

Федеральное государственное образовательное бюджетное учреждение  
высшего образования  
**«Финансовый университет при Правительстве Российской Федерации»  
(Финансовый университет)**

Институт развития профессиональных  
компетенций и квалификаций

Обсуждено и одобрено  
на Ученом совете институтов и школ  
дополнительного профессионального  
образования

Протокол № 26  
«13» мая 2023 г.

УТВЕРЖДАЮ

Проректор по дополнительному  
профессиональному образованию



Е.А. Диденко

«23» мая 2023 г.

# ПРОГРАММА

Повышения квалификации  
(вид дополнительной профессиональной программы)  
«Python и SQL для анализа данных»

Москва – 2023

## **Программа повышения квалификации «Python и SQL для анализа данных»**

### **Общая характеристика программы**

**Цель программы повышения квалификации:** совершенствование и (или) получение новых профессиональных компетенций, необходимых для профессиональной деятельности, и (или) повышение профессионального уровня в области искусственного интеллекта и больших данных при реализации проектов и разработки новых решений на основе данных, способность применять языки программирования SQL и Python для сбора, визуализации, анализа больших данных и машинного обучения, применять конкретные аналитические и продуктовые подходы при работе над реальными задачами маркетинговой и клиентской аналитики с помощью BI платформ.

**Наименование образовательных и профессиональных стандартов, используемых при разработке программы повышения квалификации.**

1. Профессиональный стандарт «Бизнес-аналитик», утвержденный приказом Министерства труда и социальной защиты РФ от 25 сентября 2018 г. N 592н;
2. Профессиональный стандарт "Специалист по большим данным", утвержденный приказом Министерства труда и социальной защиты РФ от 6 июля 2020 г. N 405н.

**Описание перечня профессиональных компетенций в рамках имеющейся квалификации, качественное изменение которых осуществляется в процессе обучения**

Программа повышения квалификации основана на требованиях профессиональных стандартов:

- «Бизнес-аналитик», утвержденный приказом Министерства труда и социальной защиты РФ от 25 сентября 2018 г. N 592н в части обобщенных трудовых и трудовых функций:

- - ОТФ С - Выявление бизнес-проблем или бизнес-возможностей;
- - ТФ С/01.5 - Сбор информации о бизнес-проблемах или бизнес-возможностях;
- - ОТФ D - Обоснование решений;
- - ТФ D/02.6 - Анализ, обоснование и выбор решения.

- "Специалист по большим данным", утвержденный приказом Министерства труда и социальной защиты РФ от 6 июля 2020 г. N 405н:

- ОТФ А - Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры;
- ТФ А/03.6 - Подготовка данных для проведения аналитических работ по исследованию больших данных;
- ТФ А/04.6 Проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика;
- ОТФ В - Управление этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации;
- ТФ В/05.7 Управление получением, хранением, передачей, обработкой больших данных.

**Профессиональные компетенции, совершенствуемые и приобретаемые слушателями в процессе освоения программы:**

ПК-1 – способность к сбору информации о бизнес-проблемах или бизнес-возможностях в области больших данных;

ПК-2 – способность проведения анализа, обоснования и выбора решения в области больших данных

ПК-3 - способность к подготовке данных для проведения аналитических работ по исследованию больших данных;

ПК-4 - способность проведения аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика;

ПК-5 – способность управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации.

### **Планируемые образовательные результаты по программе повышения квалификации:**

В результате изучения программы слушатели должны:

#### **а) знать:**

- основные модели и методы машинного обучения и разработки данных;
- основные фреймворки машинного обучения;
- возможности и особенности языков SQL и Python, платформ бизнес-аналитики используемые для решения задачи сбора и подготовки данных;
- предметную область и специфику деятельности организации в объеме, достаточном для решения задач бизнес-анализа;
- информационные технологии (программное обеспечение), применяемые в организации, в объеме, необходимом для целей бизнес-анализа;
- возможности и особенности технологий сбора, хранения и анализа больших данных;
- инструментальные средства разработки моделей машинного обучения;
- методы сбора, анализа, систематизации, хранения и поддержания в актуальном состоянии информации бизнес-анализа;
- методология по исследованию данных CRISM-DM и методику KDD (Knowledge Discovery in Databases), используемую при разработке моделей и решении стандартных задач Data Mining;
- правила и методы построения OLAP-кубов, аналитических отчетов, дашбордов, информационных панелей мониторинга и сторителлинга для получения знаний с целью поддержки принятия решений;
- правила и методы проектирования моделей данных, внедрения аналитических отчетов, дашбордов, информационных панелей мониторинга в компоненты ИТ-инфраструктуры организации для проведения мониторинга и контроллинга стратегических целей и поддержку внутренних процессов в организации;
- алгоритмические и программные методы проведения OLAP и интеллектуального анализа данных для принятия взвешенных решений;
- основные методы проведения маркетинговой и клиентской аналитики;

#### **б) уметь:**

- создавать, обучать и сравнивать модели машинного обучения, основанные на различных алгоритмах, осуществлять их оценку, используя фреймворки и платформы машинного обучения;
- применять язык SQL и библиотеки Python для сбора и подготовки данных и визуализации данных;
- решать стандартные задачи профессиональной деятельности используя OLAP-технологии аналитики;
- использовать современные информационные технологии аналитики для информационно-аналитического сопровождения деятельности организации;

- создавать программные приложения и аналитический контент, организовывать совместную аналитическую работу используя облачные аналитические платформы;
- разрабатывать рабочие процессы создания, тестирования и развертывания моделей машинного обучения в десктопных платформах и облачных средах;

**в) владеть:**

- навыками применения фреймворков машинного и платформ науки о данных для создания, обучения и оценки моделей машинного обучения;
- навыками работы с сырыми данными, использования языков SQL и Python, платформ интеграции и очистки данных, платформы бизнес-аналитики и науки о данных;
- навыками применения технологии аналитики больших данных для извлечения и добычи знаний из данных;
- инструментами и технологиями выявления, сбора и анализа информации бизнес-анализа для формирования возможных решений;
- инструментами формирования аналитических отчетов, дашбордов и «повествований» для поддержки принятия решений.



Федеральное государственное образовательное бюджетное учреждение  
высшего образования  
«Финансовый университет при Правительстве Российской Федерации»  
(Финансовый университет)

Институт развития профессиональных  
компетенций и квалификаций

Обсуждено и одобрено  
на Ученом совете институтов и школ  
дополнительного профессионального  
образования

Протокол № 26  
«23» мая 2023 г.



УТВЕРЖДАЮ

Проректор по дополнительному  
профессиональному образованию

Е.А. Диденко

«23» мая 2023 г.

**УЧЕБНЫЙ ПЛАН**  
программы повышения квалификации  
«Python и SQL для анализа данных»

|                                            |                                                                                                              |
|--------------------------------------------|--------------------------------------------------------------------------------------------------------------|
| Требования к уровню образования слушателей | лица, имеющие среднее профессиональное или высшее образование;<br>лица, получающие высшее образование        |
| Категория слушателей                       | лица, желающие освоить новый вид профессиональной деятельности в области машинного обучения и анализа данных |
| Срок обучения                              | 144 часа                                                                                                     |
| Форма обучения                             | Очно-заочная, с применением дистанционных образовательных технологий и электронного обучения                 |
| Режим занятий                              | не более 8 часов в день                                                                                      |

| № раз-дела | Наименование дисциплины                               | Трудоемкость | В том числе                     |        |    |                        | Форма контроля |
|------------|-------------------------------------------------------|--------------|---------------------------------|--------|----|------------------------|----------------|
|            |                                                       |              | Аудиторные занятия <sup>1</sup> |        |    | Самостоятельная работа |                |
|            |                                                       | В часах      | Всего часов                     | из них |    |                        |                |
| Лекции     | Практические занятия                                  |              |                                 |        |    |                        |                |
| 1          | Модуль 1. Аналитика: Инструменты, методы, показатели. | 24           | 17                              | 4      | 13 | 7                      | Зачет          |
| 2.         | Модуль 2. Python и анализ данных                      | 23           | 13                              | 4      | 9  | 10                     | Зачет          |

<sup>1</sup> С возможным применением дистанционных образовательных технологий и электронного обучения.

|    |                                                                 |            |           |           |           |           |                        |
|----|-----------------------------------------------------------------|------------|-----------|-----------|-----------|-----------|------------------------|
| 3. | Модуль 3. SQL и получение данных                                | 25         | 15        | 4         | 11        | 10        | Зачет                  |
| 4  | Модуль 4. Визуализация данных                                   | 21         | 11        | 4         | 7         | 10        | Зачет                  |
| 5. | Модуль 5. Основные модели машинного обучения                    | 25         | 17        | 6         | 11        | 8         | Зачет                  |
| 6  | Модуль 6. Глубокое машинное обучение и обработка больших данных | 24         | 17        | 4         | 13        | 7         | Зачет                  |
|    | Итоговая аттестация                                             | 2          | 2         |           | 2         |           | Защита итоговой работы |
|    | <b>ИТОГО</b>                                                    | <b>144</b> | <b>84</b> | <b>26</b> | <b>58</b> | <b>60</b> |                        |

Разработчики программы:

Бутурлин Илья Владимирович - к.ю.н., Product owner Garage Eight, брокерской платформы для 200 миллионов трейдеров ЮВА. Сооснователь криптобиржи ae.exchange, до 2021 года доцент Финансового университета при Правительстве РФ, научный руководитель бизнес-инкубатора Финансового университета. Аттестат ФСФР 5.0 специалиста по управлению инвестиционными фондами.

Анпилогов Вадим Васильевич - Департамент Рисков ПАО "СберБанк" исполнительный директор, Senior Data Analyst, Risk-Management (декабрь 2020 – октябрь 2021), Data Scientist, Risk-Management (ноябрь 2019 – декабрь 2020). Образование, Магистр, Финансы, инвестиции и банки, РЭШ (2018 – 2020). Бакалавр, Финансовый рынки и банки, Финансовый университет при Правительстве РФ (2014 - 2018)

Сафаралеев Алексей Васильевич - ML-engineer Garage-Eight, разработка ML-решений в рамках направления Engagement, разработка рекомендательной системы для проекта "Умная лента", оценка эффективности ML-решений, дизайн АБ-тестов, внедрение MLOps

Добровольский Игорь Николаевич – к.э.н., основатель и исполнительный директор глобальной технологической компании REACTOR-FZCO, до 2021 года доцент Финансового университета при Правительстве РФ, директор бизнес-инкубатора Финансового университета.

В реализации программы принимают участие профессорско-преподавательский состав Финансового университета, квалифицированные эксперты-практики.

Директор ИРПКК



Т.А. Болтенко

« \_\_\_\_ » \_\_\_\_\_ 2023 г.

 06.06.23.

**Федеральное государственное образовательное бюджетное  
учреждение высшего образования  
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ  
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»  
Институт развития  
профессиональных компетенций и квалификаций**

**УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН  
программы повышения квалификации  
«Python и SQL для анализа данных»**

| №<br>п/п | Наименование дисциплины,<br>модуля                                                                      | Трудо-<br>емкос-<br>ть | В том числе |       |                   |                     | Форма контроля             |                           |
|----------|---------------------------------------------------------------------------------------------------------|------------------------|-------------|-------|-------------------|---------------------|----------------------------|---------------------------|
|          |                                                                                                         |                        | В часах     | Всего | Контактная работа |                     |                            | Самостоятельная<br>работа |
|          |                                                                                                         |                        |             |       | Лекции            | Практич.<br>занятия |                            |                           |
| 1        | 2                                                                                                       | 3                      | 4           | 5     | 6                 | 7                   | 8                          |                           |
| М 1.     | Инструменты и методы бизнес-аналитики                                                                   | 24                     | 17          | 4     | 13                | 7                   | Зачет                      |                           |
| 1.1      | Примеры аналитических задач. Сбор данных для анализа данных                                             | 3                      | 2           | 2     |                   | 1                   | Тестирование               |                           |
| 1.2      | Инструмент работы с таблицами в режиме онлайн на примере Google-таблицы.                                | 3                      | 2           |       | 2                 | 1                   | Решение ситуационных задач |                           |
| 1.3      | Статистика и тестирование гипотез                                                                       | 3                      | 2           |       | 2                 | 1                   | Тестирование               |                           |
| 1.4      | Бизнес-показатели и точки роста. Анализ продуктовых метрик                                              | 5                      | 4           | 2     | 2                 | 1                   | Решение практических задач |                           |
| 1.5      | Финансовые метрики и Маркетинговые метрики.                                                             | 3                      | 2           |       | 2                 | 1                   | Решение практических задач |                           |
| 1.6      | Иерархия метрик и разработка отчетности. Дизайн тестов, проведение и анализ. Построение простых моделей | 3                      | 2           |       | 2                 | 1                   | Решение практических задач |                           |
| 1.7.     | Аналитика данных в Excel                                                                                | 3                      | 2           |       | 2                 | 1                   | Решение практических задач |                           |
|          | Промежуточная аттестация по модулю 1                                                                    | 1                      | 1           |       | 1                 |                     | Зачет                      |                           |
| М 2.     | Анализ данных с применением Python                                                                      | 23                     | 13          | 4     | 9                 | 10                  | Зачет                      |                           |
| 2.1.     | Python: Jupiter Notebook, комментарии в коде.                                                           | 4                      | 2           | 2     |                   | 2                   | Тестирование               |                           |

|             |                                                                                                           |           |           |          |           |           |                            |
|-------------|-----------------------------------------------------------------------------------------------------------|-----------|-----------|----------|-----------|-----------|----------------------------|
|             | арифметические операции. Управляющие конструкции и коллекции                                              |           |           |          |           |           |                            |
| 2.2.        | Функции и работа с данными, основы парсинга. Работа с файловой системой и модули                          | 4         | 2         | 2        |           | 2         | Тестирование               |
| 2.3.        | Регулярные выражения и основы синтаксического разбора. Исключения и обработка ошибок                      | 4         | 2         |          | 2         | 2         | Решение практических задач |
| 2.4.        | Функции и понятие класса. Библиотека NumPy: вычислительные задачи                                         | 3         | 2         |          | 2         | 1         | Решение практических задач |
| 2.5.        | Библиотека Pandas: возможности и оптимизация. Методы оптимизации Pandas                                   | 4         | 2         |          | 2         | 2         | Решение практических задач |
| 2.6.        | Статистическая проверка гипотез и A/B-тесты                                                               | 3         | 2         |          | 2         | 1         | Решение практических задач |
|             | <b>Промежуточная аттестация по модулю 2</b>                                                               | 1         | 1         |          | 1         |           | <b>Зачет</b>               |
| <b>М 3.</b> | <b>Применение SQL для работы с данными</b>                                                                | <b>25</b> | <b>15</b> | <b>4</b> | <b>11</b> | <b>10</b> | <b>Зачет</b>               |
| 3.1         | Введение в SQL. Установка и знакомство. Основы и работа с базами данных                                   | 4         | 2         | 2        |           | 2         | Тестирование               |
| 3.2         | Основы SQL. Базовые запросы SQL. Фильтрация данных                                                        | 4         | 2         | 2        |           | 2         | Тестирование               |
| 3.3         | Углубление в SQL. Агрегация данных. Группировка данных. Join. Оконные функции                             | 3         | 2         |          | 2         | 1         | Решение практических задач |
| 3.4         | Работа с PostgreSQL                                                                                       | 4         | 2         |          | 2         | 2         | Решение практических задач |
| 3.5         | "NoSql & MongoDB"                                                                                         | 4         | 2         |          | 2         | 2         | Тестирование               |
| 3.6         | Примеры создания дашбордов и анализ продуктовых метрик с применением SQL                                  | 5         | 4         |          | 4         | 1         | Решение практических задач |
|             | <b>Промежуточная аттестация по модулю 3</b>                                                               | 1         | 1         |          | 1         |           | <b>Зачет</b>               |
| <b>М 4.</b> | <b>Визуализация данных</b>                                                                                | <b>21</b> | <b>11</b> | <b>4</b> | <b>7</b>  | <b>10</b> | <b>Зачет</b>               |
| 4.1         | BI-инструменты и визуализация данных                                                                      | 3         | 2         | 2        |           | 1         | Тестирование               |
| 4.2         | DataLens: быстрый анализ данных с визуализацией                                                           | 3         | 2         | 2        |           | 1         | Решение практических задач |
| 4.3.        | Power BI. Построение моделей данных из разных неструктурированных источников: таблиц, сайтов и баз данных | 4         | 2         |          | 2         | 2         | Решение практических задач |
| 4.4.        | Tableau. Основы работы с расчётными полями, фильтрами, множествами и группировками.                       | 6         | 2         |          | 2         | 4         | Решение практических задач |



