Non-Profit Joint-Stock Company «L.N. Gumilyov Eurasian National University»
«Institute of Digital Sciences and Artificial Intelligence»
Department Information Systems education department

Syllabus
INFS 43002 Data analysis and processing module

on the discipline
DM 4220 Data mining for students of education program
6B06103 Information systems

Астана
2026

This document, in accordance with paragraph 1 of Article 7 of the Law of the Republic of Kazakhstan dated January 7, 2003 «On electronic document and electronic digital signature» is equivalent to a document on paper

Document signed

№	Full name	Position
1	Lamasheva Zhanar	senior lecturer
2	Mukhanova Ayagoz	Head of the cafedra
3	Davletova Ainash	Associate Professor (Associate Professor)
4	Orazaliyev Bakhytzhan	Director
5	Seilov Shakhmaran	Director of the Institute

Developer:
Ламашева Жанар Бейбутовна, senior lecturer

Syllabus on the discipline DM 4220 Data mining
(discipline code and name)
designed on the base of education program 6B061 – Information and communication technology
(education program code and name)

Considered at the department meeting Information systems
Record №11 «5» 05 2026
Approved at the meeting of the Educational and Methodological Commission of the Institute of Digital Sciences and Artificial Intelligence
Record №9 «14» 05 2026.

EXPLANATORY NOTE

1. Description

Cloud Technologies is a course designed to provide students with a comprehensive understanding of modern cloud computing concepts, architectures, and services. The course covers the fundamental principles of cloud infrastructure, virtualization, networking, data storage, security, DevOps practices, containerization, and serverless computing. Students explore leading cloud platforms and learn how to design, deploy, manage, and secure cloud-based solutions. The course combines theoretical foundations with hands-on laboratory exercises, enabling students to gain practical experience in configuring cloud resources, deploying applications, managing databases, automating infrastructure, and optimizing cloud costs.

Aim	Learning objectives*	Learning outcomes
The course introduces students to the technology of Data Mining, discusses in detail the methods, tools and application of Data Mining. The differences of Data Mining from the classical statistical methods of analysis and OLAP-systems are discussed, the types of patterns revealed by Data Mining are considered. Describes the scope of Data Mining. Data Mining methods are discussed in detail: neural networks, decision trees, bounded-search methods, genetic algorithms,	RT13 – Organize the processing of big data, provide storage of big data, create software tools for data processing, determine the data criteria for searching and extracting data, apply methods for developing SQL queries, use methods for quickly searching and processing data, apply the developed technologies for machine learning and develop software tools for data analysis. Creation and management of information systems that use artificial intelligence and machine learning. Work with various technologies such as neural networks, genetic algorithms and expert systems.	As a result of mastering the discipline, the student must: Student should know: • Basic concepts, goals, and principles of cloud technologies and cloud computing. • Main cloud service models: IaaS, PaaS, SaaS, and deployment models: public, private, hybrid cloud. • Principles of cloud infrastructure architecture, virtualization, networking, storage, and databases. • Basics of cloud security, identity and access management, monitoring, backup, and cost optimization. • Modern cloud platforms and tools, including Yandex Cloud services, Docker, Kubernetes, Terraform, and serverless technologies. Student must be able to: • Deploy and configure cloud resources such as virtual machines, networks, storage, and databases. • Design and implement scalable, reliable, and secure cloud infrastructure. • Configure access rights, security groups, monitoring, and alerts in a cloud environment. • Deploy containerized applications and manage cloud services using automation tools. • Analyze cloud resource usage and propose solutions for cost optimization. Student will gain skills:

evolutionary programming.		• Practical skills in working with cloud platforms and cloud infrastructure services. • Skills in configuring virtual machines, virtual networks, storage systems, databases, and serverless services. • Skills in using DevOps tools for infrastructure automation and application deployment. Practical experience in monitoring, securing, documenting, and optimizing cloud-based solutions.

**According to educational program 6B06103 Information systems, 2023*

2. Prerequisites

Knowledge, skills and abilities acquired during the study of the following disciplines are necessary to master this discipline: Discrete mathematics and theory of algorithms, Introduction to Machine Learning, Math statistics.

Postrequisites

The knowledge, skills and abilities acquired during the study of the discipline are necessary for the learning of the following disciplines: Completion and defense of Bachelor degree diploma project.

