

Документы
Информация о владельце: "Самарский государственный экономический университет"
ФИО: Кандрашина Елена Александровна
Должность: И.о. ректора ФГАОУ ВО «Самарский государственный экономический университет»
Дата подписания: 04.08.2026 16:19:30
Уникальный программный ключ:
2db64eb9605ce27edd3b8e8fdd32c70e0674ddd2

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ) «ИНСТРУМЕНТЫ АНАЛИЗА ДАННЫХ И МАШИННОГО ОБУЧЕНИЯ»

Уровень высшего образования: магистратура

Направление подготовки: 09.04.03 Прикладная информатика

Направленность (профиль) подготовки: Искусственный интеллект и большие данные

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Год набора (приема на обучение): 2026

Срок получения образования: 2 года

Объем:
в зачетных единицах: 4 з.е.
в академических часах: 144 ак.ч.

г. Самара, 2026

Разработчики:

Не имеет Колотилина М. А.

Рабочая программа дисциплины (модуля) составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.04.03 Прикладная информатика, утвержденного приказом Минобрнауки от 19.09.2017 № 916, с учетом трудовых функций профессиональных стандартов: "Руководитель проектов в области информационных технологий", утвержден приказом Минтруда России от 27.04.2023 № 369н.

Согласование и утверждение

№	Подразделение или коллегиальный орган	Ответственное лицо	ФИО	Виза	Дата, протокол (при наличии)
1	Кафедра прикладной информатики	Заведующий кафедрой, руководитель подразделения, реализующего ОП	Ермолина Л. В.	Рассмотрено	21.05.2026, № 9

1. Цель и задачи освоения дисциплины (модуля)

Цель освоения дисциплины - формирование у студентов системных знаний о современных инструментах анализа данных и машинного обучения, а также практических навыков их применения для решения профессиональных задач в области обработки и интерпретации данных с использованием современных информационных технологий.

Задачи изучения дисциплины:

- Изучить теоретические основы анализа данных, методы статистической обработки, алгоритмы машинного обучения (классификация, регрессия, кластеризация, снижение размерности) и современные библиотеки для их реализации.;
- Сформировать практические навыки работы с инструментами анализа данных и ML-библиотеками (Python с pandas, NumPy, scikit-learn, TensorFlow, PyTorch, а также BI-инструментами для визуализации) для решения прикладных задач обработки и прогнозирования данных.;
- Развить компетенции по постановке задач анализа данных, выбору и настройке моделей машинного обучения, оценке их качества и интерпретации результатов для принятия обоснованных решений в профессиональной деятельности..

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

Компетенции, индикаторы и результаты обучения

УК-6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки

УК-6.1 Определяет инструменты и методы управления временем при реализации приоритетов собственной деятельности для выполнения конкретных задач, проектов, достижении поставленных целей

Знать:

УК-6.1/Зн1 Современные инструменты и методы анализа данных и машинного обучения (Python с библиотеками pandas, NumPy, scikit-learn, TensorFlow, PyTorch; BI-инструменты для визуализации; методы статистической обработки, алгоритмы классификации, регрессии, кластеризации), их функциональные возможности, области применения, а также подходы к организации собственной деятельности для эффективного освоения и применения данных инструментов при решении профессиональных задач.

Уметь:

УК-6.1/Ум1 Определять приоритеты собственной деятельности при изучении и применении инструментов анализа и машинного обучения, выбирать оптимальные методы и инструменты для решения конкретных задач анализа данных, планировать время на освоение новых технологий и библиотек, оценивать собственный уровень компетенций и выстраивать траекторию профессионального развития в области Data Science и машинного обучения.

Владеть:

УК-6.1/Нв1 Навыками самоорганизации и управления временем при реализации проектов по анализу данных и построению ML-моделей, методами самооценки эффективности применения инструментов анализа, навыками планирования этапов работы над задачами машинного обучения (от постановки задачи и сбора данных до построения, оценки и интерпретации моделей) с учетом собственных приоритетов и сроков выполнения.

УК-6.2 Анализирует и использует инструменты непрерывного образования с учетом требований рынка труда и способы ее совершенствования на основе самооценки

Знать:

УК-6.2/Зн1 Современные инструменты непрерывного образования в области анализа данных и машинного обучения (онлайн-платформы: Coursera, Stepik, OpenDataScience; профессиональные сообщества: Kaggle, GitHub; актуальные требования рынка труда к специалистам в области Data Science, ML-инженерам, аналитикам данных; перечень востребованных навыков и технологий), а также методы самооценки профессиональных компетенций и способы их совершенствования.

Уметь:

УК-6.2/Ум1 Анализировать требования рынка труда к специалистам в области анализа данных и машинного обучения, выявлять дефицит собственных знаний и навыков, выбирать и использовать инструменты непрерывного образования для повышения квалификации, адаптировать траекторию профессионального развития под изменяющиеся требования рынка и технологии в сфере Data Science.

Владеть:

УК-6.2/Нв1 Навыками самооценки профессиональных компетенций в области анализа данных и ML, методами планирования непрерывного образования с использованием современных онлайн-платформ и профессиональных сообществ, навыками мониторинга актуальных трендов в сфере Data Science и адаптации собственной профессиональной деятельности к требованиям рынка труда.

ПК-2 Способен планировать и использовать методы научных исследований и инструментарий в области управления проектами в сфере ИТ малого и среднего уровня сложности

ПК-2.1 Изучает и планирует использование методов научных исследований и инструментария в области управления проектами

Знать:

ПК-2.1/Зн1 Современные методы научных исследований в области управления проектами в сфере ИТ, формальные инструменты управления рисками и проблемами проекта (матрица рисков, реестр проблем, анализ «что-если», методы Монте-Карло), методологии управления проектами (Agile, Scrum, Kanban, Waterfall), инструменты планирования конфигурационного управления (системы контроля версий Git, репозитории артефактов, CI/CD-инструменты), а также методы сбора и анализа данных для обоснования управленческих решений в условиях неопределенности.

Уметь:

ПК-2.1/Ум1 Изучать и адаптировать методы научных исследований для решения задач управления ИТ-проектами малого и среднего уровня сложности, планировать использование инструментария управления проектами (Jira, Trello, YouTrack, MS Project) с учетом специфики проекта, применять формальные инструменты управления рисками для идентификации, оценки и минимизации угроз проекту, использовать методы анализа данных для прогнозирования отклонений и принятия решений в условиях неопределенности.

Владеть:

ПК-2.1/Нв1 Навыками планирования конфигурационного управления в ИТ-проектах (настройка репозитория, управление версиями, контроль изменений), навыками применения методов научных исследований для обоснования проектных решений, навыками работы с инструментами управления рисками и проблемами, методами оценки эффективности проектных решений с использованием инструментов анализа данных и машинного обучения для прогнозирования и оптимизации проектных процессов.

ПК-2.2 Использует методы научных исследований и инструментарий в области управления проектами в сфере ИТ малого и среднего уровня сложности

Знать:

ПК-2.2/Зн1 Методы научных исследований и инструментарий управления ИТ-проектами малого и среднего уровня (Agile, Scrum, Jira, Trello, Git, CI/CD), методы анализа данных и ML для прогнозирования рисков и оптимизации решений;

Уметь:

ПК-2.2/Ум1 Применять методы исследований для управления проектами (анализ рисков, планирование ресурсов), использовать инструменты управления проектами и анализа данных (Python, pandas, scikit-learn) для обработки метрик и прогнозирования, применять системы конфигурационного управления для контроля версий и изменений;

Владеть:

ПК-2.2/Нв1 Навыками применения методов исследований в управлении проектами, работы с инструментарием управления проектами, использования анализа данных и ML для оптимизации процессов, методами конфигурационного управления и контроля изменений в условиях неопределенности.

