and satisfaction with elective courses at Kryvyi Rih State Pedagogical University in Ukraine. We investigate patterns in course selection, satisfaction determinants, and the effectiveness of the university's individual educational trajectory framework among 1,089 students. Our analysis reveals four distinct student segments with varying preferences and satisfaction profiles. Information availability before selection, alignment with career goals, teaching quality, and course relevance emerge as significant predictors of student satisfaction. We propose a data-driven framework for optimizing elective course systems that incorporates learning analytics, personalized recommendation engines, and enhanced information platforms. This research contributes to understanding how educational technology can better support student agency in curriculum customization while addressing critical issues of accessibility, equity, and educational quality. The findings align with Sustainable Development Goal 4 (Quality Education) by promoting inclusive and personalized educational opportunities that prepare students for future employment challenges.

Higher education institutions worldwide are increasingly embracing personalized learning approaches that accommodate diverse student interests, goals, and learning styles. This transformation is partly driven by the integration of digital technologies and data-informed decision-making processes[1,2]. Elective courses constitute a critical component of this personalization, allowing students to shape individual educational trajectories (IETs) aligned with their professional aspirations and personal interests. In Ukraine, recent higher education reforms have emphasized student agency – the capacity of students to make meaningful choices about their own education – in course selection, mandating that at least 25% of credits be allocated to student-selected courses[3].

The concept of educational personalization has gained prominence as institutions seek to enhance educational quality, student engagement, and learning outcomes in increasingly diverse student populations. Xu et al. [4] emphasize that this shift toward Personal Learning Environments (PLEs) represents a fundamental reconceptualization of higher education, moving from standardized curriculum approaches toward more student-centered, flexible models that acknowledge learner diversity. This transition, however, presents numerous challenges related to curriculum design, technological infrastructure, and pedagogical approaches.

While the principle of student choice in curriculum is widely endorsed, the implementation of effective elective course systems presents numerous challenges. These include ensuring sufficient course variety, providing adequate information for informed decisions, developing efficient registration processes, and maintaining educational quality across diverse offerings. Algarni and Sheldon[5] highlight that course recommender systems face significant challenges, including managing large and dynamic decision spaces and addressing the cold start problem for new students. Furthermore, Cha et al. [6] note that AI-based course recommender systems must be carefully designed to support both the search and evaluation phases of student decision-making.

Educational data mining (EDM) offers powerful tools for gaining insights by uncovering patterns in educational data that can inform policy and practice[7]. Despite growing research on EDM applications in areas

[1]Kryvyi Rih State Pedagogical University, Kryvyi Rih 50086, Ukraine. [2]Institute for Digitalisation of Education of the NAES of Ukraine, Kyiv 04060, Ukraine. [3]Academy of Cognitive and Natural Sciences, Kryvyi Rih 50086, Ukraine. [4]Zhytomyr Polytechnic State University, Zhytomyr 10005, Ukraine. [5]University of Łódź, Łódź 90-136, Poland. [6]Lviv Polytechnic National University, Lviv 79000, Ukraine. ✉email: semerikov@gmail.com

such as learning analytics and student performance prediction, relatively few studies have applied these techniques to understand elective course selection patterns in Eastern European contexts and their relationship to student satisfaction. Cai[8] notes that data mining algorithms in university information systems can help schools better understand students' learning behavior and characteristics, teachers' teaching levels, and curriculum quality, providing scientific decision-making bases for educational improvement.

This study addresses this gap by analyzing survey data collected from students at Kryvyi Rih State Pedagogical University (KSPU) regarding their experiences with the university's elective course system. Using a combination of statistical analysis, machine learning techniques, and text mining, we explore the following research questions:

- RQ1 Are there identifiable patterns in student selection of elective courses, and if so, how do these cluster into distinct student segments?
- RQ2 Which factors significantly predict student satisfaction with elective courses, and do these predictors vary across identified student segments?
- RQ3 Based on identified patterns and satisfaction predictors, what design principles should guide educational technology interventions for elective course selection?
- RQ4 What framework can integrate student segments and satisfaction predictors to support personalized course recommendation?

The remainder of this paper is organized as follows. The next section reviews relevant literature on individual educational trajectories, student decision-making in course selection, educational data mining applications, and approaches to measuring student satisfaction. The Methodology section describes the research context, survey instrument, data collection procedures, and analytical techniques. Results presents the quantitative and qualitative findings, including student segmentation, satisfaction predictors, and the derived framework for personalized course selection support. The Discussion section interprets these findings in relation to existing research and considers theoretical and practical implications. The paper concludes with recommendations for higher education institutions and directions for future research.

Literature review

Individual educational trajectories in higher education

Individual educational trajectories (IETs) represent a significant paradigm shift in higher education pedagogy and curriculum design. These personalized learning paths reflect students' unique educational needs, interests, and goals while acknowledging diverse learning approaches[4]. In higher education, IETs typically manifest through elective course systems that allow students to customize portions of their curriculum. Personalized learning strategies most commonly include tailoring instruction to individual needs, providing one-on-one support, allowing for student choice and autonomy, flexible pacing, and learner agency[9].

When effectively implemented, personalized learning strategies correlate positively with student engagement, motivation, and academic achievement[9]. However, significant variation exists in implementation quality across institutions, with many failing to provide the necessary support structures for informed decision-making. This inconsistency highlights the need for systematic approaches to supporting student choice.

The technological dimension of IETs has evolved considerably in recent years. A comprehensive ICT framework for Personal Learning Environments (PLE-ICT) in higher education encompasses five dimensions: ICT hardware, ICT software, ICT services, project development, and support teams[10]. This framework acknowledges that technology infrastructure serves as a critical enabler of personalized learning environments capable of enhancing student learning experiences and outcomes.

Student decision-making in course selection

Research on student course selection reveals complex decision-making processes influenced by multiple factors. Students consider only a narrow range of available courses when making selections – on average nine courses for their first fall term, representing less than 2% of available options at a private research university[11]. Notably, these initially considered courses predict which academic major students declare two years later, suggesting that early course exploration patterns have long-term implications for academic trajectories. This limited exploration of options indicates that course selection processes may benefit from technological support that expands students' awareness of relevant alternatives.

Collaborative recommendation systems employing data mining techniques can discover patterns between courses and improve recommendation quality[12]. Clustering students into similar groups based on their course selections plays a vital role in generating high-quality association rules compared to using entire datasets. This segmentation approach potentially overcomes the limitations of narrow course consideration by connecting students with options that similar peers have found valuable.

Recent advances have explored innovative approaches to course recommendation. Large Language Models (LLMs) show promise for assisting with course recommendation through models that utilize ChatGPT output to form rules capturing relationships among courses[13]. This approach effectively combines the rich knowledge base of LLMs with information about past students' enrollment history to generate more personalized recommendations.

In the Ukrainian context specifically, professional relevance and opportunities for skill development are particularly important to students, while administrative barriers often hinder optimal selections[14]. These findings align with broader research on student decision-making factors but highlight the importance of context-specific understanding of selection processes in different educational systems.

Educational data mining in higher education

Educational data mining (EDM) applies computational approaches to analyze educational data, detect patterns, and generate insights that can improve educational processes and outcomes[15]. The field has evolved significantly in recent years, with increasing applications in higher education decision-making.

Matrix factorization-based collaborative filtering frameworks offer one approach for course recommendations in higher education[16]. Systems analyzing data from hundreds of students can provide course recommendations based on individual preferences and past academic performance, demonstrating how machine learning techniques enable personalized course suggestions derived from historical data patterns.

Machine learning-based course enrollment recommender systems can exploit multiple types of information, including student course enrollment history and contextual features such as prerequisite restrictions, course meeting times, instructional methods, and course instructors[17]. Experimental results indicate that such recommended courses are highly relevant while still providing students with ample options, balancing personalization with breadth of choice.

The integration of learning analytics with EDM has created new opportunities for understanding and optimizing educational systems. Six key categories of learning analytics applications have been identified: sentiment analysis, questionnaire analysis, engagement analysis, topic classification, predictive modeling, and performance analysis[1]. While descriptive analysis remains the most common approach, advanced techniques such as machine learning, artificial intelligence, and social network analysis are becoming more prominent in educational research.

Data protection and privacy regulations play an important role in utilizing data mining methods in educational environments[18]. The growing volume of educational data generated and analyzed necessitates addressing privacy concerns and ensuring data security. A data ethics framework for Artificial Intelligence in Education (AIED) that maps and analyzes current policies and guidelines can promote the design, development, and implementation of ethical and trustworthy AIED systems[19].

Data mining algorithms in university information systems can help institutions better understand students' learning behavior and characteristics, teachers' teaching levels, and curriculum quality[8]. These insights provide scientific decision-making bases for educational improvement, enabling evidence-based reforms to curriculum and instruction.

Measuring and improving student satisfaction with elective systems

Student satisfaction with elective course systems has been measured using various frameworks and approaches. Personalized hybrid course recommendation systems that integrate multiple recommendation methods – item-based, user-based, and content-based filtering – can optimize parameter weights using genetic algorithms to enhance prediction accuracy[20]. Evaluation of such systems demonstrates that the majority of students (83.60%) show greater interest in courses appearing at the top of recommendation lists, suggesting improved student engagement with courses aligned to their preferences.

E-learning recommendation systems positively impact student autonomy, engagement, and learning effectiveness[21]. Structural equation model analysis involving 602 students revealed particularly strong effects on student engagement, with a path coefficient of 0.622. These findings underscore the potential of well-designed recommendation systems to enhance multiple dimensions of the student experience.

