ХАЛЫҚАРАЛЫҚ УНИВЕРСИТЕТІ · МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ · INTERNATIONAL UNIVERSITY

Faculty of «Computer technology and cybersecurity»
Department of «Mathematical and computer modeling»

APPROVED BY
Vice-rector for academic affairs
«International Information
Technology University» JSC
__________Mustafina A.K.
«19» ___03_____2024

7M061106

Data Science

CATALOGUE OF ELECTIVE DISCIPLINES

2024-2028 year of admission

2024 y.

The catalogue of elective disciplines for the EP pf the MCM department is developed on the basis of the working curriculum of the EP " 7M061106 Data Science".

The catalogue of elective disciplines was discussed at a meeting of the department of Mathematical and computer modeling

minutes No. 9 from «14» 03 2024 y.

Head of the Department	signature	Ydyrys A.Zh., PhD Full name, position, degree
CED compiler	signature	Satybaldina A.N. Full name, position, degree

1 TERMS AND ABBREVIATIONS

1. 1 Educational program is a single set of basic characteristics of education, including goals, results and content of training, the organization of educational process, ways and methods for their implementation and criteria for assessing learning outcomes.
The content of educational program of higher education consists of three cycles of disciplines - general education disciplines (hereinafter - GED), basic disciplines (hereinafter - BD) and core disciplines (hereinafter - CD).
The cycle of GED includes disciplines of the compulsory component (hereinafter - CC), the university component (hereinafter - UC) and (or) the component of choice (hereinafter - COC). BD and CD include disciplines of UC and COC.

1. 2 Catalogue of elective disciplines (CED) is a systematic annotated list of all COC disciplines, for the entire training period, containing a brief description indicating the purpose of study, a summary of main sections and expected learning outcomes. CED reflects the prerequisites and postrequisites of each academic discipline. It should provide the students with the possibility of an alternative choice of elective disciplines for the formation of an individual educational trajectory.

On the basis of educational program and CED, the students develop individual curricula with the help of advisers.

1. 3 Individual curriculum (IC) is a curriculum formed by the students independently with the help of an advisor for each academic year on the basis of the educational program, the catalogue of elective disciplines or modules;

IC defines an individual educational trajectory of each student separately. It includes disciplines and types of educational activities (internship, experimental research, forms of final certification) of the compulsory component (CC), the university component (UC) and the component of choice (COC).

1. 4 Advisor is a teacher who performs the functions of an academic mentor of a student (according to the appropriate educational program) and assists in choosing a learning path (creating an individual curriculum) and mastering the educational program during the training period.

1. 5 The university component is a list of compulsory educational disciplines determined by the university independently for the mastering of the educational program.

1. 6 The component of choice is a list of academic disciplines and the corresponding minimum amounts of academic credits offered by the university and independently chosen by students in any academic period, taking into account their prerequisites and postrequisites.

1. 7 Elective disciplines are educational disciplines that are a part of the university component and the component of choice in the framework of established academic credits, introduced by organizations of education reflecting the individual preparation of students and taking into account the specifics of socio-economic development, the needs of a particular region and established scientific schools.

1. 8 Postrequisites are the disciplines and (or) modules and other types of academic work, the study of which requires knowledge, skills and competencies acquired at the end of the study of this discipline and (or) modules;

1. 9 Prerequisites are the disciplines and (or) modules and other types of educational work containing knowledge, abilities, skills and competencies necessary for the mastering of the studied discipline and (or) modules;

1. 10 Competencies are the ability of the practical use of acquired knowledge and skills in professional activities.