3. Место дисциплины в структуре ОП

Дисциплина (модуль) «Инструменты анализа данных и машинного обучения» относится к формируемой участниками образовательных отношений части образовательной программы и изучается в семестре(ах): 2.

В процессе изучения дисциплины студент готовится к решению типов задач профессиональной деятельности, предусмотренных ФГОС ВО и образовательной программой.

Компетенция	Предшествующие дисциплины	Последующие дисциплины
ПК-2 - Способен планировать и использовать методы научных исследований и инструментарий в области управления проектами в сфере ИТ малого и среднего уровня сложности		
ПК-2.1 Изучает и планирует использование методов научных исследований и инструментария в области управления проектами	Процесс, стадии и методологии разработки решений на основе искусственного интеллекта	Выполнение и защита выпускной квалификационной работы, Производственная практика: (проектно-технологическая) практика, Производственная практика: преддипломная, Процесс, стадии и методологии разработки решений на основе искусственного интеллекта, Управление проектами создания интеллектуальных информационных систем

ПК-2.2 Использует методы научных исследований и инструментов в области управления проектами в сфере ИТ малого и среднего уровня сложности	Процесс, стадии и методологии разработки решений на основе искусственного интеллекта	Выполнение и защита выпускной квалификационной работы, Производственная практика: (проектно-технологическая) практика, Производственная практика: преддипломная, Процесс, стадии и методологии разработки решений на основе искусственного интеллекта, Управление проектами создания интеллектуальных информационных систем
УК-6 - Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки		
УК-6.1 Определяет инструменты и методы управления временем при реализации приоритетов собственной деятельности для выполнения конкретных задач, проектов, достижении поставленных целей	Методы машинного обучения, Производственная практика: научно-исследовательская работа, Процесс, стадии и методологии разработки решений на основе искусственного интеллекта, Учебная практика: технологическая (проектно-технологическая) практика	Выполнение и защита выпускной квалификационной работы, Производственная практика: (проектно-технологическая) практика, Производственная практика: научно-исследовательская работа, Процесс, стадии и методологии разработки решений на основе искусственного интеллекта
УК-6.2 Анализирует и использует инструменты непрерывного образования с учетом требований рынка труда и способы ее совершенствования на основе самооценки	Производственная практика: научно-исследовательская работа, Процесс, стадии и методологии разработки решений на основе искусственного интеллекта, Учебная практика: технологическая (проектно-технологическая) практика	Выполнение и защита выпускной квалификационной работы, Массово параллельные вычисления для ускорения машинного обучения, Производственная практика: (проектно-технологическая) практика, Производственная практика: научно-исследовательская работа, Процесс, стадии и методологии разработки решений на основе искусственного интеллекта

4. Объем дисциплины (модуля) и виды учебной работы

Период обучения	Общая трудоемкость (часы)	Общая трудоемкость (ЗЕТ)	Контактная работа (часы, всего)	Лекционные занятия (часы)	Практические занятия (часы)	Индивидуальная контактная работа (часы)	Самостоятельная работа (часы)	Промежуточная аттестация
Второй семестр	144	4	12	4	8	0,15	113,85	Зачет
Всего	144	4	12	4	8	0,15	113,85	18

5. Содержание дисциплины (модуля)

5.1. Разделы, темы дисциплины и виды занятий (часы промежуточной аттестации не указываются)

Наименование раздела, темы	Всего	Лекционные занятия	Практические занятия	Самостоятельная работа
Раздел 1. Инструменты и алгоритмы анализа данных и машинного обучения	126	4	8	113,85
Тема 1.1. Введение в анализ данных и машинное обучение Понятие анализа данных, цели и задачи. Место машинного обучения в системе наук о данных (Data Science). Основные понятия: данные, признаки, объекты, выборка, обучающая и тестовая выборки. Классификация задач машинного обучения: обучение с учителем (supervised learning), обучение без учителя (unsupervised learning), обучение с подкреплением (reinforcement learning). Типы данных: числовые, категориальные, текстовые, временные ряды. Обзор современных инструментов и библиотек для анализа данных и ML (Python, R, SAS, KNIME, RapidMiner). Роль анализа данных и машинного обучения в управлении проектами в сфере ИТ.	10			10

Тема 1.2. Методы статистического анализа данных Описательная статистика: меры центральной тенденции (среднее, медиана, мода), меры разброса (дисперсия, стандартное отклонение, размах, интерквартильный размах). Визуализация данных: гистограммы, ящики с усами (boxplot), диаграммы рассеяния (scatter plot), тепловые карты. Проверка статистических гипотез: t-тест, критерий хи-квадрат, ANOVA. Корреляционный анализ: коэффициент корреляции Пирсона и Спирмена. Анализ распределений: нормальное распределение, Z-оценка. Применение статистических методов для анализа проектных метрик и данных в сфере ИТ.	10			10
Тема 1.3. Предобработка и подготовка данных для машинного обучения Этапы подготовки данных: сбор, очистка, трансформация, интеграция. Обработка пропущенных значений (удаление, заполнение средним/медианой/модой, интерполяция). Обработка выбросов (z-score, IQR, визуальные методы). Кодирование категориальных признаков (Label Encoding, One-Hot Encoding, Target Encoding). Масштабирование и нормализация данных (StandardScaler, MinMaxScaler, RobustScaler). Разделение данных на обучающую, валидационную и тестовую выборки (train_test_split, кросс-валидация). Способы борьбы с дисбалансом классов (undersampling, oversampling, SMOTE). Инструменты предобработки данных в Python: pandas, scikit-learn.	12		2	10

Тема 1.4. Методы снижения размерности и визуализации данных Проклятие размерности (curse of dimensionality). Методы отбора признаков: фильтрация (корреляция, хи-квадрат), оберточные методы (RFE, жадный алгоритм), встроенные методы (LASSO, деревья решений). Методы извлечения признаков: главные компоненты (PCA — Principal Component Analysis), t-SNE (t-Distributed Stochastic Neighbor Embedding), UMAP (Uniform Manifold Approximation and Projection). Применение методов снижения размерности для визуализации многомерных данных и оптимизации моделей машинного обучения. Использование библиотек scikit-learn, umap-learn.	12		2	10
Тема 1.5. Классические алгоритмы машинного обучения. Регрессия Задача регрессии: прогнозирование непрерывного целевого показателя. Линейная регрессия: метод наименьших квадратов, оценка коэффициентов. Оценка качества регрессионных моделей: MAE, MSE, RMSE, R ² . Регуляризация: гребневая регрессия (Ridge), регрессия Лассо (Lasso), эластичная сеть (ElasticNet). Логистическая регрессия для задач бинарной классификации. Применение регрессионных моделей для прогнозирования сроков, бюджета и ресурсов ИТ-проектов. Библиотеки: scikit-learn, statsmodels.	12	2		10

