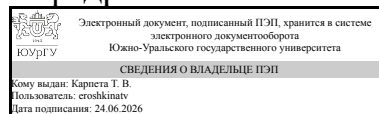


ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Заведующий выпускающей
кафедрой



Т. В. Карпета

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.М0.04 Мультиагентное обучение с подкреплением
для направления 02.04.02 Фундаментальная информатика и информационные технологии

уровень Магистратура

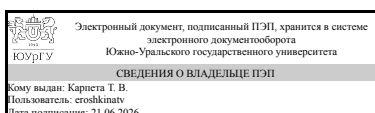
магистерская программа Инновационные методы и технологии машинного обучения и анализа данных

форма обучения очная

кафедра-разработчик Прикладная математика и программирование

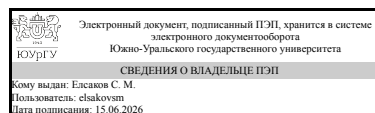
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 02.04.02 Фундаментальная информатика и информационные технологии, утверждённым приказом Минобрнауки от 23.08.2017 № 811

Зав.кафедрой разработчика,
к.физ.-мат.н.



Т. В. Карпета

Разработчик программы,
к.физ.-мат.н., доцент



С. М. Елсаков

1. Цели и задачи дисциплины

Цель дисциплины: Формирование системных теоретических знаний и практических навыков в области проектирования, обучения и анализа многоагентных систем искусственного интеллекта, функционирующих в сложных стохастических и частично наблюдаемых средах. Задачи дисциплины: Изучение математического аппарата для описания одноагентных и многоагентных взаимодействий. Изучение фундаментальных проблем MARL. Освоение современных архитектур глубокого многоагентного обучения.

Краткое содержание дисциплины

Одноагентный обучение с подкреплением: марковские процессы принятия решений (MDP), методы динамического программирования, безмодельные алгоритмы (Q-Learning) и современные архитектуры глубокого обучения с подкреплением (DQN, PPO, SAC). Многоагентные обучение с подкреплением: теория игр, Марковские игры и децентрализованным POMDP. Ключевые проблемы многоагентного обучения с подкреплением: нестационарность среды, проблему распределения вклада (Credit Assignment) и частичную наблюдаемость.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-5 Способен проводить исследования в области кооперативного и состязательного взаимодействия агентов с использованием генеративных моделей для синтеза данных и обучающих сред	Знает: Теоретико-игровые концепции кооперации и конкуренции агентов (Nash-равновесие, Pareto-оптимальность), алгоритмы self-play и population-based training, методы оценки эффективности мультиагентного взаимодействия Умеет: Проектировать эксперименты по исследованию кооперативного и состязательного поведения агентов, формулировать и проверять гипотезы об эффективности различных стратегий взаимодействия Имеет практический опыт: Постановки и проведения экспериментального исследования кооперативных и состязательных мультиагентных систем; анализа динамики взаимодействия агентов и интерпретации полученных результатов

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Нет	Не предусмотрены

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Нет

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 4 з.е., 144 ч., 58,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам
		в часах
		Номер семестра
		4
Общая трудоёмкость дисциплины	144	144
Аудиторные занятия:	48	48
Лекции (Л)	24	24
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	0	0
Лабораторные работы (ЛР)	24	24
Самостоятельная работа (СРС)	85,5	85,5
Подготовка к экзамену	85,5	85,5
Консультации и промежуточная аттестация	10,5	10,5
Вид контроля (зачет, диф.зачет, экзамен)	-	экзамен

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Одноагентное обучение с подкреплением	16	8	0	8
2	Теоретические основы многоагентного обучения (MARL)	12	4	0	8
3	Парадигмы и Алгоритмы Deep MARL	12	8	0	4
4	Продвинутое темы MARL	8	4	0	4