2 ELECTIVE DISCIPLINES

No.	Cycle of discipline	Code of discipline	Name of the discipline	Semester	Number of credits	Prerequisites
			1 year			
1	BD CC	CED7501	Elective discipline #1	2	5	
		MAT7501	Applied multivariate statistical analysis			Probability theory and mathematical statistics
		MAT7511	Bayesian statistics and analysis			Probability theory and mathematical statistics
2	BD CC	CED7502	Elective discipline #2	1	5	
		MAT7522	Deep Learning in computer vision..			Discrete Mathematics, Computational Mathematics
		MAT7532	Oracle database 11g: PLSQL Fundamentals (Oracle)..			Computer system architecture
		MAT7502	Databases: Advanced			Algorithms and Data Structures
		MAT7542	Introduction to HPC with Mpi for DS			Introduction to Programming, Object-oriented programming
		MAT7552	Numerical simulations using Python for DS			Algorithms and programming languages, numerical analysis, introduction to data science
3	ПД КВ	CED7506	Elective discipline #6	2	5	
		MAT7556	Mathematical modeling of cybersecurity			Discrete Mathematics, Number Theory, Introduction to Cryptography
		MAT7566	Advanced Data Analysis.			Algorithms and data structures
		MAT7536	Finite-difference methods for partial differential equations			Numerical methods, mathematical physics equations, programming languages and algorithms.
		MAT7576	Optimization methods for DS applications			Numerical methods
			2 year			
3	PD CC	CED7504	Elective discipline #4	3	4	
		MAT7544	Mathematical modeling of enumerative combinatorics			Linear Algebra, Discrete mathematics and Mathematical Logic, Probability Theory and Statistics
		MAT7554	Artificial Intelligence for Social Good..			Programming in Python
5	PD CC	CED7505	Elective discipline #5	3	5	
		MAT7545	Parallel computation			Introduction to Programming, Object-oriented programming

		MAT7555	Machine learning methods for solving inverse problems			Numerical methods, mathematical physics equations, programming languages and algorithms.

3 DESCRIPTION OF ELECTIVE DISCIPLINES

Description of discipline	
Code of discipline	MAT7501
Name of discipline	Applied multivariate statistical analysis
Number of credits (ESTS)	5
Prerequisites	Probability theory and mathematical statistics
Postrequisites	Exploratory data analysis and visualization
The aim of study of a discipline	The aim is to formulate and solve applied economic problems of large dimension using quantitative methods of transformation of the initial system of indicators in the framework of multidimensional statistical and econometric models.
Brief course description (main sections)	The discipline "Applied multidimensional statistical Analysis" is a systematic presentation of quantitative methods of statistical assessment of the economic condition of business entities, aimed at developing multidimensional thinking, skills and abilities to use methods of applied statistical modeling in the practical work of specialists engaged in business analytics. It is an instrumental and analytical basis for the justification and adoption of economic decisions.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	As a result of studying the discipline, the student must: know: econometrics; applied software packages for processing statistical data; be able to: build multidimensional statistical models of economic objects and indicators; solve economic problems by mathematical methods using computer and software tools based on real data; apply the knowledge gained in practical analytical work.

Description of discipline	
Code of discipline	MAT7511
Name of discipline	Bayesian statistics and analysis
Number of credits (ESTS)	5
Prerequisites	Probability theory and mathematical statistics
Postrequisites	-
The aim of study of a discipline	Introduce students to Bayesian statistics, in which conclusions about parameters or hypotheses are updated as evidence accumulates.
Brief course description (main sections)	The course will apply Bayesian methods to several practical problems to show end-to-end Bayesian analysis, which goes from asking a question to constructing models, identifying a priori probabilities, and implementing a finite a posteriori distribution in R. In addition, the course will present valid domains, Bayesian comparisons of averages and proportions, Bayesian regression and inference using multiple models, and a discussion of Bayesian forecasting.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	After completing the course, the learner will: understand and define the concepts of prior and subsequent probability and determine how they relate to each other; make optimal decisions based on Bayesian statistics and compare multiple hypotheses using Bayesian factors; implement Bayesian model averaging, interpret Bayesian multiple linear regression, and understand its relationship to the linear regression frequency approach.

