

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность:

Дата подписания: 20.06.2026 11:44:09

Уникальный программный ключ:

c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Ростовский государственный экономический университет (РИНХ)»

УТВЕРЖДАЮ

Начальник

учебно-методического управления

Т.К. Платонова

«25» мая 2026 г.

**Рабочая программа дисциплины
Обучение с подкреплением и его приложения**

Направление подготовки

01.04.02 Прикладная математика и информатика

Направленность (профиль) программы магистратуры

01.04.02.04 Искусственный интеллект: математические модели и прикладные решения

Для набора 2026 года

Квалификация
Магистр

КАФЕДРА Информационных систем и прикладной информатики**Распределение часов дисциплины по семестрам / курсам**

Семестр (<Курс>.<Семестр на курсе>)	4 (2.2)		Итого	
	12			
Неделя	12			
Вид занятий	УП	РП	УП	РП
Лекции	8	8	8	8
Лабораторные	20	20	20	20
Итого ауд.	28	28	28	28
Контактная работа	28	28	28	28
Сам. работа	143	143	143	143
Часы на контроль	9	9	9	9
Итого	180	180	180	180

ОСНОВАНИЕ

Учебный план утвержден учёным советом Университета (протокол № 9 от 03.03.2026 г.).

Программу составил(и): доцент, Хаймин Е.С.

Зав. кафедрой: д.э.н., профессор С.М. Щербаков

Методический совет направления: д.э.н., доцент Ю.Г. Чернышева

Директор института магистратуры: д.э.н., профессор Е.А. Иванова

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	получение теоретических знаний и практических навыков по применению алгоритмов обучения с подкреплением при решении практических задач профессиональной деятельности.
-----	---

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

ПК-5. Способен руководить проектами по созданию, поддержке и использованию системы искусственного интеллекта на основе нейросетевых моделей и методов

В результате освоения дисциплины обучающийся должен:**Знать:**

методы и модели машинного обучения и нейронных сетей (соотнесено с индикатором ПК-5.1)

Уметь:

проектировать системы искусственного интеллекта с использованием методов машинного обучения и нейронных сетей (соотнесено с индикатором ПК-5.2)

Владеть:

навыками разработки и использования систем искусственного интеллекта с использованием методов машинного обучения и нейронных сетей (соотнесено с индикатором ПК-5.3)

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ**Раздел 1. Алгоритмы обучения с подкреплением**

№	Наименование темы, краткое содержание	Вид занятия / работы / форма ПА	Семестр / Курс	Количество часов	Компетенции
1.1	Тема 1.1. Модель взаимодействия агента со средой Связь с оптимальным уравнением. Марковская цепь. Среда. Действия. Траектории. Марковский процесс принятия решений (MDP). Эпизодичность. Дисконтирование.	Лекционные занятия	4	2	ПК-5
1.2	Тема 1.1. Модель взаимодействия агента со средой Связь с оптимальным уравнением. Марковская цепь. Среда. Действия. Траектории. Марковский процесс принятия решений (MDP). Эпизодичность. Дисконтирование.	Лабораторные занятия	4	2	ПК-5
1.3	Тема 1.2. Алгоритмы обучения с подкреплением Условия задачи RL. On-policy. Концепция model-free алгоритмов. Классификация RL-алгоритмов. Критерии оценки RL-алгоритмов. Сложности задачи RL. Бенчмарки.	Лекционные занятия	4	2	ПК-5
1.4	Тема 1.2. Алгоритмы обучения с подкреплением Условия задачи RL. On-policy. Концепция model-free алгоритмов. Классификация RL-алгоритмов. Критерии оценки RL-алгоритмов. Сложности задачи RL. Бенчмарки.	Лабораторные занятия	4	2	ПК-5
1.5	Тема 1.3. Бэйзлайны Задача безградиентной оптимизации. Случайный поиск. Hill Climbing. Имитация отжига. Эволюционные алгоритмы. Weight Agnostic Neural Networks (WANN). Видовая спецификация. Генетические алгоритмы.	Лабораторные занятия	4	2	ПК-5
1.6	Тема 1.4. Эволюционные стратегии Идея эволюционных стратегий. Оценка вероятности редкого события. Метод Кросс-Энтропии для стохастической оптимизации. Метод Кросс-Энтропии для обучения с подкреплением (CEM). Натуральные эволюционные стратегии (NES). OpenAI-ES. Адаптация матрицы ковариации (CMA-ES).	Лабораторные занятия	4	2	ПК-5
1.7	Value-based подход. Deep Q-learning. Модификации DQN. Distributional RL.	Самостоятельная работа	4	70	ПК-5