3. Extract from the curriculum

Year 4
Semester 7
Number of ECTS 5

Types of classes	Total number of hours
Lectures	15
Laboratory practicals	30
Independent work of a student (ISW)	105
Total	150

4. Thematic plan of the discipline by modules

(in academic hours)

Module №	Name of the module
1	Pattern Discovery in Data Mining
2	Cluster Analysis in Data Mining

Lectures				
Week №	**Module №**	**Theme of the lecture**	**Number of hours**	**Types and methods of training**

1	1	Introduction to Data Mining"	1	Explanatory-illustrative method Demonstration, showing, working with presentations, diagrams, videos.
2	1	Python Libraries for Data Science	1	Interactive methods E-learning Using LMS (Moodle, Google Classroom), video lessons, online tests.
3	1	Data Analysis with Pandas	1	Explanatory-illustrative method Demonstration, showing, working with presentations, diagrams, videos.
4	1	Fundamentals of Linear Algebra	1	Interactive methods E-learning Using LMS (Moodle, Google Classroom), video lessons, online tests.
5	2	Mathematical Analysis and Optimization	1	Explanatory-illustrative method Demonstration, showing, working with presentations, diagrams, videos.
6	2	Probability Theory and Statistics	1	Explanatory-illustrative method Demonstration, showing, working with presentations, diagrams, videos.
7	2	Statistical Methods in Data Science	1	Interactive methods mini-quizzes, question-answer sessions during lectures. Using online tools (Mentimeter, Kahoot, Padlet) to engage students in the material.
8	2	Introduction to Machine Learning	1	Interactive methods mini-quizzes, question-answer sessions during lectures. Using online tools (Mentimeter, Kahoot, Padlet) to engage students in the material.
9	2	Classification Algorithms	1	Explanatory-illustrative method Demonstration, showing, working with presentations, diagrams, videos.
10	2	Decision Trees and Ensemble Methods	1	Design with comments The teacher explains the project structure step by step: from HTML markup to connecting external libraries and deployment.
11	2	Unsupervised Learning	1	Interactive methods mini-quizzes, question-answer sessions during lectures. Using online tools (Mentimeter, Kahoot, Padlet) to engage students in the material.
12	2	Time Series Analysis	1	Design with comments The teacher explains the project structure step by step: from HTML markup to connecting external libraries and deployment.
13	2	Machine Learning in Production	1	Interactive methods mini-quizzes, question-answer sessions during lectures.

				Using online tools (Mentimeter, Kahoot, Padlet) to engage students in the material.
14	2	Introduction to Deep Learning	1	Explanatory-illustrative method Demonstration, showing, working with presentations, diagrams, videos.
15	2	Deep Learning Frameworks	1	Interactive methods mini-quizzes, question-answer sessions during lectures. Using online tools (Mentimeter, Kahoot, Padlet) to engage students in the material.
TOTAL			15	

Laboratory practicals				
Week №	**Module №**	**Theme of laboratory practicals**	**Number of hours**	**Types and methods of training**
1	1	Frequent Pattern Discovery and Mining Techniques.	2	Practical methods. Practical exercises, working with code, frontend and backend development. Completing an individual task.
2	1	Pattern Evaluation and Mining a Diverse Spectrum of Patterns.	2	Step-by-step programming (algorithmization method) Practical exercises, working with code, frontend and backend development. Completing an individual task
3	1	Sequential, Spatiotemporal, and Trajectory Pattern Mining.	2	Step-by-step programming (algorithmization method). Practical exercises, working with code, frontend and backend development. Completing an individual task
4	1	Phrase Mining and Advanced Pattern Discovery Techniques.	2	Step-by-step programming (algorithmization method). Practical exercises, working with code, frontend and backend development. Completing an individual task
5	2	Mining Frequent Patterns and Association Rules	2	Step-by-step programming (algorithmization method). Practical exercises, working with code, frontend and backend development. Completing an individual task
6	2	Frequent Itemset Mining Using the Apriori Algorithm	2	Case study method Students are given real or close to real tasks: to lay out a landing page, implement a form with validation
7	2	Mining Frequent Patterns Using Vertical Data Format and Pattern-Growth Approaches	2	Step-by-step programming (algorithmization method). Practical exercises, working with code, frontend and backend development. Completing an individual task