Description of discipline	
Code of discipline	MAT7522
Name of discipline	Deep Learning in computer vision
Number of credits (ESTS)	5
Prerequisites	Discrete Mathematics, Computational Mathematics
Postrequisites	Deep learning of inverse problems
The aim of study of a discipline	Introduction to the basic ideas underlying deep learning from the point of view of applied mathematics.
Brief course description (main sections)	Multi-layer artificial neural networks are becoming an increasingly common tool in a variety of applications. At the heart of this deep learning revolution are familiar concepts from applied and computational mathematics, especially calculus, approximation theory, optimization, and linear algebra. This course focused on four fundamental questions: What is a deep neural network? How is the network trained? What is the stochastic gradient method? How to use the deep neural network algorithm to solve problems with partial differential equations. The course illustrates these ideas with a short MATLAB code that configures and trains the network. The course also demonstrates the use of state-of-the-art software to solve a large-scale image classification problem.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	They will have an understanding of neural networks; be able to train the network; be able to use the neural network algorithm to solve partial differential equations.

Description of discipline	
Code of discipline	MAT7532
Name of discipline	Oracle database 11g: PLSQL Fundamentals (Oracle)
Number of credits (ESTS)	5
Prerequisites	Architecture of computer systems
Postrequisites	-
The aim of study of a discipline	The purpose of mastering the discipline "Parallel Programming" is to form students ' theoretical knowledge and practical skills in programming parallel and distributed systems. Considerable attention is paid to issues related to the development of basic knowledge in the field of architecture of modern multiprocessor computing systems, parallel information processing, technologies for organizing parallel computing on multiprocessor computing systems with distributed or shared RAM.
Brief course description (main sections)	The following are the main topics / sections that will be covered in the course: thread creation, synchronization, common errors, profiling, thread pools and templates, clusters, memory models, linearizability.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	As a result of mastering the discipline, the student must: - Know the main approaches to the development of parallel programs; the main technologies and models of parallel programming; methods for creating parallel programs for typical tasks of multithreaded programming. - Be able to create parallel programs for computing systems with distributed, shared RAM; perform parallelization of computational algorithms; build a model for executing parallel programs; evaluate the effectiveness of parallel computing; analyze the complexity of calculations and the possibility of parallelization of the developed algorithms; apply general schemes for developing parallel programs for implementing their own algorithms; evaluate the main parameters of the resulting parallel programs, such as acceleration, efficiency and scalability. - Have the skills (gain experience) to create parallel programs for computing systems with distributed, shared RAM; build parallel analogs of computational algorithms.

Description of discipline	
Code of discipline	MAT7502
Name of discipline	Databases: Advanced
Number of credits (ESTS)	5
Prerequisites	Algorithms and data structures
Postrequisites	-
The aim of study of a discipline	Course objectives: 1. Explain database management software for data-intensive application development 2. Introduce Oracle 3. Explain the Oracle SQL development environments; 4. Demonstrate practical practice to strengthen the fundamental concept; 5. Ensure the practice of executing SQL queries; 6. Describe the built-in SQL functions;
Brief course description (main sections)	This course introduces students to database systems. The course explains what a database system is, and then proceeds to most of the training material for learning relational database systems—databases designed according to a relational (or tabular) model. From data abstraction, the course then moves on to transaction management with some additional materials to improve query performance. Finally, there is the introduction of current trends in database system design, which also reflect recent developments in the broader history of data storage technology.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	After completing the course, students should know: * The basic ideas of the theory of database systems * Defining basic SQL queries • The process of creating and processing requests * Basic SQL concepts in Oracle * Comprehensive understanding of data models, relational database management systems * SQL functions and multimedia database applications * Critical database privacy and security issues

Description of discipline	
Code of discipline	MAT7542
Name of discipline	Introduction to HPC with Mpi for DS
Number of credits (ESTS)	5
Prerequisites	Introduction to Programming, Object-oriented programming
Postrequisites	-
The aim of study of a discipline	Development of parallel algorithms for distributed memory using a standard interface called the Message Passing Interface (MPI).
Brief course description (main sections)	This soft introduction to High Performance Computing (HPC) for data science using the Message Passing Interface (MPI) standard was developed as a first course in parallel programming on distributed memory models and requires only basic programming concepts. Divided into two parts, the first part covers high-performance computing using C++ with the Message Passing Interface (MPI) standard, followed by the second part, which provides high-performance data analytics on computer clusters.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	The ability to develop parallel algorithms and then program your parallel algorithm using C ++ with C binding to MPI. Knowledge that will introduce you to the richness of the HPC and DS domains and hopefully facilitate their interaction