Тема 1.6. Классификация и деревья решений Задача классификации. Метрики качества классификации: матрица ошибок (confusion matrix), точность (accuracy), полнота (recall), точность (precision), F1-мера, AUC-ROC. Деревья решений: принципы построения, критерии разделения (Gini impurity, Information Gain, энтропия). Проблема переобучения (overfitting), методы борьбы: ограничение глубины, минимальное число выборок в листе, стрижка (pruning). Интерпретируемость деревьев решений. Применение для классификации проектных рисков и проблем в ИТ-проектах. Библиотека: scikit-learn, Graphviz для визуализации.	10			10
Тема 1.7. Ансамблевые методы машинного обучения Понятие ансамблевых методов. Бэггинг (Bagging): случайный лес (Random Forest) — принципы построения, оценка важности признаков, применение для задач классификации и регрессии. Стекинг (Stacking). Бустинг (Boosting): AdaBoost, Gradient Boosting (GBM), XGBoost, LightGBM, CatBoost — сравнительный анализ, преимущества, настройка гиперпараметров. Применение ансамблевых методов для решения задач управления ИТ-проектами (прогнозирование успешности проектов, классификация рисков, оценка трудозатрат). Библиотеки: scikit-learn, XGBoost, LightGBM, CatBoost.	10			10

Тема 1.8. Кластеризация и методы обучения без учителя Задачи кластеризации: поиск групп однородных объектов без целевой переменной. Алгоритм K-Means: принцип работы, выбор количества кластеров (метод локтя, индекс силуэта). Иерархическая кластеризация: агломеративная и дивизимная, дендрограммы. Методы на основе плотности: DBSCAN, OPTICS. Методы снижения размерности для кластеризации. Оценка качества кластеризации: коэффициент силуэта, индекс Дэвиса-Болдина. Применение методов кластеризации для сегментации проектных задач, группировки рисков и управления портфелем проектов. Библиотеки: scikit-learn, scipy.	10			10
Тема 1.9. Нейронные сети и глубокое обучение Архитектура искусственной нейронной сети. Персептрон, многослойный персептрон (MLP). Функции активации: sigmoid, tanh, ReLU, Softmax. Алгоритм обратного распространения ошибки (backpropagation). Глубокое обучение: сверточные нейронные сети (CNN) для обработки изображений, рекуррентные нейронные сети (RNN, LSTM, GRU) для последовательных данных и временных рядов. Применение нейронных сетей в управлении проектами: прогнозирование, анализ текстовых данных, обработка временных рядов проектных метрик. Библиотеки: TensorFlow, Keras, PyTorch.	12		2	10

Тема 1.10. Инструменты анализа данных и машинного обучения в Python Экосистема Python для анализа данных и машинного обучения. Библиотека NumPy для работы с массивами и математических вычислений. Библиотека Pandas для обработки и анализа табличных данных (DataFrame, Series, группировки, слияния, сводные таблицы). Библиотека Matplotlib и Seaborn для визуализации данных: построение графиков, диаграмм, тепловых карт. Инструменты для работы с Jupyter Notebook и JupyterLab. Применение инструментов для анализа проектных метрик, построения дашбордов, визуализации результатов ML-моделей.	14	2	2	10
Тема 1.11. Применение машинного обучения в управлении ИТ-проектами Практические сценарии применения ML в управлении проектами: прогнозирование сроков завершения проектов на основе исторических данных, классификация проектов по степени риска, автоматическое распределение ресурсов, выявление аномалий в проектных метриках, прогнозирование бюджетных отклонений. Построение моделей для раннего обнаружения проблемных проектов и рисков сбоев. Использование методов анализа данных для оптимизации процессов управления проектами. Инструменты и платформы для внедрения ML-решений в управление проектами. Оценка эффективности и интерпретация результатов моделей для принятия управленческих решений.	14			13,85

5.2. Контрольные мероприятия по дисциплине

Вид контроля	Форма контроля/Оценочное средство
Текущий контроль	Тестовое задание
Промежуточная аттестация	Зачет

№ п/п	Наименование раздела	Вид контроля/ используемые оценочные материалы	
		Текущий	Промежут. аттестация
1	Инструменты и алгоритмы анализа данных и машинного обучения	Тестовое задание	Зачет

6. Оценочные материалы текущего контроля

1. Инструменты и алгоритмы анализа данных и машинного обучения Тестовое задание

№ п/п	Содержание вопроса		Компетенция
		Правильный ответ (ключ ответа)	
1	Расположите в правильной последовательности этапы использования методов научных исследований и инструментария анализа данных для решения задачи предсказательной аналитики, чтобы повысить эффективность процесса А) Импорт необходимых библиотек (Pandas, Scikit-learn, Matplotlib) Б) Загрузка данных из файла с помощью Pandas В) Очистка данных от пропусков и выбросов Г) Построение модели машинного обучения и оценка качества Ответ: А → Б → В → Г		ПК-2
2	Руководитель ИТ-проекта малого уровня сложности планирует использовать методы научных исследований и инструментарий анализа данных для оптимизации управления проектами. В рамках планирования применения инструментария для обработки агрегированных данных по срокам, бюджету и ресурсам завершенных проектов, какую современную технологию хранения целесообразно применить для повышения эффективности выполнения аналитических запросов с агрегацией по нескольким полям при проведении научных исследований в области управления проектами? 1) Реляционную СУБД с транзакционной поддержкой (OLTP) 2) Колоночную СУБД (OLAP-систему) 3) Графовую базу данных 4) Файловую систему для хранения JSON-документов Ответ: 2		ПК-2
3	В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) перед вами поставлена профессиональная задача: обеспечить обработку большого набора данных (объемом 10 ГБ) с целью его очистки. Используя методы научных исследований и инструментарий в области управления проектами, обоснуйте выбор программного средства из экосистемы Python, которое позволит реализовать распределенные вычисления для повышения скорости обработки данных. 1) NumPy 2) Matplotlib 3) Dask 4) Scikit-learn Ответ: 3		ПК-2