|             |                                                                                                                                          |           |           |          |           |           |                            |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------|-----------|-----------|----------|-----------|-----------|----------------------------|
|             | Основные виды визуализаций.<br>Лучшие практики визуализации                                                                              |           |           |          |           |           |                            |
| 4.5.        | Создание интерактивных дашбордов                                                                                                         | 4         | 2         |          | 2         | 2         | Решение практических задач |
|             | <i>Промежуточная аттестация по модулю 4</i>                                                                                              | 1         | 1         |          | 1         |           | <b>Зачет</b>               |
| <b>М 5.</b> | <b>Основные методы машинного обучения</b>                                                                                                | <b>25</b> | <b>17</b> | <b>6</b> | <b>11</b> | <b>8</b>  | <b>Зачет</b>               |
| 5.1.        | Понятие машинного обучения. Регрессионный анализ. Линейная, полиномиальная и логарифмическая регрессия                                   | 5         | 4         | 2        | 2         | 1         | Решение практических задач |
| 5.2.        | Классификация. Логистическая регрессия и SVM                                                                                             | 3         | 2         |          | 2         | 1         | Решение практических задач |
| 5.3.        | Функции потерь и оптимизация. Оценка точности модели, переобучение, регуляризация                                                        | 3         | 2         |          | 2         | 1         | Решение практических задач |
| 5.4.        | Проблема качества данных. Работа с пропусками. Работа с переменными                                                                      | 5         | 4         | 2        | 2         | 1         | Решение практических задач |
| 5.5.        | Деревья решений. Ансамблирование. Поиск выбросов и генерация новых признаков                                                             | 4         | 2         | 2        |           | 2         | Тестирование               |
| 5.6.        | Алгоритмы кластеризации. Улучшение качества модели.                                                                                      | 4         | 2         |          | 2         | 2         | Решение практических задач |
|             | <i>Промежуточная аттестация по модулю 5</i>                                                                                              | 1         | 1         |          | 1         |           | <b>Зачет</b>               |
| <b>М 6.</b> | <b>Глубокое машинное обучение и обработка больших данных</b>                                                                             | <b>24</b> | <b>17</b> | <b>4</b> | <b>13</b> | <b>10</b> | <b>Зачет</b>               |
| 6.1.        | Машинное обучение и Data Science. Основные инструменты обработки больших данных. HADOOP & SPARK.                                         | 5         | 4         | 2        | 2         | 1         | Решение практических задач |
| 6.2.        | Введение и классификация рекомендательных систем. Коллаборативная фильтрация                                                             | 5         | 4         | 2        | 2         | 1         | Решение практических задач |
| 6.3.        | Знакомство с временными рядами. Элементарные методы анализа временных рядов. Модели ARIMA, GARCH                                         | 3         | 2         |          | 2         | 1         | Решение практических задач |
| 6.4.        | Обработка естественного языка (NLP). Введение в автоматическую обработку текста. Структура слова. Морфология. Тематическое моделирование | 3         | 2         |          | 2         | 1         | Решение практических задач |

|      |                                                                                                                 |            |           |           |           |           |                                   |
|------|-----------------------------------------------------------------------------------------------------------------|------------|-----------|-----------|-----------|-----------|-----------------------------------|
| 6.5. | Компьютерное зрение<br>Выделение признаков и поиск<br>похожих изображений<br>Сегментация и детекция<br>объектов | 4          | 2         |           | 2         | 2         | Решение<br>практических<br>задач  |
| 6.6. | Подготовка модели Машинное<br>обучения по практическому<br>кейсу                                                | 3          | 2         |           | 2         | 1         | Решение<br>практических<br>задач  |
|      | <i>Промежуточная аттестация<br/>по модулю 6</i>                                                                 | 1          | 1         |           | 1         |           | <b>Зачет</b>                      |
|      | <b>Итоговая аттестация</b>                                                                                      | 2          | 2         |           | 2         |           | <b>Защита итоговой<br/>работы</b> |
|      | <b>ИТОГО</b>                                                                                                    | <b>144</b> | <b>84</b> | <b>26</b> | <b>58</b> | <b>60</b> |                                   |

Федеральное государственное образовательное бюджетное учреждение  
высшего образования  
"Финансовый университет при Правительстве Российской Федерации"  
(Финансовый университет)  
Институт развития  
профессиональных компетенций и квалификаций

**Календарный учебный график**  
программы повышения квалификации  
«Python и SQL для анализа данных»

Объем программы – 144 час.

Продолжительность обучения – 16 недель (4 месяца)

Форма обучения – очно-заочная, с применением дистанционных образовательных технологий и электронного обучения

Образовательный процесс по программе может осуществляться в течение всего учебного года. Занятия проводятся по мере комплектования учебных групп. Занятия могут проводиться с каждым слушателем индивидуально.

| №  | Наименование дисциплин (модулей), тем                           | 1 месяц   | 2 месяц   | 3 месяц   | 4 месяц   | ПА | КР | СР | ИА | Всего      |
|----|-----------------------------------------------------------------|-----------|-----------|-----------|-----------|----|----|----|----|------------|
| 1  | Модуль 1. Инструменты и методы бизнес-аналитики                 | 24        |           |           |           | 1  | 17 | 7  |    | 24         |
| 2. | Модуль 2. Анализ данных с применением Python                    | 23        |           |           |           | 1  | 13 | 10 |    | 23         |
| 3. | Модуль 3. Применение SQL для работы с данными                   |           | 25        |           |           | 1  | 15 | 10 |    | 25         |
| 4  | Модуль 4. Визуализация данных                                   |           | 21        |           |           | 1  | 11 | 10 |    | 21         |
| 5. | Модуль 5. Основные методы машинного обучения                    |           |           | 25        |           | 1  | 17 | 8  |    | 25         |
| 6. | Модуль 6. Глубокое машинное обучение и обработка больших данных |           |           | 12        | 12        | 1  | 17 | 7  |    | 24         |
|    | Итоговая аттестация                                             |           |           |           |           |    |    |    | 2  | 2          |
|    | <b>ИТОГО</b>                                                    | <b>47</b> | <b>46</b> | <b>37</b> | <b>12</b> |    |    |    |    | <b>144</b> |

| Условные обозначения |                          |
|----------------------|--------------------------|
| ПА                   | Промежуточная аттестация |
| ИА                   | Итоговая аттестация      |
| КР                   | Контактная работа        |
| СР                   | Самостоятельная работа   |

Директор ИРПКК



Т.А. Болтенко

## Рабочие программы модулей учебного курса

### «Python и SQL для анализа данных»

#### Модуль 1. Инструменты и методы бизнес-аналитики.

**Цель модуля:** приобретение слушателями компетенций, необходимых для понимания и эффективной работы в области анализа больших данных, инструментов и технологий, позволяющих анализировать результаты внутренних процессов организации с помощью Google таблиц и Data Studio.

##### Планируемые образовательные результаты по модулю

В результате освоения модуля слушатели должны:

*а) знать:*

- правила создания Google-таблицы;
- правила и методы построения визуализации данных в Google Data Studio;
- правила создания аналитических отчетов и визуализация данных;
- информационные технологии (программное обеспечение), применяемые в организации, в объеме, необходимом для целей бизнес-анализа;
- теоретические основы анализа данных;
- правила и методы проектирования моделей данных, внедрения аналитических отчетов, дашбордов, информационных панелей мониторинга в компоненты ИТ-инфраструктуры организации для проведения мониторинга и контроллинга стратегических целей и поддержку внутренних процессов.

*б) уметь:*

- моделировать объем и границы работ;
- определять связи и зависимости между элементами информации бизнес-анализа;
- использовать современные информационные технологии аналитики для информационно-аналитического сопровождения деятельности;
- работать в среде Google Colaboratory;
- загружать данные в pandas и работать с ними (фильтрация, агрегация, заполнение пропущенных значений);
- моделировать требования заинтересованных сторон в соответствии с выбранными подходами;
- анализировать качество информации бизнес-анализа с точки зрения выбранных критериев;
- проводить анализ предметной области.

*в) владеть:*

- анализом потребностей заинтересованных сторон;
- инструментами и технологиями выявления, сбора и анализа информации бизнес-анализа для формирования возможных решений;
- описанием возможных решений;
- навыками работы в среде Google Colaboratory
- инструментами и технологиями выявления, сбора и анализа информации бизнес-анализа для формирования возможных решений;
- инструментами формирования аналитических отчетов, дашбордов и «повествований» для поддержки принятия решений.

##### Формируемые компетенции:

способность к сбору информации о бизнес-проблемах или бизнес-возможностях в области больших данных;

способность проведения аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика.



## УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

| №№<br>п/п | Наименование разделов,<br>модулей                                                                                | Всего<br><br>В<br>час<br>ах | В том числе           |        | Самостоятельная<br>работа | Форма<br>контроля |                                  |
|-----------|------------------------------------------------------------------------------------------------------------------|-----------------------------|-----------------------|--------|---------------------------|-------------------|----------------------------------|
|           |                                                                                                                  |                             | Аудиторные<br>занятия |        |                           |                   |                                  |
|           |                                                                                                                  |                             | Всего, часов          | из них |                           |                   |                                  |
| Лекции    | Практич<br>еская<br>раб<br>ота                                                                                   |                             |                       |        |                           |                   |                                  |
| 1         | 2                                                                                                                | 3                           | 4                     | 5      | 6                         | 7                 | 8                                |
| М 1.      | Инструменты и методы<br>бизнес-аналитики                                                                         | 24                          | 17                    | 4      | 13                        | 7                 | Зачет                            |
| 1.1       | Примеры аналитических задач.<br>Сбор данных для анализа<br>данных                                                | 3                           | 2                     | 2      |                           | 1                 | Тестирование                     |
| 1.2       | Инструмент работы с<br>таблицами в режиме онлайн на<br>примере Google-таблицы.                                   | 3                           | 2                     |        | 2                         | 1                 | Решение<br>ситуационных<br>задач |
| 1.3       | Статистика и тестирование<br>гипотез                                                                             | 3                           | 2                     |        | 2                         | 1                 | Тестирование                     |
| 1.4       | Бизнес-показатели и точки<br>роста. Анализ продуктовых<br>метрик                                                 | 5                           | 4                     | 2      | 2                         | 1                 | Решение<br>практических<br>задач |
| 1.5       | Финансовые метрики и<br>Маркетинговые метрики.                                                                   | 3                           | 2                     |        | 2                         | 1                 | Решение<br>практических<br>задач |
| 1.6       | Иерархия метрик и разработка<br>отчетности. Дизайн тестов,<br>проведение и анализ.<br>Построение простых моделей | 3                           | 2                     |        | 2                         | 1                 | Решение<br>практических<br>задач |
| 1.7.      | Аналитика данных в Excel                                                                                         | 3                           | 2                     |        | 2                         | 1                 | Решение<br>практических<br>задач |
|           | Промежуточная аттестация<br>по модулю 1                                                                          | 1                           | 1                     |        | 1                         |                   | Зачет                            |

### Тема 1.1. Примеры аналитических задач. Сбор данных для анализа данных

Понятие Аналитики. Что это такое и зачем нужно. Как работает аналитика. Данные в аналитике. Из чего состоит аналитика. Использование Аналитики и ее преимущества. Аналитические инструменты. Пример решения аналитической проблемы. Базовые инструменты аналитика, простой анализ данных.

### Содержание практических занятий

| №<br>мод<br>уля | Наименование<br>темы (раздела)<br>дисциплины                                   | Тема<br>практического<br>занятия | Содержание<br>практического<br>занятия         | Вопросы к<br>практическому<br>занятию |
|-----------------|--------------------------------------------------------------------------------|----------------------------------|------------------------------------------------|---------------------------------------|
| 1               | Тема 1.1. Примеры<br>аналитических<br>задач. Сбор данных<br>для анализа данных | Аналитические<br>инструменты     | Использование<br>аналитических<br>инструментов | Тестирование                          |

#### Тема 1.2. Инструмент работы с таблицами в режиме онлайн на примере Google-таблиц.

Понятие и назначение табличных редакторов. Знакомство с интерфейсом Google-таблиц. Google-формы как источник данных для таблиц. Формулы в таблицах. Сводные таблицы. Продвинутое фишки и формулы Google-таблиц. ВПР для обогащения данных в одной таблице из другой. Графики и визуализация в google-таблицах.

Назначение Google таблиц и их особенности. Возможности и преимущества Google таблиц для анализа данных, сравнение со возможностями Excel. Создание сводной таблицы данных с помощью автоматических рекомендаций в Таблицах. Сводные таблицы для систематизации данных, выявления закономерностей и упорядочивания информации.

### Содержание практических занятий

| №<br>мод<br>уля | Наименовани<br>е темы<br>(раздела)<br>дисциплины                      | Тема<br>практическог<br>о занятия            | Содержание<br>практического занятия   | Вопросы к<br>практическому<br>занятию     |
|-----------------|-----------------------------------------------------------------------|----------------------------------------------|---------------------------------------|-------------------------------------------|
| 1               | Тема 1.2.<br>Google-<br>таблицы.<br>Продвинутые<br>Google-<br>таблицы | Построение<br>отчета в<br>Google-<br>таблице | Построение отчета в<br>Google-таблице | Построение<br>отчета в Google-<br>таблице |

#### Тема 1.3. Статистика и тестирование гипотез.

Понятие и назначение статистики. Описательная статистика. Основы статистики. Статистические закономерности и проверка гипотез. Случайные величины. Доверительные интервалы. Выдвижение рабочих гипотез. Формулирование гипотез. Сбор требований, разработка и построение отчетности. Тестирование гипотез. Использование подхода Jobs-to-be-done при разработке гипотез и продуктов.

### Содержание практических занятий

| №<br>мод<br>уля | Наименование<br>темы<br>(раздела)<br>дисциплины                | Тема<br>практическог<br>о занятия       | Содержание<br>практического занятия | Вопросы к<br>практическому<br>занятию                                     |
|-----------------|----------------------------------------------------------------|-----------------------------------------|-------------------------------------|---------------------------------------------------------------------------|
| 1               | Тема 1.3.<br>Основы<br>статистики и<br>тестирование<br>гипотез | Статистика и<br>тестирование<br>гипотез | Система анализа на<br>основе данных | Оценка<br>статистических<br>показателей в<br>данных и<br>проверка гипотез |

#### **Тема 1.4. Бизнес-показатели и точки роста. Анализ продуктовых метрик.**

Бизнес-показатели и понимание целей бизнес. Retention. Метрики активности пользователей. Работа с данными в компании. Принципы работы с данными. Потребности пользователей и юнит экономика. Аналитические подходы в бизнесе. Ключевые метрики компаний. Аналитические фреймворки. Основные продуктовые метрики. Продуктовые гипотезы. Подход к формулированию гипотез. Методы проверки гипотез. Поиск точек роста.

##### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                               | Тема практического занятия | Содержание практического занятия                                       | Вопросы к практическому занятию          |
|----------|----------------------------------------------------------------------|----------------------------|------------------------------------------------------------------------|------------------------------------------|
| 1        | Тема 1.4. Бизнес-показатели и точки роста. Анализ продуктовых метрик | Анализ продуктовых метрик  | Бизнес-показатели и использование аналитических инструментов в бизнесе | Решение аналитической проблемы в бизнесе |

#### **Тема 1.5. Финансовые метрики и Маркетинговые метрики.**

Маркетинговые метрики и метрики продукта. Значимость и понятие маркетинговой аналитики. Понятие жизненного цикла клиента. Эффективность маркетинга. Бизнес-задачи, их эффективность и аналитика. Финансовые метрики. Бюджетирование и финансовое планирование. Связь финансовых и нефинансовых метрик. CASHFLOW. P&L

##### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                | Тема практического занятия                    | Содержание практического занятия                 | Вопросы к практическому занятию   |
|----------|-------------------------------------------------------|-----------------------------------------------|--------------------------------------------------|-----------------------------------|
| 1        | Тема 1.5. Финансовые метрики и Маркетинговые метрики. | Система финансовых и нефинансовых показателей | Подготовка финансовой и маркетинговой стратегии. | Финансовые и нефинансовые метрики |

#### **Тема 1.6. Иерархия метрик и разработка отчетности. Дизайн тестов, проведение и анализ. Построение простых моделей.**

Базовые знания об иерархии метрик. Unit-экономика. Data driven подход. Data driven гипотезы. Требования и стейкхолдеры. Сбор требований к отчетности. Проектирование и разработка отчетности. Дизайн тестов, проведение и анализ. Построение простых моделей. ARPPU. LTV. North Star Metric. Качественные и количественные методы исследования. Дизайн экспериментов. АБ-тесты. Бизнес-моделирование.

##### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины | Тема практического занятия | Содержание практического занятия | Вопросы к практическому занятию |
|----------|----------------------------------------|----------------------------|----------------------------------|---------------------------------|
|----------|----------------------------------------|----------------------------|----------------------------------|---------------------------------|

|    |                                                                                                                    |                                                                             |                                        |                                                                |
|----|--------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|----------------------------------------|----------------------------------------------------------------|
| 1. | Тема 1.6. Иерархия метрик и разработка отчетности. Дизайн тестов, проведение и анализ. Построение простых моделей. | Разработка и оптимизация отчетности. Data-driven подход в принятии решений. | Проектирование и разработка отчетности | Сбор и управление требованиями отчетности. Создание отчетности |
|----|--------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|----------------------------------------|----------------------------------------------------------------|

### Тема 1.7. Аналитика данных в Excel.

Введение. Интерфейс Excel. Книги и листы. Анализ таблиц. Печать таблиц. Сводные таблицы. Вычисления и формулы. Умные таблицы. Функции подсчёта и суммирования. Статистические функции. Функции округления. Логические функции.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины | Тема практического занятия | Содержание практического занятия | Вопросы к практическому занятию |
|----------|----------------------------------------|----------------------------|----------------------------------|---------------------------------|
| 1.       | Тема 1.7 Аналитика данных в Excel      | Аналитика в Excel          | Аналитика в Excel                | Статистический анализ в Excel   |

### Содержание самостоятельной работы слушателей

Основная цель самостоятельной работы слушателей – закрепление знаний, полученных в ходе лекционных и практических занятий.

| № темы   | Наименование (содержание) темы, по которой предусмотрена самостоятельная работа | Формы и методы проведения                                                                                                  |
|----------|---------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Тема 1.1 | Примеры аналитических задач. Сбор данных для анализа данных                     | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 1.2 | Google-таблицы. Продвинутое Google-таблицы                                      | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 1.3 | Основы статистики и тестирование гипотез                                        | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 1.4 | Бизнес-показатели и точки роста. Анализ продуктовых метрик                      | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 1.5 | Финансовые метрики и Маркетинговые метрики.                                     | Изучение видеолекций, основной и дополнительной литературы по                                                              |

|          |                                                                                                         |                                                                                                                            |
|----------|---------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
|          |                                                                                                         | программе, решение практических задач с применением ЭО и ДОТ                                                               |
| Тема 1.6 | Иерархия метрик и разработка отчетности. Дизайн тестов, проведение и анализ. Построение простых моделей | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |

### Форма контроля

В процессе обучения осуществляется текущий контроль знаний и промежуточная аттестация.

