Federal State Autonomous Educational Institution of Higher Education "Peoples' Friendship University of Russia"

Academy of Engineering

(Name of the main educational unit (MEU)- of the developer of the EP of HE)

THE WORKING PROGRAM OF THE DISCIPLINE

Big Data Mining

(Name of the discipline/module)

Recommended by Methodological Council for the Education Field for the direction of training/specialization

01.04.02 Applied mathematics and computer science

(code and name of the direction of training / specialization)

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

Ballistic Design of Space Complexes and Systems

(name (profile/specialization) EP HE)

2022 г.

1. GOAL OF MASTERING THE DISCIPLINE

The purpose of studying the discipline "Big Data Mining" is to get acquainted with the basic concepts and basic algorithms for processing big data, the peculiarities of their application.

Discipline objectives:

1) providing students with knowledge about methods of preparation for processing and analyzing information obtained in the form of certain data;

2) the formation of students' skills and abilities of practical work on the adequate application of appropriate methods, techniques and procedures in preparation for processing, analysis, scientific interpretation of information;

3) the formation of ideas among students about the use of research results in practice.

2.REQUIREMENTS FOR THE RESULTS OF MASTERING THE DISCIPLINE

Mastering the discipline Big Data Mining " is aimed at developing the following competencies:

Table 2.1. The list of competencies formed by students in the course of mastering the discipline (the results of mastering the discipline)

Code	Competence	Competence achievement indicators (within this discipline)
UC-1	The ability to search, critical analysis of problem situations based on a systematic approach, develop an action strategy	UC-1.1 - Analyzes the task, highlighting its basic components;
		UC-1.2 - Determines and ranks the information required to solve the problem;
PC-1	Able to formulate goals, objectives of scientific research in the field of applied mathematics and informatics, computer technology and modern programming technologies, choose methods and means of solving problems	PC-1.1 - Possesses fundamental knowledge gained in the field of mathematical and (or) natural sciences, programming and information technology
PC-2	Able to apply modern theoretical and experimental methods for developing mathematical models of the objects and processes under study related to professional activities in the field of training and participate in their implementation in the form of software products	PC-2.1 - Knows modern theoretical and experimental methods for developing mathematical models, innovative design tools and elements of architectural solutions for information systems

3.THE PLACE OF DISCIPLINE IN THE STRUCTURE OF EP HE

The discipline Big data processing belongs to the variable part of block 1 of the curriculum.

Table 1 shows the previous and subsequent disciplines aimed at the formation of discipline competencies in accordance with the competence matrix of EP HE.

Table 3.1. The list of components of the EP HE that contribute to the achievement of the planned results of the development of the discipline

Code	Competence	Previous disciplines/modules, practices*	Subsequent disciplines/modules, practices*
UC-1	The ability to search, critical analysis of problem situations based on a systematic approach, develop an action strategy		Artificial Neural Networks (Deep Learning) Research work
PC-1	Able to formulate goals, objectives of scientific research in the field of applied mathematics and informatics, computer technology and modern programming technologies, choose methods and means of solving problems		Artificial Neural Networks (Deep Learning) Research work
PC-2	Able to apply modern theoretical and experimental methods for developing mathematical models of the objects and processes under study related to professional activities in the field of training and participate in their implementation in the form of software products		Artificial Neural Networks (Deep Learning) Research work

4. SCOPE OF THE DISCIPLINE AND TYPES OF LEARNING ACTIVITIES

The total labor intensity of the discipline " Big data processing " is 6 credit units.

Table 4.1. Types of educational work for full-time education

Type of educational activity	Total number of hours	Module 1
Classroom lessons (total)	51	51
Including:	-	-
Lectures (L)	34	34
Practical lessons (PL)	17	17
Seminars (S)		
Laboratory work (LW)	17	17
Independent work (IW) (total)	112	112

Total labor intensity, hour	216	216
, CU	6	6

5.CONTENT OF THE DISCIPLINE

Table 5.1. Content of the discipline (module) by types of educational work

Name of discipline section	Content of the section (topics)	Types of educational work *
Section 1. Introduction to the subject "Big Data Processing"	Topic 1.1. Review of problems solved by machine learning algorithms. Classification of machine learning algorithms.	L, PL, LW
Section 2. Linear Regression Models.	Topic 2.1. Linear regression. Linear Regression Models. Basic functions. Regularization.	L, PL, LW
Section 3. Logistic regression	Topic 3.1. Objective function of logistic regression. Regularization of logistic regression.	L, PL, LW
Section 4. Cluster analysis.	Topic 4.1. The main types of cluster analysis tasks. Similarity measures and distance functions. Selection of the clustering criterion. Cluster methods based on the Euclidean metric. Hierarchical clustering. Method of K-within-group means. The use of graph theory methods in clustering problems. Clustering based on the analysis of probability densities.	L, PL, LW
Section 5. Neural networks	Topic 5.1. The structure of the neuron. Neural network structure. Training a neural network using an error backpropagation algorithm.	L, PL, LW
Section 6. Decision Trees	Topic 6.1. Decision tree structure. Types of separating functions. Decision tree training. Random Forest algorithm.	L, PL, LW