Description of discipline	
Code of discipline	MAT7552
Name of discipline	Numerical simulations using Python for DS
Number of credits (ESTS)	5
Prerequisites	Algorithms and programming languages, numerical analysis, Introduction to Data Science
Postrequisites	Machine Learning 2, Neural Network Retraining Problem, Data augmentation
The aim of study of a discipline	Students with a degree in numerical modeling using machine learning and deep learning. Students will develop appropriate programming skills. Students will demonstrate proficiency in numerical methods using machine learning and deep learning algorithms. Students will develop the ability to create and evaluate models based on data. Students will perform numerical methods as machine learning and deep learning algorithms using professional Python software. Students will demonstrate skills in data science.
Brief course description (main sections)	Do you want to learn how to solve partial differential equations by numerical methods and how to convert them into Python codes? This course gives you a basic understanding of how to apply techniques such as the finite difference method, the pseudospectral method, the linear and spectral element method to a one-dimensional (or two-dimensional) scalar wave equation. The mathematical output of the computational algorithm is accompanied by Python codes embedded in Jupyter notebooks. In a unique setup, you can see how mathematical equations are converted into computer code and the results are visualized. The emphasis is on illustrating the fundamental mathematical components of various numerical methods (e.g., Taylor series, Fourier series, differentiation, function interpolation, numerical integration) and comparing them. You will be provided with strategies on how to make sure that your solutions are correct, such as benchmarking with analytical solutions or convergence tests. The mathematical aspects are complemented by a basic introduction to wave physics, sampling, meshes, parallel programming, and computational models. Finally, we will look at the use of numerical methods for machine learning and deep learning and their applications for solving specific problems in the field of data science.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	Students who have successfully completed the course will be able to: - Understand the basic definitions of the finite difference method and its applications to acoustic wave equations. - Understand how Taylor series can be used to obtain finite difference operators of low and high order and accuracy as deep learning algorithms. - Learn how to use the Taylor series to compute the basic error terms of finite difference approximations and their applications for forward and backward propagation. - Know some scientific problems in which finite-difference approximations play an important role in gaining knowledge. - Learn about neural networks, work with optimization algorithms, train and test given models. - Learn about the special applications of Deep Neural Networks: Seismoacoustic Problems.

	Learn about the basics of Convalutional Neural Networks, special applications: seismoacoustic problems.

Description of discipline	
Code of discipline	MAT7556
Name of discipline	Mathematical modeling of cybersecurity
Number of credits (ESTS)	4
Prerequisites	Discrete Mathematics, Number Theory, Introduction to Cryptography
Postrequisites	Advanced Cybersecurity Program and Dissertation Work
The aim of study of a discipline	To introduce graduate students to important branches of cryptographic methods and their applications in network security for the use of the cybersecurity space.
Brief course description (main sections)	Discipline on protection against criminal or unauthorized use of electronic data or measures taken to do so. After human resources, information is the most important asset of an organization. Security and risk management is data-driven. All efforts to protect systems and networks are aimed at achieving three results: availability, integrity and confidentiality of data. No infrastructure security measures are 100% effective.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	Cryptography can be used to protect the privacy of financial data or personal records, whether they are stored or in transit. Cryptography can also be used to verify the integrity of data by identifying whether it has been altered and identifying the person or device that sent it. Encryption methods are crucial for the development and use of national and global information and communication networks and technologies, as well as for the development of electronic commerce.

