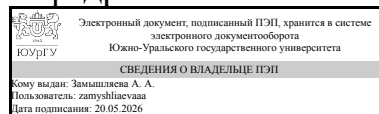


# ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:  
Заведующий выпускающей  
кафедрой



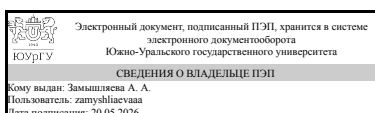
А. А. Замышляева

## 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"

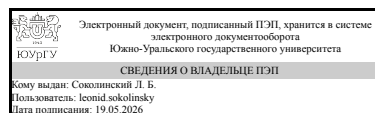
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 01.03.02 Прикладная математика и информатика, утверждённым приказом Минобрнауки от 10.01.2018 № 9

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



А. А. Замышляева

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



Л. Б. Соколинский

## 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-23 [ML-6] Способен применять алгоритмы обучения с подкреплением   | Знает: - [И-1, СУ] основные классы RL-алгоритмов: Q-обучение, SARSA, Policy Gradient и др., их достоинства и недостатки<br>Умеет: - [И-1, СУ] адаптировать стандартные RL-алгоритмы к условиям задачи, проводить аппроксимацию функции ценности агента, в том числе с помощью стратегии<br>Имеет практический опыт: - [И-1, СУ] разработки адаптивного агента |
| ПК-25 [FC-3] Способен проводить передовые 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-2, СУ] стандартные метрики оценки качества (accuracy, F1, ROC-AUC), -[И-1, СУ] принцип и алгоритмы градиентного спуска, -[И-1, СУ] TD-методы и методы Монте-Карло для обучения агента Умеет: -[И-1, БУ] проводить аппроксимацию функции ценности агента с помощью глубоких нейронных сетей; реализовать поиск по дереву методом Монте-Карло, -[И-2, СУ] использовать стандартные метрики оценки качества (accuracy, F1, ROC-AUC) и сопоставлять базовые модели между собой, -[И-1, СУ] применять регуляризацию и прореживание; выбирать размер пакета для стохастического градиентного спуска [И-2, БУ] применять основные архитектуры глубокого обучения (VGG, ResNet), -[И-1, СУ] задавать цель агента с помощью полного вознаграждения, вознаграждения с обесценением, лямбда-дохода Имеет практический опыт: -[И-2, БУ] создания агентной системы с помощью глубоких нейронных сетей на основе обучения с подкреплением, -[И-1, СУ] выбора и задания скорости обучения и функции потерь в зависимости от задачи и набора данных, -[И-1, СУ] использования TD-методов и методов Монте-Карло для обучения агента |

### 4. Объём и виды учебной работы

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

|                    |       |                            |
|--------------------|-------|----------------------------|
| Вид учебной работы | Всего | Распределение по семестрам |
|--------------------|-------|----------------------------|

|  | часов | в часах        |  |
|--|-------|----------------|--|
|  |       | Номер семестра |  |
|  |       | 7              |  |
| Общая трудоёмкость дисциплины  | 108   | 108            |  |
| <i>Аудиторные занятия:</i>   | 64    | 64             |  |
| Лекции (Л)   | 32    | 32             |  |
| Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ) | 0     | 0              |  |
| Лабораторные работы (ЛР)   | 32    | 32             |  |
| <i>Самостоятельная работа (СРС)</i>  | 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                                | 2            |

|    |   |   |   |
|----|---|---|---|
|    |   | REINFORCE.  |   |
| 11 | 7 | Актер-критик методы. Архитектура и варианты реализации.               | 2 |
| 12 | 8 | Продвинутое 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 Актер-Критик (Actor-Critic).                        | 4            |
| 12        | 8         | Решение задачи из области игр с помощью продвинутого 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   | Учитывается 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             | экзамен          |