Раздел 2. Оценочные функции и динамическое программирование

№	Наименование темы, краткое содержание	Вид занятия / работы / форма ПА	Семестр / Курс	Количество часов	Компетенции
2.1	Тема 2.1. Оценочные функции Свойства траекторий. V-функция. Уравнение Беллмана. Оптимальная стратегия. Q-функция. Принцип оптимальности Беллмана. Отказ от однородности. Вид оптимальной стратегии (доказательство через отказ от однородности). Уравнение оптимальности Беллмана. Критерий оптимальности Беллмана.	Лекционные занятия	4	2	ПК-5
2.2	Тема 2.1. Оценочные функции	Лабораторные	4	4	ПК-5

	Свойства траекторий. V-функция. Уравнение Беллмана. Оптимальная стратегия. Q-функция. Принцип оптимальности Беллмана. Отказ от однородности. Вид оптимальной стратегии (доказательство через отказ от однородности). Уравнение оптимальности Беллмана. Критерий оптимальности Беллмана.	занятия			
2.3	Тема 2.2. Динамическое программирование Метод простой итерации. Policy Evaluation. Value Iteration. Policy Iteration. Generalized Policy Iteration.	Лекционные занятия	4	2	ПК-5
2.4	Тема 2.2. Динамическое программирование Метод простой итерации. Policy Evaluation. Value Iteration. Policy Iteration. Generalized Policy Iteration.	Лабораторные занятия	4	4	ПК-5
2.5	Тема 2.3. Табличные алгоритмы Монте-Карло алгоритм. Экспоненциальное сглаживание. Стохастическая аппроксимация. Temporal Difference. Exploration-exploitation дилемма. Q-learning. SARSA.	Лабораторные занятия	4	2	ПК-5
2.6	Тема 2.4. Bias-Variance Trade-Off Дилемма смещения-разброса. N-step Temporal Difference. Backward View. Eligibility Trace. TD(lambda).	Лабораторные занятия	4	2	ПК-5
2.7	Policy Gradient подход. Policy Gradient Theorem. Схемы «Актёр-критик». Policy Gradient с переиспользованием сэмплов.	Самостоятельная работа	4	73	ПК-5
2.8	Подготовка к промежуточной аттестации	Экзамен	4	9	ПК-5

4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Структура и содержание фонда оценочных средств для проведения текущего контроля и промежуточной аттестации представлены в Приложении 1 к рабочей программе дисциплины.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1. Учебные, научные и методические издания

	Авторы, составители	Заглавие	Издательство, год	Библиотека / Количество
1	Емельянов А. А.	Прикладная информатика: журнал	Москва: Синергия ПРЕСС, 2006	ЭБС «Университетская библиотека онлайн»
2	Сараев, П. В.	Методы машинного обучения: методические указания и задания к лабораторным работам по курсу	Липецк: Липецкий государственный технический университет, ЭБС АСВ, 2017	ЭБС «IPR SMART»
3	Бутырский Е. Ю., Цехановский В. В., Жукова Н. А., Баймуратов И. Р., Куликов И. А.	Машинное обучение: учебник	Москва: Директ-Медиа, 2023	ЭБС «Университетская библиотека онлайн»

5.2. Профессиональные базы данных и информационные справочные системы

Национальная электронная библиотека (НЭБ) - <https://rusneb.ru/>

ИСС «КонсультантПлюс»

ИСС «Гарант» <http://www.internet.garant.ru/>

5.3. Перечень программного обеспечения

Операционная система РЕД ОС

LibreOffice

Python

5.4. Учебно-методические материалы для обучающихся с ограниченными возможностями здоровья

При необходимости по заявлению обучающегося с ограниченными возможностями здоровья учебно-методические материалы предоставляются в формах, адаптированных к ограничениям здоровья и восприятия информации. Для лиц с нарушениями зрения: в форме аудиофайла; в печатной форме увеличенным шрифтом. Для лиц с нарушениями слуха: в форме электронного документа; в печатной форме. Для лиц с нарушениями опорно-двигательного аппарата: в форме электронного документа; в печатной форме.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Помещения для всех видов работ, предусмотренных учебным планом, укомплектованы необходимой специализированной учебной мебелью и техническими средствами обучения:

- столы, стулья;
- персональный компьютер / ноутбук (переносной);
- проектор;

- экран / интерактивная доска.

