**ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

УТВЕРЖДАЮ:
Заведующий выпускающей кафедрой

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
ЮУрГУ
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Замышляева А. А.
Пользователь: zamyshliaevaaa
Дата подписания: 16.11.2025

А. А. Замышляева

# РАБОЧАЯ ПРОГРАММА

**дисциплины** 1.Ф.П0.09 Обучение с подкреплением
**для направления** 01.03.02 Прикладная математика и информатика
**уровень** Бакалавриат
**профиль подготовки** Искусственный интеллект, глубокое обучение и анализ данных
**форма обучения** очная
**кафедра-разработчик** Центр ОП топ-уровня в сфере ИИ "ВиртУм"

Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 01.03.02 Прикладная математика и информатика, утверждённым приказом Минобрнауки от 10.01.2018 № 9

Зав.кафедрой разработчика,
д.физ.-мат.н., проф.

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
ЮУрГУ
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Замышляева А. А.
Пользователь: zamyshliaevaaa
Дата подписания: 16.11.2025

А. А. Замышляева

Разработчик программы,
д.физ.-мат.н., проф., профессор

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
ЮУрГУ
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Соколинский Л. Б.
Пользователь: leonid.sokolinsky
Дата подписания: 13.11.2025

Л. Б. Соколинский

Челябинск

## 1. Цели и задачи дисциплины

Цели дисциплины: 1. Формирование у студентов систематизированных знаний о фундаментальных принципах и методах обучения с подкреплением. 2. Развитие практических навыков реализации и применения алгоритмов RL для решения прикладных задач. 3. Подготовка к самостоятельной работе с современными фреймворками и исследовательскими методами в области RL. Задачи курса: 1. Освоить математические основы обучения с подкреплением (MDP, уравнение Беллмана). 2. Изучить основные классы алгоритмов RL: методы временных разностей, Q-обучение, политические градиенты. 3. Сформировать понимание архитектур глубокого RL (DQN, Актор-Критик). 4. Приобрести навыки проектирования и тестирования RL-систем. 5. Научиться анализировать и интерпретировать результаты обучения RL-агентов. 6. Освоить практическое применение RL-библиотек и инструментов. 7. Выработать способность критически оценивать ограничения и перспективы методов RL.

## Краткое содержание дисциплины

Курс посвящен изучению фундаментальных принципов и современных методов обучения с подкреплением — ключевого направления искусственного интеллекта. В рамках курса рассматриваются математические основы марковских процессов принятия решений, уравнения Беллмана, а также основные классы алгоритмов: от классических методов временных разностей и Q-обучения до глубоких нейросетевых архитектур (DQN, Актор-Критик). Особое внимание уделяется практическим аспектам: проектированию систем вознаграждения, работе с симуляторами, особенностям обучения в реальных задачах. Студенты освоят популярные фреймворки и библиотеки для реализации RL-алгоритмов, научатся анализировать поведение агентов и оптимизировать гиперпараметры моделей. Курс включает выполнение лабораторных работ с постепенным усложнением — от решения классических окружений до создания собственных RL-систем.

## 2. Компетенции обучающегося, формируемые в результате освоения дисциплины

| Планируемые результаты освоения ОП ВО (компетенции) | Планируемые результаты обучения по дисциплине |
|---|---|
| ПК-22 [ML-6] Способен применять алгоритмы обучения с подкреплением | Знает: - [И-1, ПУ] основные классы RL-алгоритмов: Q-обучение, SARSA, Policy Gradient и др., их достоинства и недостатки<br>Умеет: - [И-1, ПУ] адаптировать стандартные RL-алгоритмы к условиям задачи, проводить аппроксимацию функции ценности агента, в том числе с помощью стратеги<br>Имеет практический опыт: - [И-1, ПУ] разработки адаптивного агента |
| ПК-24 [FC-3] Способен проводить фронтирные исследования в области управления, решения, агентных и мультиагентных систем | Знает: - [И-1, БУ] алгоритмы обучения с подкреплением (Q Learning, SARSA и др.;. динамическое программирование; методы Монте-Карло; принцип обучения на основе временных различий<br>Умеет: - [И-2, БУ] применять марковские |

| | |
|---|---|
| | процессы принятия решений для моделирования сред в обучении с подкреплением<br>Имеет практический опыт: - [И-1, БУ] применения алгоритмов обучения с подкреплением для решения практических задач |