8	2	Cluster Analysis and Clustering Method Selection	2	Step-by-step programming (algorithmization method). Practical exercises, working with code, frontend and backend development. Completing an individual task
9	2	Similarity and Distance Measures for Clustering	2	Explanatory and illustrative method Practical exercises, working with code, frontend and backend development. Completing an individual task
10	2	Sequential Pattern Mining Using GSP, SPADE, and PrefixSpan	2	Explanatory and illustrative method Practical exercises, working with code, frontend and backend development. Completing an individual task
11	2	Mining Closed Sequential Patterns	2	Explanatory and illustrative method Practical exercises, working with code, frontend and backend development. Completing an individual task
12	2	Spatiotemporal and Trajectory Pattern Mining	2	Explanatory and illustrative method Practical exercises, working with code, frontend and backend development. Completing an individual task
13	2	Phrase Mining Using ToPMine and SegPhrase	2	Explanatory and illustrative method Practical exercises, working with code, frontend and backend development. Completing an individual task
14	2	Mining Contiguous Sequential Patterns in Text Data	2	Project method Laboratory work is designed as a mini-project: portfolio site, online store, blog.
15	2	Advanced Applications of Pattern Discovery	2	Project method Laboratory work is designed as a mini-project: portfolio site, online store, blog.
TOTAL			30	

ISW				
Week №	**Module №**	**Theme of ISW Deadlines**	**Number of hours**	**Types and methods of training**
1	1	Applications of Frequent Pattern Discovery in Data Mining	7	Research method Students are asked to research the work of new technologies or libraries (for example, React, Vue, Bootstrap), compare them and apply them in a project
2	1	Comparative Analysis of Frequent, Closed, and Max Patterns	7	Reproductive methods Completing typical tasks - practicing skills according to a model

3		Association Rule Mining in Real-World Datasets		Problem-oriented methods, Solving practical problems – developing logic, professional thinking.
4	2	Performance Analysis of Frequent Pattern Mining Algorithms	7	Problem-oriented methods, Solving practical problems – developing logic, professional thinking.
5	2	Comparative Study of Apriori, Vertical Mining, and Pattern-Growth Methods	7	Problem-oriented methods, Solving practical problems — developing logic, professional thinking.
6		Interestingness Measures for Pattern Evaluation	7	Problem-oriented methods, Solving practical problems — developing logic, professional thinking.
7	3	Multi-Level and Multi-Dimensional Association Pattern Analysis	**7**	Reproductive methods Completing typical tasks — practicing skills according to a model
8	3	Applications of Cluster Analysis in Data Science	7	Problem-oriented methods, Solving practical problems — developing logic, professional thinking.
9		Distance and Similarity Measures: A Comparative Study	7	Problem-oriented methods, Solving practical problems – developing logic, professional thinking.
10	4	Evaluation of Partitioning and Hierarchical Clustering Methods	7	Problem-oriented methods, Solving practical problems — developing logic, professional thinking.
11	4	Applications of Sequential Pattern Mining	7	Problem-oriented methods, Solving practical problems — developing logic, professional thinking.
12		Closed Sequential Patterns and Their Practical Significance	7	Reproductive methods Completing typical tasks — practicing skills according to a model
13	5	Spatiotemporal Data Mining and Trajectory Analytics	7	Reproductive methods Completing typical tasks — practicing skills according to a model
14	6	Phrase Mining Techniques for Text Data Analysis	7	Problem-oriented methods, Solving practical problems — developing logic, professional thinking.
15		Emerging Trends and Applications of Pattern Discovery	7	Problem-oriented methods, Solving practical problems – developing logic, professional thinking.
TOTAL			**105**	

5. Brief organizational-methodological characteristics of the discipline

Types of control of educational achievements
Midterm examination 1 oral examination
Midterm examination 2 oral examination
Summative assessment: Written Exam

Course policies and procedures:
- Mandatory attendance of all classes by the trainee according to the schedule;
- Preliminary preparation for classes;
- Timely execution and delivery of IWS;
- Preparation for all types of classes should be independent and creative;
- Active work and creativity during classes;
- Participation in all types of control;
- Commitment to the University's Academic Integrity Policy.