Текущий контроль - в форме тестирования, состоящего из десяти вопросов с вариантами ответов, один из которых правильный, или решения типовых задач.

Промежуточная аттестация – выполнение промежуточного тестирования или практического задания по модулю.

### Список литературы

#### Основная литература

1. Зараменских, Е. П. Основы бизнес-информатики: учебник и практикум для бакалавриата и магистратуры / Е. П. Зараменских. – Москва: Издательство Юрайт, 2018. – 407 с. — (Высшее образование). – ISBN 978-5-9916-8210-7. – Текст : электронный // Образовательная платформа Юрайт [сайт]. – URL: <https://urait.ru/bcode/413823> (дата обращения: 10.04.2022).
2. Data Science. Наука о данных с нуля. / Билл Фрэнкс.; пер. с англ. Евстигнеева И.В. – М.: Издательство «Альпина Паблишер». – 2018. – 320 с.
3. Набатова, Д. С. Математические и инструментальные методы поддержки принятия решений: учебник и практикум для вузов / Д. С. Набатова. — Москва: Издательство Юрайт, 2022. — 292 с. — (Высшее образование). — ISBN 978-5-534-02699-3.
4. Курносков Ю.В. «Азбука аналитики», Издательство «Концептуал», 2018 -240 с.
5. Б. Марр «Ключевые инструменты бизнес-аналитики»/ пер с англ. Егоров В. Н., Издательство «Лаборатория знаний», 2018 – 339 с.
6. де Прадо М. «Машинное обучение: алгоритмы для бизнеса», Санкт Петербург: Издательский дом «Питер». – 2019. – 432 с.
7. Плас ван дер Д. «Python для сложных задач: наука о данных и машинное обучение», Санкт Петербург: Издательский дом «Питер». – 2018. – 576с.
8. Лакшманан В., Тайджани Д. «Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении», Санкт Петербург: Издательский дом «Питер». – 2021. – 496с.

#### Дополнительная литература

1. Глубокое обучение на Python. – СПб.: Питер, 2018. – 400 с.: ил. – (Серия «Библиотека программиста»).
2. Основы Data Science и Big Data. Python и наука о данных. / Силен Д., Мейсман А., Али М.; пер. с англ. – Санкт Петербург: Издательский дом «Питер». – 2018. – 336 с.
3. Радченко И.А., Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
4. Data Science. Наука о данных с нуля. / Джоэл Грас.; пер. с англ. Логунов А.В. – Санкт Петербург: Издательство «БХВ-Петербург». – 2018. – 336 с.
5. Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении. – СПб.: Питер, 2021. – 496 с.: ил.
6. Д. Битти К. Вигерс «Разработка требований к программному обеспечению», Издательство BHV, 2019 -737 с.

#### Ресурсы Internet

1. <https://www.anaconda.com/> – страница загрузки Anaconda.



2. <https://powerbi.microsoft.com/ru-ru/desktop/> – страница загрузки платформы Power BI Desktop компании Microsoft.
3. <https://public.tableau.com/en-us/s/download> – страница загрузки платформы Tableau Public.
4. <https://loginom.ru/download> – страница загрузки платформы Loginom компании BaseGroup Labs.
5. <https://www.knime.com/knime-analytics-platform> – страница загрузки платформы Knime Analytics Platform.
6. <https://community.cloud.databricks.com/login.html> – страница регистрации Databricks Community Edition
7. <https://rapidminer.com/get-started/> – страница загрузки платформы RapidMiner.
8. <http://h2o-release.s3.amazonaws.com/h2o/rel-zermelo/4/index.html> – страница загрузки платформы H2O.
9. <https://cloud.google.com/bigquery/docs/sandbox> – страница BigQuery sandbox.
10. <https://cloud.yandex.ru/services/datalens> – сервис визуализации и анализа данных Яндексе.
11. <https://lookerstudio.google.com/> - страница загрузки Looker Studio
12. <https://www.tableau.com> - страница загрузки Tableau

## Модуль 2. Анализ данных с применением Python

**Цель модуля:** приобретение слушателями компетенций, необходимых для понимания и эффективной работы в области анализа больших данных, инструментов и технологий, позволяющих анализировать результаты внутренних процессов организации с помощью инструментов языка Python.

### Планируемые образовательные результаты по модулю

В результате освоения модуля слушатели должны:

*а) знать:*

- базовые инструменты и подходы в Python, используемые для анализа с данными;
- информационные технологии (программное обеспечение), применяемые в организации, в объеме, необходимом для целей бизнес-анализа;
- теоретические основы анализа данных;
- правила и методы проектирования моделей данных, внедрения аналитических отчетов, дашбордов, информационных панелей мониторинга в компоненты ИТ-инфраструктуры организации для проведения мониторинга и контроллинга стратегических целей и поддержку внутренних процессов.

*б) уметь:*

- моделировать объем и границы работ;
- определять связи и зависимости между элементами информации бизнес-анализа;
- использовать современные информационные технологии аналитики для информационно-аналитического сопровождения деятельности;
- работать в веб блокнотах Jupyter Notebook, среде Google Colaboratory;
- загружать данные в pandas и работать с ними (фильтрация, агрегация, заполнение пропущенных значений);
- моделировать требования заинтересованных сторон в соответствии с выбранными подходами;
- анализировать качество информации бизнес-анализа с точки зрения выбранных критериев;
- проводить анализ предметной области.

*в) владеть:*

- анализом потребностей заинтересованных сторон;
- инструментами и технологиями выявления, сбора и анализа информации бизнес-анализа для формирования возможных решений;
- описанием возможных решений;
- навыками работы в веб блокнотах Jupyter Notebook, среде Google Colaboratory
- инструментами и подходами в Python, используемые для анализа данных (создание датафрейма из данных, копирование датафрейма, экспорт данных, исследование данных, вывод статистических сведений, извлечение информации, фильтрация, сортировка, агрегирование);
- инструментами и технологиями выявления, сбора и анализа информации бизнес-анализа для формирования возможных решений;
- инструментами формирования аналитических отчетов, дашбордов и «повествований» для поддержки принятия решений.

### Формируемые компетенции:

способность разрабатывать системы анализа больших данных.

способность проведения аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика.

## УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

| №<br>№<br>п/п | Наименование разделов,<br>модулей                                                                                      | Всего          | В том числе           |        | Самостоятельная<br>работа | Форма контроля |                               |
|---------------|------------------------------------------------------------------------------------------------------------------------|----------------|-----------------------|--------|---------------------------|----------------|-------------------------------|
|               |                                                                                                                        |                | Аудиторные<br>занятия |        |                           |                |                               |
|               |                                                                                                                        | В<br>час<br>ах | Всего,<br>часов       | из них |                           |                |                               |
|               |                                                                                                                        |                |                       | Лекции | Практич.<br>занятия       |                |                               |
| 1             | 2                                                                                                                      | 3              | 4                     | 5      | 6                         | 1              | 2                             |
| М 2.          | Анализ данных с<br>применением Python                                                                                  | 23             | 13                    | 4      | 9                         | 10             | Зачет                         |
| 2.1.          | Python: Jupiter Notebook,<br>комментарии в коде,<br>арифметические операции.<br>Управляющие конструкции и<br>коллекции | 4              | 2                     | 2      |                           | 2              | Тестирование                  |
| 2.2.          | Функции и работа с данными,<br>основы парсинга. Работа с<br>файловой системой и модули                                 | 4              | 2                     | 2      |                           | 2              | Тестирование                  |
| 2.3.          | Регулярные выражения и<br>основы синтаксического<br>разбора. Исключения и<br>обработка ошибок                          | 4              | 2                     |        | 2                         | 2              | Решение<br>практических задач |
| 2.4.          | Функции и понятие класса.<br>Библиотека NumPy:<br>вычислительные задачи                                                | 3              | 2                     |        | 2                         | 1              | Решение<br>практических задач |
| 2.5.          | Библиотека Pandas:<br>возможности и оптимизация.<br>Методы оптимизации Pandas                                          | 4              | 2                     |        | 2                         | 2              | Решение<br>практических задач |
| 2.6.          | Статистическая проверка<br>гипотез и A/B-тесты                                                                         | 3              | 2                     |        | 2                         | 1              | Решение<br>практических задач |
|               | Промежуточная<br>аттестация по модулю 2                                                                                | 1              | 1                     |        | 1                         |                | Зачет                         |

**Тема 2.1. Python: Jupiter Notebook, комментарии в коде, арифметические операции. Управляющие конструкции и коллекции.**

Установка и запуск Python, Jupiter Notebook: его возможности для команд. Переменные. Логические операторы. Условные конструкции. Управляющие конструкции и коллекции. Простые типы данных. Списки. Кортежи. Множества. Словари. Циклы WHILE/FOR.

### Содержание практических занятий

| № мод<br>уля | Наименование<br>темы (раздела)<br>дисциплины     | Тема<br>практического<br>занятия                | Содержание<br>практического<br>занятия        | Вопросы к<br>практическому<br>занятию      |
|--------------|--------------------------------------------------|-------------------------------------------------|-----------------------------------------------|--------------------------------------------|
| 2            | Тема 2.1.<br>Основы Python:<br>Jupiter Notebook, | Управляющие<br>конструкции и<br>коллекции, типы | Обработка разных<br>типов данных в<br>Python. | Какие типы<br>данных может<br>обрабатывать |

|  |                                                                                  |                            |  |         |
|--|----------------------------------------------------------------------------------|----------------------------|--|---------|
|  | комментарии в коде, арифметические операции. Управляющие конструкции и коллекции | данных и словари в Python. |  | Python. |
|--|----------------------------------------------------------------------------------|----------------------------|--|---------|

## **Тема 2.2. Функции и работа с данными, основы парсинга. Работа с файловой системой и модули.**

Функции. Объявление функций и ее параметры. Области видимости. Анонимные функции. Оперирование основными элементами и типами данных в Python. Работа с файловой системой и модули: чтение и запись файлов. Работа с пакетами. Пакетный менеджер. Постановка задачи. Работа с модулями. Основы парсинга.

### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                                     | Тема практического занятия | Содержание практического занятия     | Вопросы к практическому занятию                                      |
|----------|--------------------------------------------------------------------------------------------|----------------------------|--------------------------------------|----------------------------------------------------------------------|
| 2.       | Тема 2.2. Функции и работа с данными, основы парсинга. Работа с файловой системой и модули | Функции и парсинг данных   | Парсинг данных с внешних источников. | Какими способами можно осуществлять парсинг данных с помощью Python. |

## **Тема 2.3. Регулярные выражения и основы синтаксического разбора. Исключения и обработка ошибок.**

Работа с датами, регулярными выражениями, веб-запросами. Регулярные выражения: понятие. Отличие регулярного выражения от простого поиска. Базовый синтаксис. Модуль и дополнительные возможности. Модуль RE в Python. Примеры использования регулярных выражений. Код и ошибки в нем. Виды ошибок. Исключения и обработка ошибок. Строки и даты в python. Unixtime и сложности времени.

### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                                         | Тема практического занятия | Содержание практического занятия       | Вопросы к практическому занятию             |
|----------|------------------------------------------------------------------------------------------------|----------------------------|----------------------------------------|---------------------------------------------|
| 2.       | Тема 2.3. Регулярные выражения и основы синтаксического разбора. Исключения и обработка ошибок | Регулярные выражение.      | Создание регулярных выражений в Python | Примеры использования регулярных выражений. |

#### **Тема 2.4. Функции и понятие класса. Библиотека NumPy: вычислительные задачи.**

Классы и объектно-ориентированное программирование. Зачем классы в анализе данных. Работа с классами. Преимущества классов перед функциями. Наследование классов.

Метрики расстояний и векторы. Работа с матрицами. Линейные уравнения. Практика в NumPy.

##### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                      | Тема практического занятия                 | Содержание практического занятия      | Вопросы к практическому занятию       |
|----------|-----------------------------------------------------------------------------|--------------------------------------------|---------------------------------------|---------------------------------------|
| 2.       | Тема 2.4. Функции и понятие класса. Библиотека NumPy: вычислительные задачи | Классы в анализе данных. Библиотека NumPy. | Основные возможности библиотеки NumPy | Основные возможности библиотеки NumPy |

#### **Тема 2.5. Библиотека Pandas: возможности и оптимизация. Методы оптимизации Pandas.**

Обзор библиотеки Pandas. Базовые операции. Режим презентации в Jupiter Notebook. Pandas: оптимизация, сводные таблицы, методы работы со строками. Объединение датафреймов и типы. Дубликаты. Конкатенация таблиц. Оптимизация хранения данных с помощью join. Функции и группировки в Pandas. Поиск. Pivot и фильтры. Работа с падежами и склонениями.

##### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                            | Тема практического занятия | Содержание практического занятия                           | Вопросы к практическому занятию                         |
|----------|-----------------------------------------------------------------------------------|----------------------------|------------------------------------------------------------|---------------------------------------------------------|
| 2.       | Тема 2.5. Библиотека Pandas: возможности и оптимизация. Методы оптимизации Pandas | Библиотека Pandas          | Основные возможности библиотеки Pandas по работе с данными | Основные методы Pandas по работе с таблицами и данными. |

#### **Тема 2.6. Статистическая проверка гипотез и A/B-тесты**

Основные понятия статистики. Среднее. Квантиль и медиана. Стандартное отклонение. Дисперсия. Нахождение математического ожидания и дисперсии. Выбросы. Нормализация данных. Корреляция. Финансовый анализ данных. Корреляция Пирсона/Спирмена/Кендалла. Матрица корреляций. Нахождение зависимости случайных величин. Правило трех сигм. Линейная регрессия. Функция потерь. Центральная предельная теорема. Виды распределений. t- критерий Стьюдента. A/B тесты: понятие, параметры, инструменты, как проводить тестирование.



### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                | Тема практического занятия             | Содержание практического занятия | Вопросы к практическому занятию    |
|----------|-------------------------------------------------------|----------------------------------------|----------------------------------|------------------------------------|
| 2.       | Тема 2.6. Статистическая проверка гипотез и А/В-тесты | Основные понятия статистики. А/В тесты | Проведение А/В тестов            | Инструменты проведения А/В тестов. |

### Содержание самостоятельной работы слушателей

Основная цель самостоятельной работы слушателей – закрепление знаний, полученных в ходе лекционных и практических занятий.

| № темы    | Наименование (содержание) темы, по которой предусмотрена самостоятельная работа                                   | Формы и методы проведения                                                                                                  |
|-----------|-------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Тема 2.1. | Основы Python: Jupiter Notebook, комментарии в коде, арифметические операции. Управляющие конструкции и коллекции | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 2.2  | Функции и работа с данными, основы парсинга. Работа с файловой системой и модули                                  | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 2.3  | Регулярные выражения и основы синтаксического разбора. Исключения и обработка ошибок                              | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 2.4  | Функции и понятие класса. Библиотека NumPy: вычислительные задачи                                                 | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 2.5  | Библиотека Pandas: возможности и оптимизация. Методы оптимизации Pandas                                           | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 2.6  | Статистическая проверка гипотез и А/В-тесты                                                                       | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |

### Форма контроля

В процессе обучения осуществляется текущий контроль знаний и промежуточная аттестация.

Текущий контроль - в форме тестирования, состоящего из десяти вопросов с вариантами ответов, один из которых правильный, или решения типовых задач.

Промежуточная аттестация – выполнение промежуточного тестирования или практического задания по модулю.