|   |   |                     |                       |   |   |   |         |
|---|---|---------------------|-----------------------|---|---|---|---------|
|   |   |                     |                       |   |   | отсутствуют   |         |
| 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 теста.<br>0 баллов: верные ответы отсутствуют | экзамен |
| 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 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация, анализ результатов, ответы на вопросы.</li> <li>• 3-4 балла: Реализация без анализа или пропущены ответы на вопросы.</li> <li>• 0-2 балла: Не выполнены ключевые задачи.</li> </ul>   | экзамен |
| 8 | 7 | Лабораторная работа | Лабораторная работа 2 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация 3+ алгоритмов, глубокий анализ, ответы на вопросы</li> <li>• 4 балла: Реализация 2 алгоритмов с небольшими ошибками</li> <li>• 3 балла: Работает только базовый функционал</li> <li>• 0-2 балла: Код не работает или</li> </ul>                              | экзамен |

|    |   |                     |                       |   |   |   |         |
|----|---|---------------------|-----------------------|---|---|---|---------|
|    |   |                     |                       |   |   | отсутствует   |         |
| 9  | 7 | Лабораторная работа | Лабораторная работа 3 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация обоих алгоритмов, анализ сходимости, визуализация</li> <li>• 4 балла: Реализация одного алгоритма с полным анализом</li> <li>• 3 балла: Реализация с ошибками, но работающая на простых случаях</li> <li>• 0-2 балла: Алгоритмы не работают или отсутствует анализ</li> </ul>  | экзамен |
| 10 | 7 | Лабораторная работа | Лабораторная работа 4 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Реализация MC prediction и MC control, полный анализ результатов</li> <li>• 4 балла: Реализация MC prediction с визуализацией</li> <li>• 3 балла: Реализация только сбора эпизодов без оценки функции ценности</li> <li>• 0-2 балла: Код не работает или отсутствует</li> </ul>  | экзамен |
| 11 | 7 | Лабораторная работа | Лабораторная работа 5 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация обоих алгоритмов, сравнительный анализ, визуализация</li> <li>• 4 балла: Реализация одного алгоритма с полным анализом</li> <li>• 3 балла: Базовая реализация без анализа гиперпараметров</li> <li>• 0-2 балла: Алгоритмы не работают или отсутствует анализ</li> </ul>  | экзамен |
| 12 | 7 | Лабораторная работа | Лабораторная работа 6 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация Q-Learning, исследование гиперпараметров, сравнение сред</li> <li>• 4 балла: Реализация алгоритма с анализом в одной среде</li> <li>• 3 балла: Базовая реализация без исследования гиперпараметров</li> <li>• 0-2 балла: Алгоритм не работает или отсутствует анализ</li> </ul>  | экзамен |
| 13 | 7 | Лабораторная работа | Лабораторная работа 7 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полное исследование всех гиперпараметров, анализ взаимодействий, обоснованные рекомендации</li> <li>• 4 балла: Исследование 2-3 гиперпараметров с подробным анализом</li> <li>• 3 балла: Базовое исследование 1-2 гиперпараметров без анализа взаимодействий</li> <li>• 0-2 балла: Отсутствие систематического исследования или некорректные выводы</li> </ul> | экзамен |
| 14 | 7 | Лабораторная работа | Лабораторная работа 8 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация линейной аппроксимации, анализ обобщения, сравнение методов</li> <li>• 4 балла: Реализация</li> </ul>  | экзамен |

