

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Галунин Сергей Александрович  
Должность: проректор по учебной работе  
Дата подписания: 27.07.2022 11:31:08  
Уникальный программный ключ:  
08ef34338325bdb0ac5a47baa5472ce36cc3fc3b

Приложение к ОПОП  
«Безопасность и этика искус-  
ственного интеллекта»



**СПбГЭТУ «ЛЭТИ»**  
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

МИНОБРНАУКИ РОССИИ

федеральное государственное автономное образовательное учреждение высшего образования  
**«Санкт-Петербургский государственный электротехнический университет  
«ЛЭТИ» им. В.И.Ульянова (Ленина)»  
(СПбГЭТУ «ЛЭТИ»)»**

---

## **РАБОЧАЯ ПРОГРАММА**

**ДИСЦИПЛИНЫ**

**«АНАЛИЗ ДАННЫХ В ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ»**

для подготовки магистров

по направлению

09.04.01 «Информатика и вычислительная техника»

по программе

**«Безопасность и этика искусственного интеллекта»**

Санкт-Петербург

2021

## ЛИСТ СОГЛАСОВАНИЯ

Разработчики:

доцент, к.т.н. Каплун Д.И.

младший научный сотрудник Вознесенский А.С.

Рабочая программа рассмотрена и одобрена на заседании кафедры ВТ  
02.09.2021, протокол № 6

Рабочая программа рассмотрена и одобрена учебно-методической комиссией  
ФКТИ, 16.09.2021, протокол № 6

Согласовано в ИС ИОТ

Начальник ОМОЛА Загороднюк О.В.

## 1 СТРУКТУРА ДИСЦИПЛИНЫ

|   |      |
|---|------|
| Обеспечивающий факультет  | ФКТИ |
| Обеспечивающая кафедра  | АПУ  |
| Общая трудоемкость (ЗЕТ)  | 3    |
| Курс  | 2    |
| Семестр   | 3    |
| <b>Виды занятий</b>   |      |
| Лекции (академ. часов)  | 34   |
| Практические занятия (академ. часов)                                | 17   |
| Все контактные часы (академ. часов)                                 | 51   |
| Самостоятельная работа, включая часы на контроль<br>(академ. часов) | 57   |
| Всего (академ. часов)   | 108  |
| <b>Вид промежуточной аттестации</b>                                 |      |
| Дифф. зачет (курс)  | 2    |

## **2 АННОТАЦИЯ ДИСЦИПЛИНЫ**

### **«АНАЛИЗ ДАННЫХ В ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ»**

Освещаемые в курсе теоретические и прикладные вопросы полезны для студентов, специализирующихся в области управления и информационных технологий в технических системах, компьютерного и математического моделирования. В курсе даются основы Data Science, включающие вопросы сбора, хранения и подготовки данных (выявление аномалий в сырых данных, очистка данных от шума, дополнение данных, заполнение (удаление) пропущенных значений), классификации и кластеризации, поиска ассоциативных правил, интерполяции, экстраполяции и аппроксимации, корреляционного и регрессионного анализа, искусственного интеллекта и машинного обучения, нейронных сетей, сверточных нейронных сетей. Кроме того, в курсе рассматривается современная вычислительная база в контексте решения задач Data Science (CUDA, GPU, FPGA, SoC). Уделяется внимание решению практических задач с использованием современных программноаппаратных средств (MATLAB, Python, R, GPU).

### **SUBJECT SUMMARY**

#### **««DATA ANALYSIS IN ARTIFICIAL INTELLIGENCE»»**

The theoretical and applied questions covered in the course are useful for students specializing in control and information technology in technical systems, computer and mathematical modeling. The course provides the basics of Data Science, including data collection, storage and preparation (anomaly detection in raw data, data denoising, data addition, filling in (deleting) missing values), classification and clustering, association rule learning, interpolation, extrapolation and approximation, correlation and regression analysis, artificial intelligence and machine learning, neural networks, convolutional neural networks. In addition, the course examines the modern computing base in the context of solving Data Science problems (CUDA,

GPU, FPGA, SoC). Attention is paid to solving practical problems using modern software and hardware (MATLAB, Python, R, GPU).

## **3 ОБЩИЕ ПОЛОЖЕНИЯ**

### **3.1 Цели и задачи дисциплины**

1. Цель дисциплины -изучение математических моделей, методов и алгоритмов анализа данных и искусственного интеллекта и приобретения навыков их применения в профессиональной деятельности и навыков выбирать и применять методы сбора и извлечения знаний.

2. Формирование знаний и умений, необходимых для решения задач в области анализа данных и искусственного интеллекта с использованием современной вычислительной базы.

3. Приобретение теоретических знаний в области анализа данных и искусственного интеллекта.

4. Умение использования современной вычислительной базы для решения задач в области анализа данных и искусственного интеллекта.

Выбирать и разрабатывать программные компоненты систем искусственного интеллекта и ставить задачи по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области.

5. Получение практических навыков использования современной вычислительной базы для решения задач в области искусственного интеллекта.