Machine learning applications can improve course selection processes in various contexts. Analysis of historical data using algorithms such as J48 has achieved accuracy rates of 96.37% in predicting probable grades based on student records[22]. Such predictive capabilities enable students to make more informed decisions about course selections based on likely outcomes.

While these studies provide valuable insights, most have been conducted in North American or Western European contexts. Cultural differences between Eastern and Western approaches to education can significantly influence how students engage with educational systems[23]. This variation suggests the need for context-specific research on elective course systems in Eastern European settings, where institutional structures and student expectations may differ substantially.

Educational technology for enhanced course selection

Emerging educational technologies are reshaping how students interact with course selection systems. Artificial intelligence and machine learning technologies can personalize individual educational trajectories for students in engineering and technical specialties, providing extensive student data segments that can be monitored and supported regularly[24]. This capability enables institutions to offer differentiated support based on student characteristics and needs.

Visualization tools play an increasingly important role in helping students understand course relationships and make informed decisions. Familiar visualization algorithms like bar charts prove most effective in helping students and instructors quickly identify information, while more complex visualizations like sankey charts, treemaps, and violin charts are less well-known but potentially valuable for certain analyses[25]. The choice of visualization technique should match both the data complexity and user familiarity.

Learning analytics dashboards (LADs) have gained momentum as institutions seek to improve data-driven decision-making. Participatory design approaches to developing LADs for higher education have identified four crucial factors: who has access to data, the importance of time, learning analytics' role in helping students transition to university life, and discipline-specific considerations[26]. These factors highlight the need for careful attention to context and stakeholder needs in dashboard design.

Implementation of these technologies comes with challenges related to the digital divide and accessibility. Digital accessibility in computer science education requires attention to accessible educational resources, particularly given the rapid shift to online learning[27]. Multiple factors influence the digital divide, including socioeconomic status, gender, geography, age, ethnicity, parent knowledge and educational level, school type, and language barriers[28]. Addressing these factors is essential for ensuring that educational technology innovations benefit all students equitably.

Methodology

Research context

This study was conducted at Kryvyi Rih State Pedagogical University (KSPU) in Ukraine, a public institution with approximately 3,500 students across eight faculties. The university implemented a comprehensive individual educational trajectory (IET) framework in 2022, which requires that at least 25% of credits in each program be allocated to student-selected courses.

The KSPU elective course system operates according to the Regulations on Individual Educational Trajectory of Higher Education Students updated in 2024[29]. Key elements include:

- A university-wide catalog of elective courses updated annually through a systematic process involving departmental proposals, faculty council reviews, and central approval
- Both general education and profession-specific elective options
- An electronic registration system for course selection
- Minimum enrollment requirements (20 students for general courses, 12 for professional courses at the bachelor's level; 15 and 5 respectively at the master's level)
- Selection periods in April-May for continuing students and September-October for new students

The study was approved by the KSPU Ethics Committee (approval number: 49/2024, date: 16/12/2024), and all participants provided informed consent before completing the survey, in accordance with ethical guidelines for educational data mining research[30].

Research design

We employed a mixed-methods approach, combining quantitative and qualitative techniques to provide a comprehensive understanding of student experiences with elective courses. Following the methodological trends identified by Ketsman et al. [31], we designed our study to maximize methodological rigor while acknowledging the complexity of student decision-making processes.

Drawing on recommendations by Heinrich[32] for authentic qualitative research in educational technology, we explicitly addressed our research design and researcher beliefs. Our approach was primarily post-positivist, recognizing the complexity of student decision-making while seeking patterns that could inform institutional practice. We combined descriptive statistics, cluster analysis, regression modeling, and thematic analysis to triangulate findings across different data sources.

Survey instrument

The survey instrument was developed based on a review of relevant literature and input from faculty and student representatives. Following the framework by Jones et al. [33] for understanding student perspectives on educational data analytics, we ensured that privacy considerations and transparency were central to our data collection approach.

The final questionnaire included 14 items across the following sections:

1. Demographics (2 items): Degree program, educational program/major
2. Course selection awareness (3 items): Familiarity with selection procedures, attitudes toward selection process, open-ended clarification
3. Selection options (3 items): Satisfaction with quantity of electives, factors influencing selection, open-ended clarification
4. Selection process (3 items): Review of syllabi before selection, evaluation of selected courses, open-ended clarification
5. Course evaluation (2 items): Most interesting/useful courses, likelihood of selecting same courses again
6. Improvement suggestions (1 item): Recommendations for course selection policies and procedures

The complete survey instrument with all items and response options is provided as the supplementary material.

Data collection

The survey was administered electronically through the university's learning management system between October and November 2024. All currently enrolled students who had completed at least one semester and had experience selecting elective courses were invited to participate. Multiple reminder emails were sent to maximize participation.

A total of 1,089 valid responses were received, representing a 31% response rate. The sample included students from all faculties and across all years of study.

To promote open science and research reproducibility, the anonymized survey data has been converted to a FAIR (Findable, Accessible, Interoperable, Reusable) dataset and deposited in the Zenodo repository[34]. The dataset contains 1,089 valid responses with 15 variables and is available under a Creative Commons Attribution 4.0 International License.

Data analysis

We employed a mixed-methods approach for data analysis, combining statistical techniques, machine learning algorithms, and qualitative text analysis. The analysis was conducted using Python with libraries including pandas, scikit-learn, NLTK, and matplotlib. Following the systematic review framework by Ncube and Ngulube[35], we ensured that our data analysis methods aligned with ethical considerations for educational data analytics.

Descriptive and correlational analysis

Initial exploration of the dataset involved:

- Frequency distributions of demographic variables and selection patterns
- Measures of central tendency (mean, median) and dispersion (standard deviation, interquartile range) for satisfaction ratings
- Pearson correlation coefficients to assess bivariate relationships between continuous variables
- Chi-square tests for associations between categorical variables
- Visualization of key relationships through bar charts, box plots, and correlation heatmaps

Cluster analysis

To identify distinct student segments based on selection patterns and satisfaction profiles, we applied K-means clustering with the following specifications:

- *Input variables*: Seven satisfaction dimensions (course content quality, teaching quality, relevance to career goals, alignment with expectations, variety of offerings, information availability, selection process ease) standardized using z-score normalization
- *Initialization*: K-means++ algorithm to optimize initial centroid placement and improve convergence[36]
- *Number of clusters*: We tested solutions with $k = 2$ to $k = 8$ clusters, evaluating each using:
  - Elbow method: Plotting within-cluster sum of squares (WCSS) against number of clusters to identify the point of diminishing returns
  - Silhouette analysis: Computing average silhouette width for each solution, with higher values indicating better-defined clusters
  - Practical interpretability: Assessing whether cluster profiles represented meaningfully distinct student types
- *Convergence*: Maximum 300 iterations with tolerance of 10^{-4} for centroid movement
- *Validation*: The four-cluster solution was selected based on silhouette coefficient (0.47), elbow plot inflection, and interpretability of resulting segments
- *Stability*: Ten random initializations were performed, with the solution achieving lowest WCSS retained

Following the systematic review of recommendation systems by Algarni and Sheldon[5], we compared different clustering approaches and determined that K-means provided the most appropriate balance of interpretability and performance for our dataset size and structure.

Multiple regression analysis

To identify predictors of overall satisfaction with the elective course system, we conducted hierarchical multiple regression with the following specifications:

- *Dependent variable*: Overall satisfaction with the elective course system (5-point Likert scale treated as continuous)
- *Independent variables*: Seven satisfaction dimension ratings entered simultaneously
- *Assumption testing*:
  - *Multicollinearity*: Variance Inflation Factors (VIF) calculated for all predictors; all VIF values were below 3.0, indicating acceptable collinearity levels
  - *Normality of residuals*: Shapiro-Wilk test on standardized residuals ($W = 0.987$, $p = 0.062$) and visual inspection of Q-Q plots confirmed approximate normality
  - *Homoscedasticity*: Breusch-Pagan test ($\chi^2 = 4.23$, $p = 0.121$) indicated homogeneous variance of residuals
  - *Independence*: Durbin-Watson statistic (1.94) suggested no substantial autocorrelation
  - *Linearity*: Inspection of residual plots confirmed linear relationships between predictors and outcome
- *Model evaluation*: Adjusted R^2, F-statistic, and individual predictor significance ($\alpha = 0.05$)

Qualitative text analysis

Open-ended survey responses were analyzed using thematic analysis procedures:

- *Preprocessing*: Ukrainian-language responses were processed using specialized NLP tools including the pymorphy2 library for lemmatization and a custom Ukrainian stopword list. Text was tokenized, lemmatized, and cleaned of punctuation and numeric characters.
- *Word frequency analysis*: Term frequencies were calculated to identify commonly mentioned concepts and concerns.

- *Thematic coding*: Two researchers independently coded a random sample of 200 responses to develop an initial coding framework. Inter-rater reliability was assessed using Cohen's kappa (κ = 0.78), indicating substantial agreement. Disagreements were resolved through discussion, and the refined coding scheme was applied to all responses.
- *Sentiment classification*: A custom Ukrainian sentiment lexicon adapted from prior work[37] was used to classify response sentiment as positive, negative, or neutral.
- *Theme quantification*: Coded themes were quantified to identify prevalence patterns across the sample.

For visualization of cluster results, principal component analysis was applied to reduce the seven-dimensional satisfaction space to two dimensions for plotting purposes. The two principal components explained 62% of variance and were interpreted as "Course content focus" (PC1) and "Process efficiency focus" (PC2) based on variable loadings.

