

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 27.05.2026 14:42:39

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

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 «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 методов построения систем автоматического управления на основе искусственных нейронных сетей, освоение методов решения основных задач управления с использованием нейронных сетей, архитектур нейронных сетей

Целью освоения дисциплины является обучение студентов методам построения искусственных нейронных сетей.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «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)
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 Анализирует задачу, выделяя ее базовые составляющие;; УК-1.2 Определяет и ранжирует информацию, требуемую для решения поставленной задачи;; УК-1.3 Осуществляет поиск информации для решения поставленной задачи по различным типам запросов;;
УК-7	Способен искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	УК-7.1 Осуществляет поиск нужных источников информации и данных, воспринимает, анализирует, запоминает и передает информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач;; УК-7.2 Проводит оценку информации, ее достоверность, строит логические умозаключения на основании поступающих информации и данных;; УК-7.3 Владеет современными цифровыми технологиями, методами поиска, обработки, анализа, хранения и представления информации (в области управления в технических системах) в условиях цифровой экономики и современной корпоративной информационной культуры.;
ПК-2	Способен применять современные теоретические и экспериментальные методы разработки математических моделей исследуемых объектов и процессов в области управления аэрокосмическими системами	ПК-2.1 Знает современные теоретические и экспериментальные методы, применяемые для разработки математических моделей исследуемых объектов и процессов профессиональной деятельности;; ПК-2.2 Умеет определять эффективность применяемых методов для разработки математических моделей исследуемых объектов и процессов;; ПК-2.3 Владеет современными теоретическими и экспериментальными методами для разработки математических моделей объектов и процессов профессиональной деятельности по направлению подготовки.;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Искусственные нейронные сети (Обучение с подкреплением)» 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 «Искусственные нейронные сети (Обучение с подкреплением)».

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

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
УК-7	Способен искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	History and Methodology of Science; Information Technology in Mathematical Modelling; Information Databases; Research work / Научно-исследовательская работа;	Undergraduate Training;
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	History and Methodology of Science; Artificial Neural Networks (Deep Learning)**; Искусственные нейронные сети (Глубокое обучение)**; Research work / Научно-исследовательская работа;	Undergraduate Training;
ПК-2	Способен применять современные теоретические и экспериментальные методы разработки математических моделей исследуемых объектов и процессов в области управления аэрокосмическими системами	Research work / Научно-исследовательская работа; Information Technology in Mathematical Modelling; Advanced Methods of Space Flight Mechanics; History and Methodology of Science; Virtual Reality and Computer Vision; Artificial Neural Networks (Deep Learning)**; Искусственные нейронные сети (Глубокое обучение)**;	Undergraduate Training;

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

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Искусственные нейронные сети (Обучение с подкреплением)» составляет «5» зачетных единиц
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	72		72
Лекции (ЛК)	36		36
Лабораторные работы (ЛР)	36		36
Практические/семинарские занятия (СЗ)	0		0
Самостоятельная работа обучающихся, ак.ч.	72		72
Контроль (экзамен/зачет с оценкой), ак.ч.	36		36
Общая трудоемкость дисциплины ак.ч.	ак.ч.	180	180
	зач.ед.	5	5

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	Введение в обучение с подкреплением.	1.1	Структура алгоритма обучения с подкреплением.	Обучение с подкреплением как парадигма машинного обучения, где агент обучается на основе взаимодействия со средой. Структура алгоритма: агент выполняет действия, среда переходит в новое состояние и возвращает вознаграждение. Цикл взаимодействия: действие, изменение состояния, получение награды, обновление политики агента. Отличие обучения с подкреплением от обучения с учителем: отсутствие правильных ответов, обучение через пробы и ошибок.	ЛК, ЛР
		1.2	Агент. Функция политики. Функция ценности.	Агент как обучаемая сущность, принимающая решения и взаимодействующая со средой. Функция политики как правило или стратегия выбора действия агентом в каждом состоянии. Функция ценности как оценка ожидаемой суммарной награды, которую агент может получить из данного состояния или после выполнения действия. Различие между функцией ценности состояния и функцией ценности действия.	ЛК, ЛР
		1.3	Модель. Типы сред обучения с подкреплением: детерминированная, стохастическая с полной и неполной информацией, дискретная и непрерывная, эпизодическая и не эпизодическая, одноагентная и многоагентная.	Модель среды как описание её динамики, включающее вероятности переходов между состояниями и распределение наград. Типы сред: детерминированная со строго определёнными исходами действий и стохастическая со случайными переходами. Среда с полной информацией, где агент наблюдает полное состояние, и с неполной информацией с частичной наблюдаемостью. Дискретная среда с конечным набором состояний и непрерывная с бесконечным множеством. Эпизодическая среда с естественным завершением взаимодействия и непрерывная без чёткого окончания. Однолетняя среда с одним агентом и многолетняя с несколькими взаимодействующими агентами.	ЛК, ЛР
Раздел 2	Теоретические основы и методы обучения с подкреплением	2.1	Марковские цепи и Марковские процессы. Марковский процесс принятия решений.	Марковская цепь как последовательность случайных состояний, где вероятность следующего состояния зависит только от текущего. Марковский процесс как обобщение марковской цепи с непрерывным временем или пространством состояний. Марковский процесс принятия решений как формализм для описания задач обучения с подкреплением, включающий множество состояний, действий, вероятности переходов, функцию награды и коэффициент дисконтирования. Свойство марковости: будущее не зависит от прошлого при известном настоящем.	ЛК, ЛР
		2.2	Функции ценности состояния, Q-функция. Уравнение Беллмана и оптимальность. Вывод уравнения Беллмана.	Функция ценности состояния как ожидаемая сумма дисконтированных наград при следовании заданной политике из данного состояния. Q-функция как ценность выполнения конкретного действия в конкретном состоянии с последующим следованием политике. Уравнение Беллмана как рекуррентное соотношение, связывающее ценность текущего состояния с ценностью последующих состояний. Оптимальность в обучении с подкреплением: достижение максимальной ожидаемой суммарной награды. Вывод уравнения Беллмана через разложение ценности на немедленную награду и дисконтированную ценность следующего состояния.	ЛК, ЛР
		2.3	Динамическое программирование. Методы Монте-Карло и	Динамическое программирование как подход к решению марковских процессов принятия решений при известной модели среды. Итерация по политике с последовательным улучшением стратегии. Итерация по ценности с прямым	ЛК, ЛР