### **3.2 Место дисциплины в структуре ОПОП**

Дисциплина изучается на основе ранее освоенных дисциплин учебного плана:

1. «Интеллектуальные системы»

2. «Математические основания информатики»

3. «Построение и оптимизация алгоритмов»

и обеспечивает подготовку выпускной квалификационной работы.

### 3.3 Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате освоения образовательной программы обучающийся должен достичь следующие результаты обучения по дисциплине:

| <b>Код компетенции/<br/>индикатора<br/>компетенции</b> | <b>Наименование компетенции/индикатора компетенции</b>  |
|--|---|
| ПК-25  | Способен выбирать, разрабатывать и проводить экспериментальную проверку работоспособности программных компонентов систем искусственного интеллекта по обеспечению требуемых критериев эффективности и качества функционирования |
| <i>ПК-25.1</i>   | <i>Выбирает и разрабатывает программные компоненты систем искусственного интеллекта</i>   |
| ПК-26  | Способен разрабатывать и применять методы и алгоритмы машинного обучения для решения задач  |
| <i>ПК-26.1</i>   | <i>Ставит задачи по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области</i>  |

## 4 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### 4.1 Содержание разделов дисциплины

#### 4.1.1 Наименование тем и часы на все виды нагрузки

| № п/п | Наименование темы дисциплины   | Лек, ач | Пр, ач | СР, ач |
|-------|--|---------|--------|--------|
| 1     | Введение   | 1       |        | 2      |
| 2     | Искусственный интеллект и машинное обучение  | 2       | 2      | 4      |
| 3     | Сбор, хранение и подготовка данных. Дополнение данных, заполнение (удаление) пропущенных значений  | 3       | 2      | 4      |
| 4     | Подготовка данных. Выявление аномалий в сырых данных, очистка данных от шума   | 3       | 2      | 4      |
| 5     | Классификация  | 3       | 2      | 4      |
| 6     | Кластеризация  | 3       | 1      | 4      |
| 7     | Поиск ассоциативных правил   | 3       | 1      | 4      |
| 8     | Аппроксимация, интерполяция и экстраполяция. Временные ряды: тренд, сезонность, шум. Прогнозирование   | 3       |        | 3      |
| 9     | Корреляционный анализ  | 3       |        | 3      |
| 10    | Регрессионный анализ   | 3       |        | 3      |
| 11    | Нейронные сети   | 2       | 2      | 6      |
| 12    | Сверточные нейронные сети  | 2       | 2      | 8      |
| 13    | Использование современной вычислительной базы в контексте решения задач Data Science (CUDA, GPU, FPGA, SoC). Современные программно-аппаратные средства (MATLAB, Python, R, GPU) | 3       | 3      | 6      |
| 14    | Заключение   |         |        | 2      |
|       | Итого, ач  | 34      | 17     | 57     |
|       | Из них ач на контроль  | 0       | 0      | 0      |
|       | Общая трудоемкость освоения, ач/зе   | 108/3   |        |        |

#### 4.1.2 Содержание

| № п/п | Наименование темы дисциплины                | Содержание   |
|-------|---|--|
| 1     | Введение                                    |  |
| 2     | Искусственный интеллект и машинное обучение | Искусственный интеллект. Машинное обучение. Общая постановка задачи обучения по прецедентам. Типология обучения задач по прецедентам. Бустинг. Бэггинг. Области применения ИИ. |



| №<br>п/п | Наименование темы<br>дисциплины   | Содержание   |
|----------|---|--|
| 3        | Сбор, хранение и подготовка данных. Дополнение данных, заполнение (удаление) пропущенных значений | <p>Как подготовить данные к моделированию: 5 операций Data Preparation. Почему нужно готовить данные к моделированию. Как Data Scientist готовит данные: методы и средства Data Preparation. Отберем то, что нужно Data Mining: как сформировать датасет для машинного обучения. Что такое датасет для Data Mining и из чего он состоит. Где искать датасеты. Каким бывает dataset: типы выборок. Как сформировать выборку для Data Mining. Зачем нужна очистка данных для Data Mining: 10 главных проблем подготовки датасета и способы их решения. Что такое очистка данных для Data Mining. От чего надо чистить большие данные. Как Data Scientist чистит данные: детали фазы Data Preparation Это не баг, а фича: генерация признаков для Data Mining. Признаки для Data Mining: определение и виды. Как проходит генерация признаков: 3 задачи этого этапа Data Mining и способы их решения. Data Preparation: полет нормальный – что такое нормализация данных и зачем она нужна.</p> |
| 4        | Подготовка данных. Выявление аномалий в сырых данных, очистка данных от шума                      | <p>Сырые данные. Выбросы. Причины выбросов. Определение выбросов. Критерии выбросов I. Гистограмма. Ящик с усами. Критерий Шовене. Критерий Граббса. Критерий Пирса. Критерий Диксона. Правило 3 сигм. Критерии выбросов II. Статистические тесты. Модельные тесты. Итерационные методы. Метрические методы. Методы подмены задачи. Методы машинного обучения. Ансамбли алгоритмов. ROCанализ.</p> <p>Алгоритмы очистки сигналов от шума. Преобразование Фурье. Вейвлетпреобразование. Фильтр Винера. Алгоритмы Empirical Mode Decomposition и Intrinsic TimeScale Decomposition.</p> <p>Алгоритмы очистки изображений от шума. Типовые шумы. Виды шумов: аналоговый, цифровой. Методы шумоподавления: пространственные, временные, пространственновременные. Качество шумоподавления. Метрики качества: PSNR и SSIM. Усредняющий фильтр. Медианный фильтр. Гауссовский фильтр. Билатеральный фильтр. Алгоритм NLM</p>   |
| 5        | Классификация   | <p>Задача классификации. Типы входных данных. Типы классов. Области применения. Формальная постановка задачи. Алгоритмы классификации. Метод ближайших соседей. Байесовский классификатор. Деревья решений.</p>  |
| 6        | Кластеризация   | <p>Задача кластеризации. Типы входных данных. Цели кластеризации. Функции расстояния. Формальная постановка задачи. Алгоритмы кластеризации. Метод kсредних. Метод средних. Генетический алгоритм.</p>   |

