**Federal State Autonomous Educational Institution of Higher Education
"Peoples' Friendship University of Russia named after Patrice Lumumba"**

Academy of Engineering

(name of the main educational unit (MEU) that developed the educational program of higher education)

WORKING PROGRAM OF THE DISCIPLINE

MACHINE LEARNING AND BIG DATA MINING

(name of discipline/module)

Recommended for the field of study/specialty:

27.04.04 CONTROL IN TECHNICAL SYSTEMS

(code and name of the training area/specialty)

The discipline is mastered within the framework of the implementation of the main professional educational program of higher education (EP HE):

AIML and Space Sciences / Artificial Intelligence, Machine Learning and Space Sciences

(name (profile/specialization) of the educational institution of higher education)

2025

1. THE GOAL OF MASTERING THE DISCIPLINE

The course "Machine Learning and Big Data Mining" is part of the Master's program "Artificial Intelligence, Machine Learning and Space Sciences" in the direction 27.04.04 "Control in Technical Systems" and is studied in the 1st semester of the 1st year. The course is implemented by the Department of Mechanics and Control Processes. The course consists of 9 sections and 27 topics and is aimed at studying the fundamental principles of clustering and its basic techniques, classification and its basic techniques, frequent itemset mining and association rules, feature selection and dimensionality reduction, outlier detection, recommender systems and algorithms, ensemble clustering and classification, multimodal relational clustering, artificial neural methods and stochastic optimization, elements of statistical learning; analysis of the main methods for solving typical problems and familiarization with the area of their application in professional activities.

The purpose of mastering the discipline is to develop fundamental knowledge and skills in applying problem solving methods necessary for professional activities, to improve the general level of literacy of students in the discipline Machine Learning and Big Data Mining, to familiarize them with a new rapidly evolving field and provide practical knowledge experience in the analysis of real world data.

2. REQUIREMENTS TO THE RESULTS OF MASTERING THE DISCIPLINE

Mastering the discipline " Machine Learning and Big Data Mining" is aimed at developing the following competencies (parts of competencies) in students:

Table 2.1. List of competencies developed in students while mastering the discipline (results of mastering the discipline)

Cipher	Competence	Indicators of Competence Achievement (within the framework of this discipline)
GPC-5	Capable of conducting patent research, determining forms and methods of legal protection and defense of rights to the results of intellectual activity, managing rights to them to solve problems in the development of science, engineering and technology	OPC-5.1 Knows the methods and approaches to conducting patent research, forms and methods of legal protection and defense of rights to the results of intellectual activity; GPC-5.2 Able to manage rights to the results of intellectual activity to solve problems in the field of development of science, engineering and technology; OPC-5.3 Has knowledge of methods and approaches to conducting patent research, knows methods of legal protection and defense of rights to the results of intellectual activity.;
GPC-6	Capable of collecting and analyzing scientific and technical information, generalizing domestic and foreign experience in the field of automation and control equipment	OPC-6.1 Knows the basic methods of collecting and analyzing scientific and technical information; GPC-6.2 Able to analyze and generalize domestic and foreign experience in the field of automation and control equipment; GPC-6.3 Has knowledge of methods for collecting and analyzing scientific and technical information, and can also generalize domestic and foreign experience in the professional field.;

3. PLACE OF THE DISCIPLINE IN THE STRUCTURE OF THE EDUCATIONAL EDUCATIONAL EDUCATION

Discipline " Machine Learning and Big Data Mining" refers to the mandatory part of block 1 "Disciplines (modules)" of the educational program of higher education.

As part of the higher education program, students also master other disciplines and/or practices that contribute to the achievement of the planned results of mastering the discipline " Machine Learning and Big Data Mining".

Table 3.1. List of components of the educational program of higher education that contribute to the achievement of the planned results of mastering the discipline

Cipher	Name of competence	Previous courses/modules, practices*	Subsequent disciplines/modules, practices*
GPC-6	Capable of collecting and analyzing scientific and technical information, generalizing domestic and foreign experience in the field of automation and control equipment		Research work / Scientific research work; Undergraduate Training; Advanced Methods of Earth Remote Sensing;
GPC-5	Capable of conducting patent research, determining forms and methods of legal protection and defense of rights to the results of intellectual activity, managing rights to them to solve problems in the development of science, engineering and technology		Undergraduate Training; Research work / Scientific research work; Dynamics and Control of Space Systems;

* - filled in in accordance with the competency matrix and the SUP EP HE

** - elective disciplines/practices

4. SCOPE OF THE DISCIPLINE AND TYPES OF STUDY WORK

The total workload of the discipline "Machine Learning and Big Data Mining" is "5" credit units.

Table 4.1. Types of educational work by periods of mastering the educational program of higher education for full-time education.