Номер раздела	Наименование раздела дисциплины	Наименование темы		Содержание темы	Вид учебной работы*
			теория игр.	вычислением оптимальной функции ценности. Методы Монте-Карло как способ оценки функций ценности через усреднение наград от множества эпизодов взаимодействия. Применение методов Монте-Карло при отсутствии модели среды. Теория игр в контексте многолетнего обучения с подкреплением: равновесие Нэша, кооперативные и некооперативные игры.	
		2.4	Обучение на основе временных различий (Temporary Differences). TD прогнозирование. TD обучение.	Обучение на основе временных различий как комбинация методов Монте-Карло и динамического программирования. Обновление оценки ценности после каждого шага на основе разницы между текущей и следующей оценками. TD прогнозирование как задача оценки будущей суммарной награды из каждого состояния. TD ошибка как разница между новой оценкой ценности и старой. Преимущества TD-методов перед методами Монте-Карло: возможность обучения в непрерывных эпизодах и более низкая дисперсия оценок.	ЛК, ЛР
		2.5	Q обучение. Алгоритм SARSA. (State-Action-Reward-State-Action)	Q-обучение как метод обучения без модели среды, использующий обновление Q-функции по оптимальному значению независимо от текущей политики. Офф-политический характер Q-обучения: агент обучается оптимальной политике, следуя исследующей. Алгоритм SARSA как метод, обновляющий Q-функцию на основе фактически выполняемых действий агента. Он-политический характер SARSA: агент обучается той политике, которой следует. Сравнение Q-обучения и SARSA: Q-обучение более оптимистично, SARSA более безопасно в задачах с рисками.	ЛК, ЛР
Раздел 3	Программное обеспечение обучения с подкреплением	3.1	Пакеты программ для реализации нейронных сетей. Tensor Flow	Пакеты программ для реализации нейронных сетей в задачах обучения с подкреплением. TensorFlow как библиотека с открытым исходным кодом для численных вычислений и машинного обучения. Компоненты TensorFlow: тензоры как многомерные массивы, графы вычислений для описания операций, автоматическое дифференцирование. Использование TensorFlow для аппроксимации функций ценности и политик глубокими нейронными сетями. Альтернативные библиотеки: PyTorch, Keras, JAX.	ЛК, ЛР
Раздел 4	Развитие искусственных нейронных сетей. Методы символьной регрессии	4.1	Генетического программирование, декартово генетическое программирование, метод сетевого оператора, вариационные методы символьной регрессии	Генетическое программирование как эволюционный метод автоматического создания компьютерных программ, представленных в виде синтаксических деревьев. Операции генетического программирования: скрещивание для обмена поддеревьями, мутация для случайного изменения узлов. Декартово генетическое программирование с представлением программы в виде ориентированного графа узлов, расположенных на двумерной сетке. Компактность и эффективность декартова представления. Метод сетевого оператора как способ кодирования структур сложных систем в виде матрицы с последующей эволюционной оптимизацией. Вариационные методы символьной регрессии для поиска аналитических выражений, описывающих экспериментальные данные. Сравнение методов символьной регрессии с обучением нейронных сетей: интерпретируемость результатов против гибкости и масштабируемости.	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 <http://lib.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 «Троицкий мост»
2. Базы данных и поисковые системы
 - электронный фонд правовой и нормативно-технической документации <http://docs.cntd.ru/>
 - поисковая система Яндекс <https://www.yandex.ru/>

- поисковая система Google <https://www.google.ru/>

- реферативная база данных SCOPUS <http://www.elsevierscience.ru/products/scopus/>

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

1. Курс лекций по дисциплине «Искусственные нейронные сети (Обучение с подкреплением)».

* - 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