| №<br>п/п | Наименование темы<br>дисциплины  | Содержание  |
|----------|--|---|
| 7        | Поиск ассоциативных правил   | Задача поиска ассоциативных правил. Формальная постановка задачи. Концепции. Процесс построения ассоциативных правил. Алгоритмы поиска ассоциативных правил.  |
| 8        | Аппроксимация, интерполяция и экстраполяция. Временные ряды: тренд, сезонность, шум. Прогнозирование | <p>Аппроксимация. Постановка задачи. Метод наименьших квадратов. Линейная аппроксимация. Логарифмическая аппроксимация. Экспоненциальная аппроксимация. Степенная аппроксимация. Интерполяция.</p> <p>Постановка задачи. Локальная и глобальная интерполяция. Интерполяционный многочлен Лагранжа. Интерполяционный многочлен Ньютона. Экстраполяция.</p> <p>Временные ряды. Компоненты временных рядов. Объединение компонентов временного ряда. Оценка тренда. Виды функциональных зависимостей. Выделение тренда. Метод скользящего среднего. Метод экспоненциального сглаживания. Прогнозирование. Двойное экспоненциальное сглаживание. Метод ХольтВинтерса. Качество прогнозирования.</p> |
| 9        | Корреляционный анализ  | <p>Аппроксимация. Постановка задачи. Метод наименьших квадратов. Линейная аппроксимация. Логарифмическая аппроксимация. Экспоненциальная аппроксимация. Степенная аппроксимация. Интерполяция.</p> <p>Постановка задачи. Локальная и глобальная интерполяция. Интерполяционный многочлен Лагранжа. Интерполяционный многочлен Ньютона. Экстраполяция.</p> <p>Временные ряды. Компоненты временных рядов. Объединение компонентов временного ряда. Оценка тренда. Виды функциональных зависимостей. Выделение тренда. Метод скользящего среднего. Метод экспоненциального сглаживания. Прогнозирование. Двойное экспоненциальное сглаживание. Метод ХольтВинтерса. Качество прогнозирования.</p> |
| 10       | Регрессионный анализ   | Регрессия и регрессионный анализ. Линейная регрессия. Нелинейная регрессия.   |
| 11       | Нейронные сети   | Что такое нейронная сеть (НС)? Хронология. Классификация НС. Этапы решения задачи. Области применения НС. Структура нейрона. Синапс. Принципы работы НС. Функция активации. Обучение, валидация, тестирование. Итерация. Эпоха. Ошибка. Нейрон смещения. Обучение НС. Градиентный спуск. Метод обратного распространения (МОП). Гиперпараметры. Сходимость. Переобучение.   |

| № п/п | Наименование темы дисциплины   | Содержание  |
|-------|--|---|
| 12    | Сверточные нейронные сети  | Сверточные нейронные сети (СНС). Архитектура СНС. Слой свертки. Слой активации. Пулинг или субдискретизирующий слой. Полносвязная СНС. Обучение СНС. Преимущества СНС. Недостатки СНС.  |
| 13    | Использование современной вычислительной базы в контексте решения задач Data Science (CUDA, GPU, FPGA, SoC). Современные программно-аппаратные средства (MATLAB, Python, R, GPU) | Цифровые процессоры обработки сигналов (ЦПОС). Графические процессоры (GPU). Программируемые логические интегральные схемы (ПЛИС). Технология CUDA. Системы на кристалле (SoC). Языки математических и научных расчетов: MATLAB, Python, R. |
| 14    | Заключение   |   |

## 4.2 Перечень лабораторных работ

Лабораторные работы не предусмотрены.