Лабораторные занятия проводятся в компьютерных классах, рабочие места в которых оборудованы необходимыми лицензионными и/или свободно распространяемыми программными средствами и выходом в Интернет, и/или в специализированных лабораториях, предусмотренных образовательной программой.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа к электронной информационно-образовательной среде.

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Методические указания по освоению дисциплины представлены в Приложении 2 к рабочей программе дисциплины.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

1. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

1.1. Показатели и критерии оценивания компетенций:

ЗУН, составляющие компетенцию	Показатели оценивания	Критерии оценивания	Средства оценивания
ПК-5: Способен руководить проектами по созданию, поддержке и использованию системы искусственного интеллекта на основе нейросетевых моделей и методов			
3 методы и модели машинного обучения и нейронных сетей	формулирует и знает основные понятия, определения, алгоритмы и технологии	полнота и содержательность ответа умение приводить примеры	О – опрос (1-12), Э – вопросы к экзамену (1-36)
У проектировать системы искусственного интеллекта с использованием методов машинного обучения и нейронных сетей	выполняет задания, отвечает на вопросы, применяет техническое и программное обеспечение для решения задач	полнота и содержательность ответа умение приводить примеры и выполнять задания умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-8)
В навыками разработки и использования систем искусственного интеллекта с использованием методов машинного обучения и нейронных сетей	выполняет задания, проводит анализ данных и их обработку с использованием информационно-коммуникационных технологий	полнота и содержательность ответа умение приводить примеры и выполнять задания умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-8)

1.2 Шкалы оценивания:

Текущий контроль успеваемости и промежуточная аттестация осуществляется в рамках накопительной балльно-рейтинговой системы в 100-балльной шкале:

- 84-100 баллов (оценка «отлично»);
- 67-83 баллов (оценка «хорошо»);
- 50-66 баллов (оценка «удовлетворительно»);
- 0-49 баллов (оценка «неудовлетворительно»).

2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Вопросы к экзамену

- 1) Связь с оптимальным уравнением.
- 2) Марковская цепь.
- 3) Среда. Действия. Траектории.
- 4) Марковский процесс принятия решений (MDP).
- 5) Эпизодичность. Дисконтирование.
- 6) Условия задачи RL. On-policy.
- 7) Концепция model-free алгоритмов.
- 8) Классификация RL алгоритмов.
- 9) Критерии оценки RL-алгоритмов.
- 10) Сложности задачи RL. Бенчмарки.
- 11) Задача безградиентной оптимизации.
- 12) Случайный поиск. Hill Climbing. Имитация отжига.
- 13) Эволюционные алгоритмы. Weight Agnostic Neural Networks (WANN).

- 14) Видовая спецификация.
- 15) Генетические алгоритмы.
- 16) Идея эволюционных стратегий.
- 17) Оценка вероятности редкого события.
- 18) Метод КроссЭнтропии для стохастической оптимизации.
- 19) Метод Кросс-Энтропии для обучения с подкреплением (CEM).
- 20) Натуральные эволюционные стратегии (NES). OpenAI-ES.
- 21) Адаптация матрицы ковариации (CMA-ES).
- 22) Свойства траекторий. V-функция. Уравнение Беллмана.
- 23) Оптимальная стратегия. Q-функция.
- 24) Принцип оптимальности Беллмана.
- 25) Отказ от однородности. Вид оптимальной стратегии (доказательство через отказ от однородности).
- 26) Уравнение оптимальности Беллмана. Критерий оптимальности Беллмана.
- 27) Метод простой итерации. Policy Evaluation. Value Iteration. Policy Iteration. Generalized Policy Iteration.
- 28) Монте-Карло алгоритм.
- 29) Экспоненциальное сглаживание.
- 30) Стохастическая аппроксимация.
- 31) Дилемма смещения-разброса.
- 32) Q-сетка. Переход к параметрической Q-функции.
- 33) Декорреляция сэмплов.
- 34) Таргет-сеть. DQN.
- 35) Идея Distributional подхода. Z-функция. Distributional-форма.
- 36) Категориальная аппроксимация Z-функций. Categorical DQN. Квантильная аппроксимация Z-функций. Quantile Regression DQN.