## Список литературы учебного курса

### Основная литература

1. Зараменских, Е. П. Основы бизнес-информатики: учебник и практикум для бакалавриата и магистратуры / Е. П. Зараменских. – Москва: Издательство Юрайт, 2018. – 407 с. — (Высшее образование). – ISBN 978-5-9916-8210-7. – Текст : электронный // Образовательная платформа Юрайт [сайт]. – URL: <https://urait.ru/bcode/413823> (дата обращения: 10.04.2022).
2. Data Science. Наука о данных с нуля. / Билл Фрэнкс.; пер. с англ. Евстигнеева И.В. – М.: Издательство «Альпина Паблишер». – 2018. – 320 с.
3. Набатова, Д. С. Математические и инструментальные методы поддержки принятия решений: учебник и практикум для вузов / Д. С. Набатова. — Москва: Издательство Юрайт, 2022. — 292 с. — (Высшее образование). — ISBN 978-5-534-02699-3.
4. Курносов Ю.В. «Азбука аналитики», Издательство «Концептуал», 2018 -240 с.
5. Б. Марр «Ключевые инструменты бизнес-аналитики»/ пер с англ. Егоров В. Н., Издательство «Лаборатория знаний», 2018 – 339 с.
6. де Прадо М. «Машинное обучение: алгоритмы для бизнеса», Санкт Петербург: Издательский дом «Питер», – 2019. – 432 с.
7. Плас вандер Д. «Python для сложных задач: наука о данных и машинное обучение», Санкт Петербург: Издательский дом «Питер», – 2018. – 576с.
8. Лакшманан В., Тайджани Д. «Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении», Санкт Петербург: Издательский дом «Питер», – 2021. – 496с.

### Дополнительная литература

1. Глубокое обучение на Python. – СПб.: Питер, 2018. – 400 с.: ил. – (Серия «Библиотека программиста»).
2. Основы Data Science и Big Data. Python и наука о данных. / Силен Д., Мейсман А., Али М.; пер. с англ. – Санкт Петербург: Издательский дом «Питер», – 2018. – 336 с.
3. Радченко И.А., Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
4. Data Science. Наука о данных с нуля. / Джозэл Грас.; пер. с англ. Логунов А.В. – Санкт Петербург: Издательство «БХВ-Петербург». – 2018. – 336 с.
5. Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении. – СПб.: Питер, 2021. – 496 с.: ил.
6. Д. Битти К. Вигерс «Разработка требований к программному обеспечению», Издательство BHV, 2019 -737 с.

### Ресурсы Internet

1. <https://www.anaconda.com/> – страница загрузки Anaconda.
2. <https://powerbi.microsoft.com/ru-ru/desktop/> – страница загрузки платформы Power BI Desktop компании Microsoft.
3. <https://public.tableau.com/en-us/s/download> – страница загрузки платформы Tableau Public.
4. <https://loginom.ru/download> – страница загрузки платформы Loginom компании BaseGroup Labs.
5. <https://www.knime.com/knime-analytics-platform> – страница загрузки платформы Knime Analytics Platform.
6. <https://community.cloud.databricks.com/login.html> – страница регистрации Databricks Community Edition
7. <https://rapidminer.com/get-started/> – страница загрузки платформы RapidMiner.
8. <http://h2o-release.s3.amazonaws.com/h2o/rel-zermelo/4/index.html> – страница загрузки платформы H2O.
9. <https://cloud.google.com/bigquery/docs/sandbox> – страница BigQuery sandbox.
10. <https://cloud.yandex.ru/services/datalens> – сервис визуализации и анализа данных Яндекса.
11. <https://lookerstudio.google.com/> - страница загрузки Looker Studio
12. <https://www.tableau.com> - страница загрузки Tableau

### Модуль 3. Применение SQL для работы с данными

**Цель модуля:** приобретение слушателями компетенций в области локальных корпоративных хранилищ данных, современных технологий Big Data; языка программирования SQL для аналитики больших данных и облачных технологий обработки больших данных.

#### Планируемые образовательные результаты по модулю

В результате освоения модуля слушатели должны:

##### *а) знать:*

- традиционные локальные корпоративные хранилища данных;
- современные виды хранилищ больших данных;
- облачные технологии обработки больших данных;
- основы языка программирования SQL;
- возможности SQL для решения задачи сбора и подготовки данных;
- основы машинного обучения в BigQuery с использованием TensorFlow

##### *б) уметь:*

- применять язык SQL для сбора и подготовки данных и визуализации данных;
- собирать, классифицировать, систематизировать и обеспечивать хранение и актуализацию информации бизнес-анализа;
- использовать современные информационные технологии аналитики для информационно-аналитического сопровождения деятельности организации;

##### *в) владеть:*

- анализом решений с точки зрения достижения целевых показателей решений;
- навыками работы с сырыми данными, использования SQL для интеграции и очистки данных;
- навыками применения технологий машинного обучения для решения анализа данных и принятия решений.

#### Формируемые компетенции:

- способность проведения анализа больших данных с использованием существующей в организации методологической и технологической инфраструктуры;
- способность к подготовке данных для проведения аналитических работ по исследованию больших данных;
- способность управлять получением, хранением, передачей, обработкой больших данных;
- способность управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации;
- способность разрабатывать системы анализа больших данных.

# УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

| №№<br>п/п | Наименование разделов,<br>модулей                                                       | Всего      | В том числе        |        |                    | Самостоятельная<br>работа | Форма<br>контроля                |
|-----------|-----------------------------------------------------------------------------------------|------------|--------------------|--------|--------------------|---------------------------|----------------------------------|
|           |                                                                                         |            | Аудиторные занятия |        |                    |                           |                                  |
|           |                                                                                         | В<br>часах | Всего, часов       | из них |                    |                           |                                  |
|           |                                                                                         |            |                    | Лекции | Практич<br>занятия |                           |                                  |
| 1         | 2                                                                                       | 3          | 4                  | 5      | 6                  | 1                         | 2                                |
| М 3.      | Применение SQL для<br>работы с данными                                                  | 25         | 15                 | 4      | 11                 | 10                        | Зачет                            |
| 3.1       | Введение в SQL.,<br>Установка и знакомство,<br>Основы и работа с базами<br>данных       | 4          | 2                  | 2      |                    | 2                         | Тестирование                     |
| 3.2       | Основы SQL. Базовые<br>запросы SQL. Фильтрация<br>данных                                | 4          | 2                  | 2      |                    | 2                         | Тестирование                     |
| 3.3       | Углубление в SQL.,<br>Агрегация данных.<br>Группировка данных. Join.<br>Оконные функции | 3          | 2                  |        | 2                  | 1                         | Решение<br>практических<br>задач |
| 3.4       | Работа с PostgreSQL                                                                     | 4          | 2                  |        | 2                  | 2                         | Решение<br>практических<br>задач |
| 3.5       | "NoSql & MongoDB"                                                                       | 4          | 2                  |        | 2                  | 2                         | Решение<br>практических<br>задач |
| 3.6       | Примеры создания<br>дашбордов и анализ<br>продуктовых метрик с<br>применением SQL       | 5          | 4                  |        | 4                  | 1                         | Решение<br>практических<br>задач |
|           | Промежуточная<br>аттестация по модулю<br>3                                              | 1          | 1                  |        | 1                  |                           | Зачет                            |

**Тема 3.1. Введение в SQL. Установка и знакомство. Основы и работа с базами данных.**

База данных и источники данных. Инструменты работы с БД. Функции СУБД. Установка и первичная настройка ПО, подключение к БД. Основы языка программирования SQL. DE и DBEaver. Создание подключений в DBEaver. Работа с \*.backup и \*.sql Интерфейс DBEaver. ER-диаграмма. Запросы, как создавать запросы, простые запросы. Основные типы данных. Работа с разными типами данных.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                                     | Тема практического занятия             | Содержание практического занятия                        | Вопросы к практическому занятию                           |
|----------|--------------------------------------------------------------------------------------------|----------------------------------------|---------------------------------------------------------|-----------------------------------------------------------|
| 3        | Тема 3.1.<br>Введение в SQL.<br>Установка и знакомство.<br>Основы и работа с базами данных | Конструкции языка программирования SQL | Простые запросы SQL, агрегирование, массивы и структуры | Возможности языка программирования SQL для анализа данных |

#### Тема 3.2. Основы SQL. Базовые запросы SQL. Фильтрация данных.

Создание таблиц и простые SQL-запросы. Присоединение таблиц. Группировка данных. Вложенные запросы. Условия с использованием оператора CASE. Типы соединения. Соединения Join/ UNION. Команды. Создание отчетов и аналитика с помощью SQL. Усложнение выборки данных. Фильтрация

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                | Тема практического занятия | Содержание практического занятия           | Вопросы к практическому занятию                               |
|----------|-----------------------------------------------------------------------|----------------------------|--------------------------------------------|---------------------------------------------------------------|
| 3.       | Тема 3.2.<br>Основы SQL.<br>Базовые запросы SQL.<br>Фильтрация данных | Группировка данных         | Создание отчетов и аналитика с помощью SQL | Возможности SQL при создании отчетов и аналитических справок. |

#### Тема 3.3. Углубление в SQL. Агрегация данных. Группировка данных. Join. Оконные функции.

Таблицы и данные. Создание, изменение, удаление таблиц. Вставка, модификация, удаление данных. Внешние ключи. Проверяющие ограничения. Constraints. Агрегация данных. Агрегатные функции. Группировка. JOIN при работе с уникальными и неуникальными значениями.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                             | Тема практического занятия   | Содержание практического занятия | Вопросы к практическому занятию                       |
|----------|--------------------------------------------------------------------|------------------------------|----------------------------------|-------------------------------------------------------|
| 3.       | Тема 3.3.<br>Углубление в SQL.<br>Агрегация данных.<br>Группировка | Работа с данными и таблицами | Агрегация данных с помощью SQL   | С помощью каких функции SQL можно агрегировать данные |



|  |                                          |  |  |  |
|--|------------------------------------------|--|--|--|
|  | а данных.<br>Join.<br>Оконные<br>функции |  |  |  |
|--|------------------------------------------|--|--|--|

### Тема 3.4. Работа с PostgreSQL

Оконные и аналитические функции. CTE. Рекурсия. Применение с каталогами. Tablefunc. PostGIS . File\_fdw. Postgres\_fdw . Pg\_stat\_statements. Представления. Применение представлений в денормализации. Схемы и планы запросов. Ускорение запросов и индексы. Сложные типы данных. Массивы. Команды VIEW, MATERIALIZED VIEW, JSON, ARRAY.

#### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины | Тема практического занятия                         | Содержание практического занятия                 | Вопросы к практическому занятию                                   |
|----------|----------------------------------------|----------------------------------------------------|--------------------------------------------------|-------------------------------------------------------------------|
| 3        | Тема 3.4. Работа с PostgreSQL          | Ускорение запросов и индексы. Сложные типы данных. | Написание сложных запросов и работа с PostgreSQL | Какие задачи решают команды VIEW, MATERIALIZED VIEW, JSON, ARRAY. |

### Тема 3.5. "NoSql & MongoDB".

Базы данных NoSql. CAP-теорема. Классы систем. Использование базы данных NoSql. Типы данных NoSql. MongoDB. Модель устройства БД в MongoDB. Документ. Коллекция: управление и создание.

#### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины | Тема практического занятия | Содержание практического занятия | Вопросы к практическому занятию        |
|----------|----------------------------------------|----------------------------|----------------------------------|----------------------------------------|
| 3        | Тема 3.5. "NoSql & MongoDB"            | Базы данных NoSQL          | Работа с базой данных NoSql      | Какие типы данных используются в NoSQL |

### Тема 3.6. Примеры создания дашбордов и анализ продуктовых метрик с применением SQL

Основные виды дашбордов. Лучшие практики дашбордов. Основы работы с расчётными полями, фильтрами, множествами и группировками. Разработка дашбордов. Настройка взаимодействия между визуализациями. Существующие программные решения для OLAP-моделирования. Многомерный анализ. Применение OLAP при решении аналитических задач.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                             | Тема практического занятия | Содержание практического занятия         | Вопросы к практическому занятию                  |
|----------|------------------------------------------------------------------------------------|----------------------------|------------------------------------------|--------------------------------------------------|
| 3        | Тема 3.6. Примеры создания дашбордов и анализ продуктовых метрик с применением SQL | Дашборды и их виды         | Практика создания дашбордов и OLAP кубов | Применение OLAP при решении аналитических задач. |

### Содержание самостоятельной работы слушателей

Основная цель самостоятельной работы слушателей – закрепление знаний, полученных в ходе лекционных и практических занятий.

| № темы   | Наименование (содержание) темы, по которой предусмотрена самостоятельная работа | Формы и методы проведения                                                                                                  |
|----------|---------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Тема 3.1 | Введение в SQL. Установка и знакомство. Основы и работа с базами данных         | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 3.2 | Основы SQL. Базовые запросы SQL. Фильтрация данных                              | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 3.3 | Углубление в SQL. Агрегация данных. Группировка данных. Join. Оконные функции   | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 3.4 | Работа с PostgreSQL                                                             | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 3.5 | "NoSql & MongoDB"                                                               | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 3.6 | Примеры создания дашбордов и анализ продуктовых метрик с применением SQL        | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |

### Форма контроля

В процессе обучения осуществляется текущий контроль знаний и промежуточная аттестация.

Текущий контроль - в форме тестирования, состоящего из десяти вопросов с вариантами ответов, один из которых правильный, или решения типовых задач.

Промежуточная аттестация – выполнение промежуточного тестирования или практического задания по модулю.

## Список литературы

### Основная литература

1. Зараменских, Е. П. Основы бизнес-информатики: учебник и практикум для бакалавриата и магистратуры / Е. П. Зараменских. – Москва: Издательство Юрайт, 2018. – 407 с. — (Высшее образование). – ISBN 978-5-9916-8210-7. – Текст : электронный // Образовательная платформа Юрайт [сайт]. – URL: <https://urait.ru/bcode/413823> (дата обращения: 10.04.2022).
2. Data Science. Наука о данных с нуля. / Билл Фрэнкс.; пер. с англ. Евстигнеева И.В. – М.: Издательство «Альпина Паблишер». – 2018. – 320 с.
3. Набатова, Д. С. Математические и инструментальные методы поддержки принятия решений: учебник и практикум для вузов / Д. С. Набатова. — Москва: Издательство Юрайт, 2022. — 292 с. — (Высшее образование). — ISBN 978-5-534-02699-3.
4. Курносков Ю.В. «Азбука аналитики», Издательство «Концептуал», 2018 -240 с.
5. Б. Марр «Ключевые инструменты бизнес-аналитики»/ пер с англ. Егоров В. Н., Издательство «Лаборатория знаний», 2018 – 339 с.
6. де Прадо М. «Машинное обучение: алгоритмы для бизнеса», Санкт Петербург: Издательский дом «Питер». – 2019, – 432 с.
7. Плас вандер Д. «Python для сложных задач: наука о данных и машинное обучение», Санкт Петербург: Издательский дом «Питер». – 2018. – 576с.
8. Лакшманиан В., Тайджани Д. «Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении», Санкт Петербург: Издательский дом «Питер». – 2021. – 496с.

### Дополнительная литература

1. Глубокое обучение на Python. – СПб.: Питер, 2018. – 400 с.: ил. – (Серия «Библиотека программиста»).
2. Основы Data Science и Big Data. Python и наука о данных. / Силен Д., Мейсман А., Али М.; пер. с англ. – Санкт Петербург: Издательский дом «Питер». – 2018. – 336 с.
3. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
4. Data Science. Наука о данных с нуля. / Джозл Грас.; пер. с англ. Логунов А.В. – Санкт Петербург: Издательство «БХВ-Петербург». – 2018. – 336 с.
5. Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении. – СПб.: Питер, 2021. – 496 с.: ил.
6. Д. Битти К. Вигерс «Разработка требований к программному обеспечению», Издательство BHV, 2019 -737 с.

### Ресурсы Internet

1. <https://www.anaconda.com/> – страница загрузки Anaconda.
2. <https://powerbi.microsoft.com/ru-ru/desktop/> – страница загрузки платформы Power BI Desktop компании Microsoft.
3. <https://public.tableau.com/en-us/s/download> – страница загрузки платформы Tableau Public.
4. <https://loginom.ru/download> – страница загрузки платформы Loginom компании BaseGroup Labs.
5. <https://www.knime.com/knime-analytics-platform> – страница загрузки платформы Knime Analytics Platform.
6. <https://community.cloud.databricks.com/login.html> – страница регистрации Databricks Community Edition
7. <https://rapidminer.com/get-started/> – страница загрузки платформы RapidMiner.
8. <http://h2o-release.s3.amazonaws.com/h2o/rel-zermelo/4/index.html> – страница загрузки платформы H2O.
9. <https://cloud.google.com/bigquery/docs/sandbox> – страница BigQuery sandbox.
10. <https://cloud.yandex.ru/services/datalens> – сервис визуализации и анализа данных Яндекс.
11. <https://lookerstudio.google.com/> - страница загрузки Looker Studio
12. <https://www.tableau.com> - страница загрузки Tableau

## Модуль 4. Визуализация данных

**Цель модуля:** приобретение слушателями компетенций, необходимых для выработки возможных решений посредством сбора и анализа больших данных и элементами информации бизнес-анализа.

### Планируемые образовательные результаты по модулю

В результате освоения модуля слушатели должны:

*а) знать:*

- информационные технологии (программное обеспечение), применяемые в организации, в объеме, необходимом для целей бизнес-анализа;
- теоретические основы анализа данных, технологии визуализации данных Tableau, Power BI и др;
- правила и методы проектирования моделей данных, внедрения аналитических отчетов, дашбордов, информационных панелей мониторинга в компоненты ИТ-инфраструктуры организации для проведения мониторинга и контроллинга стратегических целей и поддержку внутренних процессов;

*б) уметь:*

- решать стандартные задачи профессиональной деятельности используя технологии аналитики;
- определять связи и зависимости между элементами информации бизнес-анализа;
- использовать современные информационные технологии аналитики для информационно-аналитического сопровождения деятельности;
- моделировать требования заинтересованных сторон в соответствии с выбранными подходами;
- анализировать качество информации бизнес-анализа с точки зрения выбранных критериев;
- проводить анализ предметной области.