## 4.3 Перечень практических занятий

| Наименование практических занятий  | Количество ауд. часов |
|--|-----------------------|
| 1. Искусственный интеллект и машинное обучение   | 2                     |
| 2. Сбор, хранение и подготовка данных. Дополнение данных, заполнение (удаление) пропущенных значений   | 2                     |
| 3. Подготовка данных. Выявление аномалий в сырых данных, очистка данных от шума  | 2                     |
| 4. Классификация   | 2                     |
| 5. Кластеризация   | 2                     |
| 6. Нейронные сети  | 2                     |
| 7. Сверточные нейронные сети   | 2                     |
| 8. Использование современной вычислительной базы в контексте решения задач Data Science (CUDA, GPU, FPGA, SoC). Современные программноаппаратные средства (MATLAB, Python, R, GPU) | 3                     |
| Итого  | 17                    |

## 4.4 Курсовое проектирование

Курсовая работа (проект) не предусмотрены.

#### **4.5 Реферат**

Реферат не предусмотрен.

#### **4.6 Индивидуальное домашнее задание**

Индивидуальное домашнее задание не предусмотрено.

#### **4.7 Доклад**

Доклад не предусмотрен.

#### **4.8 Кейс**

Кейс не предусмотрен.

#### **4.9 Организация и учебно-методическое обеспечение самостоятельной работы**

Изучение дисциплины сопровождается самостоятельной работой студентов с рекомендованными преподавателем литературными источниками и информационными ресурсами сети Интернет. Планирование времени для изучения дисциплины осуществляется на весь период обучения, предусматривая при этом регулярное повторение пройденного материала. Студентам необходимо регулярно дополнять сведениями из рекомендованных источников материал, законспектированный на лекциях в рамках самостоятельной работы. При этом на основе изучения рекомендованных ресурсов целесообразно составить конспект основных положений, терминов и определений, необходимых для освоения разделов учебной дисциплины. Особое место уделяется консультациям, как одной из форм обучения и контроля самостоятельной работы. Самостоятельное изучение студентами теоретических основ дисциплины обеспечено необходимыми учебно-методическими материалами (учебники, учебные пособия, кон-

спект лекций и т.п.), выполненными в печатном или электронном виде. По каждой теме содержания рабочей программы могут быть предусмотрены различные формы контроля (практические работы, тесты, индивидуальные домашние задания). Изучение студентами дисциплины сопровождается проведением регулярных консультаций преподавателей в рамках времени аудиторных занятий.

| <b>Текущая СРС</b>  | <b>Примерная<br/>трудоемкость, ач</b> |
|---|---------------------------------------|
| Работа с лекционным материалом, с учебной литературой   | 20                                    |
| Опережающая самостоятельная работа (изучение нового материала до его изложения на занятиях)                 | 10                                    |
| Самостоятельное изучение разделов дисциплины  | 0                                     |
| Выполнение домашних заданий, домашних контрольных работ   | 0                                     |
| Подготовка к лабораторным работам, к практическим и семинарским занятиям                                    | 10                                    |
| Подготовка к контрольным работам, коллоквиумам  | 0                                     |
| Выполнение расчетно-графических работ   | 0                                     |
| Выполнение курсового проекта или курсовой работы  | 0                                     |
| Поиск, изучение и презентация информации по заданной проблеме, анализ научных публикаций по заданной теме   | 7                                     |
| Работа над междисциплинарным проектом   | 0                                     |
| Анализ данных по заданной теме, выполнение расчетов, составление схем и моделей, на основе собранных данных | 0                                     |
| Подготовка к зачету, дифференцированному зачету, экзамену   | 10                                    |
| <b>ИТОГО СРС</b>  | <b>57</b>                             |

## 5 Учебно-методическое обеспечение дисциплины

### 5.1 Перечень основной и дополнительной литературы, необходимой для освоения дисциплины

| № п/п                     | Название, библиографическое описание   | К-во экз. в библи. |
|---------------------------|--|--------------------|
| Основная литература       |  |                    |
| 1                         | Вычисления и анализ данных [Текст] : метод. указания к практическим занятиям по дисциплине "Программирование в задачах приборостроения" / Санкт-Петербургский государственный электротехнический университет им. В.И. Ульянова (Ленина) "ЛЭТИ", 1998. -32 с. | 9                  |
| 2                         | Тюрин, Юрий Николаевич. Анализ данных на компьютере [Текст] : [учеб. пособие для вузов] / Ю.Н. Тюрин ; под ред. В.Э. Фигурнова, 1995. -384 с.  | 3                  |
| 3                         | Цехановский, Владислав Владимирович . Интеллектуальный анализ данных [Текст] : учеб. пособие / В. В. Цехановский, В. Д. Чертовской, 2019. - 55 с.  | 23                 |
| 4                         | Интеллектуальный анализ данных в распределенных системах [Текст] : [монография] / [М. С. Куприянов, И. И. Холод, З. А. Каршиев, И. А. Голубев], 2012. -108, [1] с.   | 9                  |
| 5                         | Монтгомери, Дуглас. Планирование эксперимента и анализ данных [Текст] / Д.К. Монтгомери ; сокр. пер. с англ. В.А. Коптяева, 1980. -382, [1] с.   | 14                 |
| Дополнительная литература |  |                    |
| 1                         | Элбон, Крис. Машинное обучение с использованием Python. Сборник рецептов [Текст] : [пер. с англ.] / К. Элбон, 2020. -369 с.  | 20                 |
| 2                         | Вандер Плас, Джейк. Python для сложных задач наука о данных: и машинное обучение [Текст] / Дж. Вандер Плас, 2020. -572, [1] с.   | 4                  |