4	Используя методы научных исследований и современный инструментарий в области управления проектами, обоснуйте выбор BI-системы для реализации данной задачи. При выборе учитывайте ограничения проекта малого уровня сложности (бюджет, сроки, доступные ресурсы) и необходимость применения формальных инструментов управления рисками и проблемами проекта при внедрении дашборда 1) Apache Airflow 2) Tableau 3) PostgreSQL 4) Jupyter Notebook	ПК-2								
Ответ: 2										
5	Используя методы научных исследований и современный инструментарий в области управления проектами, обоснуйте выбор алгоритма машинного обучения из библиотеки Scikit-learn для реализации данной задачи. При выборе учитывайте ограничения проекта малого уровня сложности (доступные вычислительные ресурсы, сроки, требуемая интерпретируемость модели для заказчика) и необходимость применения формальных инструментов управления рисками (например, риски переобучения, риски неверной интерпретации результатов) 1) KMeans 2) DecisionTreeClassifier 3) LinearRegression 4) PCA	ПК-2								
Ответ: 2										
6	В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) перед вами поставлена профессиональная задача: обеспечить полноту аналитической выборки при объединении двух таблиц базы данных. 1) INNER JOIN 2) RIGHT JOIN 3) LEFT JOIN 4) CROSS JOIN	ПК-2								
Ответ: 3										
7	В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) поставлена профессиональная задача – очистка большого набора данных (10 ГБ) с использованием распределённых вычислений. Соотнесите инструментарий обработки больших данных с его характеристиками, которые необходимо учитывать при выборе решения, используя методы научных исследований и инструментарий управления проектами. <table><tr><th>Инструмент:</th><th>Характеристика для выбора:</th></tr><tr><td>1. Apache Spark</td><td>А. Работает в оперативной памяти, обеспечивает высокую скорость, но требует значительных вычислительных ресурсов, что может повлиять на бюджет проекта малого уровня</td></tr><tr><td>2. Dask</td><td>Б. Интегрируется с экосистемой Python, масштабируется от одного узла до кластера, позволяет гибко управлять ресурсами в условиях ограничений проекта</td></tr><tr><td>3. Pandas (с векторизацией)</td><td>В. Не поддерживает распределённые вычисления, при объёме 10 ГБ может привести к ошибкам памяти, что создаёт риск срыва сроков проекта</td></tr></table>	Инструмент:	Характеристика для выбора:	1. Apache Spark	А. Работает в оперативной памяти, обеспечивает высокую скорость, но требует значительных вычислительных ресурсов, что может повлиять на бюджет проекта малого уровня	2. Dask	Б. Интегрируется с экосистемой Python, масштабируется от одного узла до кластера, позволяет гибко управлять ресурсами в условиях ограничений проекта	3. Pandas (с векторизацией)	В. Не поддерживает распределённые вычисления, при объёме 10 ГБ может привести к ошибкам памяти, что создаёт риск срыва сроков проекта	ПК-2
Инструмент:	Характеристика для выбора:									
1. Apache Spark	А. Работает в оперативной памяти, обеспечивает высокую скорость, но требует значительных вычислительных ресурсов, что может повлиять на бюджет проекта малого уровня									
2. Dask	Б. Интегрируется с экосистемой Python, масштабируется от одного узла до кластера, позволяет гибко управлять ресурсами в условиях ограничений проекта									
3. Pandas (с векторизацией)	В. Не поддерживает распределённые вычисления, при объёме 10 ГБ может привести к ошибкам памяти, что создаёт риск срыва сроков проекта									
Ответ: 1–А, 2–Б, 3–В										
8	В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) перед вами профессиональная задача – выбор BI-системы для построения интерактивного дашборда в реальном времени. Используя методы научных исследований и инструментарий управления проектами, соотнесите критерии выбора BI-инструмента с их влиянием на успех проекта. <table><tr><th>Критерий выбора:</th><th>Влияние на управление проектом:</th></tr><tr><td>1. Возможность подключения к различным источникам данных (SQL, NoSQL, API)</td><td>А. Снижает риск неполноты аналитики, но требует дополнительного времени на настройку, что должно быть учтено при планировании сроков</td></tr><tr><td>2. Наличие встроенных средств визуализации и фильтрации в реальном времени</td><td>Б. Увеличивает удобство работы для заказчика, сокращает время на разработку и повышает эффективность контроля KPI в рамках ограниченного бюджета</td></tr><tr><td>3. Стоимость лицензии и требования к серверному оборудованию</td><td>В. Критический фактор для проектов малого уровня, влияет на экономическую эффективность и распределение ресурсов, требует управления рисками превышения бюджета</td></tr></table>	Критерий выбора:	Влияние на управление проектом:	1. Возможность подключения к различным источникам данных (SQL, NoSQL, API)	А. Снижает риск неполноты аналитики, но требует дополнительного времени на настройку, что должно быть учтено при планировании сроков	2. Наличие встроенных средств визуализации и фильтрации в реальном времени	Б. Увеличивает удобство работы для заказчика, сокращает время на разработку и повышает эффективность контроля KPI в рамках ограниченного бюджета	3. Стоимость лицензии и требования к серверному оборудованию	В. Критический фактор для проектов малого уровня, влияет на экономическую эффективность и распределение ресурсов, требует управления рисками превышения бюджета	ПК-2
Критерий выбора:	Влияние на управление проектом:									
1. Возможность подключения к различным источникам данных (SQL, NoSQL, API)	А. Снижает риск неполноты аналитики, но требует дополнительного времени на настройку, что должно быть учтено при планировании сроков									
2. Наличие встроенных средств визуализации и фильтрации в реальном времени	Б. Увеличивает удобство работы для заказчика, сокращает время на разработку и повышает эффективность контроля KPI в рамках ограниченного бюджета									
3. Стоимость лицензии и требования к серверному оборудованию	В. Критический фактор для проектов малого уровня, влияет на экономическую эффективность и распределение ресурсов, требует управления рисками превышения бюджета									
Ответ: 1–А, 2–Б, 3–В										

9	В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) решается профессиональная задача – классификация клиентов по уровню лояльности с помощью интерпретируемой модели машинного обучения. Используя методы научных исследований (сравнительный анализ, эксперимент, проверка гипотез), соотнесите алгоритмы Scikit-learn с их применимостью с учётом ограничений проекта и управления рисками. <table><tr><th>Алгоритм</th><th>Применимость в проекте</th></tr><tr><td>1-Decision Tree</td><td>А. Высокая интерпретируемость, визуализация правил принятия решений, но риск переобучения требует применения методов регуляризации (управление рисками)</td></tr><tr><td>2-Logistic Regression</td><td>Б. Простая линейная модель с коэффициентами, интерпретируемость высокая, но ограничена при нелинейных зависимостях, что может снизить качество классификации</td></tr><tr><td>3-Random Forest</td><td>В. Компромисс между интерпретируемостью (важность признаков) и точностью, но требует больше ресурсов и времени обучения, что надо учитывать при планировании сроков проекта</td></tr></table> Ответ: 1–А, 2–Б, 3–В	Алгоритм	Применимость в проекте	1-Decision Tree	А. Высокая интерпретируемость, визуализация правил принятия решений, но риск переобучения требует применения методов регуляризации (управление рисками)	2-Logistic Regression	Б. Простая линейная модель с коэффициентами, интерпретируемость высокая, но ограничена при нелинейных зависимостях, что может снизить качество классификации	3-Random Forest	В. Компромисс между интерпретируемостью (важность признаков) и точностью, но требует больше ресурсов и времени обучения, что надо учитывать при планировании сроков проекта	ПК-2
Алгоритм	Применимость в проекте									
1-Decision Tree	А. Высокая интерпретируемость, визуализация правил принятия решений, но риск переобучения требует применения методов регуляризации (управление рисками)									
2-Logistic Regression	Б. Простая линейная модель с коэффициентами, интерпретируемость высокая, но ограничена при нелинейных зависимостях, что может снизить качество классификации									
3-Random Forest	В. Компромисс между интерпретируемостью (важность признаков) и точностью, но требует больше ресурсов и времени обучения, что надо учитывать при планировании сроков проекта									
10	В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) поставлена профессиональная задача – внедрение SQL-оператора для объединения двух таблиц без потери строк из левой таблицы (LEFT JOIN). Расположите в правильной последовательности шаги научно-исследовательского подхода к выбору оптимального решения с использованием инструментария управления проектами и управлением рисками: 1 Проверка гипотезы о производительности различных типов JOIN на тестовой выборке 2 Сравнительный анализ вариантов (LEFT JOIN, RIGHT JOIN, FULL JOIN) по критериям полноты и скорости 3 Выполнение эксперимента на реальном объёме данных для подтверждения корректности 4 Формулировка цели и ограничений проекта (бюджет времени, объём данных, требования к полноте выборки) 5 Выбор LEFT JOIN как оптимального для обеспечения полноты аналитической выборки 6 Документирование решения и план действий при возникновении ошибок (риск-менеджмент) Ответ: 4 → 2 → 1 → 3 → 5 → 6	ПК-2								
11	Дайте развернутый письменный ответ (3-5 предложений). В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) перед вами профессиональная задача – очистка большого набора данных (10 ГБ). Используя методы научных исследований (эксперимент, сравнительный анализ), предложите инструментарий обработки данных из экосистемы Python и обоснуйте его выбор с учётом ограничений проекта (бюджет, время, доступное оборудование) и управления рисками (например, риск нехватки памяти, риск падения производительности). Ответ: Apache Spark или Dask (за счёт распределённых вычислений); выбор обосновывается сравнением по критериям скорости, масштабируемости и стоимости ресурсов; риск нехватки памяти снижается за счёт использования кластера или облачных решений.	ПК-2								
12	Дайте развернутый письменный ответ (3-5 предложений). В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) решается профессиональная задача – выбор BI-системы для интерактивного дашборда в реальном времени. Используя инструментарий управления проектами (например, SWOT-анализ, диаграмму Ганта), опишите, какие методы научных исследований вы примените для обоснования выбора и как вы будете управлять рисками при внедрении (например, риск несовместимости с источниками данных, риск низкой скорости обновления). Ответ: SWOT-анализ для сравнения сильных и слабых сторон BI-систем; эксперимент с прототипом; риск несовместимости снижается проверкой коннекторов до выбора; риск скорости – тестированием на реальных данных; план управления рисками включается в диаграмму Ганта.	ПК-2								
13	Дайте развернутый письменный ответ (3-5 предложений). В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) поставлена профессиональная задача – классификация клиентов по уровню лояльности с использованием интерпретируемой модели машинного обучения. Применяя методы научных исследований (сравнительный анализ, проверка гипотез), обоснуйте выбор алгоритма из Scikit-learn с учётом ограничений проекта (интерпретируемость, время обучения) и опишите, как вы будете управлять риском переобучения модели. Ответ: Decision Tree или Logistic Regression – высокая интерпретируемость; риск переобучения управляется через кросс-валидацию, регуляризацию, ограничение глубины дерева; выбор подтверждается сравнительным анализом метрик.	ПК-2								