Ethical considerations
Following the frameworks proposed by Dubey et al. [30] and Jones et al. [38], we addressed several ethical considerations throughout our research process:

- *Informed consent*: All participants were provided with detailed information about the study's purpose, data collection methods, and how their data would be used.
- *Data privacy*: Personal identifiers were removed from the dataset after collection, and all analyses were conducted on anonymized data.
- *Transparency*: Participants were informed about the analytical methods used and provided with access to summary findings.
- *Benefit principle*: The research was designed to benefit students by improving the elective course system based on their feedback.
- *Algorithmic fairness*: We carefully evaluated our analytical approaches to avoid reinforcing biases or disadvantaging particular student groups.
- *Open science*: In accordance with FAIR data principles, we have made the anonymized dataset publicly available to promote research transparency and reproducibility[34].

These ethical safeguards align with emerging best practices in educational data mining research while acknowledging the unique challenges of the Ukrainian higher education context.

Results
Overall satisfaction metrics
Overall satisfaction with the elective course system showed moderate levels across the university (M = 3.42, SD = 0.86 on a 5-point scale). Analysis of different satisfaction components revealed considerable variation across different aspects of the system, as shown in Table 1.

Students were generally more satisfied with course content and teaching quality than with administrative aspects of the elective system. Particularly low ratings were observed for support during the selection process and the timing of selection periods. The relatively large standard deviations (all above 0.90) indicate considerable variation in student experiences across the sample.

Selection patterns and influencing factors
When asked about the most important factors in their selection decisions, students identified a range of influences. Fig. 1 presents these factors ranked by the percentage of students citing each as important.

Career relevance (64%), interesting topic (58%), and instructor reputation (49%) emerged as the three most frequently cited considerations. Scheduling convenience (42%) and course difficulty (38%) were also notable factors. Peer recommendations (23%) and faculty advisor guidance (18%) were cited less frequently.

The distribution of selections between general education and profession-specific electives varied by academic program, as shown in Table 2.

Aspect	Mean (1-5)	SD
Course content quality	3.78	0.92
Teaching quality	3.65	1.05
Relevance to career goals	3.54	1.12
Alignment with expectations	3.32	1.08
Variety of offerings	3.22	1.18
Information availability before selection	3.11	1.24
Selection process ease	3.08	1.16
Registration system usability	2.95	1.21
Timing of selection period	2.89	1.14
Support during selection process	2.85	1.19

Table 1. Satisfaction ratings for different aspects of the elective course system.

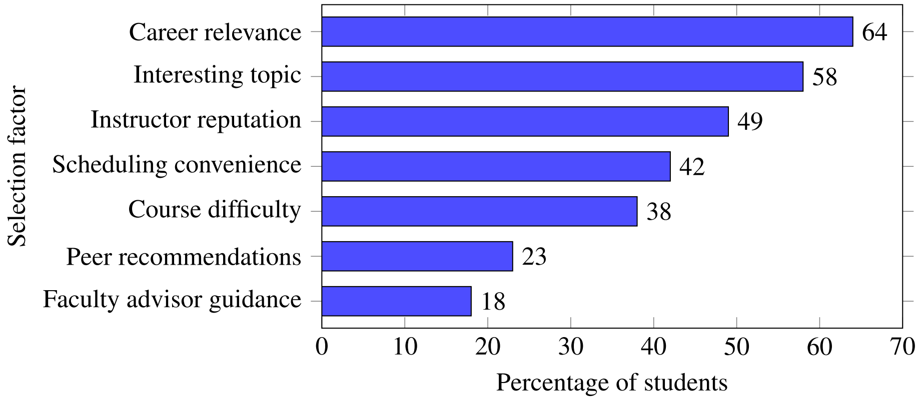

Fig. 1. Factors influencing elective course selection decisions (n = 1,089).

Academic program	General education	Profession-specific	Total
Education	55%	45%	100%
Humanities	52%	48%	100%
Arts	47%	53%	100%
Foreign Languages	36%	64%	100%
Natural Sciences	32%	68%	100%
Physical Education	29%	71%	100%

Table 2. Distribution of elective selections by academic program.

Students in Education and Humanities programs showed stronger preference for general education electives, while students in Natural Sciences, Foreign Languages, and Physical Education programs predominantly selected profession-specific electives.

Student segment identification

To identify distinct patterns in how students approach elective selection, we conducted cluster analysis using the K-means algorithm. After testing models with $k = 2$ to $k = 8$ clusters and evaluating them using silhouette scores and practical interpretability, we determined that a four-cluster solution provided the most meaningful segmentation (silhouette coefficient = 0.47).

Fig. 2 visualizes these clusters along two principal components derived from the seven satisfaction dimensions. The first component (explaining 38% of variance) loaded primarily on course content and teaching quality variables, while the second component (explaining 24% of variance) loaded on process-related variables.

The four clusters represent distinct student segments with characteristic selection approaches and satisfaction profiles:

Cluster 1: Career-focused pragmatists (32% of respondents, n = 348). This largest segment primarily selects courses based on career relevance and practical applications. Members show higher satisfaction with profession-specific electives (M = 3.87, SD = 0.78) than general electives (M = 2.98, SD = 0.94). This segment is more prevalent among upper-level students (3rd and 4th years) and disproportionately represented in Natural Sciences and Foreign Languages programs.

Cluster 2: Content enthusiasts (27% of respondents, n = 294). This segment selects courses primarily based on interest in the topic and intellectual curiosity. Members show high satisfaction with course content (M = 4.12, SD = 0.68) and teaching quality (M = 3.98, SD = 0.72). They are more evenly distributed across years of study and show higher representation in Humanities and Arts programs.

Cluster 3: Process-sensitive selectors (24% of respondents, n = 261). This segment's satisfaction is heavily influenced by the administrative aspects of course selection. Members show lower overall satisfaction (M = 2.89, SD = 0.91), particularly with registration systems (M = 2.32, SD = 1.08) and selection timing (M = 2.41, SD = 1.02). This segment is more prevalent among first and second-year students.

Cluster 4: Balanced optimizers (17% of respondents, n = 186). The smallest segment takes a multifaceted approach to course selection, balancing career relevance, personal interest, and practical considerations. Members show the highest overall satisfaction (M = 3.95, SD = 0.62) across all aspects of the elective system. This segment is more common among graduate students and shows higher average academic performance (GPA M = 3.68, SD = 0.41) compared to the full sample (GPA M = 3.42, SD = 0.58).

Table 3 summarizes the key characteristics of each segment.

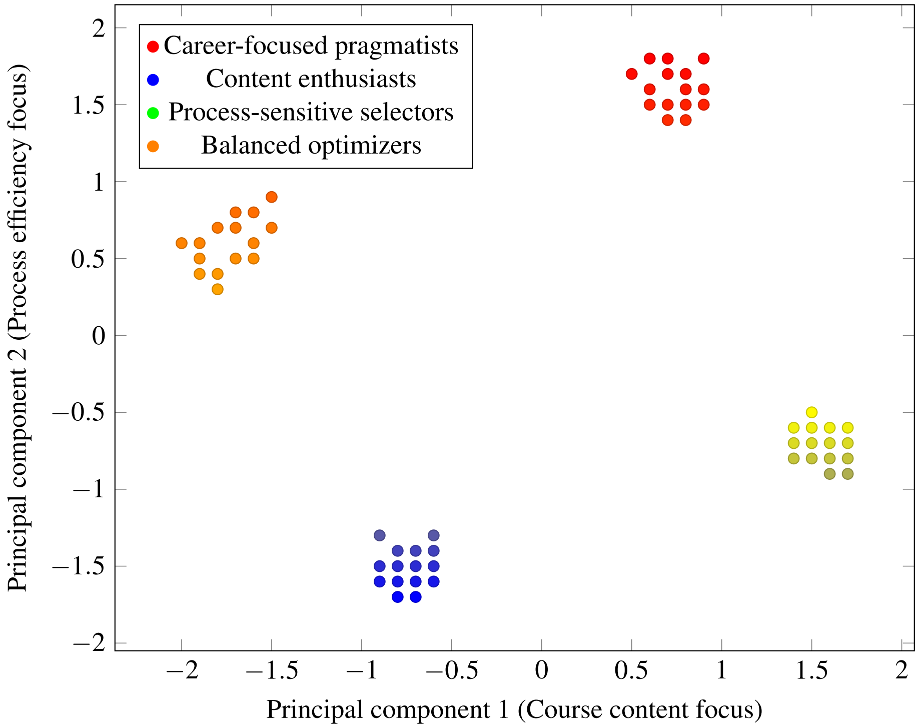

Fig. 2. Student segments visualized using principal component analysis of satisfaction dimensions (n = 1,089).

Characteristic	Career-focused	Content	Process-sensitive	Balanced
Percentage of sample	32%	27%	24%	17%
Overall satisfaction (M)	3.41	3.72	2.89	3.95
Primary selection driver	Career relevance	Topic interest	Convenience	Multiple factors
Predominant year level	3rd–4th year	All years	1st–2nd year	Graduate
Mean GPA	3.38	3.44	3.31	3.68

Table 3. Summary of student segment characteristics.

Predictor	β	SE	t	p
Alignment with expectations	0.31	0.03	9.87	<0.001
Information availability before selection	0.28	0.03	9.12	<0.001
Teaching quality	0.24	0.03	7.65	<0.001
Relevance to career goals	0.22	0.03	7.21	<0.001
Variety of offerings	0.17	0.03	5.56	<0.001
Registration system usability	0.06	0.03	1.78	0.075
Timing of selection periods	0.04	0.03	1.25	0.212

Table 4. Multiple regression analysis of factors predicting overall satisfaction. *Note.* R^2 = 0.63, Adjusted R^2 = 0.62, $F(7, 1081) = 263.4$, $p < 0.001$

Predictors of student satisfaction

To identify the factors most strongly associated with overall satisfaction with the elective system, we conducted multiple regression analysis. Table 4 presents the results.