*в) владеть:*

- инструментами и технологиями выявления, сбора и анализа информации бизнес-анализа для формирования возможных решений;
- технологиями интерактивной визуализации данных;
- инструментами формирования аналитических отчетов, дашбордов и «повествований» для поддержки принятия решений.

### Формируемые компетенции:

способность проведения анализа, обоснования и выбора решения в области больших данных;

способность к подготовке данных для проведения аналитических работ по исследованию больших данных;

способность классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта;

способность использовать инструментальные средства для решения задач машинного обучения.

# УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

| №№<br>п/п | Наименование разделов,<br>модулей                                                                                                                           | Всего      | В том числе           |        | Самостоятельная<br>работа | Форма<br>контроля |                                  |
|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|-----------------------|--------|---------------------------|-------------------|----------------------------------|
|           |                                                                                                                                                             |            | Аудиторные<br>занятия |        |                           |                   |                                  |
|           |                                                                                                                                                             | В<br>часах | Всего, часов          | из них |                           |                   | Лекции                           |
| 1         | 2                                                                                                                                                           | 3          | 4                     | 5      | 6                         | 1                 | 2                                |
| М 4.      | Визуализация данных                                                                                                                                         | 21         | 11                    | 4      | 7                         | 10                | Зачет                            |
| 4.1       | BI-инструменты и<br>визуализация данных                                                                                                                     | 3          | 2                     | 2      |                           | 1                 | Тестирование                     |
| 4.2       | DataLens: быстрый анализ<br>данных с визуализацией                                                                                                          | 3          | 2                     | 2      |                           | 1                 | Тестирование                     |
| 4.3.      | Power BI. Построение моделей<br>данных из разных<br>неструктурированных<br>источников: таблиц, сайтов и<br>баз данных                                       | 4          | 2                     |        | 2                         | 2                 | Решение<br>практических<br>задач |
| 4.4.      | Tableau. Основы работы с<br>расчётными полями,<br>фильтрами, множествами и<br>группировками. Основные<br>виды визуализаций. Лучшие<br>практики визуализации | 6          | 2                     |        | 2                         | 4                 | Решение<br>практических<br>задач |
| 4.5.      | Создание интерактивных<br>дашбордов                                                                                                                         | 4          | 2                     |        | 2                         | 2                 | Решение<br>практических<br>задач |
|           | Промежуточная<br>аттестация по модулю 4                                                                                                                     | 1          | 1                     |        | 1                         |                   | Зачет                            |

## Тема 4.1. BI-инструменты и визуализация данных.

Business Intelligence: внедрение и выгоды Выбор BI-инструментов. Виды BI-инструментов. Задачи, которые могут решить BI-инструменты.

## Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины  | Тема практического занятия | Содержание практического занятия                     | Вопросы к практическому занятию        |
|----------|-----------------------------------------|----------------------------|------------------------------------------------------|----------------------------------------|
| 4.       | Тема 4.1. BI-инструменты и визуализации | Основные BI-инструменты    | Обзор основных BI-инструментов, кейсы использования. | Сопоставить какие BI-инструменты лучше |



|  |          |  |  |                                 |
|--|----------|--|--|---------------------------------|
|  | я данных |  |  | использовать для бизнес кейсов. |
|--|----------|--|--|---------------------------------|

#### **Тема 4.2. DataLens: быстрый анализ данных с визуализацией.**

DataLens - основные возможности. Подключение к источнику данных. Описание данных. Создание визуализаций. Работа с дашбордом.

##### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                    | Тема практического занятия      | Содержание практического занятия | Вопросы к практическому занятию |
|----------|-----------------------------------------------------------|---------------------------------|----------------------------------|---------------------------------|
| 4.       | Тема 4.2. DataLens: быстрый анализ данных с визуализацией | DataLens - основные возможности | Работа с данными в DataLens      | Виды визуализаций в DataLens    |

#### **Тема 4.3. Power BI. Построение моделей данных из разных неструктурированных источников: таблиц, сайтов и баз данных.**

Назначение и функциональные блоки в Power BI Desktop. Назначение облачного сервиса аналитики PowerBI.com. Ключевые отличия и преимущества Power BI, возможности Power BI по обработке больших данных. Подключение к данным, преобразование и формирование данных, создание модели, визуализаций и отчетов, информационных панелей мониторинга, совместная работа в Power BI. DAX и язык M. Обработка естественного языка, технология вопрос и ответов Q&A в Power BI. Расширенная аналитика и интеграция технологий машинного обучения Azure.

##### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                                                              | Тема практического занятия                       | Содержание практического занятия | Вопросы к практическому занятию         |
|----------|---------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|----------------------------------|-----------------------------------------|
| 4.       | Тема 4.3. Power BI. Построение моделей данных из разных неструктурированных источников: таблиц, сайтов и баз данных | Возможности Power BI по обработке больших данных | Разработка отчетов в Power BI    | Методы преобразования данных в Power BI |

**Тема 4.4. Tableau. Основы работы с расчётными полями, фильтрами, множествами и группировками. Основные виды визуализаций. Лучшие практики визуализации.**

Tableau - основные возможности. Подключение к источнику данных. Описание данных. Создание визуализаций. Работа с дашбордом.

#### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                                                                                                 | Тема практического занятия     | Содержание практического занятия | Вопросы к практическому занятию |
|----------|--------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|----------------------------------|---------------------------------|
| 4.       | Тема 4.4. Tableau. Основы работы с расчётными полями, фильтрами, множествами и группировками. Основные виды визуализаций. Лучшие практики визуализации | Tableau - основные возможности | Работа с данными в Tableau       | Виды визуализаций в Tableau     |

#### Тема 4.5. Создание интерактивных дашбордов.

Добавление интерактивности и публикация отчетов.

#### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                                                  | Тема практического занятия          | Содержание практического занятия    | Вопросы к практическому занятию               |
|----------|---------------------------------------------------------------------------------------------------------|-------------------------------------|-------------------------------------|-----------------------------------------------|
| 4.       | Тема 4.5. Добавление интерактивности и публикация отчетов. Построение дашбордов по предложенным кейсам. | Разработка интерактивных дашбордов. | Разработка интерактивных дашбордов. | Способы добавления интерактивности в дашборды |

## Содержание самостоятельной работы слушателей

Основная цель самостоятельной работы слушателей – закрепление знаний, полученных в ходе лекционных и практических занятий.

| № темы   | Наименование (содержание) темы, по которой предусмотрена самостоятельная работа                                                              | Формы и методы проведения                                                                                                  |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Тема 4.1 | BI-инструменты и визуализация данных                                                                                                         | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 4.2 | DataLens: быстрый анализ данных с визуализацией                                                                                              | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 4.3 | Power BI. Построение моделей данных из разных неструктурированных источников: таблиц, сайтов и баз данных                                    | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 4.4 | Tableau. Основы работы с расчётными полями, фильтрами, множествами и группировками. Основные виды визуализаций. Лучшие практики визуализации | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 4.5 | Добавление интерактивности и публикация отчетов. Построение дашбордов по предложенным кейсам.                                                | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |

### Форма контроля

В процессе обучения осуществляется текущий контроль знаний и промежуточная аттестация.

Текущий контроль - в форме тестирования, состоящего из десяти вопросов с вариантами ответов, один из которых правильный, или решения типовых задач.

Промежуточная аттестация – выполнение промежуточного тестирования или практического задания по модулю.

### Список литературы

#### Основная литература

1. Зараменских, Е. П. Основы бизнес-информатики: учебник и практикум для бакалавриата и магистратуры / Е. П. Зараменских. – Москва: Издательство Юрайт, 2018. – 407 с. — (Высшее образование). – ISBN 978-5-9916-8210-7. – Текст : электронный // Образовательная платформа Юрайт [сайт]. – URL: <https://urait.ru/bcode/413823> (дата обращения: 10.04.2022).
2. Data Science. Наука о данных с нуля. / Билл Фрэнкс.; пер. с англ. Евстигнеева И.В. – М.: Издательство «Альпина Паблишер». – 2018. – 320 с.
3. Набатова, Д. С. Математические и инструментальные методы поддержки принятия решений: учебник и практикум для вузов / Д. С. Набатова. — Москва: Издательство Юрайт, 2022. — 292 с. — (Высшее образование). — ISBN 978-5-534-02699-3.
4. Курносоев Ю.В. «Азбука аналитики», Издательство «Концептуал», 2018 -240 с.

5. Б. Марр «Ключевые инструменты бизнес-аналитики»/ пер с англ. Егоров В. Н., Издательство «Лаборатория знаний», 2018 – 339 с.
6. де Прадо М. «Машинное обучение: алгоритмы для бизнеса», Санкт Петербург: Издательский дом «Питер», – 2019, – 432 с.
7. Плас вандер Д. «Python для сложных задач: наука о данных и машинное обучение», Санкт Петербург: Издательский дом «Питер», – 2018, – 576с.
8. Лакшманан В., Тайджани Д. «Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении», Санкт Петербург: Издательский дом «Питер», – 2021, – 496с.

#### **Дополнительная литература**

1. Глубокое обучение на Python. – СПб.: Питер, 2018. – 400 с.: ил. – (Серия «Библиотека программиста»).
2. Основы Data Science и Big Data. Python и наука о данных. / Силен Д., Мейсман А., Али М.; пер. с англ. – Санкт Петербург: Издательский дом «Питер», – 2018. – 336 с.
3. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
4. Data Science. Наука о данных с нуля. / Джоэл Грас.; пер. с англ. Логунов А.В. – Санкт Петербург: Издательство «БХВ-Петербург», – 2018. – 336 с.
5. Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении. – СПб.: Питер, 2021. – 496 с.: ил.
6. Д. Битти К. Вигерс «Разработка требований к программному обеспечению», Издательство BHV, 2019 -737 с.

#### **Ресурсы Internet**

1. <https://www.anaconda.com/> – страница загрузки Anaconda.
2. <https://powerbi.microsoft.com/ru-ru/desktop/> – страница загрузки платформы Power BI Desktop компании Microsoft.
3. <https://public.tableau.com/en-us/s/download> – страница загрузки платформы Tableau Public.
4. <https://loginom.ru/download> – страница загрузки платформы Loginom компании BaseGroup Labs.
5. <https://www.knime.com/knime-analytics-platform> – страница загрузки платформы Knime Analytics Platform.
6. <https://community.cloud.databricks.com/login.html> – страница регистрации Databricks Community Edition
7. <https://rapidminer.com/get-started/> – страница загрузки платформы RapidMiner.
8. <http://h2o-release.s3.amazonaws.com/h2o/rel-zermelo/4/index.html> – страница загрузки платформы H2O.
9. <https://cloud.google.com/bigquery/docs/sandbox> – страница BigQuery sandbox.
10. <https://cloud.yandex.ru/services/datalens> – сервис визуализации и анализа данных Яндекса.
11. <https://lookerstudio.google.com/> - страница загрузки Looker Studio
12. <https://www.tableau.com> - страница загрузки Tableau

## Модуль 5. Основные методы машинного обучения.

**Цель модуля:** приобретение слушателями компетенций, необходимых для эффективной работы в области анализа больших данных машинного обучения; изучение инструментов и технологий создания, обучения, оценки и развертывания моделей машинного обучения на Python.

### Планируемые образовательные результаты по модулю

В результате освоения модуля слушатели должны:

*а) знать:*

- основные модели и методы машинного обучения и разработки данных;
- основы линейных моделей, решающих деревьев, композиции моделей (случайный лес, градиентный бустинг и его имплементации), приложения в рекомендательных системах;
- инструментальные средства разработки моделей машинного обучения;

*б) уметь:*

- создавать, обучать и сравнивать модели машинного обучения, основанные на различных алгоритмах, осуществлять их оценку;
- разрабатывать рабочие процессы создания, тестирования и развертывания моделей машинного обучения в десктопных платформах и облачных средах;
- применять фреймворки и платформы науки о данных для создания моделей машинного обучения;

*в) владеть:*

- навыками создания и обучения моделей машинного обучения, используя библиотеки, фреймворки и платформы науки о данных;
- инструментами и подходами в Python, используемые для анализа данных (создание датафрейма из данных, копирование датафрейма, экспорт данных, исследование данных, вывод статистических сведений, извлечение информации, фильтрация, сортировка, агрегирование);
- анализом решений с точки зрения достижения целевых показателей решений;
- навыками применения технологий машинного обучения для решения анализа данных и принятия решений.

### Формируемые компетенции:

способность классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта;

способность разрабатывать и применять методы машинного обучения для решения задач;

способность использовать инструментальные средства для решения задач машинного обучения.



# УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

| №№<br>п/п | Наименование разделов,<br>модулей                                                                               | Всего<br><br>В<br>часа<br>х | В том числе        |        |                    | Самостоятельная<br>работа | Форма<br>контроля                |
|-----------|-----------------------------------------------------------------------------------------------------------------|-----------------------------|--------------------|--------|--------------------|---------------------------|----------------------------------|
|           |                                                                                                                 |                             | Аудиторные занятия |        |                    |                           |                                  |
|           |                                                                                                                 |                             | Всего, часов       | из них |                    |                           |                                  |
|           |                                                                                                                 |                             |                    | Лекции | Практич<br>занятия |                           |                                  |
| 1         | 2                                                                                                               | 3                           | 4                  | 5      | 6                  | 1                         | 2                                |
| М 5.      | Основные методы<br>машинного обучения                                                                           | 25                          | 17                 | 6      | 11                 | 8                         | Зачет                            |
| 5.1.      | Понятие машинного обучения.<br>Регрессионный анализ.<br>Линейная, полиномиальная и<br>логарифмическая регрессия | 5                           | 4                  | 2      | 2                  | 1                         | Решение<br>практических<br>задач |
| 5.2.      | Классификация.<br>Логистическая регрессия и<br>SVM                                                              | 3                           | 2                  |        | 2                  | 1                         | Решение<br>практических<br>задач |
| 5.3.      | Функции потерь и<br>оптимизация<br>Оценка точности модели,<br>переобучение, регуляризация                       | 3                           | 2                  |        | 2                  | 1                         | Решение<br>практических<br>задач |
| 5.4.      | Проблема качества данных.<br>Работа с пропусками<br>Работа с переменными                                        | 5                           | 4                  | 2      | 2                  | 1                         | Решение<br>практических<br>задач |
| 5.5.      | Деревья решений<br>Ансамблирование<br>Поиск выбросов и генерация<br>новых признаков                             | 4                           | 2                  | 2      |                    | 2                         | Тестирование                     |
| 5.6.      | Алгоритмы кластеризации<br>Улучшение качества модели.                                                           | 4                           | 2                  |        | 2                  | 2                         | Решение<br>практических<br>задач |
|           | Промежуточная<br>аттестация по модулю 5                                                                         | 1                           | 1                  |        | 1                  |                           | Зачет                            |

**Тема 5.1. Понятие машинного обучения. Регрессионный анализ. Линейная, полиномиальная и логарифмическая регрессия.**

Основная задача машинного обучения. Приложения на основе машинного обучения. Жизненный цикл машинного обучения. Понятие data science, big data, как это работает и где применяется. Обзор методов машинного обучения в библиотеке scikit-learn. Виды и алгоритмы обучения.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                                                                    | Тема практического занятия          | Содержание практического занятия                                   | Вопросы к практическому занятию                                   |
|----------|---------------------------------------------------------------------------------------------------------------------------|-------------------------------------|--------------------------------------------------------------------|-------------------------------------------------------------------|
| 5.       | Тема 5.1.<br>Понятие машинного обучения.<br>Регрессионный анализ.<br>Линейная, полиномиальная и логарифмическая регрессия | Основные задачи машинного обучения. | Какие задачи машинное обучение решает в бизнесе.<br>Примеры кейсов | Какие библиотеки Python используются в задачах машинного обучения |

#### Тема 5.2. Классификация. Логистическая регрессия и SVM.

Задача классификации: постановка и примеры. Метрики качества в задачах классификации. Оценка качества классификации. Алгоритмы классификации. Измерений качества решений в задачах классификации. Преимущества и недостатки линейных моделей. Требования к линейным моделям. Логическая регрессия. Градиентный спуск. SVM & Kernel trick

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                      | Тема практического занятия        | Содержание практического занятия                                                                        | Вопросы к практическому занятию                                     |
|----------|-------------------------------------------------------------|-----------------------------------|---------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|
| 5.       | Тема 5.2.<br>Классификация<br>Логистическая регрессия и SVM | Задачи классификации и регрессии. | Метрики качества в задачах классификации.<br>Оценка качества классификации.<br>Алгоритмы классификации. | Как оценить качество проведенной классификации в машинном обучении. |

#### Тема 5.3. Функции потерь и оптимизация.

##### Оценка точности модели, переобучение, регуляризация.

Метрики качества модели и переобучение. Обучение и функция потерь. Обучающая и тестовая выборка. Обучающие, тестовые и валидационные множества. Кросс-валидация модели. Метрики качества: accuracy, precision, recall. Оценки моделей, метрики. Оценка качества разных версий модели по AUC. Смещение и разброс (bias-variance tradeoff). Параметры модели. Признаки переобучения и регуляризации.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                                                | Тема практического занятия           | Содержание практического занятия                                                                           | Вопросы к практическому занятию                                     |
|----------|-------------------------------------------------------------------------------------------------------|--------------------------------------|------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|
| 5.       | Тема 5.3.<br>Функции потерь и оптимизация.<br><br>Оценка точности модели, переобучение, регуляризация | Обучение моделей машинного обучения. | Обучение моделей машинного обучения. Разработка и валидацией моделей. Определение метрик и оценка моделей. | Дайте характеристику метрикам качества: accuracy, precision, recall |