|    |   |                     |                        |   |   |  |         |
|----|---|---------------------|------------------------|---|---|--|---------|
|    |   |                     |                        |   |   | <p>аппроксимации с тестированием на обучающих данных</p> <ul style="list-style-type: none"> <li>• 3 балла: Базовая реализация без анализа обобщающей способности</li> <li>• 0-2 балла: Алгоритм не работает или отсутствует анализ</li> </ul>  |         |
| 15 | 7 | Лабораторная работа | Лабораторная работа 9  | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация DQN, успешное обучение, глубокий анализ</li> <li>• 4 балла: Реализация DQN с обучением, но без достижения критерия успеха</li> <li>• 3 балла: Частичная реализация (только сеть или только буфер)</li> <li>• 0-2 балла: Алгоритм не работает или отсутствует анализ</li> </ul>  | экзамен |
| 16 | 7 | Лабораторная работа | Лабораторная работа 10 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация REINFORCE, анализ градиентов, сравнение с базовой линией</li> <li>• 4 балла: Реализация базового REINFORCE с полным анализом результатов</li> <li>• 3 балла: Частичная реализация без анализа градиентов</li> <li>• 0-2 балла: Алгоритм не работает или отсутствует анализ</li> </ul>                                     | экзамен |
| 17 | 7 | Лабораторная работа | Лабораторная работа 11 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация Actor-Critic и A2C, сравнение архитектур, глубокий анализ</li> <li>• 4 балла: Реализация базового Actor-Critic с анализом взаимодействия компонентов</li> <li>• 3 балла: Частичная реализация без полного анализа</li> <li>• 0-2 балла: Алгоритм не работает или отсутствует анализ взаимодействия компонентов</li> </ul> | экзамен |
| 18 | 7 | Лабораторная работа | Лабораторная работа 12 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Реализация сложного алгоритма, глубокий анализ, сравнение методов</li> <li>• 4 балла: Полная реализация с анализом, но без сравнения</li> <li>• 3 балла: Базовая реализация без полного анализа</li> <li>• 0-2 балла: Неполная реализация или отсутствие анализа</li> </ul>   | экзамен |
| 19 | 7 | Лабораторная работа | Лабораторная работа 13 | 1 | 5 | <ul style="list-style-type: none"> <li>• 5 баллов: Полная реализация системы тестирования, комплексный анализ, артефакты</li> <li>• 4 балла: Реализация основных тестов без некоторых улучшений</li> <li>• 3 балла: Базовое тестирование без комплексного анализа</li> <li>• 0-2 балла: Неполная реализация или отсутствие анализа</li> </ul>  | экзамен |

|    |   |                          |   |   |    |   |         |
|----|---|--------------------------|---|---|----|---|---------|
| 20 | 7 | Промежуточная 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 обучению с подкреплением | - | 15 | За выступление с докладом начисляется 10 баллов. За участие в семинаре начисляется 1 балл.  | экзамен |

## 6.2. Процедура проведения, критерии оценивания

| Вид промежуточной аттестации | Процедура проведения  | Критерии оценивания                     |
|------------------------------|---|---|
| экзамен                      | <p>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 согласия 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.</p> | 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-23       | Знает: - [И-1, СУ] основные классы RL-алгоритмов: Q-обучение, SARSA, Policy Gradient и др., их достоинства и недостатки   |      |   | + | + | + |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    | +  |
| ПК-23       | Умеет: - [И-1, СУ] адаптировать стандартные RL-алгоритмы к условиям задачи, проводить аппроксимацию функции ценности агента, в том числе с помощью стратегии                      |      |   | + | + | + |   |   | + |   | +  | +  | +  | +  | +  | +  | +  |    |    |    |    | +  |
| ПК-23       | Имеет практический опыт: - [И-1, СУ] разработки адаптивного агента  |      |   |   |   |   |   |   | + |   | +  | +  | +  | +  | +  | +  | +  |    |    |    |    | +  |
| ПК-25       | 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-25       | Умеет: - [И-2, БУ] применять марковские процессы принятия решений для моделирования сред в обучении с подкреплением   |      | + |   |   | + | + | + |   | + |    |    |    |    |    |    |    |    |    | +  | +  | +  |
| ПК-25       | 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. Иванов С. Конспект по обучению с подкреплением.  
[<https://education.yandex.ru/handbook/ml/article/obuchenie-s-podkrepleniem>]

2. 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>

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

2. 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>

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<br>издательства<br>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. <a href="https://e.lanbook.com/book/179453">https://e.lanbook.com/book/179453</a> |
| 2 | Дополнительная литература | ЭБС<br>издательства<br>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. <a href="https://e.lanbook.com/book/179493">https://e.lanbook.com/book/179493</a>         |

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 |
|----------------------|-------------|-------------------------|
| Лабораторные занятия | 110<br>(3г) | Компьютерный класс      |
| Экзамен              | 110<br>(3г) | Компьютерный класс      |
| Лекции               | 484<br>(3)  | Проектор, компьютер     |