14	Дайте развернутый письменный ответ (3-5 предложений) В рамках управления проектом в сфере ИТ среднего уровня сложности (руководитель проекта) решается профессиональная задача – объединение двух таблиц с помощью SQL-оператора для обеспечения полноты аналитической выборки. Используя инструментарий управления проектами (прототипирование, резервирование времени, peer-review), опишите, как вы примените методы научных исследований для выбора между LEFT JOIN, RIGHT JOIN и FULL JOIN и как вы учтёте ограничения проекта (объём данных, время выполнения запроса).		ПК-2			
	Ответ:	сравнительный анализ производительности на тестовой выборке; эксперимент с замерами времени выполнения; прототипирование на малом объёме данных; резервирование времени на доработку; peer-review для проверки корректности; выбор LEFT JOIN для сохранения всех строк левой таблицы.				
15	Дайте развернутый письменный ответ (3-5 предложений) В рамках управления проектом в сфере ИТ малого уровня сложности (руководитель проекта) перед вами профессиональная задача – комплексный выбор инструментов для анализа больших данных (обработка, визуализация, моделирование). Используя инструментарий управления проектами (матрица принятия решений, риск-реестр) и методы научных исследований, опишите ваш подход к выбору стека технологий с учётом ограничений проекта (бюджет, компетенции команды, сроки) и управления рисками (технические, кадровые, организационные).		ПК-2			
	Ответ:	матрица решений по критериям (стоимость, скорость, сложность освоения); сравнительный анализ технологий (Spark vs Dask, Tableau vs Power BI, Scikit-learn vs XGBoost); риск-реестр включает риски нехватки компетенций (решение – обучение или аутсорсинг), риски несовместимости (решение – прототипирование), риски превышения сроков (решение – резервирование времени).				
16	a) Microsoft Excel b) Apache Spark c) Adobe Photoshop d) CorelDRAW a) Microsoft Excel b) Apache Spark c) Adobe Photoshop d) CorelDRAW		УК-6			
	Ответ:	b				
17	Для совершенствования профессиональных навыков в области визуализации данных и самооценки уровня владения BI-системами наиболее эффективным инструментом непрерывного образования является: a) чтение художественной литературы b) прохождение онлайн-курсов с выполнением практических кейсов и получением обратной связи c) просмотр видеороликов без практики d) участие в спортивных соревнованиях		УК-6			
	Ответ:	b				
18	При определении приоритетов собственной деятельности в области машинного обучения специалисту по данным с учетом требований рынка труда целесообразно в первую очередь освоить библиотеку a) Scikit-learn (для построения интерпретируемых моделей) b) Pygame (для разработки игр) c) Tkinter (для создания десктопных приложений) d) Django (для веб-разработки)		УК-6			
	Ответ:	a				
19	Способом совершенствования навыков работы с SQL и самооценки своего уровня подготовки является: a) регулярное решение практических задач на платформах типа LeetCode или SQLZoo с последующим анализом ошибок b) разовое чтение документации без выполнения запросов c) заучивание синтаксиса без применения d) написание стихов на тему баз данных		УК-6			
	Ответ:	a				
20	С учетом требований рынка труда к аналитику данных наиболее востребованным инструментом непрерывного образования для освоения современных методов анализа больших данных является a) изучение устаревших версий программного обеспечения b) участие в профессиональных сообществах, хакатонах и открытых вебинарах, с последующей рефлексией и корректировкой приоритетов собственной деятельности c) игнорирование новых технологий d) работа только с бумажными носителями информации		УК-6			
	Ответ:	b				
21	Соотнесите инструменты непрерывного образования с их вкладом в способы совершенствования профессиональных компетенций в области анализа больших данных, используя самооценку как механизм контроля прогресса:		УК-6			
	<table><tr><th>Инструмент</th><th>Вклад в совершенствование</th></tr><tr><td>1. Онлайн-курсы с сертификацией (Coursera, Stepik)</td><td>А. Позволяет систематизировать знания, получить обратную связь от экспертов и подтвердить уровень компетенции, что важно для самооценки</td></tr></table>			Инструмент	Вклад в совершенствование	1. Онлайн-курсы с сертификацией (Coursera, Stepik)
Инструмент	Вклад в совершенствование					
1. Онлайн-курсы с сертификацией (Coursera, Stepik)	А. Позволяет систематизировать знания, получить обратную связь от экспертов и подтвердить уровень компетенции, что важно для самооценки					