#### Тема 5.4. Проблема качества данных. Работа с пропусками. Работа с переменными.

Feature engineering. Анализ и подготовка данных. Потенциальные проблемы. Статистический анализ. Feature selection&engineering. Уменьшение размерности и методы декомпозиции.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                           | Тема практического занятия  | Содержание практического занятия                                                                             | Вопросы к практическому занятию         |
|----------|----------------------------------------------------------------------------------|-----------------------------|--------------------------------------------------------------------------------------------------------------|-----------------------------------------|
| 5.       | Тема 5.4.<br>Проблема качества данных. Работа с пропусками. Работа с переменными | Анализ и подготовка данных. | Проблемы с которыми сталкиваются Аналитики данных при отборе данных для создания моделей машинного обучения. | Методы декомпозиции в машинном обучении |

#### Тема 5.5. Деревья решений. Ансамблирование. Поиск выбросов и генерация новых признаков.

Дерево решений: построение. Обзор sklearn, treeDecisionTreeClassifier. Достоинство и недостатки деревьев решений. Визуализируем принятие решений и предсказания алгоритма. Переобучение при помощи ансамблей. Интерпретировать ML модель, автоматическая оценка важности признаков. Бэггинг и переобучение. Случайный лес. Композиции моделей разных классов.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины | Тема практического занятия | Содержание практического занятия | Вопросы к практическому занятию |
|----------|----------------------------------------|----------------------------|----------------------------------|---------------------------------|
|----------|----------------------------------------|----------------------------|----------------------------------|---------------------------------|

|    |                                                                                                             |                     |                                                         |                                   |
|----|-------------------------------------------------------------------------------------------------------------|---------------------|---------------------------------------------------------|-----------------------------------|
| 5. | Тема 5.5.<br>Деревья<br>решений.<br>Ансамблирование.<br>Поиск выбросов<br>и генерация<br>новых<br>признаков | Деревья<br>решений. | Деревья решений -<br>применение в<br>машинном обучении. | Что такое<br>переобучение модели. |
|----|-------------------------------------------------------------------------------------------------------------|---------------------|---------------------------------------------------------|-----------------------------------|

#### **Тема 5.6. Алгоритмы кластеризации. Улучшение качества модели.**

Кластеризация данных. Задача кластеризации: постановка и примеры. Основные алгоритмы. Метрики качества кластеризации. Алгоритмы для решения задач.

#### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                              | Тема практического занятия | Содержание практического занятия | Вопросы к практическому занятию                         |
|----------|---------------------------------------------------------------------|----------------------------|----------------------------------|---------------------------------------------------------|
| 5.       | Тема 5.5.<br>Алгоритмы кластеризации.<br>Улучшение качества модели. | Кластеризация данных.      | Примеры кластеризации данных     | Какие метрики качества кластеризации данных существуют. |

#### **Содержание самостоятельной работы слушателей**

Основная цель самостоятельной работы слушателей – закрепление знаний, полученных в ходе лекционных и практических занятий.

| № темы   | Наименование (содержание) темы, по которой предусмотрена самостоятельная работа                        | Формы и методы проведения                                                                                                  |
|----------|--------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Тема 5.1 | Понятие машинного обучения. Регрессионный анализ. Линейная, полиномиальная и логарифмическая регрессия | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 5.2 | Классификация. Логистическая регрессия и SVM                                                           | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 5.3 | Функции потерь и оптимизация<br>Оценка точности модели, переобучение, регуляризация                    | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |

|          |                                                                                  |                                                                                                                            |
|----------|----------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Тема 5.4 | Проблема качества данных.<br>Работа с пропусками<br>Работа с переменными         | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 5.5 | Деревья решений<br>Ансамблирование<br>Поиск выбросов и генерация новых признаков | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 5.6 | Алгоритмы кластеризации<br>Улучшение качества модели.                            | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |

### Форма контроля

В процессе обучения осуществляется текущий контроль знаний и промежуточная аттестация.

Текущий контроль - в форме тестирования, состоящего из десяти вопросов с вариантами ответов, один из которых правильный, или решения типовых задач.

Промежуточная аттестация – выполнение промежуточного тестирования или практического задания по модулю.

### Список литературы

#### Основная литература

1. Зараменских, Е. П. Основы бизнес-информатики: учебник и практикум для бакалавриата и магистратуры / Е. П. Зараменских. – Москва: Издательство Юрайт, 2018. – 407 с. — (Высшее образование). – ISBN 978-5-9916-8210-7. – Текст : электронный // Образовательная платформа Юрайт [сайт]. – URL: <https://urait.ru/bcode/413823> (дата обращения: 10.04.2022).
2. Data Science. Наука о данных с нуля. / Билл Фрэнкс.; пер. с англ. Евстигнеева И.В. – М.: Издательство «Альпина Паблишер», – 2018. – 320 с.
3. Набатова, Д. С. Математические и инструментальные методы поддержки принятия решений: учебник и практикум для вузов / Д. С. Набатова. — Москва: Издательство Юрайт, 2022. — 292 с. — (Высшее образование). — ISBN 978-5-534-02699-3.
4. Курносков Ю.В. «Азбука аналитики», Издательство «Концептуал», 2018 -240 с.
5. Б. Марр «Ключевые инструменты бизнес-аналитики»/ пер с англ. Егоров В. Н., Издательство «Лаборатория знаний», 2018 – 339 с.
6. де Прадо М. «Машинное обучение: алгоритмы для бизнеса», Санкт Петербург: Издательский дом «Питер». – 2019. – 432 с.
7. Плас вандер Д. «Python для сложных задач: наука о данных и машинное обучение», Санкт Петербург: Издательский дом «Питер». – 2018. – 576с.
8. Лакшманан В., Тайджани Д. «Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении», Санкт Петербург: Издательский дом «Питер». – 2021. – 496с.

#### Дополнительная литература

1. Глубокое обучение на Python. – СПб.: Питер, 2018. – 400 с.: ил. – (Серия «Библиотека программиста»).
2. Основы Data Science и Big Data. Python и наука о данных. / Силен Д., Мейсман А., Али М.; пер. с англ. – Санкт Петербург: Издательский дом «Питер». – 2018. – 336 с.
3. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
4. Data Science. Наука о данных с нуля. / Джоэл Грас.; пер. с англ. Логунов А.В. – Санкт Петербург: Издательство «БХВ-Петербург». – 2018. – 336 с.



5. Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении. – СПб.: Питер, 2021. – 496 с.: ил.
6. Д. Битти К. Вигерс «Разработка требований к программному обеспечению», Издательство BHV, 2019 -737 с.

### **Ресурсы Internet**

1. <https://www.anaconda.com/> – страница загрузки Anaconda.
2. <https://powerbi.microsoft.com/ru-ru/desktop/> – страница загрузки платформы Power BI Desktop компании Microsoft.
3. <https://public.tableau.com/en-us/s/download> – страница загрузки платформы Tableau Public.
4. <https://loginom.ru/download> – страница загрузки платформы Loginom компании BaseGroup Labs.
5. <https://www.knime.com/knime-analytics-platform> – страница загрузки платформы Knime Analytics Platform.
6. <https://community.cloud.databricks.com/login.html> – страница регистрации Databricks Community Edition
7. <https://rapidminer.com/get-started/> – страница загрузки платформы RapidMiner.
8. <http://h2o-release.s3.amazonaws.com/h2o/rel-zermelo/4/index.html> – страница загрузки платформы H2O.
9. <https://cloud.google.com/bigquery/docs/sandbox> – страница BigQuery sandbox.
10. <https://cloud.yandex.ru/services/datalens> – сервис визуализации и анализа данных Яндексе.
11. <https://lookerstudio.google.com/> - страница загрузки Looker Studio
12. <https://www.tableau.com> - страница загрузки Tableau

## Модуль 6. Глубокое машинное обучение и обработка больших данных.

**Цель модуля:** приобретение слушателями компетенций, необходимых для эффективной работы в области продвинутого анализа и обработки больших данных машинного обучения; изучение инструментов и технологий обработки больших данных, построения рекомендательных систем и обработки естественного языка (NLP).

### Планируемые образовательные результаты по модулю

В результате освоения модуля слушатели должны:

*а) знать:*

- Основные инструменты обработки больших данных, HADOOP & SPARK;
- Элементарные методы анализа временных рядов. Модели ARIMA, GARCH.
- инструментальные средства разработки рекомендательных систем;

*б) уметь:*

- создавать, обучать и сравнивать модели машинного обучения, основанные на алгоритмах, использующих временные ряды, компьютерное зрение, осуществлять их оценку;
- разрабатывать рабочие процессы создания, тестирования и развертывания моделей машинного обучения в десктопных платформах и облачных средах;
- применять фреймворки и платформы науки о данных для создания моделей машинного обучения;

*в) владеть:*

- навыками создания и обучения моделей машинного обучения, используя библиотеки, фреймворки и платформы науки о данных;
- инструментами и подходами автоматической обработки текстов, сегментации и детекции объектов;
- моделями ARIMA, GARCH и прогнозирование значений на их основе
- анализом решений с точки зрения достижения целевых показателей решений;
- навыками применения технологий машинного обучения для решения анализа данных и принятия решений.

### Формируемые компетенции:

способность проведения анализа больших данных с использованием существующей в организации методологической и технологической инфраструктуры;

способность управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации;

способность управлять получением, хранением, передачей, обработкой больших данных;

способность разрабатывать и применять методы машинного обучения для решения задач;

способность использовать инструментальные средства для решения задач машинного обучения;

способность создавать и поддерживать системы искусственного интеллекта на основе нейросетевых моделей и методов;

способность создавать и внедрять одну или несколько сквозных цифровых субтехнологий искусственного интеллекта.

# УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

| №№<br>п/п | Наименование разделов,<br>модулей                                                                                                            | Всего      | В том числе           |        | Самостоятельная<br>работа | Форма<br>контроля |                            |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------|------------|-----------------------|--------|---------------------------|-------------------|----------------------------|
|           |                                                                                                                                              |            | Аудиторные<br>занятия |        |                           |                   |                            |
|           |                                                                                                                                              | В<br>часах | Всего, часов          | из них |                           |                   |                            |
| 1         | 2                                                                                                                                            | 3          | 4                     | 5      | 6                         | 1                 | 2                          |
| М 6.      | Глубокое машинное обучение и обработка больших данных                                                                                        | 24         | 17                    | 4      | 13                        | 10                | Зачет                      |
| 6.1.      | Машинное обучение и Data Science.<br>Основные инструменты обработки больших данных.<br>HADOOP & SPARK.                                       | 5          | 4                     | 2      | 2                         | 1                 | Тестирование               |
| 6.2.      | Введение и классификация рекомендательных систем.<br>Коллаборативная фильтрация                                                              | 5          | 4                     | 2      | 2                         | 1                 | Решение практических задач |
| 6.3.      | Знакомство с временными рядами<br>Элементарные методы анализа временных рядов<br>Модели ARIMA, GARCH                                         | 3          | 2                     |        | 2                         | 1                 | Решение практических задач |
| 6.4.      | Обработка естественного языка (NLP). Введение в автоматическую обработку текста<br>Структура слова. Морфология<br>Тематическое моделирование | 3          | 2                     |        | 2                         | 1                 | Решение практических задач |
| 6.5.      | Компьютерное зрение<br>Введение в компьютерное зрение<br>Выделение признаков и поиск похожих изображений<br>Сегментация и детекция объектов  | 4          | 2                     |        | 2                         | 2                 | Решение практических задач |
| 6.6.      | Подготовка модели<br>Машинное обучения по практическому кейсу                                                                                | 3          | 2                     |        | 2                         | 1                 | Решение практических задач |
|           | Промежуточная аттестация по модулю 6                                                                                                         | 1          | 1                     |        | 1                         |                   | Зачет                      |

### **Тема 6.1. Машинное обучение и Data Science. Основные инструменты обработки больших данных. HADOOP & SPARK.**

Механизмы распределённого хранения больших данных на базе Hadoop и Spark. Основные паттерны реализации их распределённой обработки. Поточковая обработка данных, методы и средства мониторинга и профилирования заданий Spark.

#### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                                                     | Тема практического занятия                    | Содержание практического занятия                                                                                | Вопросы к практическому занятию              |
|----------|------------------------------------------------------------------------------------------------------------|-----------------------------------------------|-----------------------------------------------------------------------------------------------------------------|----------------------------------------------|
| 6.       | Тема 6.1. Машинное обучение и Data Science. Основные инструменты обработки больших данных. HADOOP & SPARK. | Основные инструменты обработки больших данных | Мотивация и технологии работы с большими данными. Основные инструменты обработки больших данных. HADOOP & SPARK | Методы и средства мониторинга заданий Spark. |

### **Тема 6.2. Введение и классификация рекомендательных систем. Коллаборативная фильтрация.**

Введение в рекомендательные системы. Неперсонализированные рекомендации. Персонализированные рекомендации. Развитие рекомендательных систем. Коллаборативная фильтрация. Гибридные рекомендательные системы. Рекомендации на основе содержания. Рекомендации на основе скрытых факторов

#### **Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                                 | Тема практического занятия | Содержание практического занятия                                                   | Вопросы к практическому занятию                                 |
|----------|----------------------------------------------------------------------------------------|----------------------------|------------------------------------------------------------------------------------|-----------------------------------------------------------------|
| 6.       | Тема 6.2. Введение и классификация рекомендательных систем. Коллаборативная фильтрация | Рекомендательные системы   | Виды рекомендательных систем. Разработка и требования к рекомендательным системам. | Факторы, которые могут лежать в основе рекомендательных систем. |

### **Тема 6.3. Знакомство с временными рядами. Элементарные методы анализа временных рядов. Модели ARIMA, GARCH**

Элементарные алгоритмы обработки временных рядов. Элементарные методы анализа временных рядов. Модели ARIMA и GARCH. Прогнозирование значений на их основе. Модели авторегрессии условной гетероскедастичности. Сингулярный спектральный анализ.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                                                                      | Тема практического занятия | Содержание практического занятия                               | Вопросы к практическому занятию |
|----------|-----------------------------------------------------------------------------------------------------------------------------|----------------------------|----------------------------------------------------------------|---------------------------------|
| 6.       | Тема 6.3.<br>Знакомство с временными рядами.<br><br>Элементарные методы анализа временных рядов.<br><br>Модели ARIMA, GARCH | Обработка временных рядов. | Модели ARIMA и GARCH.<br>Прогнозирование значений на их основе | Методы прогнозирования цен      |

### Тема 6.4. Обработка естественного языка (NLP). Введение в автоматическую обработку текста. Структура слова. Морфология. Тематическое моделирование.

Введение в обработку текста. Морфологический и синтаксический анализ, дистрибутивная семантика и информационный поиск. Классификация в AOT. Снижение размерности в векторной модели, классификация, извлечение информации. Генерация текстов. Трансформеры. Токенизация: BPE, BERT. Словари. Подкрепление знаний. Языковые модели. Языковые модели. Sequence to sequence-алгоритмы. Информационный поиск. Синтаксический анализ. Обзор существующих библиотек, их использование и доработка. Использование внешних ресурсов. Грязные тексты: что это такое и как с ними работать. Дистрибутивная семантика. Чатботы: разбор генерации текстов. Нейросети для NLP.

### Содержание практических занятий

| № модуля | Наименование темы (раздела) дисциплины                                                                                                                                 | Тема практического занятия          | Содержание практического занятия                                                          | Вопросы к практическому занятию               |
|----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|-------------------------------------------------------------------------------------------|-----------------------------------------------|
| 6.       | Тема 6.4.<br>Обработка естественного языка (NLP).<br>Введение в автоматическую обработку текста.<br>Структура слова.<br>Морфология.<br><br>Тематическое моделирование. | Обработка естественного языка (NLP) | Морфологический и синтаксический анализ, дистрибутивная семантика и информационный поиск. | Разработка чатботов для отчетности и анализа. |



**Тема 6.5. Компьютерное зрение. Выделение признаков и поиск похожих изображений. Сегментация и детекция объектов.**

Введение в компьютерное зрение. Техники машинного зрения. Базовая теория. Обзор кейсов применения. Разбор реальных задач: рукописный ввод, детекция и сегментация объектов на изображениях. Задачи детекции и сегментации. Свёрточные нейронные сети. Обучение свёрточной сети на практике. Рекуррентные нейронные сети в задачах компьютерного зрения: Image Captioning. Выделение признаков и поиск похожих изображений. Порождающие модели.

**Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                                                                 | Тема практического занятия | Содержание практического занятия | Вопросы к практическому занятию      |
|----------|------------------------------------------------------------------------------------------------------------------------|----------------------------|----------------------------------|--------------------------------------|
| 6.       | Тема 6.5.<br>Компьютерное зрение<br>Выделение признаков и поиск похожих изображений<br>Сегментация и детекция объектов | Компьютерное зрение        | Техники машинного зрения         | Обучение свёрточной сети на практике |

**Тема 6.6. Подготовка модели. Машинное обучение по практическому кейсу.**

Разработка и внедрение собственного ML-решения/проекта либо разработка предложенного нами кейса.