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Марковские процессы принятия решений (MDP). Уравнения Беллмана. Безмодельные методы (Q-Learning). Проблема Exploration (UCB, Thompson Sampling).	4
2	1	Глубокое RL (DQN, PER). Методы градиентов политик (REINFORCE, Actor-Critic, PPO, SAC).	4
3	2	Нормальные и экстенсивные игры. Равновесие Нэша. Марковские игры (Markov Games). Минимаксные алгоритмы. POMDP и Децентрализованные POMDP (Dec-POMDP), их вычислительная сложность.	4
4	3	Проблема нестационарности и Independent Q-Learning (IQL). Парадигма CTDE (Centralized Training with Decentralized Execution).	4
5	3	Декомпозиция ценностей (VDN, QMIX, QPLEX). Многоагентные градиенты	4

		политик: MADDPG, Counterfactual Baselines (COMA), MAPPO.	
6	4	Обучение коммуникации (CommNet, TarMAC). Opponent Modeling и Theory of Mind. Many-Agent RL и Mean-Field MARL. Эволюционные методы, Population-Based Training (PBT) и PSRO. Безопасность и Robust MARL.	4

5.2. Практические занятия, семинары

Не предусмотрены

5.3. Лабораторные работы

№ занятия	№ раздела	Наименование или краткое содержание лабораторной работы	Кол-во часов
1	1	Многорукие бандиты	4
2	1	Глубокое RL	4
3	2	Матричные игры	4
4	2	Разработка своей MARL среды	4
5	3	Многоагентные Actor-Critic методы	4
6	4	Освоение коммуникации	4

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Подготовка к экзамену	Саттон, Р. С. Обучение с подкреплением: введение : руководство / Р. С. Саттон, Э. Д. Барто ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2020. — 552 с. — ISBN 978-5-97060-097-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/179453 (дата обращения: 15.06.2026). — Режим доступа: для авториз. пользователей.	4	85,5

6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учитывается в ПА
1	4	Текущий контроль	ЛР1	1	20	Студентом предоставляется оформленный отчет. Оценивается качество оформления,	экзамен

						соответствие нормативным документам. Общий балл при оценке складывается из следующих показателей: Выполнено задание - 10 баллов. Нет замечаний к отчету - 10 баллов. Максимальное количество баллов – 20. Весовой коэффициент мероприятия – 1.	
2	4	Текущий контроль	ЛР2	1	20	Студентом предоставляется оформленный отчет. Оценивается качество оформления, соответствие нормативным документам. Общий балл при оценке складывается из следующих показателей: Выполнено задание - 10 баллов. Нет замечаний к отчету - 10 баллов. Максимальное количество баллов – 20. Весовой коэффициент мероприятия – 1.	экзамен
3	4	Текущий контроль	ЛР3	1	20	Студентом предоставляется оформленный отчет. Оценивается качество оформления, соответствие нормативным документам. Общий балл при оценке складывается из следующих показателей: Выполнено задание - 10 баллов. Нет замечаний к отчету - 10 баллов. Максимальное количество баллов – 20. Весовой коэффициент мероприятия – 1.	экзамен
4	4	Текущий контроль	ЛР4	1	20	Студентом предоставляется оформленный отчет. Оценивается качество оформления, соответствие нормативным документам. Общий балл при оценке складывается из следующих показателей: Выполнено задание - 10 баллов. Нет замечаний к отчету - 10 баллов. Максимальное количество баллов – 20. Весовой коэффициент мероприятия – 1.	экзамен
5	4	Текущий контроль	ЛР5	1	20	Студентом предоставляется оформленный отчет. Оценивается качество оформления, соответствие нормативным документам. Общий балл при оценке складывается из следующих показателей: Выполнено задание - 10 баллов. Нет замечаний к отчету - 10 баллов. Максимальное количество баллов – 20. Весовой коэффициент мероприятия – 1.	экзамен
6	4	Текущий контроль	ЛР6	1	20	Студентом предоставляется оформленный отчет. Оценивается качество оформления, соответствие нормативным документам. Общий балл при оценке складывается из следующих показателей: Выполнено задание - 10 баллов. Нет замечаний к отчету - 10 баллов. Максимальное количество баллов – 20. Весовой коэффициент мероприятия – 1.	экзамен
7	4	Промежуточная аттестация	Экзамен	-	120	Контрольное мероприятие промежуточной аттестации включает устный ответ на билет и проводятся во время экзамена.	экзамен