### 5.2 Перечень ресурсов информационно-телекоммуникационной сети «Интернет», используемых при освоении дисциплины

| № п/п | Электронный адрес  |
|-------|--|
| 1     | Как научиться Data Science онлайн: 12 шагов от новичка до профи: <a href="https://proglib.io/p/kak-nauchitsya-data-science-onlayn-12-shagov-ot-novichka-do-profi-2020-06-30">https://proglib.io/p/kak-nauchitsya-data-science-onlayn-12-shagov-ot-novichka-do-profi-2020-06-30</a> |
| 2     | ШАД Яндекс: <a href="https://yandexdataschool.ru/">https://yandexdataschool.ru/</a>  |
| 3     | Как стать экспертом в Data Science: пошаговый план обучения <a href="https://tproger.ru/curriculum/data-science-expert-plan/">https://tproger.ru/curriculum/data-science-expert-plan/</a>  |
| 4     | Экспонента: <a href="https://exponenta.ru/mathworks">https://exponenta.ru/mathworks</a>  |

| № п/п | Электронный адрес  |
|-------|--|
| 5     | Чем занимается специалист по Data Science и как начать работать в этой области?: <a href="https://academy.yandex.ru/posts/chem-zanimaetsya-spetsialist-po-data-science-i-kak-nachat-rabotat-v-etoj-oblasti">https://academy.yandex.ru/posts/chem-zanimaetsya-spetsialist-po-data-science-i-kak-nachat-rabotat-v-etoj-oblasti</a> |
| 6     | MathWorks: <a href="https://www.mathworks.com/">https://www.mathworks.com/</a>   |
| 7     | MATLAB: инструмент будущего или дорогая игрушка И кому это нужно.: <a href="https://geekbrains.ru/posts/how_to_matlab">https://geekbrains.ru/posts/how_to_matlab</a>   |
| 8     | Машинное обучение: <a href="http://www.machinelearning.ru">http://www.machinelearning.ru</a>   |
| 9     | Машинное обучение для людей. Разбираемся простыми словами: <a href="https://vas3k.ru/blog/machine_learning/">https://vas3k.ru/blog/machine_learning/</a>   |
| 10    | Искусственный интеллект: <a href="https://ai.google/">https://ai.google/</a>   |
| 11    | Технологии анализа данных: <a href="https://basegroup.ru/">https://basegroup.ru/</a>   |
| 12    | Анализ данных и машинное обучение: <a href="https://www.kaggle.com/">https://www.kaggle.com/</a>   |
| 13    | AI, практический курс. Предобработка и дополнение данных с изображениями: <a href="https://habr.com/ru/company/intel/blog/414635/">https://habr.com/ru/company/intel/blog/414635/</a>  |
| 14    | Data Science с нуля: подробный гайд для начинающих: <a href="https://tproger.ru/translations/a-beginners-guide-to-data-science/">https://tproger.ru/translations/a-beginners-guide-to-data-science/</a>  |
| 15    | Как научиться Data Science онлайн: 12 шагов от новичка до профи: <a href="https://proglib.io/p/kak-nauchitsya-data-science-onlayn-12-shagov-ot-novichka-do-profi-2020-06-30">https://proglib.io/p/kak-nauchitsya-data-science-onlayn-12-shagov-ot-novichka-do-profi-2020-06-30</a>   |
| 16    | Какие навыки нужны специалисту по Data Science: <a href="https://vc.ru/skillfactory/81058-kakie-navyki-nuzhny-specialistu-po-data-science">https://vc.ru/skillfactory/81058-kakie-navyki-nuzhny-specialistu-po-data-science</a>  |
| 17    | Лучшие датасеты для машинного обучения и анализа данных: <a href="https://tproger.ru/translations/the-best-datasets-for-machine-learning-and-data-science/">https://tproger.ru/translations/the-best-datasets-for-machine-learning-and-data-science/</a>   |
| 18    | Как работает сверточная нейронная сеть: архитектура, примеры, особенности: <a href="https://neurohive.io/ru/osnovy-data-science/glubokaya-svertochnaja-nejronnaja-set/">https://neurohive.io/ru/osnovy-data-science/glubokaya-svertochnaja-nejronnaja-set/</a>   |
| 19    | Поиск датасетов: <a href="https://datasetsearch.research.google.com/">https://datasetsearch.research.google.com/</a>   |
| 20    | Что такое свёрточная нейронная сеть: <a href="https://habr.com/ru/post/309508/">https://habr.com/ru/post/309508/</a>   |
| 21    | Сверточные нейронные сети: <a href="https://neerc.ifmo.ru/wiki/index.php?title=Сверточные_нейрон">https://neerc.ifmo.ru/wiki/index.php?title=Сверточные_нейрон</a>   |
| 22    | Подготовка данных: <a href="https://www.bigdataschool.ru/blog/data-preparation-operations.html">https://www.bigdataschool.ru/blog/data-preparation-operations.html</a>   |
| 23    | Computer Science Center: <a href="https://compscicenter.ru/">https://compscicenter.ru/</a>   |

### 5.3 Адрес сайта курса

Адрес сайта курса: <https://vec.etu.ru/moodle/course/view.php?id=7826>

## 6 Критерии оценивания и оценочные материалы

### 6.1 Критерии оценивания

Для дисциплины «Анализ данных в искусственном интеллекте» формой промежуточной аттестации является зачет с оценкой.

