

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Ястребов Олег Александрович  
Должность: Ректор  
Дата подписания: 01.06.2025 14:30:54  
Уникальный программный ключ:  
sa953a0120d891083f939673078ef1a989dae18a

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

**Инженерная академия**

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

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **ТЕОРИЯ И ПРАКТИКА ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ**

(наименование дисциплины/модуля)

**Рекомендована МССН для направления подготовки/специальности:**

**13.04.03 ЭНЕРГЕТИЧЕСКОЕ МАШИНОСТРОЕНИЕ /  
27.04.04 УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ**

(код и наименование направления подготовки/специальности)

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

### **ИНТЕЛЛЕКТУАЛЬНЫЕ ЭНЕРГЕТИЧЕСКИЕ СИСТЕМЫ**

(наименование (профиль/специализация) ОП ВО)

**2025 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Теория и практика обучения с подкреплением» входит в программу магистратуры «Интеллектуальные энергетические системы» по направлениям 13.04.03 «Энергетическое машиностроение» / 27.04.04 «Управление в технических системах» и изучается в 3 семестре 2 курса. Дисциплину реализует Кафедра механики и процессов управления. Дисциплина состоит из 7 разделов и 21 тема и направлена на изучение - базовых принципов обучения с подкреплением; - основных алгоритмов обучения с подкреплением. - терминологического аппарата обучения с подкреплением. - библиотеки глубокого обучения Tensorflow.

Целью освоения дисциплины является познакомиться и овладеть базовыми принципами и основными алгоритмами обучения с подкреплением для применения их в своей профессиональной деятельности.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Теория и практика обучения 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<br>(в рамках данной дисциплины)                                                                                                                                                                                                                                                                                                                                   |
|-------|---------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ОПК-8 | Способен выбирать методы и разрабатывать системы управления сложными техническими объектами и технологическими процессами | ОПК-8.1 Знает основные методы, применяемые для разработки систем управления сложными техническими объектами и технологическими процессами;;<br>ОПК-8.2 Умеет разрабатывать системы управления сложными техническими объектами и технологическими процессами;;<br>ОПК-8.3 Имеет навыки выбора методов и разработки систем управления сложными техническими объектами и технологическими процессами.; |
| ПК-3  | Способен решать прикладные задачи в области искусственного интеллекта                                                     | ПК-3.1 Знаком с основными методами и подходами, применяемыми для решения задач в области искусственного интеллекта;;<br>ПК-3.2 Владеет методами решения профессиональных задач в области искусственного интеллекта ;;<br>ПК-3.3 Умеет применять математические методы и современные информационные технологии при проведении научных исследований.;                                                 |

## 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Теория и практика обучения 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 «Теория и практика обучения 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-8 | Способен выбирать методы и разрабатывать системы управления сложными техническими объектами и технологическими процессами | Современные методы машинного обучения;<br>Глубокое обучение и генеративные модели; |                                          |
| ПК-3  | Способен решать прикладные задачи в области искусственного интеллекта                                                     | <i>Большие языковые модели и агенты**;</i><br><i>Виртуальная реальность**;</i>     |                                          |

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

\*\* - элективные дисциплины /практики

#### 4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Теория и практика обучения с подкреплением» составляет «5» зачетных единиц.

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) |
|-------------------------------------------|--------------|-----|-------------|
|                                           |              |     | 3           |
| Контактная работа, ак.ч.                  | 36           |     | 36          |
| Лекции (ЛК)                               | 18           |     | 18          |
| Лабораторные работы (ЛР)                  | 18           |     | 18          |
| Практические/семинарские занятия (СЗ)     | 0            |     | 0           |
| Самостоятельная работа обучающихся, ак.ч. | 117          |     | 117         |
| Контроль (экзамен/зачет с оценкой), ак.ч. | 27           |     | 27          |
| Общая трудоемкость дисциплины             | ак.ч.        | 180 | 180         |
|                                           | зач.ед.      | 5   | 5           |