Five predictors showed statistically significant associations with overall satisfaction. In order of standardized coefficient magnitude, these were: alignment with expectations ($\beta = 0.31$, $p < 0.001$), information availability before selection ($\beta = 0.28$, $p < 0.001$), teaching quality ($\beta = 0.24$, $p < 0.001$), relevance to career goals ($\beta = 0.22$, $p < 0.001$), and variety of offerings ($\beta = 0.17$, $p < 0.001$).

Registration system usability and timing of selection periods did not reach statistical significance as predictors, despite receiving relatively low satisfaction ratings in descriptive analyses. The regression model explained 63% of variance in overall satisfaction ($R^2 = 0.63$, adjusted $R^2 = 0.62$).

To examine whether predictors varied across student segments, we conducted separate regression analyses for each cluster. Table 5 presents the top three predictors for each segment.

The predictor profiles differed substantially across segments. Career-focused pragmatists showed the strongest association between satisfaction and career relevance ($\beta = 0.42$). Content enthusiasts' satisfaction was most strongly predicted by teaching quality ($\beta = 0.38$) and course content quality ($\beta = 0.31$). Process-sensitive selectors showed strong associations with administrative factors including information availability ($\beta = 0.36$) and selection process ease ($\beta = 0.29$). Balanced optimizers showed a more even distribution of predictor weights, with alignment with expectations ($\beta = 0.29$), information availability ($\beta = 0.27$), and variety of offerings ($\beta = 0.25$) all contributing substantially.

Variation across academic programs

Significant differences in satisfaction were found across academic programs (one-way ANOVA: $F(5, 1083) = 8.74$, $p < 0.001$, $\eta^2 = 0.039$). Post-hoc comparisons using Tukey's HSD test revealed that students in Arts programs reported significantly higher satisfaction (M = 3.78, SD = 0.74) than students in all other programs except Humanities. Students in Natural Sciences programs reported the lowest satisfaction levels (M = 3.12, SD = 0.92), significantly lower than students in Arts, Education, and Humanities programs.

Fig. 3 displays the distribution of satisfaction ratings across academic programs.

Further analysis indicated that Arts programs offered the highest ratio of elective courses per student (0.18 courses per student compared to university average of 0.11), which may partially account for the higher satisfaction levels observed. The segment distribution also varied by academic program: Career-focused pragmatists were overrepresented in Natural Sciences (41% vs. 32% overall) and Foreign Languages (38%), while Content enthusiasts were overrepresented in Humanities (35% vs. 27% overall) and Arts (33%).

Qualitative analysis of student feedback

Analysis of open-ended survey responses ($n = 847$ non-empty responses) revealed several patterns in student feedback about course selection procedures.

Response categories were distributed as follows:

- *Satisfaction with current system* (28%, $n = 237$): Responses indicating general satisfaction, including variations of "Everything is fine" or "I am satisfied"
- *No suggestions* (19%, $n = 161$): Explicit statements of having no proposals for change
- *Requests for more course options* (15%, $n = 127$): Suggestions to expand available course selection
- *Requests for better information* (11%, $n = 93$): Calls for improved descriptions and guidance
- *Other substantive feedback* (27%, $n = 229$): Various specific suggestions and concerns

Fig. 4 displays the distribution of response categories.

Among responses containing substantive feedback ($n = 449$, 53% of non-empty responses), thematic coding identified five primary themes:

1. *Course relevance* (32% of substantive responses, $n = 144$): Students expressed desire for courses more closely aligned with their professional specialization and future careers. Representative response: "So that they are directed toward our future profession."
2. *Information quality* (27% of substantive responses, $n = 121$): Requests for more detailed course descriptions, clearer explanations of content, and better presentation of expected outcomes. Representative responses: "More meaningful explanation" and "Before choosing a discipline, learn about it, its content and program."

Segment	Predictor	β
Career-focused pragmatists	Relevance to career goals	0.42
	Alignment with expectations	0.28
	Teaching quality	0.19
Content enthusiasts	Teaching quality	0.38
	Course content quality	0.31
	Alignment with expectations	0.22
Process-sensitive selectors	Information availability	0.36
	Selection process ease	0.29
	Registration system usability	0.24
Balanced optimizers	Alignment with expectations	0.29
	Information availability	0.27
	Variety of offerings	0.25

Table 5. Top three satisfaction predictors by student segment. *Note.* All coefficients significant at $p < 0.01$

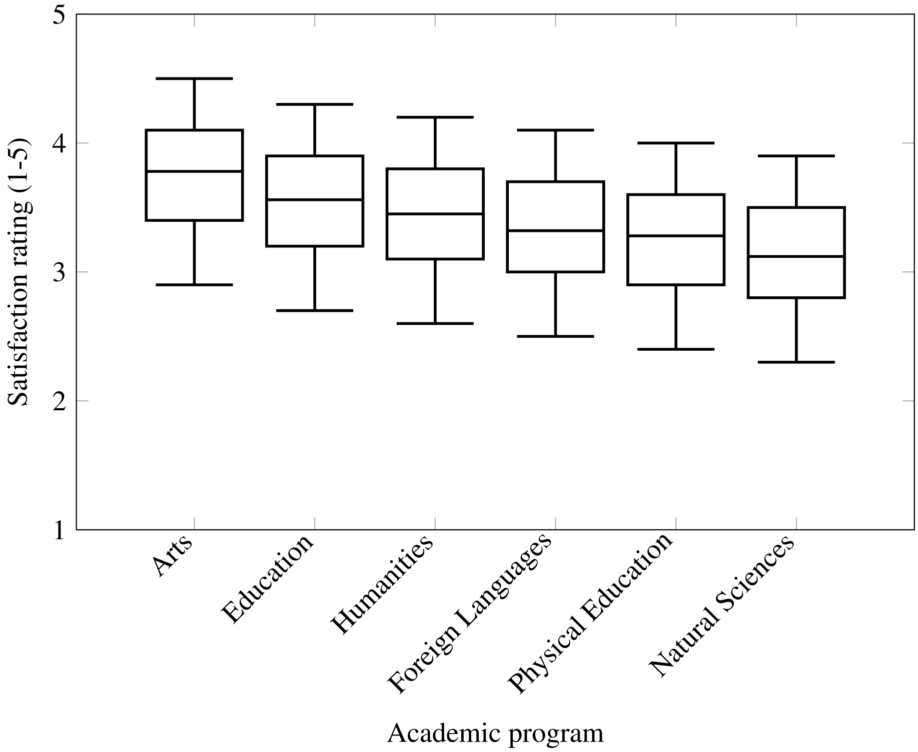

Fig. 3. Distribution of satisfaction ratings by academic program.

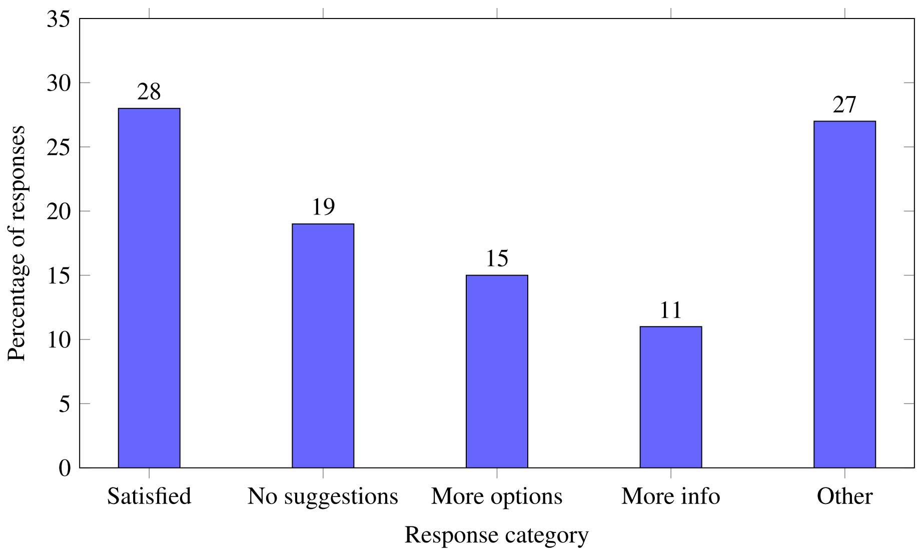

Fig. 4. Distribution of response categories in open-ended student feedback (n = 847).

3. *Selection process mechanics* (18% of substantive responses, n = 81): Suggestions to improve technical aspects of course selection, including timing, platform usability, and registration procedures. Representative responses: "Give more time for selection" and "Would like to have the opportunity to reselect the discipline after the first class."
4. *Course variety* (13% of substantive responses, n = 58): Requests for wider range of course options, including interdisciplinary courses and specialized topics. Representative response: "Add more disciplines related to learning programs such as Adobe Photoshop, Adobe Illustrator, etc."
5. *Selection agency* (10% of substantive responses, n = 45): Concerns about limited actual choice when courses are canceled due to low enrollment or when institutional decisions override student preferences. Represent-

ative response: "Give students free choice of disciplines they would like to study, because unfortunately the university decides too much for us."

Sentiment analysis of substantive responses indicated that 42% were neutral in tone (primarily informational suggestions), 35% were negative (expressing frustration or criticism), and 23% were positive (expressing appreciation while suggesting improvements).

