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

ARTIFICIAL NEURAL NETWORKS (REINFORCEMENT LEARNING)
(name of discipline/module)

Recommended for the field of study/specialty:

27.04.04 CONTROL IN TECHNICAL SYSTEMS
(code and name of the field of study/specialty)

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

Artificial Intelligence, Machine Learning, and Space Science
(name (profile/specialization) of the educational institution of higher education)

2026

1. THE GOAL OF MASTERING THE DISCIPLINE

The course "Artificial Neural Networks (Reinforcement Learning)" is part of the Master's program "Artificial Intelligence, Machine Learning, and Space Sciences" in the 27.04.04 "Control in Technical Systems" program and is studied in the third semester of the second year. The course is offered by the Department of Mechanics and Control Processes. It consists of four sections and 10 topics and focuses on methods for constructing automatic control systems based on artificial neural networks, mastering methods for solving basic control problems using neural networks, and neural network architectures.

The purpose of mastering the discipline is to teach students methods of constructing artificial neural networks.

2. REQUIREMENTS FOR THE RESULTS OF MASTERING THE DISCIPLINE

Mastering the course "Artificial Neural Networks (Reinforcement Learning)" 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 Competency Achievement** (within this discipline)
PC-1	Able to formulate goals and objectives of scientific research in the field of aerospace systems management, and select methods and means for solving professional problems	PC-1.1 Knows the methods and means of solving scientific research problems in the field of artificial intelligence systems and robotic systems; PC-1.2 Able to formulate the goals and objectives of scientific research in the professional field; PC-1.3 Proficient in techniques for formulating the goals and objectives of scientific research, and knows how to select methods and means for solving problems of professional activity;
PC-4	Capable of participating in scientific research and development of design solutions in the field of ballistics, dynamics and flight control of spacecraft	PC-4.1 Familiar with the basic methods and approaches used to solve problems in the field of artificial intelligence and robotic systems; PC-4.2 Proficient in methods for solving professional problems in the field of artificial intelligence and robotic systems; PC-4.3 Able to apply mathematical methods and modern information technologies when conducting scientific research;

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

Course "Artificial Neural Networks (Reinforcement Learning)" refers to the part formed by the participants of educational relations 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 "Artificial Neural Networks (Reinforcement Learning)".

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, practical training*	Subsequent disciplines/modules, practices*
PC-1	Able to formulate goals and objectives of scientific research in the field of aerospace systems management, and select methods and means for solving professional problems	Research work / Scientific research work (acquiring primary skills in scientific research work); Introduction to Natural Language Processing;	Undergraduate practice / Pre-graduation practice;
PC-4	Capable of participating in scientific research and development of design solutions in the field of ballistics, dynamics and flight control of spacecraft	Research work / Scientific research work (acquiring primary skills in scientific research work); Artificial Intelligence;	Undergraduate practice / Pre-graduation practice;

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

** - elective courses/practices

4. SCOPE OF THE DISCIPLINE AND TYPES OF EDUCATIONAL WORK

The total workload of the course "Artificial Neural Networks (Reinforcement Learning)" is 3 credits.

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, academic hours		Semester(s)
			3
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	*47*		*47*
Control (exam/test with assessment), academic hours	*27*		*27*
Total complexity of the discipline	**academic hours**	***108***	*108*
	credit	***3***	*3*

5. CONTENT OF THE DISCIPLINE

*Table 5.1. Content of the discipline (module) by type of educational*oh work

Section number	Name of the discipline section	Topic Title		Topic Contents	Type of academic work*
Section 1	Introduction to Reinforcement Learning.	1.1	Structure of the reinforcement learning algorithm.	Reinforcement learning as a machine learning paradigm where an agent learns through interaction with an environment. Structure of the algorithm: the agent performs actions, the environment transitions to a new state and returns a reward. The interaction cycle: action, state change, reward reception, policy update. Difference between reinforcement learning and supervised learning: absence of correct answers, learning through trial and error.	LC, LW
		1.2	Agent. Policy function. Value function.	The agent as a learnable entity that makes decisions and interacts with the environment. Policy function as a rule or strategy for selecting actions in each state. Value function as an estimate of the total expected reward the agent can obtain from a given state or after performing a given action. Difference between state value function and action value function.	LC, LW
		1.3	Model. Types of reinforcement learning environments: deterministic, stochastic with complete and incomplete information, discrete and continuous, episodic and non-episodic, single-agent and multi-agent.	Environment model as a description of its dynamics, including transition probabilities between states and reward distributions. Environment types: deterministic with strictly defined action outcomes versus stochastic with random transitions. Full information environments where the agent observes the complete state versus partial information environments with limited observability. Discrete environments with a finite set of states versus continuous environments with infinite possibilities. Episodic environments with natural termination versus continuing environments without a clear end. Single-agent environments versus multi-agent environments with interacting agents.	LC, LW
Section 2	Theoretical foundations and methods of reinforcement learning	2.1	Markov chains and Markov processes. Markov decision process.	Markov chain as a sequence of random states where the probability of the next state depends only on the current state. Markov process as a generalization of Markov chains with continuous time or state space. Markov decision process as a formalism for describing reinforcement learning problems, including states, actions, transition probabilities, reward function, and discount factor. Markov property: the future does not depend on the past given the present.	LC, LW
		2.2	State value functions, Q-function. Bellman equation and optimality. Derivation of the	State value function as the expected sum of discounted rewards when following a given policy from a given state. Q-function as	LC, LW