## 3. Место дисциплины в структуре ОП ВО

| Перечень предшествующих дисциплин, видов работ учебного плана | Перечень последующих дисциплин, видов работ |
|---|---|
| Глубокие нейронные сети | Проектно-исследовательский семинар, Производственная практика (преддипломная, стажировка) (8 семестр) |

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

| Дисциплина | Требования |
|---|---|
| Глубокие нейронные сети | Знает: -[И-1, ПУ] принцип и алгоритмы градиентного спуска Умеет: -[И-1, ПУ] применять регуляризацию и прореживание; выбирать размер пакета для стохастического градиентного спуска[И-2, БУ] применять основные архитектуры глубокого обучения (VGG, ResNet), -[И-1, БУ] проводить аппроксимацию функции ценности агента с помощью глубоких нейронных сетей Имеет практический опыт: -[И-1, ПУ] выбора и задания скорости обучения и функции потерь в зависимости от задачи и набора данных, -[И-2, БУ] создания агентной системы с помощью глубоких нейронных сетей на основе обучения с подкреплением |

## 4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 3 з.е., 108 ч., 72,5 ч. контактной работы

| Вид учебной работы | Всего часов | Распределение по семестрам в часах |
|---|---|---|
| | | Номер семестра |
| | | 7 |
| Общая трудоёмкость дисциплины | 108 | 108 |
| *Аудиторные занятия:* | 64 | 64 |
| Лекции (Л) | 32 | 32 |
| Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ) | 0 | 0 |
| Лабораторные работы (ЛР) | 32 | 32 |
| *Самостоятельная работа (СРС)* | 35,5 | 35,5 |

| | | |
|---|---|---|
| Студенческий научный семинар по мультиагентному обучению с подкреплением | 25 | 25 |
| Подготовка к экзамену | 10,5 | 10.5 |
| Консультации и промежуточная аттестация | 8,5 | 8,5 |
| 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 | Введение в обучение с подкреплением | 4 | 2 | 0 | 2 |
| 2 | Марковские процессы принятия решений | 6 | 4 | 0 | 2 |
| 3 | Динамическое программирование | 4 | 2 | 0 | 2 |
| 4 | Методы Монте-Карло | 4 | 2 | 0 | 2 |
| 5 | Временные разности и Q-обучение | 10 | 4 | 0 | 6 |
| 6 | Аппроксимация функций и глубокое обучение с подкреплением | 10 | 4 | 0 | 6 |
| 7 | Градиент политики и методы прямого поиска стратегии | 10 | 4 | 0 | 6 |
| 8 | Продвинутые темы и современные методы | 10 | 6 | 0 | 4 |
| 9 | Инженерия окружений и практические аспекты | 6 | 4 | 0 | 2 |

### 5.1. Лекции

| № лекции | № раздела | Наименование или краткое содержание лекционного занятия | Кол-во часов |
|---|---|---|---|
| 1 | 1 | Введение в обучение с подкреплением. Ключевые понятия. Примеры задач | 2 |
| 2 | 2 | Формализация задачи как Марковского процесса принятия решений. | 2 |
| 3 | 2 | Функция ценности. Уравнение Беллмана. | 2 |
| 4 | 3 | Динамическое программирование: Итерация по ценности. Алгоритмы Policy Iteration и Value Iteration. | 2 |
| 5 | 4 | Методы Монте-Карло для предсказания ценности и управления | 2 |
| 6 | 5 | Методы временных разностей: TD-обучение и SARSA. | 2 |
| 7 | 5 | Алгоритм Q-Learning. Сравнение методов и условия сходимости. | 2 |
| 8 | 6 | Аппроксимация функции ценности. Линейные и нелинейные аппроксиматоры. | 2 |
| 9 | 6 | Глубокие Q-сети: архитектура, алгоритм обучения и проблемы. | 2 |
| 10 | 7 | Параметризация политики. Теорема о политическом градиенте. Алгоритм REINFORCE. | 2 |
| 11 | 7 | Актор-критик методы. Архитектура и варианты реализации. | 2 |
| 12 | 8 | Продвинутые методы исследования. Обратное обучение с подкреплением. | 2 |
| 13 | 8 | Имитационное обучение. Мультиагентное обучение с подкреплением. | 2 |
| 14 | 8 | Иерархическое обучение с подкреплением. | 2 |
| 15 | 9 | Инженерия окружений: проектирование состояний, действий и награды. | 2 |
| 16 | 9 | Практические аспекты: симуляторы, перенос моделей, обзор фреймворков. | 2 |