6. TECHNICAL EQUIPMENT FOR THE DIDISCIPLINE

Table 6.1. Technical equipment for the discipline

Audience type	Audience equipment	Specialized educational / laboratory equipment, software and materials for mastering the discipline (if needed)
Lection	An auditorium for lecture-type classes, equipped with a set of specialized furniture; board (screen) and technical means of multimedia presentations.	Computer classroom equipped with 25 workstations with a personal computer, specialized software for laboratory work and practical lessons
Labor	An auditorium for laboratory work, individual	Classroom equipped with 30 workstations for lectures and group lessons

	consultations, current control and intermediate certification, equipped with a set of specialized furniture and equipment.	
Seminar	An auditorium for conducting seminar-type classes, group and individual consultations, current control and intermediate certification, equipped with a set of specialized furniture and technical means for multimedia presentations.	Computer classroom equipped with 25 workstations with a personal computer, specialized software for laboratory work and practical lessons
Computer class	Computer class for conducting classes, group and individual consultations, current control and intermediate certification, equipped with personal computers (in the amount of ____ pcs.), Board (screen) and multimedia equipment презентаций.	
Individual work	An auditorium for independent work of students (can be used for seminars and consultations), equipped with a set of specialized furniture and computers with access to the EIES.	

7.INFORMATION SUPPORT OF THE DISCIPLINE

a) software:
- space mission general analysis software GMAT;
- low level space dynamics library Orekit;
- flight simulator and information support program (MIOP);
- software development tools Python, C++, etc.

b) databases, reference and information, and search systems:
- electronic fund of legal, and normative and technical documentation http://docs.cntd.ru/;
- search system Yandex https://www.yandex.ru/;
- search system Google https://www.google.ru/;
- abstract database SCOPUS http://www.elsevierscience.ru/products/scopus/.

a) main literature
1. Jones, MT Programming artificial intelligence in applications [Electronic resource] / M. Tim Jones; Per. from English A.I. Osipov. - 2nd ed. - M.: DMK Press, 2011 .-- 312 p .: ill. - ISBN 978-5-94074-746-8. http://znanium.com/bookread2.php?book=410211
2. Luengo, J., García-Gil, D., Ramírez-Gallego, S., García, S., & Herrera, F. (2020). Big data preprocessing: enabling smart data. Springer Nature.
3. Rutkowski, L., Jaworski, M., & Duda, P. (2020). Stream data mining: algorithms and their probabilistic properties. Cham: Springer.
4. Yang, X. S. (2019). Introduction to algorithms for data mining and machine learning. Academic press.

b) additional literature
1. Information analytical systems [Electronic resource]: textbook / T. V. Alekseeva, Yu. V. Amiridi, V. V. Dick and others; ed. V. V. Dick. - M .: MFPU Synergy, 2013 .-- 384 p. Access mode: http://znanium.com/bookread2.php?book=451186
2. Kupriyanov M. S., Barsegyan A. A. Data analysis technologies: Data Mining, Visual Mining, Text Mining, OLAP [Electronic resource] / A. A. Barsegyan, M. S. Kupriyanov, V. V. Stepanenko ... - 2nd ed., Rev. and add. - SPb .: BHV-Petersburg, 2007 .-- 384 p. Access mode: http: //znanium.com/bookread2.php? Book = 489445
for various purposes, and predicting the motion of spacecraft.

8. EVALUATION MATERIALS AND SCORE-RATING SYSTEM FOR ASSESSING THE LEVEL OF FORMATION OF COMPETENCES IN THE DISCIPLINE

In accordance with the requirements of the OS VO RUDN University, for attestation of students for compliance of their personal achievements with the planned discipline learning outcomes, assessment tools funds have been created (VF is presented in Annex 1).

The teacher has the right to change the number and content of assignments given to students (student), based on the contingent (their level of preparedness).
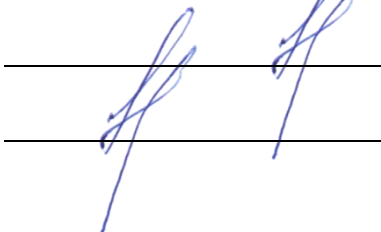
Developer:

docent		Saltykova O.A.

Head of Programm

Prof.		Razoumny Yu.N.

Head of Department

Prof.		Razoumny Yu.N.