

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 29.05.2024 10:35:38
Уникальный программный ключ:
ca953a0120d891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов имени Патриса Лумумбы»**

Факультет искусственного интеллекта

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

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

02.03.02 ФУНДАМЕНТАЛЬНАЯ ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

(код и наименование направления подготовки/специальности)

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

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: РАЗРАБОТКА И ОБУЧЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

(наименование (профиль/специализация) ОП ВО)

2024 г.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Обучение с подкреплением» входит в программу бакалавриата «Искусственный интеллект: разработка и обучение интеллектуальных систем» по направлению 02.03.02 «Фундаментальная информатика и информационные технологии» и изучается в 6 семестре 3 курса. Дисциплину реализует NOT DEFINED. Дисциплина состоит из 11 разделов и 11 тем и направлена на изучение

Целью освоения дисциплины является изучение основных методов разработки и моделирования среды и агента, взаимодействующего со средой, современных алгоритмов выбора и формирования политики и стратегии взаимодействия агента со средой, а также приобретение практических умений и прикладных навыков применения математического обучения с подкреплением для решения современных задач математического моделирования и обработки данных. Задачи дисциплины: приобретение знаний и формирование практических умений и навыков в области разработки современных алгоритмов обучения с подкреплением, модели рующих действия интеллектуального агента в смоделированной среде, их применения к практическим задачам; формирование компетенций в вопросах подготовки и разработки современного программного продукта, на базе алгоритмов и архитектур современного обучения с подкреплением.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «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)
ОПК-2	Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий, компьютерных/суперкомпьютерных методов и современного программного обеспечения, в том числе отечественного происхождения, с учетом основных требований информационной безопасности	ОПК-2.1 Знает основные положения и концепции в области программирования, архитектуру языков программирования, знает основную терминологию, знаком с содержанием Единого Реестра Российских программ; ОПК-2.2 Умеет анализировать типовые языки программирования, составлять программы; ОПК-2.3 Имеет практический опыт решения задач анализа, интеграции различных типов программного обеспечения;
ПК-3	Способен разрабатывать и оптимизировать алгоритмы обучения с подкреплением и системы автоматизированного принятия решений для создания интеллектуальных агентов и роботов	ПК-3.1 Демонстрирует умение разрабатывать и реализовывать алгоритмы обучения с подкреплением для обучения интеллектуальных агентов и роботов; ПК-3.2 Умеет интегрировать системы автоматизированного принятия решений в интеллектуальные агенты и роботы;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Обучение с подкреплением» относится к обязательной части 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 «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*
ОПК-2	Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий, компьютерных/суперкомпьютерных методов и современного программного обеспечения, в том числе отечественного происхождения, с учетом основных требований информационной безопасности	Технологическая (проектно-технологическая) практика (учебная); Практикум по программированию; История и теория программирования; Программирование на языке Python; Методы машинного обучения; Глубинное обучение;	Технологическая (проектно-технологическая) практика (производственная); Практикум по программированию; Этика и юридические аспекты в искусственном интеллекте;
ПК-3	Способен разрабатывать и оптимизировать алгоритмы обучения с подкреплением и системы автоматизированного принятия решений для создания интеллектуальных агентов и роботов	Нейронные сети; Инструменты обработки и визуализации данных; Практикум по программированию; <i>Типовые задачи и применение ИИ в лингвистике**;</i> <i>Типовые задачи и применение ИИ в финансах**;</i>	Преддипломная практика; Технологическая (проектно-технологическая) практика (производственная); Практикум по программированию;

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

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Обучение с подкреплением» составляет «4» 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)
			6
<i>Контактная работа, ак.ч.</i>	54		54
Лекции (ЛК)	18		18
Лабораторные работы (ЛР)	18		18
Практические/семинарские занятия (СЗ)	18		18
<i>Самостоятельная работа обучающихся, ак.ч.</i>	63		63
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	27		27
Общая трудоемкость дисциплины	ак.ч.	144	144
	зач.ед.	4	4