Framework derivation for personalized course selection support

The preceding analyses identified four distinct student segments (RQ1), five significant satisfaction predictors (RQ2), and segment-specific predictor profiles. These findings provide an empirical basis for deriving design principles and a framework to guide educational technology interventions (RQ3, RQ4).

Design principles derived from findings

Based on the pattern of results, five design principles emerge for educational technology interventions supporting elective course selection:

1. *Segment-aware personalization*: The identification of four distinct student segments with different selection approaches and satisfaction drivers indicates that uniform interventions will be suboptimal. Technology solutions should incorporate mechanisms to identify student segment membership and tailor information presentation accordingly.
2. *Information quality prioritization*: Information availability before selection emerged as the second strongest predictor of overall satisfaction ($\beta = 0.28$) and was the top predictor for process-sensitive selectors ($\beta = 0.36$). This indicates that enhanced course information systems should be a primary intervention target.
3. *Expectation-outcome alignment*: Alignment with expectations was the strongest overall predictor ($\beta = 0.31$). Technology systems should help students form accurate expectations by providing detailed syllabi, student reviews, and outcome data.
4. *Career pathway integration*: Career relevance was cited by 64% of students as an important selection factor and was the top predictor for career-focused pragmatists ($\beta = 0.42$). Course selection systems should explicitly connect elective options to career pathways and professional skill development.
5. *Process accessibility*: While administrative factors (registration usability, timing) were not significant predictors in the overall model, they were central to satisfaction among process-sensitive selectors (24% of students). Interventions should ensure that process improvements benefit this substantial minority.

Integrated framework

Fig. 5 presents a framework integrating these findings and principles into a system architecture for personalized course selection support.

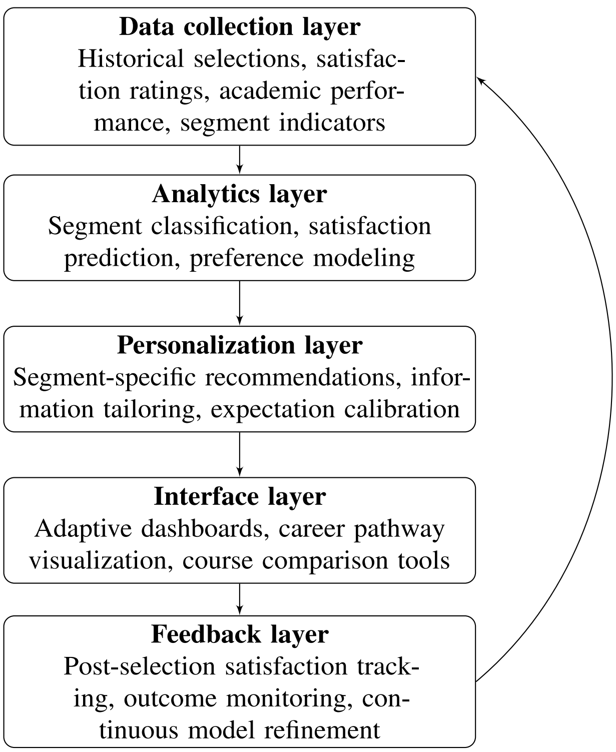

Fig. 5. Framework for personalized course selection support derived from empirical findings.

Segment	Priority information	Interface features	Key metrics
Career-focused pragmatists	Career outcomes, skill alignment, professional applications	Career pathway visualizations, alumni outcome data	Employment relevance score
Content enthusiasts	Syllabus depth, instructor profiles, intellectual scope	Detailed content previews, instructor ratings, topic maps	Content quality rating
Process-sensitive selectors	Clear procedures, timeline guidance, support resources	Simplified selection wizard, progress tracking, help integration	Process completion rate
Balanced optimizers	Multi-dimensional comparisons, comprehensive data	Advanced filtering, comparison matrices, personalized rankings	Overall fit score

Table 6. Framework feature mapping by student segment.

The framework comprises five interconnected layers:

1. *Data collection layer*: Gathers student data including historical course selections, satisfaction ratings across dimensions identified in this study, academic performance indicators, and behavioral signals that enable segment classification.
2. *Analytics layer*: Applies the analytical approaches validated in this study – cluster analysis for segment identification and regression modeling for satisfaction prediction – to classify incoming students and predict their satisfaction drivers.
3. *Personalization layer*: Uses segment membership and predicted satisfaction drivers to tailor the selection experience. For career-focused pragmatists, this layer prioritizes career relevance information; for content enthusiasts, it emphasizes teaching quality and content depth; for process-sensitive selectors, it streamlines administrative procedures; for balanced optimizers, it provides comprehensive multi-factor comparisons.
4. *Interface layer*: Presents personalized information through adaptive dashboards. Based on findings that 64% of students prioritize career relevance and that alignment with expectations is the strongest satisfaction predictor, interfaces should prominently feature career pathway connections and outcome expectations.
5. *Feedback layer*: Collects post-selection satisfaction data to validate predictions and refine models. The feedback loop enables continuous improvement of segment classification accuracy and recommendation quality.

Table 6 maps specific framework features to each student segment based on their identified characteristics and predictor profiles.

This framework directly addresses RQ3 by specifying design principles grounded in empirical findings and RQ4 by providing an integrated architecture that connects student segments to personalized interventions.

Discussion

Key findings and theoretical implications

This study employed educational data mining techniques to analyze student experiences with elective course selection at a Ukrainian university, addressing four research questions about selection patterns, satisfaction predictors, design principles, and an integrative framework. The findings reveal a complex picture of student engagement with elective course systems that both supports and extends previous research.

Student segments and selection patterns (RQ1)

The identification of four distinct student segments – career-focused pragmatists, content enthusiasts, process-sensitive selectors, and balanced optimizers – demonstrates that students approach elective course selection through fundamentally different lenses. This segmentation aligns partially with findings from research on course recommendation systems, where different student response patterns to recommendations have been observed[20]. However, our four-segment solution provides greater specificity than binary or three-category models prevalent in earlier work.

The existence of these segments suggests that theoretical models of student choice in higher education need to account for heterogeneity in decision-making approaches. Current frameworks, such as those proposed by Babad and Tayeb[39], may oversimplify the variation in student selection strategies. Our findings indicate that selection behaviors are not uniformly distributed across the student population but are instead shaped by a combination of individual characteristics, program context, and stage of study.

Career relevance (64%), interesting topic (58%), and instructor reputation (49%) emerged as the most frequently cited selection factors, echoing findings from collaborative recommendation system research regarding the primary drivers of student course selection[12]. However, the relatively weak influence of peer recommendations in our sample (cited by only 23% of students) contrasts with studies conducted in Western contexts, where social influences play a more prominent role in course selection decisions[40]. This difference may reflect cultural variation in educational decision-making in the Ukrainian context, where structured pathways and institutional guidance traditionally carry greater weight than peer-influenced choices. Kaiser and Blömeke[23] have noted distinctive Eastern European educational approaches that may account for such differences.

Satisfaction predictors and segment variation (RQ2)

The regression analysis revealed that five factors significantly predicted overall satisfaction: alignment with expectations (β = 0.31), information availability before selection (β = 0.28), teaching quality (β = 0.24), relevance to career goals (β = 0.22), and variety of offerings (β = 0.17). The model explained 63% of variance in satisfaction, indicating that these factors collectively provide a robust account of what shapes student experiences with elective systems.

The strong predictive value of information availability before selection aligns with findings emphasizing the importance of information quality in helping students make appropriate educational choices[41]. Students who reported having adequate information to make informed decisions showed substantially higher satisfaction, regardless of their eventual course choices. This finding underscores the value of transparency and comprehensive course descriptions in elective systems.

Notably, students were generally more satisfied with course content and teaching quality than with administrative aspects of the system. This pattern aligns with findings from research on e-learning recommendation systems, which found that content and teaching quality have stronger influence on student satisfaction than technical or administrative factors[21]. The relatively large standard deviations observed across all satisfaction dimensions (all above 0.90) indicate considerable variation in student experiences, reinforcing the value of segment-specific approaches.

The segment-specific regression analyses revealed that satisfaction predictors varied substantially across groups. Career-focused pragmatists showed the strongest association between satisfaction and career relevance ($\beta = 0.42$), while content enthusiasts' satisfaction was most strongly predicted by teaching quality ($\beta = 0.38$). Process-sensitive selectors showed uniquely strong associations with administrative factors. These differential patterns suggest that uniform interventions will have unequal effects across the student population.

Design principles and framework (RQ3, RQ4)

The derivation of five design principles and an integrated framework directly from empirical findings (see Results section) represents a contribution to bridging the gap between descriptive educational data mining research and actionable system design. Previous studies have often proposed frameworks based on theoretical considerations or technical capabilities rather than empirically identified student needs[5]. Our approach grounds framework components in specific findings: the five-layer architecture maps directly to identified segments and their distinct predictor profiles.

The framework's emphasis on segment-aware personalization responds to findings that a one-size-fits-all approach to supporting elective selection is suboptimal. The 24% of students classified as process-sensitive selectors, for instance, have satisfaction profiles dominated by administrative factors that are largely irrelevant to the 27% classified as content enthusiasts. Technology solutions that fail to account for this heterogeneity risk optimizing for one segment while neglecting others.

Implications for learning analytics applications

Our findings suggest several promising applications of learning analytics to improve elective course systems, building on recent developments in the field[1,42].

Segment-specific analytics dashboards

The identification of distinct student segments creates opportunities for personalized analytics dashboards that address the specific needs of different student groups. Following human-centered design principles for learning analytics dashboards[26], these interfaces could present different information types and visualizations based on segment characteristics.