	<table><tr><td>2. Участие в соревнованиях Kaggle</td><td>Б. Развивают навык решения реальных задач, дают возможность сравнить свои результаты с лучшими специалистами, помогают скорректировать приоритеты собственной деятельности</td></tr><tr><td>3. Чтение профессиональных блогов и документации (Stats, Medium, официальная документация)</td><td>В. Обеспечивают актуальную информацию о новых технологиях, позволяют самостоятельно изучать точечные темы с учетом требований рынка труда</td></tr></table>	2. Участие в соревнованиях Kaggle	Б. Развивают навык решения реальных задач, дают возможность сравнить свои результаты с лучшими специалистами, помогают скорректировать приоритеты собственной деятельности	3. Чтение профессиональных блогов и документации (Stats , Medium , официальная документация)	В. Обеспечивают актуальную информацию о новых технологиях, позволяют самостоятельно изучать точечные темы с учетом требований рынка труда					
2. Участие в соревнованиях Kaggle	Б. Развивают навык решения реальных задач, дают возможность сравнить свои результаты с лучшими специалистами, помогают скорректировать приоритеты собственной деятельности									
3. Чтение профессиональных блогов и документации (Stats , Medium , официальная документация)	В. Обеспечивают актуальную информацию о новых технологиях, позволяют самостоятельно изучать точечные темы с учетом требований рынка труда									
	Ответ: 1–А, 2–Б, 3–В									
22	Соотнесите этапы самооценки профессионального уровня с соответствующими способами совершенствования в области BI-аналитики: <table><tr><th>Этап самооценки</th><th>Способ совершенствования</th></tr><tr><td>1. Выявление пробелов в знаниях (например, недостаток навыков работы с Power BI)</td><td>А. Прохождение специализированного модуля или курса по данному инструменту с выполнением практического проекта</td></tr><tr><td>2. Определение приоритетов собственной деятельности (что учить в первую очередь)</td><td>Б. Анализ вакансий и требований рынка труда для выбора наиболее востребованных навыков</td></tr><tr><td>3. Оценка эффективности обучения (достигнут ли прогресс)</td><td>В. Повторное решение тестовых заданий или выполнение нового кейса и сравнение результатов с предыдущими</td></tr></table>	Этап самооценки	Способ совершенствования	1. Выявление пробелов в знаниях (например, недостаток навыков работы с Power BI)	А. Прохождение специализированного модуля или курса по данному инструменту с выполнением практического проекта	2. Определение приоритетов собственной деятельности (что учить в первую очередь)	Б. Анализ вакансий и требований рынка труда для выбора наиболее востребованных навыков	3. Оценка эффективности обучения (достигнут ли прогресс)	В. Повторное решение тестовых заданий или выполнение нового кейса и сравнение результатов с предыдущими	УК-6
Этап самооценки	Способ совершенствования									
1. Выявление пробелов в знаниях (например, недостаток навыков работы с Power BI)	А. Прохождение специализированного модуля или курса по данному инструменту с выполнением практического проекта									
2. Определение приоритетов собственной деятельности (что учить в первую очередь)	Б. Анализ вакансий и требований рынка труда для выбора наиболее востребованных навыков									
3. Оценка эффективности обучения (достигнут ли прогресс)	В. Повторное решение тестовых заданий или выполнение нового кейса и сравнение результатов с предыдущими									
	Ответ: 1–А, 2–Б, 3–В									
23	Соотнесите инструменты непрерывного образования с их ролью в совершенствовании навыков работы с SQL и большими данными с учетом самооценки: <table><tr><th>Инструмент</th><th>Роль в совершенствовании</th></tr><tr><td>1. Платформы с интерактивными задачами (SQL Academy, LeetCode)</td><td>А. Позволяют оценить текущий уровень в реальном времени через автоматическую проверку и рейтинг, что способствует объективной самооценке</td></tr><tr><td>2. Вебинары и мастер-классы от практиков индустрии</td><td>Б. Дают возможность узнать о реальных кейсах и современных практиках, помогают скорректировать приоритеты собственной деятельности</td></tr><tr><td>3. Создание собственного проекта-портфолио на GitHub</td><td>В. Служит доказательством компетенций для работодателя, а также позволяет рефлексировать над своими ошибками и этапами роста</td></tr></table>	Инструмент	Роль в совершенствовании	1. Платформы с интерактивными задачами (SQL Academy, LeetCode)	А. Позволяют оценить текущий уровень в реальном времени через автоматическую проверку и рейтинг, что способствует объективной самооценке	2. Вебинары и мастер-классы от практиков индустрии	Б. Дают возможность узнать о реальных кейсах и современных практиках, помогают скорректировать приоритеты собственной деятельности	3. Создание собственного проекта-портфолио на GitHub	В. Служит доказательством компетенций для работодателя, а также позволяет рефлексировать над своими ошибками и этапами роста	УК-6
Инструмент	Роль в совершенствовании									
1. Платформы с интерактивными задачами (SQL Academy, LeetCode)	А. Позволяют оценить текущий уровень в реальном времени через автоматическую проверку и рейтинг, что способствует объективной самооценке									
2. Вебинары и мастер-классы от практиков индустрии	Б. Дают возможность узнать о реальных кейсах и современных практиках, помогают скорректировать приоритеты собственной деятельности									
3. Создание собственного проекта-портфолио на GitHub	В. Служит доказательством компетенций для работодателя, а также позволяет рефлексировать над своими ошибками и этапами роста									
	Ответ: 1–А, 2–Б, 3–В									
24	Расположите в правильной последовательности шаги самооценки и совершенствования профессиональных навыков в области анализа больших данных с учетом приоритетов собственной деятельности и требований рынка труда: 1 Регулярный пересмотр приоритетов собственной деятельности на основе анализа новых требований рынка труда 2 Выполнение практических заданий и проектов для закрепления навыков 3 Прохождение тестирования или решения кейса для объективной самооценки текущего уровня 4 Выбор конкретных инструментов непрерывного образования (курсы, книги, вебинары) на основе выявленных пробелов	УК-6								
	Ответ: 3 → 4 → 2 → 1									
25	Расположите в правильной последовательности этапы построения индивидуальной траектории непрерывного образования в области BI-аналитики на основе самооценки и приоритетов собственной деятельности: 1 Оценка доступных инструментов непрерывного образования (бесплатные/платные курсы, литература, сообщества) 2 Анализ текущего уровня владения BI-инструментами через выполнение диагностического задания (самооценка) 3 Корректировка траектории обучения на основе промежуточных результатов и изменений требований рынка труда 4 Определение приоритетов собственной деятельности (какие BI-инструменты наиболее востребованы и критичны для карьерного роста) 5 Составление плана обучения с указанием сроков и контрольных точек	УК-6								
	Ответ: 2 → 4 → 1 → 5 → 3									
26	Дайте развернутый ответ Опишите, как вы проводите самооценку уровня владения инструментами анализа больших данных (например, Spark, SQL, Scikit-learn). Укажите, какие инструменты непрерывного образования вы используете для выявления пробелов и какие приоритеты собственной деятельности устанавливаете на основе результатов самооценки и требований рынка труда.	УК-6								

	Ответ: Самооценка проводится через выполнение тестовых заданий, анализ ошибок, участие в соревнованиях; пробелы выявляются через сравнение с вакансиями; приоритеты корректируются в сторону освоения Spark и BI-систем; используются курсы (Coursera, Stepik), практические проекты, сообщества.	
27	Дайте развернутый ответ Предложите способ совершенствования навыков работы с SQL для аналитика данных, учитывая современные требования рынка труда. Опишите, как вы оцениваете эффективность выбранного способа (механизм самооценки) и как корректируете приоритеты собственной деятельности при появлении новых технологий (например, NoSQL, NewSQL). Ответ: Ежедневное решение задач на LeetCode + анализ планов выполнения запросов; эффективность оценивается через рост скорости и точности; приоритеты корректируются путем изучения новых типов БД (MongoDB, ClickHouse) через курсы и документацию.	УК-6
28	Дайте развернутый ответ На основе анализа требований рынка труда (изучите 3–5 вакансий на позицию "Аналитик данных" или "Специалист по анализу больших данных") определите приоритеты собственной деятельности на ближайшие 6 месяцев. Опишите, какие инструменты непрерывного образования вы выберете для их реализации и как проведете самооценку прогресса. Ответ: Приоритеты: SQL (сложные запросы), Python (Pandas, Scikit-learn), BI (Power BI/Tableau), Spark; инструменты: курс на Coursera, Kaggle, YouTube-каналы; самооценка через контрольные работы и проекты, сравнение с требованиями вакансий через 6 мес.	УК-6
29	Дайте развернутый ответ Опишите ситуацию, когда результаты самооценки ваших компетенций в области машинного обучения показали необходимость изменения приоритетов собственной деятельности. Какие инструменты непрерывного образования вы использовали для устранения пробелов и как изменился ваш подход к совершенствованию навыков в соответствии с требованиями рынка труда? Ответ: Выявлен пробел в интерпретируемых моделях; приоритеты смещены на Decision Trees и логистическую регрессию; пройден курс по Scikit-learn и статистике; подход изменился в сторону регулярной практики с реальными данными и участия в хакатонах.	УК-6
30	Дайте развернутый ответ Разработайте краткий план непрерывного образования на 1 год для специалиста по анализу больших данных, начинающего карьеру. В плане укажите приоритеты собственной деятельности, конкретные инструменты непрерывного образования, механизмы самооценки и способы учета изменений требований рынка труда (например, появление новых версий библиотек или трендов в области ИИ). Ответ: План: 1–3 мес. – SQL + основы Python; 4–6 мес. – Pandas, визуализация, BI; 7–9 мес. – Scikit-learn, статистика; 10–12 мес. – Spark, большие данные; самооценка – тесты и проекты; мониторинг рынка – подписка на профильные каналы, анализ вакансий, участие в конференциях.	УК-6