**Содержание практических занятий**

| № модуля | Наименование темы (раздела) дисциплины                                     | Тема практического занятия             | Содержание практического занятия  | Вопросы к практическому занятию              |
|----------|----------------------------------------------------------------------------|----------------------------------------|-----------------------------------|----------------------------------------------|
| 6.       | Тема 6.6.<br>Подготовка модели<br>Машинное обучение по практическому кейсу | Построение ML-модели для решения задач | Разработка и внедрение ML-решения | ML-модель для решения профессиональных задач |

## Содержание самостоятельной работы слушателей

Основная цель самостоятельной работы слушателей – закрепление знаний, полученных в ходе лекционных и практических занятий.

| № темы   | Наименование (содержание) темы, по которой предусмотрена самостоятельная работа                                                              | Формы и методы проведения                                                                                                  |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Тема 6.1 | Машинное обучение и Data Science. Основные инструменты обработки больших данных. HADOOP & SPARK.                                             | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 6.2 | Введение и классификация рекомендательных систем. Коллаборативная фильтрация                                                                 | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 6.3 | Знакомство с временными рядами<br>Элементарные методы анализа временных рядов<br>Модели ARIMA, GARCH                                         | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 6.4 | Обработка естественного языка (NLP). Введение в автоматическую обработку текста<br>Структура слова. Морфология<br>Тематическое моделирование | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 6.5 | Компьютерное зрение<br>Выделение признаков и поиск похожих изображений<br>Сегментация и детекция объектов                                    | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |
| Тема 6.6 | Подготовка модели. Машинное обучение по практическому кейсу.                                                                                 | Изучение видеолекций, основной и дополнительной литературы по программе, решение практических задач с применением ЭО и ДОТ |

## Форма контроля

В процессе обучения осуществляется текущий контроль знаний и промежуточная аттестация.

Текущий контроль - в форме тестирования, состоящего из десяти вопросов с вариантами ответов, один из которых правильный, или решения типовых задач.

Промежуточная аттестация – выполнение промежуточного тестирования или практического задания по модулю.

## Список литературы

### Основная литература

1. Зараменских, Е. П. Основы бизнес-информатики: учебник и практикум для бакалавриата и магистратуры / Е. П. Зараменских. – Москва: Издательство Юрайт, 2018. – 407 с. — (Высшее образование). – ISBN 978-5-9916-8210-7. – Текст : электронный // Образовательная платформа Юрайт [сайт]. – URL: <https://urait.ru/bcode/413823> (дата обращения: 10.04.2022).

2. Data Science. Наука о данных с нуля. / Билл Фрэнкс.; пер. с англ. Евстигнеева И.В. – М.: Издательство «Альпина Паблишер». – 2018. – 320 с.
3. Набатова, Д. С. Математические и инструментальные методы поддержки принятия решений: учебник и практикум для вузов / Д. С. Набатова. — Москва: Издательство Юрайт, 2022. — 292 с. — (Высшее образование). — ISBN 978-5-534-02699-3.
4. Курносоев Ю.В. «Азбука аналитики», Издательство «Концептуал», 2018 -240 с.
5. Б. Марр «Ключевые инструменты бизнес-аналитики»/ пер с англ. Егоров В. Н., Издательство «Лаборатория знаний», 2018 – 339 с.
6. де Прадо М. «Машинное обучение: алгоритмы для бизнеса», Санкт Петербург: Издательский дом «Питер». – 2019. – 432 с.
7. Плас вандер Д. «Python для сложных задач: наука о данных и машинное обучение», Санкт Петербург: Издательский дом «Питер». – 2018. – 576с.
8. Лакшманан В., Тайджани Д. «Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении», Санкт Петербург: Издательский дом «Питер». – 2021. – 496с.

#### **Дополнительная литература**

1. Глубокое обучение на Python. – СПб.: Питер, 2018. – 400 с.; ил. – (Серия «Библиотека программиста»).
2. Основы Data Science и Big Data. Python и наука о данных. / Силен Д., Мейсман А., Али М.; пер. с англ. – Санкт Петербург: Издательский дом «Питер». – 2018. – 336 с.
3. Паклин Н.Б. Орешков В.И. Бизнес-аналитика: от данных к знаниям, СПб: Питер 2013. – 706 с.
4. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
5. Data Science. Наука о данных с нуля. / Джозл Грас.; пер. с англ. Логунов А.В. – Санкт Петербург: Издательство «БХВ-Петербург». – 2018. – 336 с.
6. Google BigQuery. Всё о хранилищах данных, аналитике и машинном обучении. – СПб.: Питер, 2021. – 496 с.; ил.
7. Д. Битти К. Вигерс «Разработка требований к программному обеспечению», Издательство BHV, 2019 -737 с.

#### **Ресурсы Internet**

1. <https://www.anaconda.com/> – страница загрузки Anaconda.
2. <https://powerbi.microsoft.com/ru-ru/desktop/> – страница загрузки платформы Power BI Desktop компании Microsoft.
3. <https://public.tableau.com/en-us/s/download> – страница загрузки платформы Tableau Public.
4. <https://loginom.ru/download> – страница загрузки платформы Loginom компании BaseGroup Labs.
5. <https://www.knime.com/knime-analytics-platform> – страница загрузки платформы Knime Analytics Platform.
6. <https://community.cloud.databricks.com/login.html> – страница регистрации Databricks Community Edition
7. <https://rapidminer.com/get-started/> – страница загрузки платформы RapidMiner.
8. <http://h2o-release.s3.amazonaws.com/h2o/rel-zermelo/4/index.html> – страница загрузки платформы H2O.
9. <https://cloud.google.com/bigquery/docs/sandbox> – страница BigQuery sandbox.
10. <https://cloud.yandex.ru/services/datalens> – сервис визуализации и анализа данных Яндекс.
11. <https://lookerstudio.google.com/> - страница загрузки Looker Studio
12. <https://www.tableau.com> - страница загрузки Tableau

### 1. Применяемые образовательные технологии, формы и методы обучения, в том числе интерактивные.

Образовательная программа рассчитана на 144 академических часов обучения и включает перечень учебных тем, тематику, виды занятий, предназначенных для повышения уровня знаний, совершенствования практических умений и навыков, необходимых для решения задач, предусмотренных программой обучения.<sup>2</sup>

Образовательная деятельность слушателей предусматривает следующие виды учебных занятий:

- лекции;
- практические занятия;
- контрольно-проверочные занятия.<sup>3</sup>

В образовательном процессе применяются разнообразные формы работы со слушателями, к которым относятся:

- лекционный курс - систематизирование основ теоретических знаний и передового практического опыта;
- практические занятия - анализ нормативных правовых актов и практических ситуаций, решение кейсов и практико-ориентированных заданий;
- кейс-стадии – изучение конкретных ситуаций из практики (case-study). Для выполнения данного вида заданий обучающимся должна быть представлена в письменной форме информация относительно реальной ситуации (профессиональной или жизненной) и поставлены конкретные задачи её изучения проблемы, обучающиеся анализируют различные аспекты проблемы и предлагают выработанные решения;
- тестирование – это стандартизированный метод оценки знаний, умений, навыков обучающихся. В программе применяется электронное тестирование на базе системы дистанционного обучения (СДО) Финансового университета.

### 2. Перечень информационных технологий, используемых при осуществлении образовательного процесса по курсу.

Программа реализуется по очно-заочной форме обучения с использованием дистанционных образовательных технологий и электронного обучения посредством информационно-телекоммуникационных сетей при опосредованном взаимодействии обучающихся и педагогических работников.

#### Электронное обучение с использованием дистанционных образовательных технологий

Обучение слушателей проводится в СДО Финуниверситета с использованием сети Интернет с рабочих мест или личных компьютеров слушателей.

Все слушатели программы, до начала обучения получают на электронную почту информационные письма с логином и паролем для доступа к личному кабинету на СДО и к расписанию проведения учебных занятий.

#### Электронный учебный контент по программе

В рамках обучения по программе для слушателей на платформе СДО будут размещены учебные материалы, состоящие из:

- учебной программы дополнительной программы повышения квалификации;
- расписания занятий по программе для слушателей;
- учебно-методических материалов, необходимых для освоения образовательной программы;

<sup>2</sup> предусматривается возможность проведения занятий в режиме вебинаров

<sup>3</sup> промежуточная и итоговая аттестации проводятся через СДО



- вопросов для самопроверки;
- контрольных заданий;
- презентационных материалов;
- видеолекций (вебинаров) по разбору практических задания и ответов на сложные вопросы.

В период обучения всем слушателем будет доступна техническая и информационная поддержка по электронной почте, программ обмена сообщениями для смартфонов и других устройств и в виде «горячей линии» по телефону, которая включает в себя:

- консультации слушателей по содержательной части курса и работе;
- взаимодействие слушателя с ответственными лицами для ответов на обращения слушателей.

**В рамках сопровождения процесса обучения Слушателей Финиуниверситет:**

- предоставляет логины и пароли для регистрации обучаемых в СДО;
- администрирует СДО;
- создает встроенную систему взаимодействия Слушателей в режиме реального времени посредством чатов и в асинхронном режиме посредством электронной почты и форумов;
- получает и анализирует результаты обучения Слушателей.

**Контактная форма взаимодействия со слушателями:**

- Лекции реализуются в формате Online вебинаров с возможностью контакта преподавателя со слушателями, вопросов и ответов в Online режиме, в дальнейшем записи проведенных вебинаров выкладываются в СДО и/или на иные платформы видеохостинга, что позволяет слушателям лучше усвоить материал.
- Практические занятия проводятся в формате Online вебинаров с возможностью контакта преподавателя со слушателями, вопросов и ответов в Online режиме.

**Условия для функционирования электронной информационно-образовательной среды**

| Электронные информационные ресурсы                           | Вид занятий                                                                                                                                                                                        | Наименование оборудования, Программного обеспечения                     |
|--------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| Система дистанционного обучения, система видеоконференцсвязи | <ul style="list-style-type: none"> <li>- Все виды контактной работы</li> <li>- Выполнения практического задания: практические задания по модулям курса</li> <li>- Итоговое тестирование</li> </ul> | Компьютер, подключенный к сети Интернет; интернет-браузер; Adobe Reader |

**3. Описание материально-технической базы, необходимой для осуществления образовательного процесса.**

| Наименование специализированных учебных помещений | Вид занятий          | Наименование оборудования, программного обеспечения                                                                           |
|---------------------------------------------------|----------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Учебный интерактивный тренажерный класс           | Практические занятия | Мультимедийное оборудование, компьютеры, МФУ. Компьютер, подключенный к сети Интернет, интернет-браузер, Adobe Reader, и т.д. |

Материально-технические условия соответствуют действующим санитарным и противопожарным правилам и нормам.



При проведении учебных занятий с применением дистанционных образовательных технологий (ДОТ) у слушателя должен быть персональный компьютер, оснащенный средствами ввода и вывода звуковой информации, с доступом в сеть интернет.

#### **4. Кадровое обеспечение образовательного процесса.**

В реализации программы принимают участие профессорско-преподавательский состав Финансового университета, а также иные квалифицированные практики и эксперты.<sup>4</sup>

#### **5. Выдаваемый документ.**

Удостоверение о повышении квалификации Финансового университета при Правительстве Российской Федерации.

### **Описание системы оценки качества освоения программы**

Контроль результатов освоения программы повышения квалификации осуществляется в ходе текущего контроля успеваемости, промежуточной аттестации и итоговой аттестации.

**1. Текущий контроль успеваемости** осуществляется в процессе изучения слушателями учебного материала в форме выполнения практических заданий, разбора практических ситуаций или тестирования по каждой теме в личном кабинете слушателя.

#### **Возможные виды текущего контроля:**

- Контроль системой СДО количества, изученных слушателем видеолекций, текстовых материалов.
- Контроль системой СДО выполнения практических заданий по темам.
- Тестирование – слушателям предлагается в формате самостоятельной работы ответить на вопросы.
- Посещение и активность на аудиторных занятиях.

**2. Промежуточная аттестация** проводится в форме зачета в виде выполнения практического или тестового задания с выставлением оценки в СДО – «зачтено»; «незачтено».

**Порядок выполнения практического задания:** практические задания по модулям курса выполняются слушателями на личных компьютерах, в требуемом программном обеспечении и выкладываются слушателями в личном кабинете СДО в виде ссылок.

**Критерии оценивания:** критерии выставления оценки за выполнение практического задания приведены в таблице:

| <b>Критерии к выставлению оценки</b>                                                                                                                  | <b>Оценка за выполнение практического задания по модулям</b> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|
| оценка выставляется при наличии серьезных ошибок и пробелов в знаниях при выполнении практического задания.                                           | незачтено                                                    |
| оценка выставляется при выполнении практического задания в соответствующем программном обеспечении в полном соответствии со всеми пунктами задания; а | зачтено                                                      |

<sup>4</sup> преподавательский состав, участвующий в обучении, имеет опыт преподавания по темам программы и опыт практической работы по тематике обучения

|                                                                            |  |
|----------------------------------------------------------------------------|--|
| также работы имеющие некоторые неточности в выполнении и описании задания. |  |
|----------------------------------------------------------------------------|--|

**Порядок проведения тестирования:** тестирование проводится с личного компьютера слушателя, 20 тестовых вопросов по отдельным модулям, 60 мин., количество попыток – 2 по каждому модулю.

Критерии выставления оценки за промежуточное тестирование приведены в таблице:

| Количество правильных ответов при тестировании | Оценка за промежуточное тестирование по дисциплине |
|------------------------------------------------|----------------------------------------------------|
| ≥50%                                           | зачтено                                            |
| <50%                                           | незачтено                                          |

### 3. Итоговая аттестация

После успешного освоения всех модулей программы и успешного прохождения промежуточной аттестации, для слушателей, завершающих обучение обязательной, является итоговая аттестация.

**Проведение итоговой аттестации.** Итоговая аттестация проводится в форме защиты итоговой практической работы, выполнения практического задания и отправки на проверку записи видеозащиты.

Практическая работа состоит из выбора индивидуального кейса (набор данных) и выполнении заданий, которые охватывают все практические методы, подходы в соответствующих программных продуктах, применяемые в машинном обучении и анализе больших данных рассмотренные и изученные в соответствующих темах в рамках учебной программы.

В результирующую оценку по практической работе входит оценка уровня сформированности у слушателя универсальных и профессиональных компетенций, а также оценки собственно результата/продукта, полученного в ходе выполнения работы.

Порядок проведения итоговой аттестации: письменный ответ на практическую работу и запись защиты работы с выставлением оценки по 4 балльной шкале: "неудовлетворительно"; "удовлетворительно", "хорошо", "отлично".

**Критерии оценивания:** для выставления оценки по итоговой аттестации необходимо пользоваться следующими критериями, приведенными в таблице.

| Критерии оценки итоговой аттестационной работы.                                                                                                                                                                                                                                        | Критерии к выставлению оценки / Оценка за итоговую аттестационную работу по программе |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| оценка выставляется при наличии серьезных ошибок и пробелов в знаниях в практической работе                                                                                                                                                                                            | неудовлетворительно                                                                   |
| оценка выставляется при наличии отдельных неточностей в ответах, при неполных ответах на задания, частичном выполнении заданий или при наличии замечаний к практической работе не принципиального характера (опеки, случайные ошибки арифметического характера, грамматические ошибки) | удовлетворительно                                                                     |
| оценка выставляется при наличии верных ответов на все задания в ходе выполнения практической работы, при грамотном выполнении всех заданий, но при отсутствии отличительных признаков, как, например: детальных                                                                        | хорошо                                                                                |

|                                                                                                                                                                                                                                                                                                                     |         |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| выкладок или пояснений, качественного оформления работы                                                                                                                                                                                                                                                             |         |
| оценка выставляется при четком достижении цели и выполнении всех заданий практической работы, то есть при наличии полных (с детальными пояснениями выкладок и выводов), оригинальных и правильных решений, а также при полных и развернутых ответах на комментарии преподавателя и итоговой аттестационной комиссии | отлично |

По окончании обучения каждый слушатель заполняет Анкету оценки обучения.