For career-focused pragmatists (32% of students), dashboards should highlight career outcome data, skills development trajectories, and professional applications of course content. These students showed the strongest satisfaction association with career relevance ($\beta = 0.42$), suggesting that explicit connections between courses and career pathways would be particularly valuable.

For content enthusiasts (27% of students), visualizations emphasizing concept relationships, knowledge maps, and thematic connections between courses would align with their interest-driven selection approach. Their high satisfaction with course content (M = 4.12) and teaching quality (M = 3.98) suggests receptivity to detailed academic information.

For process-sensitive selectors (24% of students), simplified interfaces with clear timelines, step-by-step guidance, and readily accessible support resources could address their primary satisfaction drivers. Their lower overall satisfaction (M = 2.89) and strong associations with administrative factors indicate that process improvements could yield substantial benefits for this group.

For balanced optimizers (17% of students), comprehensive dashboards integrating multiple information types and advanced comparison tools would match their multifaceted approach to course selection. Their high overall satisfaction (M = 3.95) suggests that current systems may already serve this segment reasonably well, though enhanced tools could further support their sophisticated decision-making.

Predictive modeling for course recommendation

Building on collaborative filtering approaches to course recommendation[16], our findings suggest that predictive models could incorporate segment membership as a key parameter. The segment-specific predictor profiles identified in this study could inform recommendation algorithms that weight different course attributes according to individual student characteristics.

Research on large language models for course recommendation[13] provides a promising direction for enhancing such systems with natural language understanding capabilities. Combining segment-aware weighting with LLM-based course description analysis could generate recommendations that simultaneously match student preferences and provide rich explanatory information.

Machine learning approaches to course selection have achieved high accuracy in predicting student outcomes[17,22]. Integrating our segment classifications with outcome prediction could enable recommendations that consider both preference alignment and likely success, potentially reducing course withdrawals and improving completion rates.

Ethical considerations in analytics implementation
Following data ethics frameworks for AI in education[19], we recognize the importance of addressing ethical considerations in implementing learning analytics for elective systems. Several key issues emerge from our findings.

Privacy and consent remain paramount. Students should be informed about data collection practices and how their information supports personalization. The segment classifications derived in this study rely on satisfaction and behavioral data that students may not expect to be used for ongoing personalization.

Algorithmic fairness requires careful attention. Our finding that satisfaction varies significantly across academic programs (with Natural Sciences students reporting lower satisfaction than Arts students) raises questions about whether recommendation systems might inadvertently perpetuate these disparities. Systems should be designed and monitored to avoid reinforcing existing inequities.

Transparency in recommendation logic supports informed decision-making. Students benefit from understanding why particular courses are suggested, enabling them to evaluate recommendations critically rather than accepting them passively.

Student autonomy must be preserved. While analytics can support decision-making, they should not replace student agency in course selection. The concerns about selection agency expressed by some students in our qualitative data (10% of substantive responses) highlight the importance of positioning technology as supportive rather than directive.

These considerations align with findings on student perspectives regarding educational data analytics, which emphasize the importance of transparency and consent in maintaining student trust[33].

Addressing digital divide and accessibility concerns
Our findings reveal potential issues related to the digital divide and accessibility in elective course systems. The process-sensitive selectors segment (24% of students), which showed the lowest overall satisfaction, was particularly concerned with technological barriers in the selection process. This finding aligns with research emphasizing the need for accessible educational resources, particularly as institutions increasingly rely on digital platforms[27].

The low satisfaction ratings for registration system usability ($M = 2.95$) and support during selection ($M = 2.85$) suggest that current technological implementations may create barriers for some students. Research has identified multiple factors influencing the digital divide, including socioeconomic status, prior technology experience, and institutional support[28]. Our finding that process-sensitive selectors are disproportionately represented among first and second-year students suggests that experience with university systems may mitigate some technological challenges over time, but this places newer students at a disadvantage during a critical period of academic integration.

Several strategies emerge from our findings to address these concerns. Multi-modal information provision – offering course information in text, video, and interactive formats – could accommodate different learning preferences and access situations. Simplified interface options for students with limited digital literacy or technical resources would ensure core functionality remains accessible. Maintaining non-digital alternatives for critical selection processes ensures participation regardless of technology access. Progressive enhancement approaches could provide basic functionality to all users while offering advanced features to those with greater technological capabilities.

These approaches align with recommendations for bridging the digital divide in educational contexts[43], emphasizing that technological solutions should enhance rather than limit educational accessibility.

Practical implications for higher education institutions
Our findings have several practical implications for universities implementing or refining elective course systems.

Differentiated information provision
The identification of different student segments suggests that uniform information provision is insufficient. Following models for enhancing student decision-making through dashboards[44], institutions should consider developing differentiated information resources addressing the specific concerns of different student types.

For career-focused pragmatists, detailed information about professional applications, skill development, and career outcomes would be most beneficial. For content enthusiasts, comprehensive syllabi with intellectual scope and pedagogical approach would better support their decisions. For process-sensitive selectors, clear procedural guidance and accessible support resources take priority.

Timing and process optimization
The low satisfaction with selection timing ($M = 2.89$) indicates a need to reconsider when students make selection decisions. Our qualitative analysis revealed frequent mentions of insufficient time for decision-making and conflicts with other academic demands during selection periods. Adaptive interface models that accommodate varying student needs and timelines[45] could help address these concerns by providing more flexible selection processes.

The finding that process-sensitive selectors are overrepresented among first and second-year students suggests particular attention to supporting newer students during their initial selection experiences. Early positive experiences with elective selection may influence long-term engagement with personalized learning opportunities.

Enhanced decision support
The gap between selection support satisfaction (M = 2.85) and its importance to certain segments suggests an opportunity for enhanced guidance systems. Research indicates that faculty engagement with learning analytics is critical for providing effective student support[46]. Institutions should consider developing technologically-enhanced advising systems that connect students with knowledgeable faculty who can help them navigate the selection process.

The qualitative finding that 27% of substantive student responses addressed information quality suggests substantial demand for better pre-selection guidance. Students requested "more meaningful explanation" of courses and opportunities to "learn about [courses], content and program" before making binding decisions.

Program-specific strategies
The significant differences in satisfaction across academic programs highlight the need for context-specific approaches. Students in Arts programs reported significantly higher satisfaction (M = 3.78) than those in Natural Sciences (M = 3.12), a difference partially explained by variation in the ratio of available electives per student. Rather than implementing identical systems across all departments, universities should consider program-specific strategies addressing unique disciplinary contexts and student expectations.

These disciplinary differences align with observations about variation between Eastern and Western educational approaches[23], suggesting that even within a single institution, different academic cultures may require tailored solutions. Engineering and technical programs may have different requirements for personalized learning compared to humanities programs[24], and elective systems should accommodate this variation.

Limitations and future research
This study has several limitations that suggest directions for future research.

First, our study relied on self-reported data collected at a single point in time. While the sample size was substantial (n = 1,089), cross-sectional designs cannot establish causal relationships between satisfaction predictors and outcomes. Longitudinal studies tracking students throughout their academic careers would provide more robust insights into how selection patterns and satisfaction evolve over time and how early choices impact later academic and career outcomes. Research on longitudinal career tracking[47] demonstrates the potential of such approaches for understanding long-term effects of educational choices.

Second, while our response rate (31%) is acceptable for voluntary online surveys, it raises the possibility of selection bias. Students with stronger opinions – whether positive or negative – about the elective system may have been more likely to participate. Future studies could employ stratified sampling techniques to ensure balanced representation across student segments.

Third, our study focused on a single institution in Ukraine, limiting generalizability. The distinctive patterns we observed – particularly the relatively weak influence of peer recommendations – may reflect cultural and institutional factors specific to this context. Comparative studies across multiple institutions and national contexts would enhance understanding of how cultural and structural factors influence elective course systems. Frameworks for cross-national educational research[48] could be adapted for studying elective systems across diverse settings.

Fourth, our analysis focused primarily on student perspectives. While student satisfaction is an important outcome, it represents only one dimension of elective system effectiveness. Future research should incorporate faculty viewpoints, administrative data on actual enrollment patterns, and educational outcomes such as course completion, grades, and skill development to develop a more comprehensive understanding of effective elective systems. Industry-institute collaboration in elective design[49] represents another dimension worthy of investigation.

Fifth, the four-segment solution, while statistically supported by silhouette analysis, represents one possible categorization of student approaches. Alternative clustering solutions or different analytical approaches might reveal additional meaningful distinctions. Future research could explore whether finer-grained segmentation provides additional practical value for personalization.

Finally, the framework derived from our findings awaits empirical validation. While the framework components are grounded in observed patterns, their effectiveness for improving student outcomes must be tested through implementation studies. Randomized controlled trials comparing segment-aware interventions with uniform approaches would provide evidence for the practical value of personalization based on our segmentation model.

Conclusion
This study applied educational data mining techniques to analyze student experiences with elective course selection at a Ukrainian university, contributing empirical insights and a practical framework for enhancing personalized learning in higher education.

Regarding student selection patterns (RQ1), we identified four distinct segments – career-focused pragmatists (32%), content enthusiasts (27%), process-sensitive selectors (24%), and balanced optimizers (17%) – each characterized by different selection approaches and satisfaction profiles. Career relevance, topic interest, and instructor reputation emerged as the primary selection drivers, though their relative importance varied substantially across segments.