Description of discipline	
Code of discipline	MAT7566
Name of discipline	Advanced Data Analysis
Number of credits (ESTS)	5
Prerequisites	Algorithms and data structures
Postrequisites	-
The aim of study of a discipline	Course objectives: 1. Explain database management software for data-intensive application development 2. Introduce Oracle 3. Explain the Oracle SQL development environments; 4. Demonstrate practical practice to strengthen the fundamental concept; 5. Ensure the practice of executing SQL queries; 6. Describe the built-in SQL functions;
Brief course description (main sections)	This course introduces students to database systems. The course explains what a database system is, and then proceeds to most of the training material for learning relational database systems—databases designed according to a relational (or tabular) model. From data abstraction, the course then moves on to transaction management with some additional materials to improve query performance. Finally, there is the introduction of current trends in database system design, which also reflect recent developments in the broader history of data storage technology.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	After completing the course, students should know: * The basic ideas of the theory of database systems * Defining basic SQL queries • The process of creating and processing requests * Basic SQL concepts in Oracle * Comprehensive understanding of data models, relational database management systems * SQL functions and multimedia database applications * Critical database privacy and security issues

Description of discipline	
Code of discipline	MAT7536
Name of discipline	Finite-difference methods for partial differential equations
Number of credits (ESTS)	5
Prerequisites	Numerical methods, mathematical physics equations, programming languages and algorithms.
Postrequisites	Special subjects of the educational program.
The aim of study of a discipline	Students ' development of finite-difference methods for solving the problem of partial differential equations.
Brief course description (main sections)	With the advent of electronic computing processes, it became possible to solve complex problems of production and technology. The initial material for electronic computers was initially the grid method, then the finite-difference scheme. The advantages of finite-difference schemes over other methods were discovered long ago, but there are also problems associated with the properties of difference derivatives, the so-called argument spreading. In a word, there are schemes that lead to the goal, there is a group of difference schemes that are unstable. It is necessary to be able to distinguish between these schemes, because many direct and inverse problems are solved using finite-difference schemes. In the discipline, different classes of difference schemes are proposed for the same problem, the stability of which is determined using numerical calculations. Students will develop an algorithm for solving the problem, make programs and perform numerical calculations. Analyzing the results and finding a stable scheme is a prerequisite for the discipline.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	Skills of drawing up a program for solving grid equation problems.

Description of discipline	
Code of discipline	MAT7576
Name of discipline	Optimization methods for DS applications
Number of credits (ESTS)	5
Prerequisites	Numerical methods
Postrequisites	Machine Learning
The aim of study of a discipline	* basic iterative methods of nonlinear optimization; * convergence of iterative methods of nonlinear optimization; • Development of algorithms and programming for nonlinear optimization methods; * Conducting computational experiments and data analysis.
Brief course description (main sections)	Students will learn how to use the method of constructing an iterative scheme of nonlinear optimization; get acquainted with the methodology of proving the convergence of iterative optimization schemes; develop algorithms for solving the problem of nonlinear optimization, as well as build a software product. They will learn how to conduct computational experiments by choosing the damping parameters of the computational process. Also, students will analyze the result and evaluate the accuracy of the developed method.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	Knowledge: Development of iterative methods of nonlinear optimization; Understanding: convergence of iterative nonlinear optimization methods; Application: software development and numerical calculations; Analysis: Analysis of the results of numerical calculations; Synthesis: using the damping parameter to increase the convergence rate of nonlinear optimization methods; Evaluation: based on numerical experiments, evaluate the accuracy of the developed numerical methods of nonlinear optimization.