## 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      | Введение в обучение с подкреплением       | 1.1                       | Постановка задачи обучения с подкреплением. Отличия обучения с подкреплением от обучения с учителем и обучения без учителя. Области и примеры использования задачи обучения с подкреплением                                             | ЛК, ЛР              |
|               |                                           | 1.2                       | Основные компоненты обучения с подкреплением: функция поведения агента, функция ценности действий агента и модель представления окружения агентом                                                                                       | ЛК, ЛР              |
| Раздел 2      | Основные понятия обучения с подкреплением | 2.1                       | Понятие рационального агента. Понятие окружения                                                                                                                                                                                         | ЛК, ЛР              |
|               |                                           | 2.2                       | Понятие действия агента в обучении подкреплением. Понятие награды агента в обучении подкреплением. Понятие истории в обучении подкреплением. Понятие состояния агента в обучении подкреплением                                          | ЛК, ЛР              |
|               |                                           | 2.3                       | Алгоритмы принятия решения и награждения агента в обучении с подкреплением                                                                                                                                                              | ЛК, ЛР              |
| Раздел 3      | Задача "многоруких бандитов"              | 3.1                       | Постановка задачи "многоруких бандитов". Области и примеры использования задачи "многоруких бандитов"                                                                                                                                   | ЛК, ЛР              |
|               |                                           | 3.2                       | Обзор методов решения задачи "многоруких бандитов": жадный метод, эпсилон жадный (англ. epsilon-greedy) метод, метод верхней доверительной границы (англ. upper confidence bound), метод gradient-bandit и др. Методы ценности действия | ЛК, ЛР              |
| Раздел 4      | Марковский процесс принятия решений       | 4.1                       | Постановка задачи марковского процесса принятия решений                                                                                                                                                                                 | ЛК, ЛР              |
|               |                                           | 4.2                       | Основные компоненты марковского процесса принятия решений: множество состояний, множество действий, функции моментальных поощрений, функции перехода между состояниями                                                                  | ЛК, ЛР              |
|               |                                           | 4.3                       | Эпизодические и непрерывные задачи. Примеры марковских процессов принятия решений                                                                                                                                                       | ЛК, ЛР              |
| Раздел 5      | Динамическое программирование             | 5.1                       | Постановка задачи динамического программирования. Применение динамического программирования в обучении подкреплением                                                                                                                    | ЛК, ЛР              |
|               |                                           | 5.2                       | Применение уравнения Беллмана в обучении подкреплением                                                                                                                                                                                  | ЛК, ЛР              |
|               |                                           | 5.3                       | Оценка политики рационального агента. Методы Policy Iteration, Policy Evaluation, Value Iteration и Value Evaluation. Проблемы динамического программирования                                                                           | ЛК, ЛР              |
| Раздел 6      | Методы Монте-Карло                        | 6.1                       | Постановка задачи обучения с подкреплением на основе полного эпизода действий агента. Метод Монте-Карло в обучении с подкреплением                                                                                                      | ЛК, ЛР              |
|               |                                           | 6.2                       | Оценка ценности действий с помощью метода Монте-Карло. Контроль Монте-Карло. Преимущества и недостатки метода метода Монте-Карло                                                                                                        | ЛК, ЛР              |
|               |                                           | 6.3                       | Примеры задач обучения с подкреплением использующие метод Монте-Карло                                                                                                                                                                   | ЛК, ЛР              |

| Номер раздела | Наименование раздела дисциплины           | Содержание раздела (темы) |                                                                                                           | Вид учебной работы* |
|---------------|-------------------------------------------|---------------------------|-----------------------------------------------------------------------------------------------------------|---------------------|
| Раздел 7      | Нейронные сети и обучение с подкреплением | 7.1                       | Понятие Q-обучения (англ. Q-learning)                                                                     | ЛК, ЛР              |
|               |                                           | 7.2                       | Архитектуры нейронных сетей на основе сверточных нейронных сетей и на основе рекуррентных нейронных сетей | ЛК, ЛР              |
|               |                                           | 7.3                       | Целевые функции                                                                                           | ЛК, ЛР              |
|               |                                           | 7.4                       | Особенности обучения: исследования и эксплуатации, политика принятия решений, параллельное обучение       | ЛК, ЛР              |
|               |                                           | 7.5                       | Нейронная сеть Value Iteration Network                                                                    | ЛК, ЛР              |