Regarding satisfaction predictors (RQ2), five factors significantly predicted overall satisfaction: alignment with expectations, information availability before selection, teaching quality, relevance to career goals, and variety of offerings. The regression model explained 63% of variance in satisfaction. Critically, predictor profiles differed across segments: career-focused pragmatists' satisfaction depended most strongly on career relevance, while process-sensitive selectors' satisfaction was driven primarily by administrative factors.

Regarding design principles (RQ3), five empirically grounded principles emerged: segment-aware personalization, information quality prioritization, expectation-outcome alignment, career pathway integration, and process accessibility. These principles address the heterogeneity in student needs revealed by segment analysis.

Regarding the integrative framework (RQ4), we derived a five-layer architecture – comprising data collection, analytics, personalization, interface, and feedback layers – that maps specific features to each student segment based on their identified characteristics and predictor profiles.

Based on these findings, we offer the following recommendations for higher education institutions. First, institutions should move beyond uniform information provision toward segment-differentiated resources. Career-focused students benefit from explicit career pathway connections, while content-oriented students value detailed syllabi and instructor information. Process-sensitive students require clear procedural guidance and accessible support.

Second, the strong predictive value of information availability (β = 0.28) suggests that investment in comprehensive, accessible course information systems will yield substantial returns in student satisfaction. Pre-selection information quality deserves priority attention.

Third, the variation in satisfaction across academic programs indicates that institution-wide standardization may be counterproductive. Program-specific approaches that account for disciplinary cultures and student expectations are more likely to succeed.

Fourth, the concentration of process-sensitive selectors among first and second-year students highlights the importance of supporting newer students during initial selection experiences. Early interventions may have lasting effects on engagement with personalized learning opportunities.

This research aligns with Sustainable Development Goal 4 (Quality Education) by promoting inclusive and equitable educational opportunities. The framework and recommendations presented here aim to enhance student agency in curriculum customization while maintaining educational quality and addressing accessibility concerns. As higher education continues to embrace personalization, evidence-based approaches to supporting student choice become increasingly important for preparing graduates for diverse career pathways.

The dataset supporting this study is openly available following FAIR principles, enabling replication and extension of this work across different institutional and cultural contexts.

Data Availability

The dataset used in this study, "Elective Disciplines Survey Data from Kryvyi Rih State Pedagogical University", is openly available in the Zenodo repository at https://zenodo.org/records/15098020. The dataset follows FAIR principles and is provided in multiple formats (CSV, TSV, JSON, and Excel) with comprehensive metadata to facilitate reuse.

Received: 7 October 2025; Accepted: 23 January 2026
Published online: 02 February 2026

References

1. Wong, B. T. M., Li, K. C. & Liu, M. The Role of Learning Analytics in Evaluating Course Effectiveness. *Sustainability* **17**, 559. https://doi.org/10.3390/su17020559 (2025).
2. Deri, M. N., Singh, A., Zaazie, P. & Anandene, D. Leveraging Artificial Intelligence in Higher Educational Institutions: A Comprehensive Overview. *Revista de Educacion y Derecho* https://doi.org/10.1344/REYD2024.30.45777 (2024).
3. Ministry of Education and Science of Ukraine. Zatverdzheni standarty vyshchoi osvity [Approved higher education standards] (2025). https://mon.gov.ua/osvita-2/vishcha-osvita-ta-osvita-doroslikh/naukovo-metodichna-rada-ministerstva-osviti-i-nauki-ukraini/zatverdzheni-standarti-vishchoi-osviti.
4. Xu, X., Sun, Y., Weng, J. & Zhang, Y. Theoretical Framework of Personal Learning Environments: SPET Model. In Cheung, S. K. S. *et al.* (eds.) *Technology in Education. Innovative Practices for the New Normal*, vol. 1974 CCIS of *Communications in Computer and Information Science*, 139–156, https://doi.org/10.1007/978-981-99-8255-4_13 (Springer Nature Singapore, Singapore, 2024).
5. Algarni, S. & Sheldon, F. Systematic Review of Recommendation Systems for Course Selection. *Mach. Learn. Knowledge Extract.* **5**, 560–596. https://doi.org/10.3390/make5020033 (2023).
6. Cha, S., Loeser, M. & Seo, K. The Impact of AI-Based Course-Recommender System on Students' Course-Selection Decision-Making Process. *Appl. Sci.* **14**, 3672. https://doi.org/10.3390/app14093672 (2024).
7. Baker, R. S. & Inventado, P. S. Educational Data Mining and Learning Analytics. In Larusson, J. A. & White, B. (eds.) *Learning Analytics: From Research to Practice*, 61–75, https://doi.org/10.1007/978-1-4614-3305-7_4 (Springer, New York, NY, 2014).
8. Cai, C. Exploration on Data Mining Algorithms for University Information Systems Based on Big Data Environment. In *Proceedings - 2023 International Conference on Computer Simulation and Modeling, Information Security, CSMIS 2023*, 626–632, https://doi.org/10.1109/CSMIS60634.2023.00117 (2023).
9. Alenezi, A. Personalized Learning Strategies in Higher Education in Saudi Arabia: Identifying Common Approaches and Conditions for Effective Implementation. *TEM J.***12**, 2023–2037, https://doi.org/10.18421/TEM124-13 (2023).
10. Xu, X., Wang, X., Zhang, Y., Yue, Y. & Wu, Y. ICT Framework for the Personal Learning Environment (PLE-ICT) in Higher Education: Results from Experts' Interview. *Front. Artif. Intell. Appl.* **370**, 16–29. https://doi.org/10.3233/FAIA230166 (2023).
11. Chaturapruek, S. et al. Studying Undergraduate Course Consideration at Scale. *AERA Open* **7**, 2332858421991148. https://doi.org/10.1177/2332858421991148 (2021).
12. Obeidat, R., Duwairi, R. & Al-Aiad, A. A Collaborative Recommendation System for Online Courses Recommendations. In *Proceedings - 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications, Deep-ML 2019*, 49–54, https://doi.org/10.1109/Deep-ML.2019.00018 (2019).
13. Khan, M. A. Z., Polyzou, A. & Bennamane, N. How Can We Use LLMs for EDM Tasks? The Case of Course Recommendation. In Pinto, J. D. *et al.* (eds.) *Joint Proceedings of the Human-Centric eXplainable AI in Education and the Leveraging Large Language Models for Next Generation Educational Technologies Workshops (HEXED-L3MNGET 2024) co-located with 17th International Conference on Educational Data Mining (EDM 2024), Atlanta, Georgia, USA, July 14, 2024*, vol. 3840 of *CEUR Workshop Proceedings* (CEUR-WS.org, 2024). https://ceur-ws.org/Vol-3840/L3MNGET24_paper12.pdf.
14. Trotsko, A., Korotkova, Y., Rybalko, L., Kirichok, A. & Perepelytsia, K. Trends in the development of tertiary education in Ukraine under the COVID-19 pandemic. *Int. J. Higher Educ.* **10**, 241–251. https://doi.org/10.5430/ijhe.v10n3p241 (2021).