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

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

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
Раздел 1	Задача обучения с подкреплением. Модель взаимодействия агента со средой. Алгоритмы обучения с подкреплением	1.1	Модель взаимодействия агента со средой. Оптимальное управление. Марковские цепи. Среда, действия и траектории. Марковский процесс принятия решений. Эпизодичность и дисконтирование. Алгоритмы обучения с подкреплением. Условия задачи обучения с подкреплением. Отличие от задач обучения с учителем. Концепция modelfree алгоритмов. Onpolicy и Offpolicy. Классификация алгоритмов обучения с подкреплением. Критерии оценки алгоритмов. Выбор функции награды.	ЛК, ЛР
Раздел 2	Эволюционный подход и без градиентная оптимизация.	2.1	Задача безградиентной оптимизации. Случайный поиск. Имитация отжига. Эволюционные алгоритмы. Weight Agnostic Neural Networks. Генетические алгоритмы. Метод КроссЭнтропии для стохастической оптимизации. Метод КроссЭнтропии для стохастической оптимизации (СЕМ). Натуральные эволюционные стратегии (NES). Натуральные эволюционные стратегии (NES).	ЛК, ЛР
Раздел 3	Классическая теория. Оценочные функции. Улучшение политики. Динамическое программирование	3.1	Оценочные функции. Vфункция. Уравнения Беллмана. Оптимальная стратегия. Qфункция. Принцип оптимальности Беллмана. Уравнения оптимальности Беллмана. Критерий оптимальности Беллмана. Улучшение политики. Вид оптимальной стратегии. Динамическое программирование. Метод простой итерации. Value Iteration. Policy Iteration.	ЛК, ЛР
Раздел 4	Табличные алгоритмы. BiasVariance TradeOf.	4.1	Табличные алгоритмы. МонтеКарло алгоритм. Экспоненциальное сглаживание. Экспоненциальное сглаживание. Qlearning. Explorationexploitation дилемма. SARSA. Дилемма смещенияразброса. Интерпретация через Credit Assingment. Backward View. Eligibility Trace.	ЛК, ЛР
Раздел 5	Deep QNetwork (DQN) и его модификации.	5.1	Deep Qlearning. Qсетка. Переход к параметрической Qфункции. Таргетсеть. Декорреляция сэмплов. Double DQN, приоритизированный буфер, дуэльная архитектура, шумные сети, многошаговый DQN.	ЛК, ЛР
Раздел 6	Distributional RL.	6.1	Идея Distributional подхода. Zфункция. Distributional форма уравнения Беллмана. Категориальная аппроксимация Zфункций. Categorical DQN. Квантильная аппроксимация Zфункций. Quantile Regression DQN.	ЛК, ЛР
Раздел 7	Policy gradient подход. Advantage ActorCritic (A2C). REINFORCE. Продвинутое Policy Gradient	7.1	Подход Policy gradient. REINFORCE. Схемы «Актёркритик». Biasvariance tradeoff. Advantage Actor Critic (A2C). Trust Region Policy Optimization (TRPO). Proximal Policy Optimization (PPO).	ЛК, ЛР
Раздел 8	Modelbased подход. Планирование для дискретного управления	8.1	Задача многоруких бандитов. Теорема ЛаиРоббинса. Upper Confidence Bound (UCB). Внутренняя мотивация. Сэмплирование Томпсона. Обобщение на табличные MDP. Планирование. MonteCarlo Tree Search (MCTS). AlphaZero. MuZero.	ЛК, ЛР

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
Раздел 9	Непрерывное управление.	9.1	Linear Quadratic Regulator (LQR). Iterative LQR (iLQR). Soft ActorCritic (SAC).	ЛК, ЛР
Раздел 10	Имитационное обучение и обратное обучение с подкреплением.	10.1	Имитационное обучение. Клонирование поведения. Обратное обучение с подкреплением (Inverse RL). Схема Guided Cost Learning. Генеративно-состязательное имитационное обучение (GAIL).	ЛК
Раздел 11	Заключение.	11.1	Обзор пройденного материала. Дальнейшие темы для самостоятельного изучения. Обсуждение возможностей применения пройденного материала в дальнейшей научной и практической деятельности	ЛК

* - 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 20 шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели 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 Р. С. Обучение с подкреплением: введение [Электронный ресурс], 2020. 552 с.
2. Лю Ю. Обучение с подкреплением на PyTorch. Сборник рецептов : руководство / Ю. Лю, 2020. 282 с. Текст : электронный.
3. Лонца А. Алгоритмы обучения с подкреплением на Python [Электронный ресурс] : практическое руководство, 2020. 286 с

Дополнительная литература:

1. Давичандиран Судхарсан Глубокое обучение с подкреплением на Python. OpenAI Gym и TensorFlow для профи / Судхарсан Равичандиран, 2020. 320 с. Текст : электронный
2. Уиндер Ф. Обучение с подкреплением для реальных задач: Пер. с англ. / Ф. Уиндер, 2022. 400 с. Текст : электронный.
3. Максим Лапань Глубокое обучение с подкреплением. AlphaGo и другие технологии / Лапань Максим, 2021. 496 с. Текст : электронный.

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 «Обучение с подкреплением».

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

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ

Оценочные материалы и балльно-рейтинговая система* оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Обучение с подкреплением» представлены в Приложении к настоящей Рабочей программе дисциплины.

* - ОМ и БРС формируются на основании требований соответствующего локального нормативного акта РУДН.

РАЗРАБОТЧИК:

<hr/> <i>Должность, БУП</i>	<hr/> <i>Подпись</i>	<hr/> Баум Валентина Владимировна <i>Фамилия И.О.</i>
-----------------------------	----------------------	---

РУКОВОДИТЕЛЬ БУП:

<hr/> Заведующий кафедрой <i>Должность БУП</i>	<hr/> <i>Подпись</i>	<hr/> <i>Фамилия И.О.</i>
---	----------------------	---------------------------

РУКОВОДИТЕЛЬ ОП ВО:

<hr/> <i>Должность, БУП</i>	<hr/> <i>Подпись</i>	<hr/> <i>Фамилия И.О.</i>
-----------------------------	----------------------	---------------------------