#### Зачет с оценкой

| <b>Оценка</b>       | <b>Описание</b>   |
|---------------------|---|
| Неудовлетворительно | Курс не освоен. Студент испытывает серьезные трудности при ответе на ключевые вопросы дисциплины                  |
| Удовлетворительно   | Студент в целом овладел курсом, но некоторые разделы освоены на уровне определений и формулировок теорем          |
| Хорошо              | Студент овладел курсом, но в отдельных вопросах испытывает затруднения. Умеет решать задачи                       |
| Отлично             | Студент демонстрирует полное овладение курсом, способен применять полученные знания при решении конкретных задач. |



## Особенности допуска

Условием допуска является выполнение 8 онлайн тестов с количеством правильных ответов не менее 70% MathWorks и итогового тестирования -не менее 70% правильных ответов. Для получения оценки ”удовлетворительно” достаточно пройти тестирования, для повышения оценки до ”хорошо” и ”отлично” обучающиеся проходят собеседование по вопросам дифференцированного зачета п. 6.2.

## 6.2 Оценочные материалы для проведения текущего контроля и промежуточной аттестации обучающихся по дисциплине

### Примерные вопросы к дифф.зачету

| № п/п | Описание  |
|-------|---|
| 1     | Искусственный интеллект и машинное обучение. Искусственный интеллект. Машинное обучение. Общая постановка задачи обучения по прецедентам. Типология обучения задач по прецедентам. Бустинг. Бэггинг. Области применения ИИ  |
| 2     | Сбор, хранение и подготовка данных. Дополнение данных, заполнение (удаление) пропущенных значений. Как подготовить данные к моделированию: 5 операций Data Preparation. Почему нужно готовить данные к моделированию. Как Data Scientist готовит данные: методы и средства Data Preparation. Отберем то, что нужно Data Mining: как сформировать датасет для машинного обучения. Что такое датасет для Data Mining и из чего он состоит. Где искать датасеты. Каким бывает dataset: типы выборок. Как сформировать выборку для Data Mining. Зачем нужна очистка данных для Data Mining: 10 главных проблем подготовки датасета и способы их решения. Что такое очистка данных для Data Mining. От чего надо чистить большие данные. Как Data Scientist чистит данные: детали фазы Data Preparation Это не баг, а фича: генерация признаков для Data Mining. Признаки для Data Mining: определение и виды. Как проходит генерация признаков: 3 задачи этого этапа Data Mining и способы их решения. Data Preparation: полет нормальный – что такое нормализация данных и зачем она нужна. Что такое нормализация данных и чем она отличается от нормировки и нормирования. Зачем нормализовать датасет для Data Mining и Machine Learning. Выделение признаков: зачем отбирать предикторы и как это правильно сделать – готовим датасет к Data Mining и Machine Learning. Что такое отбор признаков и зачем он нужен. Как отбирать признаки: методы Feature Selection. |