15. Romero, C. & Ventura, S. Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **10**, e1355. https://doi.org/10.1002/widm.1355 (2020).
16. Muzdybayeva, G., Khashimova, D., Amirzhanov, A. & Kadyrov, S. A Matrix Factorization-based Collaborative Filtering Framework for Course Recommendations in Higher Education. In *Proceedings - 2023 17th International Conference on Electronics Computer and Computation, ICECCO 2023*, https://doi.org/10.1109/ICECCO58239.2023.10147152 (2023).
17. Wang, X. *et al.* A Machine Learning-based Course Enrollment Recommender System. In *Proceedings of the 14th International Conference on Computer Supported Education - Volume 1: CSEDU*, 436–443, INSTICC (SciTePress, 2022).https://doi.org/10.5220/0011109100003182
18. Polydorou, E. Towards a Secure and Privacy Compliant Framework for Educational Data Mining. In Nurcan, S., Opdahl, A. L., Mouratidis, H. & Tsohou, A. (eds.) *Research Challenges in Information Science: Information Science and the Connected World*, vol. 476 LNBIP of *Lecture Notes in Business Information Processing*, 534–541, https://doi.org/10.1007/978-3-031-33080-3_35 (Springer Nature Switzerland, Cham, 2023).
19. Hong, Y., Nguyen, A., Dang, B. & Nguyen, B.-P. T. Data Ethics Framework for Artificial Intelligence in Education (AIED). In *Proceedings - 2022 International Conference on Advanced Learning Technologies, ICALT 2022*, 297–301,https://doi.org/10.1109/ICALT55010.2022.00095 (2022).
20. Chang, H.-T. *et al.* AI, Please Help Me Choose a Course: Building a Personalized Hybrid Course Recommendation System to Assist Students in Choosing Courses Adaptively. *Educational Technology and Society***26**, 203–217, https://doi.org/10.30191/ETS.202301_26(1).0015 (2023).
21. Majjate, H. *et al.* The Impact of E-Learning Recommendation System on Student Autonomy, Engagement, and Learning Effectiveness. In Mahboub, O., Haddouch, K., Omara, H. & Hefnawi, M. (eds.) *Big Data and Internet of Things*, vol. 887 LNNS of *Lecture Notes in Networks and Systems*, 775–788, https://doi.org/10.1007/978-3-031-74491-4_60 (Springer Nature Switzerland, Cham, 2024).
22. Shannaq, B. The Role of AI in University Course Registration in the Middle East: AI and Machine Learning Approaches to Improve Academic Performance. In *2024 2nd International Conference on Computing and Data Analytics, ICCDA 2024 - Proceedings*, https://doi.org/10.1109/ICCDA64887.2024.10867316 (2024).
23. Kaiser, G. & Blömeke, S. Learning from the Eastern and the Western debate: The case of mathematics teacher education. *ZDM - Int. J. Mathemat. Educ.* **45**, 7–19. https://doi.org/10.1007/s11858-013-0490-x (2013).
24. Starostin, V. S., Arzhanova, K. A. & Dolgopolov, D. V. Marketing Priorities for Artificial Intelligence Technologies Implementation in Engineering and Technical Universities. In Mantulenko, V. (ed.) *Proceedings of the 3rd International Conference Engineering Innovations and Sustainable Development*, vol. 540 LNCE of *Lecture Notes in Civil Engineering*, 339–350, https://doi.org/10.1007/978-3-031-67372-6_43 (Springer Nature Switzerland, Cham, 2024).
25. Hansen, L., Holanda, M., Borges, V. R. P. & Da Silva, D. Visual Analysis of Educational Data: a Case Study of Introductory Programming courses at the University of Brasília. In *Proceedings - Frontiers in Education Conference, FIE*, 2022, https://doi.org/10.1109/FIE56618.2022.9962427 (IEEE, 2022).
26. Revano, T. F. & Garcia, M. B. Designing Human-Centered Learning Analytics Dashboard for Higher Education Using a Participatory Design Approach. In *2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management, HNICEM 2021*, https://doi.org/10.1109/HNICEM54116.2021.9731917 (2021).
27. McKie, M. H. & Strong, A. C. Mapping the Landscape of Digital Accessibility in Computer Science Education: A Mapping Literature Review. In *2024 ASEE Annual Conference & Exposition*, DOI:https://doi.org/10.18260/1-2–47760 (ASEE Conferences, Portland, Oregon, 2024).
28. Martin, F., Ceviker, E. & Gezer, T. From digital divide to digital equity: Systematic review of two decades of research on educational digital divide factors, dimensions, and interventions. *Journal of Research on Technology in Education* 1–25, https://doi.org/10.1080/15391523.2024.2425442 (2024).
29. Kryvyi Rih State Pedagogical University. Polozhennia pro indyvidualnu osvitniu traiektoriiu zdobuvachiv vyshchoi osvity u Kryvorizkomu derzhavnomu pedahohichnomu universyteti (nova redaktsiia) [Regulations on Individual Educational Trajectory of Higher Education Students in Kryvyi Rih State Pedagogical University (New Edition)] (2024). https://drive.google.com/file/d/1ZxjonYBmhiCTm2XuLZHu52stCIZ2JK9V/view.
30. Dubey, S. K., Gupta, S., Pandey, S. K. & Mittal, S. Ethical Considerations in Data Mining and Database Research. In *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2024*, https://doi.org/10.1109/ICRITO61523.2024.10522200 (IEEE, 2024).
31. Ketsman, O., Droog, A. & Qazi, S. Mapping the prevalence of mixed methods research in educational technology journals. *Comput. Educ.* **226**, 105207. https://doi.org/10.1016/j.compedu.2024.105207 (2025).
32. Heinrich, E. Revolutionising educational technology: The imperative for authentic qualitative research. *Soc. Sci. Humanit. Open* **10**, 101073. https://doi.org/10.1016/j.ssaho.2024.101073 (2024).
33. Jones, K. M. L. et al. Transparency and Consent: Student Perspectives on Educational Data Analytics Scenarios. *Portal* **23**, 485–515. https://doi.org/10.1353/pla.2023.a901565 (2023).
34. Semerikov, S. & Bondarenko, O. & Kryvyi Rih State Pedagogical University. *Elective Disciplines Survey Data from Kryvyi Rih State Pedagogical Univ.* https://doi.org/10.5281/zenodo.15098020 (2025).
35. Ncube, M. M. & Ngulube, P. A Systematic Review of Postgraduate Programmes Concerning Ethical Imperatives of Data Privacy in Sustainable Educational Data Analytics. *Sustainability* **16**, 6377. https://doi.org/10.3390/su16156377 (2024).
36. Arthur, D. & Vassilvitskii, S. k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, 1027–1035 (Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2007).
37. Liashenko, R. O. & Semerikov, S. O. Training Specialised Chatbots on Ukrainian Scientific Text Corpora Using Transfer Learning. In *2024 IEEE 19th International Conference on Computer Science and Information Technologies (CSIT)*, 1–4, https://doi.org/10.1109/CSIT65290.2024.10982573 (2024).
38. Jones, K. M. L. et al. "We're being tracked at all times": Student perspectives of their privacy in relation to learning analytics in higher education. *J. Am. Soc. Inf. Sci.* **71**, 1044–1059. https://doi.org/10.1002/asi.24358 (2020).
39. Babad, E. & Tayeb, A. Experimental analysis of students' course selection. *Br. J. Educ. Psychol.* **73**, 373–393. https://doi.org/10.1348/000709903322275894 (2003).
40. Parks-Stamm, E. J., Zafonte, M. & Palenque, S. M. The effects of instructor participation and class size on student participation in an online class discussion forum. *Br. J. Edu. Technol.* **48**, 1250–1259. https://doi.org/10.1111/bjet.12512 (2017).
41. Liang, Y., Zou, D., Wang, F. L., Xie, H. & Cheung, S. K. S. Investigating Demographics and Behavioral Engagement Associated with Online Learning Performance. In Li, C., Cheung, S. K. S., Wang, F. L., Lu, A. & Kwok, L. F. (eds.) *Blended Learning : Lessons Learned and Ways Forward*, vol. 13978 LNCS of *Lecture Notes Comput. Sci.*, 124–136, https://doi.org/10.1007/978-3-031-35731-2_12 (Springer Nature Switzerland, Cham, 2023).
42. Khanipoor, F. & Karimian, Z. Unleashing the power of data: the promising future of learning analytics in medical education: a commentary. *Educ. Inf. Technol.* https://doi.org/10.1007/s10639-024-13273-y (2024).
43. Ahuja, V. Equity and Access in Digital Education: Bridging the Divide. In Arinushkina, A. A., Morozov, A. V. & Robert, I. V. (eds.) *Contemporary Challenges in Education: Digitalization, Methodology, and Management*, 45–59, https://doi.org/10.4018/979-8-3693-1826-3.ch005 (IGI Global, Hershey, PA, 2023).

44. Li, T. S., Sinal, M. S. b., Omar, M. & Anuardi, M. N. A. b. M. University Student Dashboard: Enhancing Student Trend Analysis and Decision-Making Processes. In Zakaria, N. H., Mansor, N. S., Husni, H. & Mohammed, F. (eds.) *Computing and Informatics*, vol. 2002 CCIS of *Communi. Comput. Inf. Sci.*, 139–153, https://doi.org/10.1007/978-981-99-9592-9_11 (Springer Nature Singapore, Singapore, 2024).
45. Hashiura, Y., Matsuo, T. & Terashima, T. Adaptive interface model for lecture allocation system. In *Proceedings - 9th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2010*, 526–528,https://doi.org/10.1109/ICIS.2010.149 (2010).
46. Shepard, L., Rehrey, G. & Groth, D. Faculty Engagement with Learning Analytics: Advancing a Student Success Culture in Higher Education. In Shah, M., Kift, S. & Thomas, L. (eds.) *Student Retention and Success in Higher Education: Institutional Change for the 21st Century*, 89–107, https://doi.org/10.1007/978-3-030-80045-1_5 (Springer International Publishing, Cham, 2021).
47. Poole, P. *et al.* Developing New Zealand's medical workforce: Realising the potential of longitudinal career tracking. *New Zealand Medical Journal* **132**, 65–73 (2019). https://nzmj.org.nz/media/pages/journal/vol-132-no-1495/developing-new-zealand-s-medical-workforce-realising-the-potential-of-longitudinal-career-tracking/0a667d1c89-1696475276/developing-new-zealand-s-medical-workforce-realising-the-potential-of-longitudinal-career-tracking.pdf.
48. Troman, G. & Jeffrey, B. Providing a framework for a 'shared repertoire' in a cross-national research project. *Studies Educ. Ethnography* **11**, 207–225. https://doi.org/10.1016/S1529-210X(05)11012-2 (2006).
49. Hegade, P., Patil, N. & Bidari, I. Principles of elective design with industry-institute collaboration. *Journal of Engineering Education Transformations* **34**, 384–390, https://doi.org/10.16920/jeet/2021/v34i0/157184 (2021).

Acknowledgements
The authors would like to express their sincere gratitude to the administration of Kryvyi Rih State Pedagogical University for supporting this research initiative. We are particularly grateful to the Faculty Deans who facilitated access to student participants and provided valuable institutional context. Special thanks to the 1,089 students who generously contributed their time and insights by completing the survey, making this study possible. We acknowledge the technical assistance provided by the University's IT Department in implementing the electronic survey platform and ensuring data security. Finally, we appreciate the guidance from the Ethics Committee members who ensured our research adhered to appropriate ethical standards for educational research.

Author contributions
SOS conceived and designed the study, developed the conceptual framework for analyzing individual educational trajectories, and drafted the primary manuscript sections. OVB designed the survey instrument, coordinated the data collection process, and contributed to the ethics compliance procedures. TAV conducted the literature review on educational data mining applications, contributed to the discussion of practical implications, and contextualized findings within the Ukrainian higher education framework. ISM performed the advanced statistical analyses including cluster analysis and regression modeling, assisted in developing the theoretical models, and validated the analytical approaches. PPN implemented the data preprocessing procedures, developed the text mining algorithms for qualitative data analysis, and created the data visualizations. All authors contributed to the interpretation of results, critically reviewed multiple manuscript drafts, and approved the final version for submission.

Funding
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to S.O.S.