Экзаменационное задание включает три вопроса – два теоретических вопроса и одно практико-ориентированное задание из числа приведенных ниже лабораторных заданий.

Критерии оценивания:

- 84-100 баллов (оценка «отлично») – изложенный материал фактически верен, наличие глубоких исчерпывающих знаний в объеме пройденной программы дисциплины в соответствии с поставленными программой курса целями и задачами обучения; правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при ответе, усвоение основной и знакомство с дополнительной литературой;

- 67-83 баллов (оценка «хорошо») – наличие твердых и достаточно полных знаний в объеме пройденной программы дисциплины в соответствии с целями обучения, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности, обучающийся усвоил основную литературу, рекомендованную в рабочей программе дисциплины;

- 50-66 баллов (оценка «удовлетворительно») – наличие твердых знаний в объеме пройденного курса в соответствии с целями обучения, изложение ответов с отдельными ошибками, уверенно исправленными после дополнительных вопросов; правильные в целом действия по применению знаний на практике;

- 0-49 баллов (оценка «неудовлетворительно») – ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы.

Задания для опроса

Вариант 1

Связь с оптимальным уравнением.

Марковская цепь.

Среда. Действия. Траектории.

Вариант 2

Марковский процесс принятия решений (MDP).

Эпизодичность. Дисконтирование.

Условия задачи RL. On-policy.

Вариант 3

Концепция model-free алгоритмов.

Классификация RL-алгоритмов.

Критерии оценки RL-алгоритмов.

Вариант 4

Сложности задачи RL. Бенчмарки.

Задача безградиентной оптимизации.

Случайный поиск. Hill Climbing. Имитация отжига.

Вариант 5

Эволюционные алгоритмы. Weight Agnostic Neural Networks (WANN).

Видовая спецификация.

Генетические алгоритмы.

Вариант 6

Идея эволюционных стратегий.

Оценка вероятности редкого события.

Метод Кросс-Энтропии для стохастической оптимизации.

Вариант 7

Метод Кросс-Энтропии для обучения с подкреплением (CEM).

Натуральные эволюционные стратегии (NES). OpenAI-ES.

Адаптация матрицы ковариации (CMA-ES).

Вариант 8

Свойства траекторий. V-функция. Уравнение Беллмана.

Оптимальная стратегия. Q-функция.

Принцип оптимальности Беллмана.

Вариант 9

Отказ от однородности. Вид оптимальной стратегии (доказательство через отказ от однородности).

Уравнение оптимальности Беллмана. Критерий оптимальности Беллмана.

Метод простой итерации. Policy Evaluation. Value Iteration. Policy Iteration. Generalized Policy Iteration.

Вариант 10

Монте-Карло алгоритм.

Экспоненциальное сглаживание.

Стохастическая аппроксимация.

Вариант 11

Дилемма смещения-разброса.

Q-сетка. Переход к параметрической Q-функции.

Декорреляция сэмплов.

Вариант 12

Таргет-сеть. DQN.

Идея Distributional подхода. Z-функция. Distributional-форма.

Категориальная аппроксимация Z-функций. Categorical DQN. Квантильная аппроксимация Z-функций. Quantile Regression DQN.

Критерии оценивания (для каждого варианта):

17-20 б. – ответы на все три вопроса варианта даны верно;

13-16 б. – один ответ из 3-х с неточностями;

9-12 б. – 2 ответа из 3-х с неточностями;

5-8 б. – 3 ответа с неточностями;

1-4 б. – нет ответа на один вопрос из 3-х;
Максимальное количество баллов за опрос – 20.

Лабораторные задания

Лабораторное задание 1.

Тема 1.1. Модель взаимодействия агента со средой Связь с оптимальным уравнением. Марковская цепь. Среда. Действия. Траектории. Марковский процесс принятия решений (MDP). Эпизодичность. Дисконтирование.

Лабораторное задание 2.

