

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 28.05.2026 15:21:30

Уникальный программный ключ:

ca953a0120d891083f939673078ef1a989dae18a

Федеральное государственное автономное образовательное учреждение высшего образования

«Российский университет дружбы народов имени Патриса Лумумбы»

Инженерная академия

(наименование основного учебного подразделения (ОУП) – разработчика ОП ВО)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

ARTIFICIAL NEURAL NETWORKS (REINFORCEMENT LEARNING) / ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ (ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ)

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

27.04.04 УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

(код и наименование направления подготовки/специальности)

Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, МАШИННОЕ ОБУЧЕНИЕ И КОСМИЧЕСКИЕ НАУКИ

(наименование (профиль/специализация) ОП ВО)

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Artificial Neural Networks (Reinforcement Learning) / Искусственные нейронные сети (Обучение с подкреплением)» входит в программу магистратуры «Искусственный интеллект, машинное обучение и космические науки» по направлению 27.04.04 «Управление в технических системах» и изучается в 3 семестре 2 курса. Дисциплину реализует Кафедра механики и процессов управления. Дисциплина состоит из 4 разделов и 10 тем и направлена на изучение *methods for constructing automatic control systems based on artificial neural networks, mastering methods for solving basic control problems using neural networks, neural network architectures.*

Целью освоения дисциплины является *teaching students methods of constructing artificial neural networks.*

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Artificial Neural Networks (Reinforcement Learning) / Искусственные нейронные сети (Обучение с подкреплением)» направлено на формирование у обучающихся следующих компетенций (части компетенций):

Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
ПК-1	Способен формулировать цели, задачи научных исследований в области управления аэрокосмическими системами, выбирать методы и средства решения профессиональных задач	ПК-1.1 Знает методы и средства решения задач научных исследований в области систем искусственного интеллекта и робототехнических систем;; ПК-1.2 Умеет формулировать цель и задачи научных исследований в профессиональной области;; ПК-1.3 Владеет приемами для формулировки цели и задач научных исследований, умеет выбирать методы и средства решения задач профессиональной деятельности;
ПК-4	Способен участвовать в проведении научных исследований и разработке проектных решений в области баллистики, динамики и управления полетами космических аппаратов	ПК-4.1 Знаком с основными методами и подходами, применяемыми для решения задач в области искусственного интеллекта и робототехнических систем;; ПК-4.2 Владеет методами решения профессиональных задач в области искусственного интеллекта и робототехнических систем;; ПК-4.3 Умеет применять математические методы и современные информационные технологии при проведении научных исследований;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Artificial Neural Networks (Reinforcement Learning)» относится к части, формируемой участниками образовательных отношений блока 1 «Дисциплины (модули)» образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Artificial Neural Networks (Reinforcement Learning)».

Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
------	--------------------------	---	--

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
ПК-1	Способен формулировать цели, задачи научных исследований в области управления аэрокосмическими системами, выбирать методы и средства решения профессиональных задач	Research work / Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы); Introduction to Natural Language Processing;	Undergraduate practice / Преддипломная практика;
ПК-4	Способен участвовать в проведении научных исследований и разработке проектных решений в области баллистики, динамики и управления полетами космических аппаратов	Research work / Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы); Artificial Intelligence;	Undergraduate practice / Преддипломная практика;

* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Artificial Neural Networks (Reinforcement Learning) / Искусственные нейронные сети (Обучение с подкреплением)» составляет «3» зачетные единицы.

Таблица 4.1. Виды учебной работы по периодам освоения образовательной программы высшего образования для очной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			3
Контактная работа, ак.ч	34		34
Лекции (ЛК)	17		17
Лабораторные работы (ЛР)	17		17
Практические/семинарские занятия (СЗ)	0		0
Самостоятельная работа обучающихся, ак.ч.	47		47
Контроль (экзамен/зачет с оценкой), ак.ч.	27		27
Общая трудоемкость дисциплины ак.ч.	ак.ч.	108	108
	зач.ед.	3	3

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы*

Номер раздела	Наименование раздела дисциплины	Наименование темы		Содержание темы	Вид учебной работы*
Раздел 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.	ЛК, ЛР
		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.	ЛК, ЛР
		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.	ЛК, ЛР
Раздел 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.	ЛК, ЛР
		2.2	State value functions, Q-function. Bellman equation and optimality. Derivation of the Bellman equation.	State value function as the expected sum of discounted rewards when following a given policy from a given state. Q-function as 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.	ЛК, ЛР
		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: Nash equilibrium, cooperative and non-cooperative games.	ЛК, ЛР
		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	ЛК, ЛР

Номер раздела	Наименование раздела дисциплины	Наименование темы		Содержание темы	Вид учебной работы*
				Monte Carlo methods: ability to learn in continuing episodes and lower estimate variance.	
		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.	ЛК, ЛР
Раздел 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 TensorFlow to approximate value functions and policies with deep neural networks. Alternative libraries: PyTorch, Keras, JAX.	ЛК, ЛР
Раздел 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.	ЛК, ЛР

* - заполняется только по ОЧНОЙ форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – практические/семинарские занятия.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве ____ шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	
Для самостоятельной работы	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	

* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

1. Саттон Ричард С., Барто Эндрю Г. Обучение с подкреплением = Reinforcement Learning. — 2-е издание. — М.: ДМК пресс, 2020. — 552 с. — ISBN 978-5-97060-097-9.
2. Розенблатт, Ф. Принципы нейродинамики: Перцептроны и теория механизмов мозга = Principles of Neurodynamic: Perceptrons and the Theory of Brain Mechanisms. — М.: Мир, 1965. — 480 с.3.
3. А.Н.Васильев, Д.А.Тархов. Нейростроевое моделирование. Принципы. Алгоритмы. Приложения. СПб.: Изд-во Политехн. Ун-та, 2009. ISBN 978-5-7422-2272-9
4. С.С.Аггарвал. Neural Networks and Deep Learning. A Textbook. Springer International Publishing
5. Д.А.Тархов. Нейронные сети. Модели и алгоритмы. М., Радиотехника, 2005. (Научная серия "Нейрокомпьютеры и их применение", ред. А.И.Галушкин. Кн.18.)

Дополнительная литература:

1. D.E.Rumelhardt, G.E.Hinton, R.J.Williams. 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.

Ресурсы информационно-телекоммуникационной сети «Интернет»:

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров
 - Электронно-библиотечная система РУДН – ЭБС РУДН <https://mega.rudn.ru/MegaPro/Web>
 - ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>
 - ЭБС Юрайт <http://www.biblio-online.ru>
 - ЭБС «Консультант студента» www.studentlibrary.ru
 - ЭБС «Знаниум» <https://znanium.ru/>
2. Базы данных и поисковые системы
 - Sage <https://journals.sagepub.com/>
 - Springer Nature Link <https://link.springer.com/>

- Wiley Journal Database <https://onlinelibrary.wiley.com/>

- Научно-метрическая база данных Lens.org <https://www.lens.org>

Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля:*

1. Курс лекций по дисциплине «Artificial Neural Networks (Reinforcement Learning)».

* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**

РАЗРАБОТЧИКИ

Доцент

Должность

РУКОВОДИТЕЛЬ БУП

Заведующий кафедрой

Должность

РУКОВОДИТЕЛЬ ОП ВО

Профессор

Должность

Салтыкова О.А.

Фамилия И.О

Разумный Ю.Н.

Фамилия И.О

Разумный Ю.Н.

Фамилия И.О