|    |   |
|----|---|
| 3  | Подготовка данных. Выявление аномалий в сырых данных, очистка данных от шума. Сырые данные. Выбросы. Причины выбросов. Определение выбросов. Критерии выбросов I. Гистограмма. Ящик с усами. Критерий Шовене. Критерий Граббса. Критерий Пирса. Критерий Диксона. Правило 3 сигм. Критерии выбросов II. Статистические тесты. Модельные тесты. Итерационные методы. Метрические методы. Методы подмены задачи. Методы машинного обучения. Ансамбли алгоритмов. ROC-анализ. Алгоритмы очистки сигналов от шума. Преобразование Фурье. Вейвлетпреобразование. Фильтр Винера. Алгоритмы Empirical Mode Decomposition и Intrinsic TimeScale Decomposition. Алгоритмы очистки изображений от шума. Типовые шумы. Виды шумов: аналоговый, цифровой. Методы шумоподавления: пространственные, временные, пространственновременные. Качество шумоподавления. Метрики качества: PSNR и SSIM. Усредняющий фильтр. Медианный фильтр. Гауссовский фильтр. Билатеральный фильтр. Алгоритм NLM. |
| 4  | Классификация. Задача классификации. Типы входных данных. Типы классов. Области применения. Формальная постановка задачи. Алгоритмы классификации. Метод ближайших соседей. Байесовский классификатор. Деревья решений.   |
| 5  | Кластеризация. Задача кластеризации. Типы входных данных. Цели кластеризации. Функции расстояния. Формальная постановка задачи. Алгоритмы кластеризации. Метод k-средних. Метод средних. Генетический алгоритм  |
| 6  | Поиск ассоциативных правил. Задача поиска ассоциативных правил. Формальная постановка задачи. Концепции. Процесс построения ассоциативных правил. Алгоритмы поиска ассоциативных правил   |
| 7  | Аппроксимация, интерполяция и экстраполяция. Временные ряды: тренд, сезонность, шум. Прогнозирование. Аппроксимация. Постановка задачи. Метод наименьших квадратов. Линейная аппроксимация. Логарифмическая аппроксимация. Экспоненциальная аппроксимация. Степенная аппроксимация. Интерполяция. Постановка задачи. Локальная и глобальная интерполяция. Интерполяционный многочлен Лагранжа. Интерполяционный многочлен Ньютона. Экстраполяция. Временные ряды. Компоненты временных рядов. Объединение компонентов временного ряда. Оценка тренда. Виды функциональных зависимостей. Выделение тренда. Метод скользящего среднего. Метод экспоненциального сглаживания. Прогнозирование. Двойное экспоненциальное сглаживание. Метод Холт-Винтерса. Качество прогнозирования.  |
| 8  | Корреляционный анализ. Корреляция и корреляционный анализ. Коэффициенты корреляции. Коэффициент корреляции Пирсона. Коэффициент корреляции Спирмена. Коэффициент корреляции Кендалла. Коэффициент корреляции Фехнера. R-анализ. Показатель Херста   |
| 9  | Регрессионный анализ. Регрессия и регрессионный анализ. Линейная регрессия. Нелинейная регрессия.   |
| 10 | Нейронные сети. Что такое нейронная сеть (НС)? Хронология. Классификация НС. Этапы решения задачи. Области применения НС. Структура нейрона. Синапс. Принципы работы НС. Функция активации. Обучение, валидация, тестирование. Итерация. Эпоха. Ошибка. Нейрон смещения. Обучение НС. Градиентный спуск. Метод обратного распространения (МОП). Гиперпараметры. Сходимость. Переобучение  |
| 11 | Сверточные нейронные сети. Сверточные нейронные сети (СНС). Архитектура СНС. Слой свертки. Слой активации. Пулинг или субдискретизирующий слой. Полносвязная НС. Обучение СНС. Преимущества СНС. Недостатки СНС.  |

|    |  |
|----|--|
| 12 | Использование современной вычислительной базы в контексте решения задач Data Science (CUDA, GPU, FPGA, SoC). Современные программноаппаратные средства (MATLAB, Python, R, GPU). Цифровые процессоры обработки сигналов (ЦПОС). Графические процессоры (GPU). Программируемые логические интегральные схемы (ПЛИС). Технология CUDA. Системы на кристалле (SoC). Языки математических и научных расчетов: MATLAB, Python, R. |
|----|--|

## Образцы задач (заданий) для контрольных (проверочных) работ

Контрольная работа проводится в виде тестирования.

Сигнал, определенный в фиксированные (заданные) моменты времени и являющийся непрерывным по состоянию, называется:

1. **Дискретным сигналом**
2. Аналоговым сигналом
3. Цифровым сигналом

Реакция системы на цифровой единичный импульс при нулевых начальных условиях называется:

1. Переходной характеристикой
2. **Импульсной характеристикой**
3. Амплитудно-частотной характеристикой

Растекание спектра – это:

1. **Появление дополнительных составляющих в спектральном составе последовательности при вычислении ДПФ**
2. Повышение частоты следования спектральных линий (уменьшение периода дискретизации по частоте) за счет дополнения последовательности нулями
3. Расширение спектра за счет умножения сигнала на специальную расширяющую последовательность

Фильтр, пропускающий спектральные составляющие сигнала в заданной

полосе частот и подавляющий спектральные составляющие сигнала за пределами данной полосы, называется:

1. Фильтр нижних частот
2. **Полосно-пропускающий (полосовой) фильтр**
3. Полосно-задерживающий (режекторный) фильтр

Укажите, какая из нижеприведенных задач решается с использованием обратной идентификации и адаптивных фильтров:

1. Оценивание импульсной характеристики неизвестной системы
2. Очистка сигнала от шума
3. **Компенсация искажений, вносимых неизвестной системой**

Сигнал, определенный в фиксированные (заданные) моменты времени и являющийся квантованным по состоянию, называется:

1. Дискретным сигналом
2. Аналоговым сигналом
3. **Цифровым сигналом**

Реакция системы на цифровой единичный скачок при нулевых начальных условиях называется:

1. **Переходной характеристикой**
2. Импульсной характеристикой
3. Амплитудно-частотной характеристикой

Фильтр, подавляющий спектральные составляющие сигнала в заданной полосе частот и пропускающий спектральные составляющие сигнала за пределами данной полосы, называется:

1. Фильтр верхних частот

2. Полосно-пропускающий (полосовой) фильтр
3. **Полосно-задерживающий (режекторный) фильтр**

Изменение частоты дискретизации сигнала в нецелое число раз называется:

1. Интерполяцией
2. Децимацией
3. **Передискретизацией**

Сигнал, непрерывный по времени и по состоянию называется:

1. Дискретным сигналом
2. **Аналоговым сигналом**
3. Цифровым сигналом

ASIC – это

1. Класс ПЛИС
2. Фирма-изготовитель ПЛИС
3. **Заказная микросхема на основе программируемой логики**