## Оценочные материалы

### Примеры вопросов для текущего контроля знаний

#### **1. Модуль 1. Инструменты и методы бизнес-аналитики**

- 1.1. Как создать Гугл –таблицу
- 1.2. Как сделать настраиваемый фильтр в Гугл –таблице
- 1.3. Что такое бизнес-аналитика
- 1.4. Как удалить и восстановить таблицу
- 1.5. Что является наиболее важным аспектом аналитической отчетности
- 1.6. Сколько Google-таблицы поддерживает типов данных
- 1.7. С какого ключевого слова начинается определение функции
- 1.8. Каковы преимущества использования Google-таблицы
- 1.9. Если функция должна возвращать результат, какая используется инструкция
- 1.10. Рабочий лист Google-таблицы это
- 1.11. Какие типы данных поддерживаются в Google-таблицы
- 1.12. От чего зависит результат импорта CSV файла
- 1.13. Как получить доступ к значениям в словаре
- 1.14. С помощью какой функции можно подсчитать количество строк
- 1.15. К параметрам сводных таблиц не относятся
- 1.16. Как обрабатывать входные данные в Google-таблицы
- 1.17. В чем разница между тем, когда функция range()

#### **2. Модуль 2. Анализ данных с применением Python**

- 2.1. Сколько Python поддерживает типов данных
- 2.2. С какого ключевого слова начинается определение функции
- 2.3. Каковы преимущества использования Python
- 2.4. Если функция должна возвращать результат, какая используется инструкция
- 2.5. Что такое локальные и глобальные переменные в Python
- 2.6. Какие типы данных поддерживаются в Python
- 2.7. Что такое словарь в Python
- 2.8. Как получить доступ к значениям в словаре
- 2.9. Что такое абстракции у словаря и списка в Python
- 2.10. Как создать пустой класс в Python
- 2.11. Как обрабатывать входные данные в Python
- 2.12. В чем разница между тем, когда функция range()

#### **3. Модуль 3. Применение SQL для работы с данными**

- 3.1. Как выглядит запрос, для вывода BCEX значений из таблицы Orders
- 3.2. Какие данные мы получим из этого запроса: `select id, date, customer_name from Orders;`

- 3.3. Есть ли ошибка в запросе: `select id, date, customer_name from Orders where customer_name = Mike;`
- 3.4. Что покажет следующий запрос: `select * from Orders where date between '2017-01-01' and '2017-12-31'`
- 3.5. Что неверно в запросе: `select id, date from Orders where seller_id = NULL;`
- 3.6. Порядок выполнения операторов AND и OR
- 3.7. Что покажет следующий запрос: `select DISTINCT seller_id order by seller_id from Orders;`
- 3.8. Выберите корректный пример использования функции CONCAT
- 3.9. Что покажет следующий запрос: `select concat('index', " ", 'city') AS delivery_address from Orders;`
- 3.10. Что покажет следующий запрос: `select id from Orders where year (date) > 2018;`
- 3.11. Для чего используется LIMIT: `select * from Orders limit 10;`
- 3.12. Выберите пример правильно составленного запроса с использованием агрегирующей функции SUM
- 3.13. Выберите корректно составленный запрос с функцией GROUP BY
- 3.14. Что покажет следующий запрос: `select seller_id, count(*) from Orders GROUP BY seller_id HAVING seller_id IN (2,4,6).`

#### **Модуль 4. Визуализация данных**

- 4.1. Общий поток работы с Power BI - ...
- 4.2. Стандартные блоки Power BI ...
- 4.3. Коллекция готовых визуальных элементов, заранее упорядоченных в панелях мониторинга и отчетах в Power BI - это ...
- 4.4. Расширенные возможности аналитики с помощью отчетов и визуализаций обеспечивает ...
- 4.5. Задача аналитика данных имеет важное влияние на производительность при составлении отчетов и анализе данных ...
- 4.6. Для извлечения дней недели из набора дат используется ...
- 4.7. Для создания визуального элемента нужно ...
- 4.8. Чтобы импортировать табличные данные с веб-сайта в Power BI нужно ...
- 4.9. Для обнаружения связи в данных после их импорта в Power BI Desktop ...
- 4.10. Чтобы в Power BI распределить данные дат по годам, кварталам, месяцам и дням нужно ...
- 4.11. Чтобы создать в Power BI промежуточные вычисления и данные, которые требуется хранить как часть модели, а не запроса ...
- 4.12. Данная функций DAX позволяет использовать неактивную связь при вычислении выражения меры ...
- 4.13. Данный тип диаграммы является наиболее подходящим для сравнения нескольких значений, например, продаж по регионам - ...

#### **Модуль 5. Основные методы машинного обучения**

- 5.1. Задача кластеризации заключается в ...
- 5.2. Задача регрессии сводится к ...
- 5.3. Линейная, полиномиальная и логарифмическая регрессия отличаются ...
- 5.4. SVM это ...
- 5.5. Классификация помогает решить задачи ....
- 5.6. Проблема переобучения в машинном обучении заключается ....
- 5.7. Увеличение какого из этих гиперпараметров может привести к переобучению случайного леса ...
- 5.8. Правильная последовательность этапов Knowledge Discovery in Databases – процесса обнаружения знаний в базах данных ...
- 5.9. Большинство методов Data Mining были разработаны в рамках ...
- 5.10. Машинное обучение – это ...



- 5.11 Целью поиска ассоциативных правил является ...
- 5.12 К технологии потоковой обработки событий в режиме реального времени относится ...
- 5.13 В этом этапе процесса CRISP-DM происходит проверка гипотез – это ...
- 5.14 Первую модель искусственного нейрона разработал ...
- 5.15 Регрессионные модели описывают ...
- 5.16 Для распределенного глубокого машинного обучения (Deep Learning) больше подходит фреймворк ...
- 5.17 Алгоритм k-средних предназначен для решения задачи ...
- 5.18 Ошибка обучения – ...
- 5.19 В кластерном анализе для определения близости между кластерами используются метрики ...
- 5.20 Задачи типичные для машинного обучения с учителем:

## **Модуль 6. Глубокое машинное обучение и обработка больших данных**

- 6.1. Hadoop – это ...
- 6.2. Выявление клиентов, которые, вероятно, перестанут пользоваться услугами компании в ближайшую неделю, относится к задаче ...
- 6.3. Алгоритмы, используемые для решения задачи, объекты в которой нужно разделить на классы:
- 6.4. Рекомендательные системы подразделяются на следующие виды
- 6.5. Классификации рекомендательных систем
- 6.6. Обработка естественного языка это
- 6.7. Методы анализа временных рядов
- 6.8. Исследование и обнаружение машиной (алгоритмами, средствами искусственного интеллекта) в сырых данных скрытых знаний, которые: ранее не были известны, нетривиальны, практически полезны, доступны для интерпретации человеком, называется
- 6.9. Обработка естественного языка это -
- 6.10. Морфологический и синтаксический анализ позволяют
- 6.11. Компьютерное зрение это
- 6.12. Задачами детекции и сегментации являются

### **Практико-ориентированных задания и кейсы по модулям.**

В состав практико-ориентированных заданий и кейсов включены задачи от организаций, работающих в области искусственного интеллекта и в смежных областях и/или организаций, применяющих технологии интеллектуальной обработки данных для решения своих производственных задач.

## **Модуль 1. Инструменты и методы бизнес-аналитики**

### **1. Введение в Google-таблицы, сводные таблицы Excel**

По индивидуальному заданию (датасет) провести быструю аналитика в таблицах Google:

- найдите закономерности в данных с помощью автоматически формируемых диаграмм;
- создайте сводные таблиц на основе своих данных;
- визуализируйте подмножества данных, выбирая только нужные столбцы или ячейки.

### **2. Создание отчетов в Google Data Studio**

По индивидуальному заданию (датасет) создать интегративный отчет (дашборд):

- создать новый пустой отчет, добавить источник данных в отчет, добавить таблицу (с помощью панели инструментов), добавить график временных рядов (с помощью меню), настроить стиль отчета, добавить баннер и добавить заголовок к отчету;
- добавить диаграмму и настроить различные диаграммы, гистограммы и карты;
- настроить режим просмотра отчета и поделиться им.

### **3. Применение pandas для анализа и обработки данных для пользователей Excel**



Особенности языка Python для аналитики, работа с окружением Python и основными требуемыми инструментами, и библиотеками; работу с базовыми концепциями, понятиями, принципами и возможностями Python.

## Модуль 2. Анализ данных с применением Python

### 1. Библиотеки визуализации данных Matplotlib, Seaborn, Altair, Plotly Express

Основные структуры данных в Pandas классы Series и DataFrame. Импорт и чтение данных, обобщенная информация и краткая статистика, изменение типов и сортировка данных, индексация и извлечение данных, сводные таблицы, группирование данных, преобразование датафреймов, визуализации.

### 2. Исследовательский анализ данных (EDA) с использованием pandas

Использование библиотеки Matplotlib для создания графиков, гистограмм, спектров мощности, круговых диаграмм и др. Особенности библиотеки Seaborn для визуализации данных. Использование декларативной библиотеки Altair для создания эффективных визуализаций с минимальным количеством кода.

По индивидуальному заданию (датасет) выполнить исследовательский анализ данных (EDA):

- изучить распределения данных в виде статистик и визуализаций;
- обработать отсутствующие значения набора данных (наиболее частая проблема с каждым набором данных);
- найти и обработать выбросы и аномалии;
- найти и удалить повторяющиеся данные (устранить дубликаты);
- провести кодирование категориальных переменных;
- выполнить нормализацию и масштабирование числовых признаков.

## Модуль 3. Применение SQL для работы с данными

### 1. Большие данные

Извлечение записей с помощью SELECT. Создание псевдонимов столбцов с помощью AS. Фильтрация с WHERE. SELECT \*, EXCEPT, REPLACE. Агрегирование с GROUP BY. Подсчет записей с COUNT. Поиск уникальных значений с DISTINCT. Основы соединения таблиц.

### 2. Облачные технологии обработки больших данных

Типы данных, функции и операторы SQL в Google BigQuery:

Числовые типы и функции. Математические функции. Стандартное вещественное деление. Сравнение. Условные выражения. Строковые функции. Парсинг и форматирование отметок времени. Арифметические операции с временными метками. Функции для работы с географическими Координатами.

### 3. Доступ к облачному сервису BigQuery из сервиса Count.co и Colab.

Доступ к наборам данных BigQuery для их изучения. Визуализация данных BigQuery в записной книжке Jupyter, преимущества совместного использования Google BigQuery и Colab для решения задач по науке о данных. Использование BigQuery на Kaggle.

### 4. Машинное обучение на SQL в BigQuery

Машинное обучение в BQ ML. Ограничений BQ ML на виды моделей: Линейная регрессия (LINEAR\_REG). Логистическая регрессия (LOGISTIC\_REG). KMEANS. TENSORFLOW, решение задач Data Minig.

### 5. Платформа Databricks Lakehouse.

Загрузка данных с помощью Databricks – импорт данных из файла. Доступ к импортированным данным. Создание Spark Dataframe, преобразование в Frame pandas. Подготовка и очистка данных, исследовательский анализ данных (EDA), корреляционный анализ, создание визуализаций. Расширенный исследовательский анализ данных: создание простой базовой модели, OneHotEncoder и масштабирование функций, уменьшение размерности и изучение важности признаков. Создание озера данных и управление им с помощью Delta Lake: импорт необработанных данных и сохранение их в таблице Delta Lake (Bronze), сохранение подготовленных и очищенных данных в Silver table in Delta Lake, создание для предоставления чистых и надежных данных для конкретного бизнес-подразделения или варианта использования – Gold table. Машинное обучение в платформе

Databricks: библиотека Apache Spark (Mllib), Mlflow, использование сторонних фреймворков с открытым кодом.

#### **Модуль 4. Визуализация данных**

##### **1. Платформа Yandex Data lens**

Использование веб интерфейса Yandex Data lens для анализа данных и визуализации.

##### **2.. Платформа Tableau**

Используя технологии Tableau создать модели исследования данных, визуализацию и дашборд.

##### **3. Платформа Power Bi.**

Использование веб интерфейса Power Bi для бизнес-анализа, сегментации клиентов и проведение количественных исследований.

#### **Модуль 5. Основные методы машинного обучения**

1. Машинное обучение для решения задач Data Mining. Градиентный спуск в машинном обучении

Выполнение линейной регрессии с помощью NumPy. Выполнение линейной регрессии с помощью scikit-learn, визуализация результатов с использованием библиотеки Seaborn и их анализ.

##### **2. Фреймворки машинного обучения**

Сравнение результатов машинного обучения (метрики качества моделей) для задач классификации и регрессии в популярных фреймворках: LightGBM, XGBoost, CatBoost. Особенности работы.

##### **3.. Кластерный анализ, алгоритм k-means**

Решение задачи бинарной классификации на датасете Titanic, используя алгоритм кластерного анализа k-means. Выбор признаков, создание и обучение модели. Оценка качества модели, сравнение с результатами алгоритма Random Forest и Gradient Boosting for classification библиотеке scikit-learn.

##### **4. Поиск ассоциативных правил**

Использование алгоритма Apriori для выявления часто встречающихся наборов элементов из набора данных - решение задачи анализа рыночной корзины на датасете с Kaggle. Реализации алгоритма Apriori в различных фреймворках машинного обучения.

##### **5. Алгоритмы построения ансамблей деревьев решений. Бэггинг, бустинг, стекинг**

Использование алгоритма Random Forest для решения задачи бинарной классификации на датасете Titanic: импорт и чтение данных, обобщенная информация и краткая статистика, изменение типов данных и заполнение пропущенных значений. Выбор признаков, создание и обучение модели. Оценка качества модели. Настройка гиперпараметров. Предсказание цены недвижимости на датасете Boston, используя алгоритм градиентного бустинга библиотеки scikit-learn, сравнение результатов с алгоритмом Random Forest. Оценка качества моделей. Настройка гиперпараметров.

##### **6. Фреймворки машинного обучения**

Сравнение результатов машинного обучения (метрики качества моделей) для задач классификации и регрессии в популярных фреймворках: LightGBM, XGBoost, CatBoost. Особенности работы.

#### **Модуль 6. Глубокое машинное обучение и обработка больших данных**

##### **1. Введение в нейронные сети, глубокие нейронные сети**

Решение задач классификации и регрессии с помощью многослойного перцептрона. Сравнение с результатами алгоритма Random Forest и Gradient Boosting for classification библиотеке scikit-learn. Решение задач классификации изображений, обнаружения объектов, анализ настроений с помощью глубоких нейронных сетей реализуемых во фреймворках AutoGluon, Auto Keras.

## 2. Анализ временных рядов

Применение библиотек XGBoost, PyCaret, Prophet и глубоких нейронных сетей для прогнозирования временных рядов на датасете с Kaggle.

## 3. Автоматическое машинное обучение (AutoML)

Применение технологий автоматического машинного обучения во фреймворках Auto Sklearn, Tree-Based Pipeline Optimization Tool (TPOT), Auto Keras. для решения задач классификации и регрессии на структурированных данных.

### Практические задания к итоговой аттестации

Для проведения итоговой аттестации включены практические задания, включающие вопросы практико-ориентированные задач от организаций, работающих в области искусственного интеллекта и в смежных областях и/или организаций, применяющих технологии интеллектуальной обработки данных для решения своих производственных задач.

Количество практико-ориентированных задач, включенных в итоговую аттестацию, рассчитывается исходя из количества слушателей, обучающихся на одном потоке обучения. Одну практико-ориентированную задачу решает не более 10 человек.

#### 1. Выберите модель машинного обучения, для создания модели, предсказывающей, какие пассажиры выжили после кораблекрушения "Титаника".

**Компетенция:** Способен классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта.

##### Индикаторы достижения компетенции

- Классифицирует и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей
- Выбирает методы и инструментальные средства искусственного интеллекта для решения задач в зависимости от особенностей проблемной области

#### 2. Разработайте классификацию и создайте на его основании нейросеть для распознавания объектов или предметов.

**Компетенция:** Способен разрабатывать и применять методы машинного обучения для решения задач.

##### Индикаторы достижения компетенции

- Проводит анализ требований и определяет необходимые классы задач машинного обучения
- Принимает участие в оценке и выборе используемых методов машинного обучения

#### 3. Построить модель LDA предсказывающую вероятность наступления события.

**Компетенция:** Способен использовать инструментальные средства для решения задач машинного обучения.

##### Индикаторы достижения компетенции

- Осуществляет оценку и выбор инструментальных средств для решения поставленной задачи
- Разрабатывает модели машинного обучения для решения задач
- Создает, поддерживает и использует системы искусственного интеллекта, включающие разработанные модели и методы с применением выбранных инструментов машинного обучения



**4. Построить персонализированные и неперсонализированные рекомендательные системы для сервиса внутреннего туризма**

**Компетенция:** Способен создавать и поддерживать системы искусственного интеллекта на основе нейросетевых моделей и методов.

**Индикаторы достижения компетенции**

- Осуществляет оценку и выбор моделей искусственных нейронных сетей и инструментальных средств для решения поставленной задачи
- Разрабатывает системы искусственного интеллекта на основе моделей искусственных нейронных сетей и инструментальных средств

**5. Создайте OLAP куб, настройте атрибуты и иерархии, добавьте измерения в куб. Создайте отчет на основании данных.**

**Компетенция:** Способен разрабатывать системы анализа больших данных.

**Индикаторы достижения компетенции**

- Разрабатывает программные компоненты извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных
- Разрабатывает программные компоненты обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и достоверностью больших данных

**6. Создайте генератор картинок или текстов на основе предложенного датасета. Обучить нейросеть для генерации контента.**

**Компетенция:** Способен создавать и внедрять одну или несколько сквозных цифровых субтехнологий искусственного интеллекта.

**Индикаторы достижения компетенции:**

- Участвует в реализации проектов в области сквозной цифровой субтехнологии «Компьютерное зрение»
- Участвует в реализации проектов в области сквозной цифровой субтехнологии «Обработка естественного языка»
- Участвует в реализации проектов в области сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений»
- Участвует в реализации проектов в области сквозной цифровой субтехнологии «Распознавание и синтез речи»
- Участвует в реализации проектов в области перспективных сквозных цифровых субтехнологий искусственного интеллекта

Обсуждено и одобрено на заседании Научно-методического совета Института развития профессиональных компетенций и квалификаций, протокол № 17 от 17 мая 2023 г.

Директор ИРПКК



Т.А. Болтенко