Section number	Name of the discipline section	Topic Title		Topic Contents	Type of academic work*
			Bellman equation.	the value of taking a specific action in a specific state and then following the policy. Bellman equation as a recursive relationship linking the value of the current state to the values of subsequent states. Optimality in reinforcement learning: achieving the maximum expected total reward. Derivation of the Bellman equation by decomposing value into immediate reward and discounted future value.	
		2.3	Dynamic programming. Monte Carlo methods and game theory.	Dynamic programming as an approach to solving Markov decision processes with a known environment model. Policy iteration with sequential policy improvement. Value iteration for directly computing the optimal value function. Monte Carlo methods as a way to estimate value functions by averaging rewards from multiple interaction episodes. Application of Monte Carlo methods when no environment model is available. Game theory in the context of multi-agent reinforcement learning: Our equilibrium, cooperative and non-cooperative games.	LC, LW
		2.4	Learning based on temporal differences. TD forecasting. TD learning.	Temporal difference learning as a combination of Monte Carlo methods and dynamic programming. Value update after each step based on the difference between current and next estimates. TD prediction as the task of estimating future total reward from each state. TD error as the difference between new and old value estimates. Advantages of TD methods over Monte Carlo methods: ability to learn in continuing episodes and lower estimate variance.	LC, LW
		2.5	Q learning. SARSA algorithm. (State-Action-Reward-State-Action)	Q-learning as a model-free method that updates Q-function based on the optimal value regardless of the current policy. Off-policy nature of Q-learning: the agent learns the optimal policy while following an exploratory policy. SARSA algorithm as a method that updates Q-function based on the actual actions taken by the agent. On-policy nature of SARSA: the agent learns the policy it follows. Comparison of Q-learning and SARSA: Q-learning is more optimistic, SARSA is safer in risk-sensitive tasks.	LC, LW
Section 3	Reinforcement learning software	3.1	Software packages for implementing neural networks. Tensor Flow	Software packages for implementing neural networks in reinforcement learning tasks. TensorFlow as an open-source library for numerical computation and machine learning. Components of TensorFlow: tensors as multi-dimensional arrays, computation graphs for describing operations, automatic differentiation. Using	LC, LW

Section number	Name of the discipline section	Topic Title		Topic Contents	Type of academic work*
				TensorFlow to approximate value functions and policies with deep neural networks. Alternative libraries: PyTorch, Keras, JAX.	
Section 4	Development of artificial neural networks. Symbolic regression methods	4.1	Genetic programming, Cartesian genetic programming, network operator method, variational methods of symbolic regression	Genetic programming as an evolutionary method for automatic creation of computer programs represented as syntax trees. Genetic programming operations: crossover for swapping subtrees, mutation for randomly changing nodes. Cartesian genetic programming with program representation as a directed graph of nodes arranged on a two-dimensional grid. Compactness and efficiency of Cartesian representation. Network operator method as a way to encode structures of complex systems as matrices followed by evolutionary optimization. Variational methods of symbolic regression for finding analytical expressions describing experimental data. Comparison of symbolic regression methods with neural network training: interpretability of results versus flexibility and scalability.	LC, LW

* - to be completed 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 for the discipline

Audience type	**Equipment of the auditorium**	**Specialized educational/laboratory equipment, software and materials for mastering the discipline (if necessary)**
Lecture	A lecture hall equipped with specialized furniture, a whiteboard (screen), and multimedia presentation equipment.	
Computer class	A computer room for conducting classes, group and individual consultations, ongoing monitoring and midterm assessment, equipped with personal computers (in the amount of ____ units), a board (screen) and technical means for multimedia presentations.	
For independent work	A classroom for independent student work (can be used for seminars and consultations), equipped with a set of specialized furniture and computers with access to the Electronic Information System.	

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

7. EDUCATIONAL, METHODOLOGICAL AND INFORMATIONAL SUPPORT OF THE DISCIPLINE

Main literature:

1. Sutton Richard S., Barto Andrew G. Reinforcement Learning. - 2nd edition. - M .: DMK press, 2020. - 552 p. - ISBN 978-5-97060-097-9.
2. Rosenblatt, F. Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Moscow: Mir, 1965. 480 p.3.
3. A.N.Vasiliev, D.A.Tarkhov. Neuronal modeling. Principles. Algorithms. Applications. SPb.: Publishing House Polytechnic. Univ., 2009. ISBN 978-5-7422-2272-9
4. CCAggarwal. Neural Networks and Deep Learning. A Textbook. Springer International Publishing
5. D.A. Tarkhov. Neural Networks. Models and Algorithms. Moscow, Radio Engineering, 2005. (Scientific series "Neurocomputers and Their Applications", ed. A.I. Galushkin. Book 18.)

Further reading:

1. DERumelhardt, GEHinton, RJWilliams. Learning representations by back-propagating errors. Nature, 1986, V.323, pp.533-536.
2. Caudill, M. The Kohonen Model. Neural Network Primer. AI Expert, 1990, 25-31.
3. J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. Proceedings of National Academy of Sciences of USA, 1982, V.79, No.8, pp.2554-2558.

Resources of the information and telecommunications network "Internet":

1. RUDN University Electronic Library System and third-party electronic library systems to which university students have access based on concluded agreements

- RUDN University Electronic Library System – RUDN University Electronic Library System https://mega.rudn.ru/MegaPro/Web
- Electronic Library System "University Library Online" http://www.biblioclub.ru
- EBS Yurayt http://www.biblio-online.ru
- Electronic Library System "Student Consultant" www.studentlibrary.ru
- EBS “Knowledge” 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 "Artificial Neural Networks (Reinforcement Learning)".

* - all teaching 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 DEPART-MENT:

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

HEAD OF THE EP HE:

Professor		Razumny Yuri Nikolaevich
Position, DEPARTMENT	*Signature*	*Surname I.O.*