7. Оценочные материалы промежуточной аттестации

Зачет второй семестр

№ п/п	Содержание вопроса		Компетенция												
		Правильный ответ (ключ ответа)													
1	Дайте развернутый письменный ответ ДЛЯ ОПТИМИЗАЦИИ ВРЕМЕНИ ВЫПОЛНЕНИЯ ЗАПРОСА В БАЗЕ ДАННЫХ С 50 МЛН ЗАПИСЕЙ ПРИ ЧАСТЫХ ВЫБОРКАХ ПО ПОЛЮ EMAIL И СОРТИРОВКЕ ПО ПОЛЮ REGISTRATION_DATE, КАКОЕ КОНКРЕТНОЕ ПРОГРАММНОЕ РЕШЕНИЕ (ОБЪЕКТ В SQL) НЕОБХОДИМО СОЗДАТЬ И КАКИЕ ПОЛЯ В НЕГО ВКЛЮЧИТЬ? СФОРМУЛИРУЙТЕ ОТВЕТ В ОДНОМ ПРЕДЛОЖЕНИИ. Ответ: Необходимо создать составной индекс (CREATE INDEX) по полям (email, registration_date) для обеспечения быстрого доступа к данным и ускорения сортировки, что повысит эффективность выполнения поисковых запросов.		ПК-2, УК-6												
2	Дайте развернутый письменный ответ ПРИ ИСПОЛЬЗОВАНИИ СОВРЕМЕННЫХ ПРОГРАММНЫХ СРЕДСТВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАСТЕРИЗАЦИИ КЛИЕНТОВ ВАМИ БЫЛ ПРИМЕНЕН АЛГОРИТМ K-MEANS. КАКОЙ МЕТРИКОЙ (НАЗВАНИЕ И КРАТКОЕ ОПИСАНИЕ) СЛЕДУЕТ ВОСПОЛЬЗОВАТЬСЯ ДЛЯ ОПТИМАЛЬНОГО ВЫБОРА ЧИСЛА КЛАСТЕРОВ K? НАПИШИТЕ НАЗВАНИЕ МЕТОДА. Ответ: Метод локтя (Elbow Method), основанный на вычислении суммы квадратов расстояний внутри кластеров (WCSS) для различных значений K и выборе точки изгиба графика.		ПК-2, УК-6												
3	Дайте развернутый письменный ответ В РАМКАХ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ БИЗНЕС-ПРОЦЕССА ПРОВОДИТСЯ А/В-ТЕСТИРОВАНИЕ. ДЛЯ КОНТРОЛЬНОЙ ГРУППЫ СРЕДНЯЯ КОНВЕРСИЯ СОСТАВИЛА 12%, ДЛЯ ТЕСТОВОЙ — 15%. РАЗМЕР КАЖДОЙ ГРУППЫ — 1000 НАБЛЮДЕНИЙ. РАССЧИТАЙТЕ Z-СТАТИСТИКУ ДЛЯ ПРОВЕРКИ ГИПОТЕЗЫ О РАЗЛИЧИИ ДОЛЕЙ (ПРОПОРЦИЙ) МЕЖДУ ДВУМЯ ГРУППАМИ. В ОТВЕТЕ УКАЖИТЕ ТОЛЬКО ЧИСЛОВОЕ ЗНАЧЕНИЕ, ОКРУГЛЕННОЕ ДО ДВУХ ЗНАКОВ ПОСЛЕ ЗАПЯТОЙ. <table><tr><th>Этап</th><th>Бюджет (тыс. руб.)</th><th>Валовое освоение (%)</th></tr><tr><td>1. Академия</td><td>500</td><td>100%</td></tr><tr><td>2. Разработка платформы</td><td>2000</td><td>70%</td></tr><tr><td>3. Сметы-контракты</td><td>1500</td><td>50%</td></tr></table>		Этап	Бюджет (тыс. руб.)	Валовое освоение (%)	1. Академия	500	100%	2. Разработка платформы	2000	70%	3. Сметы-контракты	1500	50%	ПК-2, УК-6
Этап	Бюджет (тыс. руб.)	Валовое освоение (%)													
1. Академия	500	100%													
2. Разработка платформы	2000	70%													
3. Сметы-контракты	1500	50%													