### 5.2. Практические занятия, семинары

Не предусмотрены

## 5.3. Лабораторные работы

| № занятия | № раздела | Наименование или краткое содержание лабораторной работы | Кол-во часов |
|---|---|---|---|
| 1 | 1 | Знакомство со средой разработки и библиотеками RL. Простейшая агент-среда. | 2 |
| 2 | 2 | Реализация и исследование Multi-Armed Bandit. | 2 |
| 3 | 3 | Решение GridWorld с помощью динамического программирования. | 2 |
| 4 | 4 | Реализация методов Монте-Карло для игры Blackjack. | 2 |
| 5 | 5 | Реализация и сравнение алгоритмов TD-обучения и SARSA. | 2 |
| 6 | 5 | Реализация и эксперименты с Q-Learning в среде FrozenLake. | 2 |
| 7 | 5 | Исследование влияния гиперпараметров на сходимость Q-Learning. | 2 |
| 8 | 6 | Реализация аппроксимации Q-функции с помощью линейной модели. | 2 |
| 9 | 6 | Разработка и обучение Deep Q-Network (DQN) для игры CartPole. | 4 |
| 10 | 7 | Реализация алгоритма Policy Gradient (REINFORCE). | 2 |
| 11 | 7 | Реализация архитектуры Актор-Критик (Actor-Critic). | 4 |
| 12 | 8 | Решение задачи из области игр с помощью продвинутых методов. | 4 |
| 13 | 9 | Интеграция обученной модели в тестовое окружение и финальное тестирование. | 2 |

## 5.4. Самостоятельная работа студента

| Выполнение СРС | | | |
|---|---|---|---|
| Подвид СРС | Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс | Семестр | Кол-во часов |
| Студенческий научный семинар по мультиагентному обучению с подкреплением | 1. Li Z. et al. A Comprehensive Review of Multi-Agent Reinforcement Learning in Video Games // IEEE Trans. Games. 2025. P. 1–21. DOI:10.1109/TG.2025.3588809. 2. Hu K. et al. An overview: Attention mechanisms in multi-agent reinforcement learning // Neurocomputing. 2024. Vol. 598. P. 128015. DOI:10.1016/J.NEUCOM.2024.128015. 3. Sun C., Huang S., Pompili D. LLM-based Multi-Agent Reinforcement Learning: Current and Future Directions // arXiv:2405.11106 [cs.MA]. 2024. DOI:10.48550/arXiv.2405.11106. | 7 | 25 |
| Подготовка к экзамену | Основная литература. Дополнительная литература. | 7 | 10,5 |