Type of academic work	TOTAL, ac.h.		Semester(s) 1
Contact work, academic hours	*34*		*34*
Lectures (LC)	17		17
Laboratory work (LW)	17		17
Practical/seminar classes (SC)	0		0
Independent work of students, academic hours	*110*		*110*
Control (exam/test with assessment), academic hours	*36*		*36*
General complexity of the discipline	**ac.h.**	**180**	*180*
	credit.ed.	**5**	*5*

5. CONTENT OF THE DISCIPLINE

Table 5.1. Contents of the discipline (module) by types of academic work

Section number	Name of the discipline section		Section Contents (Topics)	Type of academic work*
Section 1	Introduction to Machine Learning and Data Mining	1.1	Introduction to modern data analysis	LC, LW
		1.2	Machine learning. Data Mining and Knowledge Discovery in Data Bases	LC, LW
Section 2	Clustering and its basic techniques	2.1	The task of clustering	LC, LW
		2.2	K-means and its modifications (k-medoids and fuzzy cmeans clustering)	LC, LW
		2.3	Density-based methods: DB-scan and Mean Shift	LC, LW
		2.4	Hierarchical clustering	LC, LW
		2.5	Quality criteria	LC, LW
Section 3	Classification and its basic techniques	3.1	The task of classification	LC, LW
		3.2	1-Rules. K-Nearest Neighbors approach	LC, LW
		3.3	Naïve Bayes. Decision Trees. Logistic Regression	LC, LW
		3.4	Quality assessment: precision, recall, F - measure, loss-function, confusion-matrix, cross- validation and learning curves (ROC, lift etc.)	LC, LW
		3.5	Multi-class and multi-label classification	LC, LW
Section 4	Frequent Itemset Mining and Association Rules	4.1	Frequent itemsets. Apriori and FP-growth algorithms	LC, LW
		4.2	Association rules. Interestingness measures: support and confidence. Closed itemsets	LC, LW
		4.3	Connection with Lattice Theory and Formal Concept Analysis. Applications	LC, LW
Section 5	Feature Selection and Dimensionality Reduction. Outlier detection	5.1	Feature selection versus feature extraction and generation	LC, LW
		5.2	Singular Value Decomposition, Latent Semantic Analysis and Principal Component Analysis. Boolean Matrix Factorization	LC, LW
		5.3	Outlier and novelty detection techniques	LC, LW
Section 6	Recommender Systems and Algorithms	6.1	Collaborative filtering. User-based and item-based methods. Slope one	LC, LW
		6.2	Association rules based and bicluster-based techniques. Quality: MAE, precision and recall assessment	LC, LW
		6.3	SVD-based approaches: pureSVD, SVD++ and time-SVD. Factorization machines	LC, LW
Section 7	Ensemble Clustering and Classification	7.1	Ensemble methods of clusterization for k-means partitions' aggregation	LC, LW
		7.2	Ensemble methods of classification: Bagging, Boosting, and Random Forest	LC, LW
Section 8	Multimodal relational clustering	8.1	Biclustering. Spectral co-clustering. Triclustering	LC, LW
		8.2	Two-mode networks. FoLCsonomies and resource-sharing systems. Multimodal approaches	LC, LW
		8.3	Applications: Community detection in Socail Network Analysis and gene expression analysis	LC, LW
Section 9	Artificial Neural Methods and Stochastic Optimization. Elements of Statistical Learning	9.1	Artificial Neural Networks. Basic ideas of Deep Learning. (Stochastic) gradient descent. Statistical (Bayesian) view on Machine learning	LC, LW

* - filled in only for FULL-TIME education: LC – lectures; LW – laboratory work; SC – practical/seminar classes.

6. LOGISTIC AND TECHNICAL SUPPORT OF DISCIPLINE

Table 6.1. Material and technical support of the discipline

Audience type	Equipping the auditorium	Specialized educational/laboratory equipment, software and materials for mastering the discipline (if necessary)
Lecture		
Computer class		
For independent work		

* - the audience for independent work of students MUST be indicated!

7. EDUCATIONAL, METHODOLOGICAL AND INFORMATIONAL SUPPORT OF THE DISCIPLINE

Main literature:
1. Han, J., Kamber, M., Pei, J. Data Mining: Concepts and Techniques, Third Edition. – Morgan Kaufmann Publishers, 2011. – 740 pp.
Further reading:
1. Hall, M., Witten, Ian H., Frank, E. Data Mining: practical machine learning tools and techniques. – 2011. – 664 pp
Resources of the information and telecommunications network "Internet":
1. RUDN University EBS and third-party EBSs to which university students have access on the basis of concluded agreements
- Electronic library system of RUDN - ELS RUDN
https://mega.rudn.ru/MegaPro/Web
- Electronic library system "University library online"http://www.biblioclub.ru
- EBS Yuraithttp://www.biblio-online.ru
- Electronic Library System "Student Consultant" www.studentlibrary.ru
- EBS "Znanium"https://znanium.ru/
2. Databases and search engines
- Sage https://journals.sagepub.com/
- Springer Nature Link https://link.springer.com/
- Wiley Journal Database https://onlinelibrary.wiley.com/
- Scientometric database Lens.org https://www.lens.org
Educational and methodological materials for independent work of students in mastering a discipline/module*:
1. Lecture course on the subject "Machine learning and big data analysis".

* - all educational and methodological materials for independent work of students are posted in accordance with the current procedure on the discipline page in TUIS!

DEVELOPER:

Associate Professor		Saltykova Olga Alexandrovna
Position, Department	*Signature*	*Surname I.O.*

HEAD OF THE DEPARTMENT:

Head of Department		Razumny Yuri Nikolaevich
Position of the Department	*Signature*	*Surname I.O.*

HEAD OF THE EP HE:

Head of Department		Razumny Yuri Nikolaevich
Position, Department	*Signature*	*Surname I.O.*