Открытым стандартом для написания программ для GPU в настоящее время является

1. **OpenCL**
2. C++
3. JAVA

Именно эти две фирмы являются мировыми лидерами в производстве ПЛИС

1. Intel FPGA (Altera), Atmel
2. **Intel FPGA (Altera), Xilinx**

3. Xilinx, Atmel

Именно это семейство ПЛИС фирмы Intel FPGA (Altera) относится к SOC

1. **CycloneV**

2. ArriaII
3. MaxV

К семействам ПЛИС фирмы Intel FPGA (Altera) не относится следующее

1. Cyclone
2. **Virtex**
3. Arria

Такую разрядность имеют DSP блоки в FPGA фирмы Intel FPGA (Altera)

1. 9x9
2. **18x18**
3. 20x20

Помимо фирмы Nvidia технология распределённых вычислений на графических процессорах есть у фирмы

1. Intel
2. **AMD**
3. Apple

Барьер для группы процессов или потоков означает, что

1. **Каждый поток (процесс) должен остановиться в этой точке и подождать достижения барьера всеми потоками (процессами) группы**
2. Выполнение программы дальше невозможно
3. Все потоки после барьера объединяются

IDE среда фирмы Intel FPGA (Altera) для проектирования на ПЛИС назы-

вается

1. **QuartusII**

2. Vivaldo

3. AVR-studio

Утилита в среде QuartusII для отладки в реальном времени называется

1. **SignalTap**

2. SignalProbe

3. ChipPlanner

Весь комплект контрольно-измерительных материалов для проверки сформированности компетенции (индикатора компетенции) размещен в закрытой части по адресу, указанному в п. 5.3

### 6.3 График текущего контроля успеваемости

| Неделя | Темы занятий  | Вид контроля |
|--------|---|--------------|
| 2      | Искусственный интеллект и машинное обучение   | Тест         |
| 3      | Подготовка данных. Выявление аномалий в сырых данных, очистка данных от шума  | Тест         |
| 4      | Подготовка данных. Выявление аномалий в сырых данных, очистка данных от шума  | Тест         |
| 5      | Классификация   | Тест         |
| 6      | Кластеризация   | Тест         |
| 12     | Сверточные нейронные сети   | Тест         |
| 13     | Использование современной вычислительной базы в контексте решения задач Data Science (CUDA, GPU, FPGA, SoC). Современные программно-аппаратные средства (MATLAB, Python, R, GPU)                      | Тест         |
| 11     | Нейронные сети<br>Поиск ассоциативных правил<br>Аппроксимация, интерполяция и экстраполяция. Временные ряды: тренд, сезонность, шум. Прогнозирование<br>Регрессионный анализ<br>Корреляционный анализ | Тест         |

### 6.4 Методика текущего контроля

Условием допуска является выполнение 8 онлайн тестов с количеством правильных ответов не менее 70% MathWorks и итогового тестирования - не менее 70% правильных ответов. Для получения оценки "удовлетворительно" достаточно пройти тестирования, для повышения оценки до "хорошо" и "отлично" обучающиеся проходят собеседование по вопросам дифференцированного зачета п. 6.2.



## 7 Описание информационных технологий и материально-технической базы

| Тип занятий            | Тип помещения                        | Требования к помещению   | Требования к программному обеспечению  |
|------------------------|--------------------------------------|--|--|
| Лекция                 | Лекционная аудитория                 | Количество посадочных мест по численности группы, рабочее место преподавателя (ПК), маркерная доска, проектор, экран;  | 1. Windows 10 и выше;<br>2. Microsoft Office 2010 и выше<br>3. MATLAB R2020 и выше |
| Практические занятия   | Аудитория                            | Количество посадочных мест по численности группы, рабочее место преподавателя (ПК), маркерная доска;   | 1. Windows 10 и выше;<br>2. Microsoft Office 2010 и выше<br>3. MATLAB R2020 и выше |
| Самостоятельная работа | Помещение для самостоятельной работы | Оснащено компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета. | 1. Windows 10 и выше;<br>2. Microsoft Office 2010 и выше<br>3. MATLAB R2020 и выше |

## **8 Адаптация рабочей программы для лиц с ОВЗ**

Адаптированная программа разрабатывается при наличии заявления со стороны обучающегося (родителей, законных представителей) и медицинских показаний (рекомендациями психолого-медико-педагогической комиссии). Для инвалидов адаптированная образовательная программа разрабатывается в соответствии с индивидуальной программой реабилитации.

## ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ

| <b>№ п/п</b> | <b>Дата</b> | <b>Изменение</b>   | <b>Дата и номер протокола заседания УМК</b> | <b>Автор</b> | <b>Начальник ОМОЛА</b> |
|--------------|-------------|--|---|--------------|------------------------|
| 1            | 23.12.2021  | Внесены изменения в компетентностную модель образовательной программы, на основании письма Минобрнауки России от 21.12.2021 № МН-5/22720 | 23.12.2021 №9                               |              |                        |