Description of discipline	
Code of discipline	MAT7544
Name of discipline	Mathematical modeling of enumerative combinatorics
Number of credits (ESTS)	4
Prerequisites	Linear Algebra, Discrete mathematics and Mathematical Logic, Probability Theory and Statistics
Postrequisites	Machine Learning courses
The aim of study of a discipline	The main goal of this paper is to prove the applicability of the result on asymptotic stability on discrete - time SIS epidemics in complex networks using the combinatorial enumeration argument. Here, SIS means that nodes can be susceptible to infection, can go into an infection state, and once they recover, they can become susceptible to infection again
Brief course description (main sections)	Enumerative combinatoricity is related to counting the number of elements of a finite set S. This definition in its current form tells us little about the subject, since almost any mathematical problem can be solved in these terms. In the present enumeration problem, the elements of S usually have a fairly simple combinatorial definition and very little additional structure. It will be clear that there are many elements in S, and the main problem will be to count (or evaluate) them all, rather than, for example, finding a specific element.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	It is expected that after completing this course, the student will be introduced to the basics of the theory of enumerative combinatorics and the methods used in it. In particular, the student must - know the various standard concepts and sequences in combinatorics and their properties, - be able to reformulate and thus solve problems in terms of the concepts mentioned above, - be able to manipulate and derive properties of formal power series, - be able to derive recursions, generating functions, and provide explicit formulas for various combinatorial sequences, - be able to construct combinatorial proofs of identities and inequalities, - Be able to use Mobius inversion, inclusion-exclusion, and related sieve methods to solve enumerative problems, - be able to define and extract properties of different position classes, - be able to describe and perform calculations in the incidence algebra of cast iron, - Be able to calculate the Mobius function of the order by various methods and interpret such problems in topological terms.

Description of discipline	
Code of discipline	MAT7554
Name of discipline	Artificial Intelligence for Social Good
Number of credits (ESTS)	4
Prerequisites	Python Programming
Postrequisites	-
The aim of study of a discipline	The goal of this course is to provide students with a basic understanding of modern neural networks and their applications in computer vision and natural language understanding.
Brief course description (main sections)	Deep learning has added a huge boost to the already rapidly developing field of computer vision. With deep learning, many new applications of computer vision techniques have been introduced and are now becoming part of our daily lives. These include face recognition and indexing, photo styling, or machine vision in self-driving cars. We will cover both image and video recognition, including image classification and annotation, object recognition and image search, various object detection techniques, motion estimation, object tracking in video, human action recognition, and finally image styling, editing, and a new generation of images. During the project, students will learn how to create a facial recognition and manipulation system to understand the internal mechanics of this technology, probably the most famous and often demonstrated in films and TV shows example of computer vision and AI.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	Chatterbot, Tensorflup, Deep Learning, Natural Language Processing

Description of discipline	
Code of discipline	MAT7545
Name of discipline	Parallel computation
Number of credits (ESTS)	5
Prerequisites	Introduction to programming, object-oriented programming
Postrequisites	-
The aim of study of a discipline	The purpose of the course is to learn the basics of parallel algorithms, programming using OpenMP and MPI technologies for solving practical problems using the C ++language.
Brief course description (main sections)	The course "Parallel Computing" covers topics such as design methodologies for parallel programming systems, performance of parallel processing systems, and the use of parallel programming to create graphical interfaces based on active processes. The focus will be on practical examples of parallel programming scenarios and solutions.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	The ability to develop parallel algorithms and then program your parallel algorithm using C ++ with C binding to MPI. Knowledge that will introduce you to the richness of the HPC and DS domains and hopefully facilitate their interaction

Description of discipline	
Code of discipline	MAT7555
Name of discipline	Machine learning methods for solving inverse problems
Number of credits (ESTS)	5
Prerequisites	Numerical methods, mathematical physics equations, programming languages and algorithms.
Postrequisites	Special subjects of the educational program.
The aim of study of a discipline	students ' mastering the methods of machine learning in solving the inverse problem.
Brief course description (main sections)	With the advent of high-speed electronic computing tools, it became possible to solve inverse problems using machine learning methods. That is, knowing the consequence, you need to determine the cause. If we can determine the reason for the appearance of a particular process with a probability acceptable for practice, then in the future using this technique and the corresponding software product, we can predict the future of this process with a certain probability. In the discipline, various inverse engineering problems are considered and the problem is solved by the machine learning meta-method. Students will develop an algorithm for solving the problem, make programs and perform numerical calculations. Analyzing the results and finding the accuracy of the method is a prerequisite for the discipline.
Expected Learning Outcomes (knowledge, abilities, skills and competencies acquired by students)	Skills of applying machine learning methods in solving the inverse problem.