## 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 | Учи-тыва-ется в ПА |
|---|---|---|---|---|---|---|---|
| 1 | 7 | Текущий контроль | Контрольный опрос 1 | 1 | 3 | Контрольный опрос оформлен в виде теста из 3 вопросов. Каждый правильный ответ оценивается 1 баллом. Продолжительность теста 5 мин.<br>3 балла: даны верные ответы на все вопросы теста.<br>2 балла: даны верные ответы на 2 вопроса теста.<br>1 балл: дан верный ответ на 1 вопрос теста.<br>0 баллов: верные ответы отсутствуют | экзамен |
| 2 | 7 | Текущий контроль | Контрольный опрос 2 | 1 | 3 | Контрольный опрос оформлен в виде теста из 3 вопросов. Каждый правильный ответ оценивается 1 баллом. Продолжительность теста 5 мин.<br>3 балла: даны верные ответы на все вопросы теста.<br>2 балла: даны верные ответы на 2 вопроса теста.<br>1 балл: дан верный ответ на 1 вопрос теста.<br>0 баллов: верные ответы отсутствуют | экзамен |
| 3 | 7 | Текущий контроль | Контрольный опрос 3 | 1 | 3 | Контрольный опрос оформлен в виде теста из 3 вопросов. Каждый правильный ответ оценивается 1 баллом. Продолжительность теста 5 мин.<br>3 балла: даны верные ответы на все вопросы теста.<br>2 балла: даны верные ответы на 2 вопроса теста.<br>1 балл: дан верный ответ на 1 вопрос теста.<br>0 баллов: верные ответы отсутствуют | экзамен |
| 4 | 7 | Текущий контроль | Контрольный опрос 4 | 1 | 3 | Контрольный опрос оформлен в виде теста из 3 вопросов. Каждый правильный ответ оценивается 1 баллом. Продолжительность теста 5 мин.<br>3 балла: даны верные ответы на все вопросы теста.<br>2 балла: даны верные ответы на 2 вопроса теста.<br>1 балл: дан верный ответ на 1 вопрос | экзамен |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | теста.<br>0 баллов: верные ответы отсутствуют | |
| 5 | 7 | Текущий контроль | Контрольный опрос 5 | 1 | 3 | Контрольный опрос оформлен в виде теста из 3 вопросов. Каждый правильный ответ оценивается 1 баллом. Продолжительность теста 5 мин.<br>3 балла: даны верные ответы на все вопросы теста.<br>2 балла: даны верные ответы на 2 вопроса теста.<br>1 балл: дан верный ответ на 1 вопрос теста.<br>0 баллов: верные ответы отсутствуют | экзамен |
| 6 | 7 | Текущий контроль | Контрольный опрос 6 | 1 | 3 | Контрольный опрос оформлен в виде теста из 3 вопросов. Каждый правильный ответ оценивается 1 баллом. Продолжительность теста 5 мин.<br>3 балла: даны верные ответы на все вопросы теста.<br>2 балла: даны верные ответы на 2 вопроса теста.<br>1 балл: дан верный ответ на 1 вопрос теста.<br>0 баллов: верные ответы отсутствуют | экзамен |
| 7 | 7 | Лабораторная работа | Лабораторная работа 1 | 1 | 5 | • 5 баллов: Полная реализация, анализ результатов, ответы на вопросы.<br>• 3-4 балла: Реализация без анализа или пропущены ответы на вопросы.<br>• 0-2 балла: Не выполнены ключевые задачи. | экзамен |
| 8 | 7 | Лабораторная работа | Лабораторная работа 2 | 1 | 5 | • 5 баллов: Полная реализация 3+ алгоритмов, глубокий анализ, ответы на вопросы<br>• 4 балла: Реализация 2 алгоритмов с небольшими ошибками<br>• 3 балла: Работает только базовый функционал<br>• 0-2 балла: Код не работает или отсутствует | экзамен |
| 9 | 7 | Лабораторная работа | Лабораторная работа 3 | 1 | 5 | • 5 баллов: Полная реализация обоих алгоритмов, анализ сходимости, визуализация<br>• 4 балла: Реализация одного алгоритма с полным анализом<br>• 3 балла: Реализация с ошибками, но работающая на простых случаях<br>• 0-2 балла: Алгоритмы не работают или отсутствует анализ | экзамен |
| 10 | 7 | Лабораторная | Лабораторная | 1 | 5 | • 5 баллов: Реализация MC prediction | экзамен |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | работа | работа 4 | | | и MC control, полный анализ результатов<br>• 4 балла: Реализация MC prediction с визуализацией<br>• 3 балла: Реализация только сбора эпизодов без оценки функции ценности<br>• 0-2 балла: Код не работает или отсутствует | |