For students with special needs, lecture presentations are provided, additional time is allocated for passing knowledge rating control and intermediate certification

To conduct online consultations, an online channel has been created in TIMS

A separate package of exam papers and individual assignments for midterm control has been developed for students with special educational needs

Information for students with special educational needs.

Students are required to notify the teacher in a timely manner to discuss the conditions necessary for successful completion of the course and achievement of educational goals. Students with special educational needs are provided with access to information in adapted formats using both distance and face-to-face educational technologies.

Such students are provided with lecture materials in the form of presentations, and are also allocated additional time to pass the midterm assessment and midterm certification. The forms of current and final certification are selected taking into account the individual psychophysical characteristics of the student (oral, written, computerized, testing, etc.). If necessary, additional time is provided to prepare an answer to the exam. Additionally, you can contact the Department of Social and Civil Development or the university library, where assistance in working with educational literature is also provided.

Academic Integrity Students and the teacher must adhere to the principles of academic integrity prescribed in the internal document of the university "Academic Integrity Policy of the L.N. Gumilyov Eurasian National University. Gumilyov", please read the university's plagiarism policy https://enu.kz/ru/o-enu/osnovopolagavushchie dokumenty/

Students who commit an act of academic dishonesty may receive 0 points for an assignment or course.

6. Educational and methodological coverage of the discipline

№	Author, name, publisher, year of publication	Information carrier	Available in stock (p)		
			In the library	Global digital libraries	At the department
1	2	3	4		5
Basic literature					

1	Богданов, Е. П. Интеллектуальный анализ данных : практикум для магистрантов направления 09.04.03 «Прикладная информатика» профиль подготовки «Информационные системы и технологии корпоративного управления» / Е. П. Богданов. - Волгоград : ФГБОУ ВО Волгоградский ГАУ, 2019. - 112 с.	e-Book	-	https://znanium.ru/catalog/product/1087885	-
2	Ланских, Ю. В. Интеллектуальный анализ данных : учебное пособие / Ю. В. Ланских, В. Г. Ланских. — Киров : ВятГУ, 2023. — 240 с.	e-Book	-	https://e.lanbook.com/book/408569	-
3	Митяков, Е. С. Искусственный интеллект и машинное обучение : учебное пособие для вузов / Е. С. Митяков, А. Г. Шмелева, А. И. Ладынин. — 2-е изд., стер. — Санкт-Петербург : Лань, 2026. — 252 с. — ISBN 978-5-507-51198-3. — Текст : электронный // Лань : электронно-библиотечная система.	e-Book	-	https://e.lanbook.com/book/507451	-
4	Баланов, А. Н. Машинное обучение и искусственный интеллект : учебное пособие для вузов / А. Н. Баланов. — 3-е изд., стер. — Санкт-Петербург : Лань, 2026. — 172 с. — ISBN 978-5-507-54962-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL:	e-Book	-	https://e.lanbook.com/book/513580	-
5	Ozdemir Sinan. Principles of data science : a beginner's guide to essential math and coding skills for data fluency and machine learning / Sinan Ozdemir. - 3rd edition. - Birmingham : Packt Publishing, 2024. - XIX, 305 с. : ил. - IEEE. Packt Publishing. Computing science. - Указ.: с. 293-300. - ISBN 978-1-83763-630-3.	Monograph	-	https://ieeexplore.ieee.org/document/10460884	-
Additional literature					
6	Data Science и интеллектуальный анализ данных / Д.М. Назаров, С.В. Бегичева, Д.Б. Ковтун, А.Д. Назаров. - Москва : Ай Пи Эр Медиа, 2023. - 305, [1] с. : ил., табл. - Библиогр.: с. 301-305. - ISBN 978-5-4497-2095-5.	Monograph	1	https://library.enu.kz/MegaPro/Download/MObject/105636	-
7	The definitive guide to data integration : unlock the power of data integration to efficiently manage, transform, and analyze data / Pierre-Yves Bonnefoy, Emeric Chaize, Raphael Mansuy, Mehdi Tazi. - Birmingham : Packt Publishing, 2024. - XXII, 467 с. : ил. - IEEE. Packt Publishing. Computing science. - Указ.: с. 445-462. - ISBN 978-1-83763-191-9. - Текст : электронный.	Study guide	-	https://library.enu.kz/MegaPro/UserEntry?Action=Link_FindDoc&id=181600&idb=0	-
Electronic and Internet resources					
8	https://stepik.org/course/62243/promo				

The list of basic literature includes basic textbooks and manuals (usually three or four titles) in the disciplines of the socio-humanitarian profile for the last 5 years, in other areas - for the last 10 years.