\* - 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 [Параметр] шт.), доской (экраном) 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. Ростовцев, В. С. Искусственные нейронные сети : учебник / В. С. Ростовцев. - Санкт-Петербург : Лань, 2019. - 216 с. - ISBN 978-5-8114-3768-9. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/122180> (дата обращения: 06.03.2020). - Режим доступа: для авториз. пользователей
2. Осипов, Г. С. Методы искусственного интеллекта : монография / Г. С. Осипов. - Москва : Физматлит, 2011. - 296 с. - ISBN 978-5-9221-1323-6. - Текст : электронный. -

URL: <https://znanium.com/catalog/product/544787> (дата обращения: 06.03.2020). - Режим доступа: по подписке

3. Жданов, А. А. Автономный искусственный интеллект : учебное пособие / А. А. Жданов. - 5-е изд. - Москва : Лаборатория знаний, 2020. - 362 с. - ISBN 978-5-00101-655-7. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/135544> (дата обращения: 06.03.2020). - Режим доступа: для авториз. пользователей

4. Вьюгин, В. В. Математические основы машинного обучения и прогнозирования : учебное пособие / В. В. Вьюгин. - Москва : МЦНМО, 2014. - 304 с. - ISBN 978-5-4439-2014-6. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/56397> (дата обращения: 06.03.2020). - Режим доступа: для авториз. пользователей

*Дополнительная литература:*

1. Алексеева, Т. В. Информационные аналитические системы [Электронный ресурс] : учебник / Т. В. Алексеева, Ю. В. Амириди, В. В. Дик и др.; под ред. В. В. Дика. - Москва : МФПУ Синергия, 2013. - 384 с. - (Университетская серия). - ISBN 978-5-4257-0092-6. - Текст : электронный. - URL: <https://znanium.com/catalog/product/451186> (дата обращения: 06.03.2020). - Режим доступа: по подписке

2. Смолин, Д. В. Введение в искусственный интеллект: конспект лекций : учебное пособие / Д. В. Смолин. - 2-е изд., перераб. - Москва : ФИЗМАТЛИТ, 2007. - 264 с. - ISBN 978-5-9221-0862-1. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/2325> (дата обращения: 06.03.2020). - Режим доступа: для авториз. пользователей

3. Гвоздева, В. А. Введение в специальность программиста : учебник / В.А. Гвоздева. - 2-е изд., испр. и доп. - М. : ИД 'ФОРУМ' : ИНФРА-М, 2019. - 208 с. : ил. - (Профессиональное образование). - ISBN 978-5-8199-0297-4. - Текст : электронный. - URL: <https://znanium.com/catalog/product/988422> (дата обращения: 06.03.2020). - Режим доступа: по подписке

*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

<https://mega.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](http://www.studentlibrary.ru)

- ЭБС «Знаниум» <https://znanium.ru/>

2. Базы данных и поисковые системы

- Sage <https://journals.sagepub.com/>

- Springer Nature Link <https://link.springer.com/>

- Wiley Journal Database <https://onlinelibrary.wiley.com/>

- Научометрическая база данных Lens.org <https://www.lens.org>

*Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля\*:*

1. Курс лекций по дисциплине «Теория и практика обучения с подкреплением».

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

## РАЗРАБОТЧИКИ:

|                       |                |                                  |
|-----------------------|----------------|----------------------------------|
| Доцент                |                | Салтыкова Ольга<br>Александровна |
| <i>Должность, БУП</i> | <i>Подпись</i> | <i>Фамилия И.О.</i>              |
| Доцент                |                | Иванюхин Алексей<br>Викторович   |
| <i>Должность, БУП</i> | <i>Подпись</i> | <i>Фамилия И.О.</i>              |

## РУКОВОДИТЕЛЬ БУП:

|                       |                |                             |
|-----------------------|----------------|-----------------------------|
| Заведующий кафедрой   |                | Разумный Юрий<br>Николаевич |
| <i>Должность, БУП</i> | <i>Подпись</i> | <i>Фамилия И.О.</i>         |

## РУКОВОДИТЕЛИ ОП ВО:

|                       |                |                             |
|-----------------------|----------------|-----------------------------|
| Заведующий кафедрой   |                | Разумный Юрий<br>Николаевич |
| <i>Должность, БУП</i> | <i>Подпись</i> | <i>Фамилия И.О.</i>         |
| Заведующий кафедрой   |                | Радин Юрий Анатольевич      |
| <i>Должность, БУП</i> | <i>Подпись</i> | <i>Фамилия И.О.</i>         |