| 11 | 7 | Лабораторная работа | Лабораторная работа 5 | 1 | 5 | • 5 баллов: Полная реализация обоих алгоритмов, сравнительный анализ, визуализация<br>• 4 балла: Реализация одного алгоритма с полным анализом<br>• 3 балла: Базовая реализация без анализа гиперпараметров<br>• 0-2 балла: Алгоритмы не работают или отсутствует анализ | экзамен |
| 12 | 7 | Лабораторная работа | Лабораторная работа 6 | 1 | 5 | • 5 баллов: Полная реализация Q-Learning, исследование гиперпараметров, сравнение сред<br>• 4 балла: Реализация алгоритма с анализом в одной среде<br>• 3 балла: Базовая реализация без исследования гиперпараметров<br>• 0-2 балла: Алгоритм не работает или отсутствует анализ | экзамен |
| 13 | 7 | Лабораторная работа | Лабораторная работа 7 | 1 | 5 | • 5 баллов: Полное исследование всех гиперпараметров, анализ взаимодействий, обоснованные рекомендации<br>• 4 балла: Исследование 2-3 гиперпараметров с подробным анализом<br>• 3 балла: Базовое исследование 1-2 гиперпараметров без анализа взаимодействий<br>• 0-2 балла: Отсутствие систематического исследования или некорректные выводы | экзамен |
| 14 | 7 | Лабораторная работа | Лабораторная работа 8 | 1 | 5 | • 5 баллов: Полная реализация линейной аппроксимации, анализ обобщения, сравнение методов<br>• 4 балла: Реализация аппроксимации с тестированием на обучающих данных<br>• 3 балла: Базовая реализация без анализа обобщающей способности<br>• 0-2 балла: Алгоритм не работает или отсутствует анализ | экзамен |
| 15 | 7 | Лабораторная работа | Лабораторная работа 9 | 1 | 5 | • 5 баллов: Полная реализация DQN, успешное обучение, глубокий анализ<br>• 4 балла: Реализация DQN с обучением, но без достижения | экзамен |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | критерия успеха<br>• 3 балла: Частичная реализация (только сеть или только буфер)<br>• 0-2 балла: Алгоритм не работает или отсутствует анализ | |
| 16 | 7 | Лабораторная работа | Лабораторная работа 10 | 1 | 5 | • 5 баллов: Полная реализация REINFORCE, анализ градиентов, сравнение с базовой линией<br>• 4 балла: Реализация базового REINFORCE с полным анализом результатов<br>• 3 балла: Частичная реализация без анализа градиентов<br>• 0-2 балла: Алгоритм не работает или отсутствует анализ | экзамен |
| 17 | 7 | Лабораторная работа | Лабораторная работа 11 | 1 | 5 | • 5 баллов: Полная реализация Actor-Critic и A2C, сравнение архитектур, глубокий анализ<br>• 4 балла: Реализация базового Actor-Critic с анализом взаимодействия компонентов<br>• 3 балла: Частичная реализация без полного анализа<br>• 0-2 балла: Алгоритм не работает или отсутствует анализ взаимодействия компонентов | экзамен |
| 18 | 7 | Лабораторная работа | Лабораторная работа 12 | 1 | 5 | • 5 баллов: Реализация сложного алгоритма, глубокий анализ, сравнение методов<br>• 4 балла: Полная реализация с анализом, но без сравнения<br>• 3 балла: Базовая реализация без полного анализа<br>• 0-2 балла: Неполная реализация или отсутствие анализа | экзамен |
| 19 | 7 | Лабораторная работа | Лабораторная работа 13 | 1 | 5 | • 5 баллов: Полная реализация системы тестирования, комплексный анализ, артефакты<br>• 4 балла: Реализация основных тестов без некоторых улучшений<br>• 3 балла: Базовое тестирование без комплексного анализа<br>• 0-2 балла: Неполная реализация или отсутствие анализа | экзамен |
| 20 | 7 | Проме-жуточная аттестация | Итоговый тест | - | 18 | Тест содержит 18 равноценных вопросов и рассчитан на 45 мин. За каждый правильный ответ начисляется 1 балл. Максимальное количество баллов за итоговый тест равно 18. | экзамен |
| 21 | 7 | Бонус | Участие в студенческом научном семинаре по мультиагентному | - | 15 | За выступление с докладом начисляется 10 баллов.<br>За участие в семинаре начисляется 1 балл. | экзамен |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | обучению с подкреплением | | | | |