					При оценивании результатов мероприятия используется балльно-рейтинговая система оценивания результатов. В билете два вопроса. Критерии оценивания выполнения зачетной работы: - ответ на один вопрос из билета без замечаний – 30 баллов; - ответ на один вопрос из билета с недочетами – 20 баллов; - ответ на один вопрос из билета с грубыми замечаниями – 10 баллов; - нет ответа на один вопрос из билета – 0 баллов; - ответ на один дополнительный вопрос без замечаний – 30 баллов; - ответ на один дополнительный вопрос с недочетами – 20 баллов; - ответ на один дополнительный вопрос с грубыми замечаниями – 10 баллов; - нет ответа на один дополнительный вопрос – 0 баллов; Максимальное количество баллов за промежуточную аттестацию – 120.	
--	--	--	--	--	---	--

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
экзамен	Прохождение контрольных мероприятий промежуточной аттестации не обязательно. Экзамен проводится по билетам. В билете два вопроса. Билет выбирается случайным образом. Студенту дается 30 минут на подготовку. После этого он рассказывает ответы на вопросы билета. Студенту задается дополнительный вопрос по каждому вопросу.	В соответствии с пп. 2.5, 2.6 Положения

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ						
		1	2	3	4	5	6	7
ПК-5	Знает: Теоретико-игровые концепции кооперации и конкуренции агентов (Nash-равновесие, Pareto-оптимальность), алгоритмы self-play и population-based training, методы оценки эффективности мультиагентного взаимодействия	+	+	+	+	+	+	+
ПК-5	Умеет: Проектировать эксперименты по исследованию кооперативного и состязательного поведения агентов, формулировать и проверять гипотезы об эффективности различных стратегий взаимодействия	+	+	+	+	+	+	+
ПК-5	Имеет практический опыт: Постановки и проведения экспериментального исследования кооперативных и состязательных мультиагентных систем; анализа динамики взаимодействия агентов и интерпретации полученных результатов	+	+	+	+	+	+	+

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) основная литература:

Не предусмотрена

б) дополнительная литература:

Не предусмотрена

в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:

Не предусмотрены

г) методические указания для студентов по освоению дисциплины:

1. Методические указания по организации самостоятельной работы студента

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Методические указания по организации самостоятельной работы студента

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	ЭБС издательства Лань	Саттон, Р. С. Обучение с подкреплением: введение : руководство / Р. С. Саттон, Э. Д. Барто ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2020. — 552 с. — ISBN 978-5-97060-097-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/179453 (дата обращения: 15.06.2026). — Режим доступа: для авториз. пользователей.
2	Дополнительная литература	ЭБС издательства Лань	Жукова, Н. А. Обучение с подкреплением: лабораторные работы : учебно-методическое пособие / Н. А. Жукова, И. А. Куликов, А. Н. Субботин. — Санкт-Петербург : НИУ ИТМО, 2022. — 95 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/283880 (дата обращения: 15.06.2026). — Режим доступа: для авториз. пользователей.
3	Основная литература	ЭБС издательства Лань	Лю, Ю. Обучение с подкреплением на PyTorch. Сборник рецептов : руководство / Ю. Лю ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2020. — 282 с. — ISBN 978-5-97060-853-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/179493 (дата обращения: 15.06.2026). — Режим доступа: для авториз. пользователей.

Перечень используемого программного обеспечения:

1. -Python(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Лабораторные занятия	333 (36)	проектор, компьютеры
Лекции	333 (36)	проектор