Additional literature should include no more than 10 titles.

When specifying electronic and Internet resources, it is necessary to specify the short name of the information to which the link is given, then place the active link. The number of links to Internet resources should not exceed 10 titles.

7. System of learning outcomes assessment of a student

Students' knowledge, skills and abilities are evaluated according to the following system

Grade by letter system	Digital equivalent	Percentage content	Grade by traditional system	Grade criteria
A	4,0	95-100	Excellent	**Grade A** is awarded when the student provides a complete, detailed, and technically accurate answer, demonstrates a deep understanding of data mining concepts, pattern discovery, association analysis, clustering, sequential pattern mining, text mining, and advanced pattern discovery techniques. The student confidently applies theoretical knowledge to practical data mining problems, explains relationships between algorithms and data characteristics, evaluates mining results using appropriate measures, and justifies methodological choices. Knowledge is demonstrated within the broader context of data science, machine learning, statistics, and database systems. The answer is logical, well-structured, evidence-based, and presented using professional data mining terminology.
A-	3,67	90-94		**Grade A-** is awarded when the student provides a complete and detailed answer, demonstrates a strong understanding of data mining methods and algorithms, and correctly explains key concepts and applications. The answer is logically structured and uses appropriate technical terminology. Minor inaccuracies may occur but are independently corrected during the response.
B+	3,33	85-89	Good	**Grade B+** is awarded when the student provides a complete answer and demonstrates a good understanding of frequent pattern mining, association rules, clustering methods, sequential pattern mining, and pattern evaluation measures. The answer is generally accurate and logically organized. Minor shortcomings may be corrected with instructor guidance.
B	3,0	80-84		**Grade B** is awarded when the student demonstrates a solid understanding of core data mining techniques and is able to explain major concepts and practical applications. The answer is structured and technically sound but may contain minor errors or omissions that are corrected with instructor assistance.
B-	2,67	75-79		**Grade B-** is awarded when the student provides a detailed answer and demonstrates understanding of the fundamental concepts of data mining, including pattern discovery,

				association analysis, clustering, and sequential pattern mining. Some inaccuracies or omissions are present but can be corrected through guiding questions.
C+	2,33	70-74		**Grade C+** is awarded when the student provides a generally complete answer but demonstrates limited consistency in explaining data mining concepts or algorithmic approaches. The student shows an understanding of essential principles but may make one or two errors in terminology or technical explanations.
C	2,0	65-69	Satisfactory	**Grade C** is awarded when the student demonstrates only a basic understanding of data mining concepts. The answer lacks sufficient detail, contains inaccuracies in terminology, and shows difficulty in explaining relationships between data mining tasks, algorithms, and evaluation measures. Practical examples can be provided only with instructor support.
C-	1,67	60-64		**Grade C-** is awarded when the answer is incomplete and demonstrates significant gaps in understanding data mining concepts. Major errors are made in explaining mining algorithms, pattern evaluation methods, clustering techniques, or sequential pattern mining approaches. Conclusions are absent, and practical application of knowledge is not demonstrated.
D+	1,33	55-59		**Grade D+** is awarded when the student provides an incomplete and poorly structured answer with substantial errors in data mining terminology and concepts. Understanding of pattern discovery, association analysis, clustering, and mining applications is very limited. Connections between concepts become apparent only after significant instructor prompting.
D	1,0	50-54		**Grade D** is awarded when the student's answer consists of fragmented and disconnected knowledge with serious technical errors. The student is unable to explain basic data mining concepts, algorithms, or applications and cannot relate them to practical data analysis tasks. Additional questions do not significantly improve the response.
FX	0,5	25-49		**Grade FX/F** is awarded when the student demonstrates insufficient knowledge of the fundamental concepts of Data Mining, including pattern discovery, association rule mining, clustering, sequential pattern mining, text mining, and advanced pattern discovery techniques. The student has not mastered a significant portion of the course material, makes fundamental errors, fails to complete required practical assignments and laboratory work, and has not adequately studied the recommended course resources.
F	0	0-24		