Тема 1.2. Алгоритмы обучения с подкреплением Условия задачи RL. On-policy. Концепция model-free алгоритмов. Классификация RL-алгоритмов. Критерии оценки RL-алгоритмов. Сложности задачи RL. Бенчмарки.

Лабораторное задание 3.

Тема 1.3. Бэйзлайны Задача безградиентной оптимизации. Случайный поиск. Hill Climbing. Имитация отжига. Эволюционные алгоритмы. Weight Agnostic Neural Networks (WANN). Видовая спецификация. Генетические алгоритмы.

Лабораторное задание 4.

Тема 1.4. Эволюционные стратегии Идея эволюционных стратегий. Оценка вероятности редкого события. Метод КроссЭнтропии для стохастической оптимизации. Метод Кросс-Энтропии для обучения с подкреплением (CEM). Натуральные эволюционные стратегии (NES). OpenAI-ES. Адаптация матрицы ковариации (CMA-ES).

Лабораторное задание 5.

Тема 2.1. Оценочные функции Свойства траекторий. V-функция. Уравнение Беллмана. Оптимальная стратегия. Q-функция. Принцип оптимальности Беллмана. Отказ от однородности. Вид оптимальной стратегии (доказательство через отказ от однородности). Уравнение оптимальности Беллмана. Критерий оптимальности Беллмана.

Лабораторное задание 6.

Тема 2.2. Динамическое программирование Метод простой итерации. Policy Evaluation. Value Iteration. Policy Iteration. Generalized Policy Iteration.

Лабораторное задание 7.

Тема 2.3. Табличные алгоритмы Монте-Карло алгоритм. Экспоненциальное сглаживание. Стохастическая аппроксимация. Temporal Difference. Exploration-exploitation дилемма. Q-learning. SARSA.

Лабораторное задание 8.

Тема 2.4. Bias-Variance Trade-Off Дилемма смещения-разброса. N-step Temporal Difference. Backward View. Eligibility Trace. TD(λ).

Критерии оценивания (для каждого задания):

9-10 б. – задание выполнено верно;

6-8 б. – при выполнении задания были допущены неточности, не влияющие на результат;

3-5 б. – при выполнении задания были допущены ошибки;

0-2 б. – при выполнении задания были допущены существенные ошибки.

Максимальное количество баллов за лабораторные задания – 80 (8 заданий по 10 баллов).

3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Процедуры оценивания включают в себя текущий контроль и промежуточную аттестацию.

Текущий контроль успеваемости проводится с использованием оценочных средств, представленных в п. 2 данного приложения. Результаты текущего контроля доводятся до сведения студентов до промежуточной аттестации.

Промежуточная аттестация проводится в форме экзамена.

Экзамен проводится по расписанию промежуточной аттестации. Количество вопросов в экзаменационном задании – 3 (два теоретических вопроса и одно практико-ориентированное задание). Объявление результатов производится в день экзамена. Результаты аттестации заносятся в ведомость и зачетную книжку студента. Студенты, не прошедшие промежуточную аттестацию по графику сессии, должны ликвидировать задолженность в установленном порядке.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Учебным планом предусмотрены следующие виды занятий:

- лекции;
- лабораторные занятия.

В ходе лекционных занятий рассматриваются основные теоретические вопросы, даются рекомендации для самостоятельной работы и подготовке к лабораторным занятиям.

В ходе лабораторных занятий углубляются и закрепляются знания студентов по ряду рассмотренных на лекциях вопросов, развиваются навыки практической работы.

При подготовке к лабораторным занятиям каждый студент должен:

- изучить рекомендованную учебную литературу;
- изучить конспекты лекций;
- подготовить ответы на все вопросы по изучаемой теме.

В процессе подготовки к лабораторным занятиям студенты могут воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на лекциях, лабораторных занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом опроса и выполнения лабораторных заданий. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме, дополнить конспекты лекций недостающим материалом, выписками из рекомендованных первоисточников, выделить непонятные термины, найти их значение в энциклопедических словарях.

Студент должен готовиться к предстоящему лабораторному занятию по всем обозначенным в рабочей программе дисциплины вопросам.

Для подготовки к занятиям, текущему контролю и промежуточной аттестации студенты могут воспользоваться электронно-библиотечными системами. Также обучающиеся могут взять на дом необходимую литературу на абонементе университетской библиотеки или воспользоваться читальными залами.