## 6.2. Процедура проведения, критерии оценивания

| Вид промежуточной аттестации | Процедура проведения | Критерии оценивания |
|---|---|---|
| экзамен | Оценка за дисциплину формируется на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля следующим образом: • Отлично: Величина рейтинга обучающегося по дисциплине 85…100 %. • Хорошо: Величина рейтинга обучающегося по дисциплине 75…84 %. • Удовлетворительно: Величина рейтинга обучающегося по дисциплине 60…74 %. • Неудовлетворительно: Величина рейтинга обучающегося по дисциплине 0…59 %. Если студент согласен с оценкой, полученной по результатам текущего контроля, то он может в день, предшествующий промежуточной аттестации дать свое согласие в личном кабинете. В случае явки студента на промежуточную аттестацию, давшего свое согласие на оценку в личном кабинете, студент имеет право пройти мероприятия текущего контроля по дисциплине на промежуточной аттестации для улучшения своего рейтинга в день ее проведения. Снижение оценки в этом случае запрещено. Если студент не дал согласия в личном кабинете, то он может согласиться с оценкой лично на промежуточной аттестации в день ее проведения. Если студент не согласен с оценкой, то он имеет право пройти контрольно-рейтинговые мероприятия на промежуточной аттестации для улучшения своего рейтинга в день ее проведения. Фиксация результатов учебной деятельности по дисциплине проводится в день промежуточной аттестации на основе согласия студента, данного им в личном кабинете. При отсутствии согласия в журнале дисциплины фиксация результатов происходит при личном присутствии студента. Если студент не дал согласие в личном кабинете и не явился на промежуточную аттестацию – ему выставляется «неявка». Промежуточная аттестация проводится в форме тестирования. Тестирование проводится в системе edu.susu.ru. Тест содержит 18 вопросов. На выполнение теста дается 45 минут. В этом случае оценка за дисциплину рассчитывается на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля и промежуточной аттестации. | 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 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| ПК-22 | Знает: - [И-1, ПУ] основные классы RL-алгоритмов: Q-обучение, SARSA, Policy Gradient и др., их достоинства и недостатки | + | + | + | + | | | + | | | | + | + | + | + | + | + | + | | | + | |
| ПК-22 | Умеет: - [И-1, ПУ] адаптировать стандартные RL-алгоритмы к условиям задачи, проводить аппроксимацию | | | + | + | | | + | | | | + | + | + | + | + | + | + | | | | |

| | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | функции ценности агента, в том числе с помощью стратеги | | | | | | | | | | | | | | | | | | | | | |
| ПК-22 | Имеет практический опыт: - [И-1, ПУ] разработки адаптивного агента | | | | | | + | | | + | + | + | + | + | + | + | | | | | | |
| ПК-24 | Знает: - [И-1, БУ] алгоритмы обучения с подкреплением (Q Learning, SARSA и др.;. динамическое программирование; методы Монте-Карло; принцип обучения на основе временных различий | | + | | + | + | | + | + | + | | | | | | | | + | + | + | + | |
| ПК-24 | Умеет: - [И-2, БУ] применять марковские процессы принятия решений для моделирования сред в обучении с подкреплением | + | | | | | | + | + | + | | | | | | | | + | + | + | + | |
| ПК-24 | Имеет практический опыт: - [И-1, БУ] применения алгоритмов обучения с подкреплением для решения практических задач | | | | | | | + | + | + | | | | | | | | + | + | | + | |

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:*

1. Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика / Юж.-Урал. гос. ун-т; ЮУрГУ. - Челябинск : Издательство ЮУрГУ, 2012-. -. URL: http://vestnik.susu.ac.ru/

*г) методические указания для студентов по освоению дисциплины:*

1. 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
2. Иванов С. Конспект по обучению с подкреплением. [https://education.yandex.ru/handbook/ml/article/obuchenie-s-podkrepleniem]
3. Уиндер Ф. Обучение с подкреплением для реальных задач. Инженерный подход. СПб.: БХВ-Петербург, 2023. 400 р.

*из них: учебно-методическое обеспечение самостоятельной работы студента:*

1. 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

2. Иванов С. Конспект по обучению с подкреплением. [https://education.yandex.ru/handbook/ml/article/obuchenie-s-podkrepleniem]

3. Уиндер Ф. Обучение с подкреплением для реальных задач. Инженерный подход. СПб.: БХВ-Петербург, 2023. 400 р.

**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. https://e.lanbook.com/book/179453 |
| 2 | Дополнительная 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. https://e.lanbook.com/book/179493 |

Перечень используемого программного обеспечения:

1. Microsoft-Office(бессрочно)
2. -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 |
|---|---|---|
| Экзамен | 110 (3г) | Компьютерный класс |
| Лекции | 484 (3) | Проектор, компьютер |
| Лабораторные занятия | 110 (3г) | Компьютерный класс |