	<table border="1"> <tr> <td>4. Тестирование</td><td>1200</td><td>20%</td></tr> <tr> <td>5. Внедрение</td><td>800</td><td>5%</td></tr> <tr> <td>Этап</td><td>Бюджет (тыс. руб.)</td><td>Фактически освоено (%)</td></tr> <tr> <td>1. Анализа</td><td>500</td><td>100%</td></tr> <tr> <td>2. Разверстка платформы</td><td>2000</td><td>70%</td></tr> <tr> <td>3. Сквозь контенты</td><td>1500</td><td>50%</td></tr> </table>	4. Тестирование	1200	20%	5. Внедрение	800	5%	Этап	Бюджет (тыс. руб.)	Фактически освоено (%)	1. Анализа	500	100%	2. Разверстка платформы	2000	70%	3. Сквозь контенты	1500	50%	
4. Тестирование	1200	20%																		
5. Внедрение	800	5%																		
Этап	Бюджет (тыс. руб.)	Фактически освоено (%)																		
1. Анализа	500	100%																		
2. Разверстка платформы	2000	70%																		
3. Сквозь контенты	1500	50%																		
	Ответ: Формула: $Z = (p_1 - p_2) / \sqrt{p * (1-p) * (1/n_1 + 1/n_2)}$, где $p = (x_1+x_2)/(n_1+n_2)$. $p = (120+150)/(1000+1000) = 270/2000 = 0,135$. $Z = (0,12 - 0,15) / \sqrt{0,135 * 0,865 * (1/1000 + 1/1000)} = -0,03 / \sqrt{0,116775 * 0,002} = -0,03 / \sqrt{0,00023355} = -0,03 / 0,015283 = -1,963$. Числовой ответ (по модулю): 1,96.																			
4	Дайте развернутый письменный ответ С ИСПОЛЬЗОВАНИЕМ СОВРЕМЕННОГО ИНСТРУМЕНТАРИЯ ETL (НА ПРИМЕРЕ PYTHON И БИБЛИОТЕКИ PANDAS) ОПИШИТЕ ПОСЛЕДОВАТЕЛЬНОСТЬ ДЕЙСТВИЙ ДЛЯ ИЗВЛЕЧЕНИЯ ДАННЫХ ИЗ CSV-ФАЙЛА, ИХ ТРАНСФОРМАЦИИ (УДАЛЕНИЕ ДУБЛИКАТОВ) И ЗАГРУЗКИ В НОВЫЙ CSV-ФАЙЛ. НАПИШИТЕ НЕОБХОДИМЫЕ СТРОКИ ПРОГРАММНОГО КОДА. Ответ: <code>import pandas as pd df = pd.read_csv('source.csv') df = df.drop_duplicates() df.to_csv('cleaned_data.csv', index=False)</code>	ПК-2, УК-6																		
5	Дайте развернутый письменный ответ ПРИ РЕШЕНИИ ПРОФЕССИОНАЛЬНОЙ ЗАДАЧИ ПО ВИЗУАЛИЗАЦИИ ДАННЫХ НАЗОВИТЕ МИНИМУМ ДВЕ БИБЛИОТЕКИ PYTHON И УКАЖИТЕ ДЛЯ КАЖДОЙ ТИП ГРАФИКА, КОТОРЫЙ ОПТИМАЛЬНО ИСПОЛЬЗОВАТЬ ДЛЯ ОТОБРАЖЕНИЯ РАСПРЕДЕЛЕНИЯ КОЛИЧЕСТВЕННОГО ПРИЗНАКА. ОТВЕТ СФОРМУЛИРУЙТЕ В ВИДЕ ПЕРЕЧИСЛЕНИЯ. Ответ: Matplotlib — гистограмма (hist); Seaborn — ящик с усами (boxplot).	ПК-2, УК-6																		
6	Дайте развернутый письменный ответ ПРИ ИСПОЛЬЗОВАНИИ СОВРЕМЕННЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ (SQL) НАПИШИТЕ ЗАПРОС ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПОСТРОЕНИЯ ОТЧЕТА О ДИНАМИКЕ ПРОДАЖ: РАССЧИТАТЬ СКОЛЬЗЯЩЕЕ СРЕДНЕЕ ЗА 7 ДНЕЙ ПО ЕЖЕДНЕВНОЙ ВЫРУЧКЕ ИЗ ТАБЛИЦЫ DAILY_REVENUE (ПОЛЯ: REVENUE_DATE, TOTAL_AMOUNT). ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ИСПОЛЬЗУЙТЕ ОКОННЫЕ ФУНКЦИИ. Ответ: <code>SELECT revenue_date, total_amount, AVG(total_amount) OVER (ORDER BY revenue_date ROWS BETWEEN 6 PRECEDING AND CURRENT ROW) AS moving_avg_7d FROM daily_revenue;</code>	ПК-2, УК-6																		
7	Дайте развернутый письменный ответ С ИСПОЛЬЗОВАНИЕМ СОВРЕМЕННЫХ ПРОГРАММНЫХ СРЕДСТВ (SCIKIT-LEARN) ВЫ ПОСТРОИЛИ МОДЕЛЬ КЛАССИФИКАЦИИ И ПОЛУЧИЛИ ЗНАЧЕНИЕ F1-MEASURE = 0,87. ОПИШИТЕ, ЧТО ОЗНАЧАЕТ ДАННЫЙ ПОКАЗАТЕЛЬ С ТОЧКИ ЗРЕНИЯ ЭФФЕКТИВНОСТИ РЕШЕНИЯ ПРОФЕССИОНАЛЬНОЙ ЗАДАЧИ, И КАКИЕ КОМПОНЕНТЫ (PRECISION И RECALL) НА НЕГО ВЛИЯЮТ. СФОРМУЛИРУЙТЕ ОТВЕТ В 2-3 ПРЕДЛОЖЕНИЯХ. Ответ: F1-measure = 0,87 является средним гармоническим между полнотой (Recall) и точностью (Precision) и свидетельствует о высокой эффективности модели, обеспечивающей хороший баланс между этими метриками. Данный показатель позволяет оптимизировать решение задачи классификации в условиях несбалансированных классов, что повышает надежность принимаемых решений.	ПК-2, УК-6																		
8	Дайте развернутый письменный ответ ДЛЯ ОПТИМИЗАЦИИ ХРАНЕНИЯ БОЛЬШИХ ОБЪЕМОВ ИСТОРИЧЕСКИХ ДАННЫХ И ПОВЫШЕНИЯ СКОРОСТИ АНАЛИТИЧЕСКИХ ЗАПРОСОВ ПРЕДЛОЖИТЕ СОВРЕМЕННУЮ ТЕХНОЛОГИЮ ХРАНЕНИЯ ДАННЫХ (ТИП СУБД), КОТОРАЯ ИСПОЛЬЗУЕТ КОЛОНОЧНУЮ АРХИТЕКТУРУ. ОБОСНУЙТЕ СВОЙ ВЫБОР В ОДНОМ ПРЕДЛОЖЕНИИ. Ответ: Целесообразно использовать колоночную СУБД, например, ClickHouse или Apache Cassandra, поскольку она обеспечивает высокую эффективность агрегации и фильтрации данных за счет хранения значений каждого столбца последовательно, что значительно ускоряет выполнение OLAP-запросов.	ПК-2, УК-6																		
9	Дайте развернутый письменный ответ (4-6 предложений) ПРИ РЕШЕНИИ ЗАДАЧИ ПРЕДОБРАБОТКИ ДАННЫХ НАЗОВИТЕ БИБЛИОТЕКУ PYTHON И КОНКРЕТНЫЙ МЕТОД, КОТОРЫЙ ПОЗВОЛИТ АВТОМАТИЗИРОВАТЬ ПРОЦЕСС ЗАМЕНЫ ПРОПУЩЕННЫХ ЗНАЧЕНИЙ В ЧИСЛОВОМ СТОЛБЦЕ DATAFRAME НА МЕДИАННОЕ ЗНАЧЕНИЕ. НАПИШИТЕ СООТВЕТСТВУЮЩУЮ СТРОКУ ПРОГРАММНОГО КОДА. Ответ: Библиотека Pandas, метод fillna(). Строка кода: <code>df['column_name'].fillna(df['column_name'].median(), inplace=True).</code>	ПК-2, УК-6																		
10	Дайте развернутый письменный ответ (4-6 предложений) ИСПОЛЬЗУЯ СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ (SQL), НАПИШИТЕ ЗАПРОС ДЛЯ РЕШЕНИЯ ПРОФЕССИОНАЛЬНОЙ ЗАДАЧИ ПО АНАЛИЗУ ПРОДАЖ: ВЫБРАТЬ ТОП-5 ТОВАРОВ ПО СУММАРНОЙ ВЫРУЧКЕ ЗА ПОСЛЕДНИЙ КВАРТАЛ ИЗ ТАБЛИЦЫ SALES (ПОЛЯ: PRODUCT_ID, AMOUNT, SALE_DATE). ЗАПРОС ДОЛЖЕН ОБЕСПЕЧИТЬ ОПТИМИЗАЦИЮ ВЫЧИСЛЕНИЙ И ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ АНАЛИТИЧЕСКОЙ ОТЧЕТНОСТИ.	ПК-2, УК-6																		

Ответ:	SELECT product_id, SUM(amount) AS total_revenue FROM sales WHERE sale_date >= CURRENT_DATE - INTERVAL '3 month' GROUP BY product_id ORDER BY total_revenue DESC LIMIT 5;
--------	---

7.1. Уровни овладения

Компетенция: УК-6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки.

Индикатор достижения компетенции: УК-6.1 Определяет инструменты и методы управления временем при реализации приоритетов собственной деятельности для выполнения конкретных задач, проектов, достижения поставленных целей.

Уровень	Характеристика	Оценка в баллах
Повышенный	Достигнуто полное овладение знаниями, умениями и навыками. Студент свободно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	81-100
Базовый	Достигнуто достаточное овладение знаниями, умениями и навыками. Студент уверенно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	61-80
Пороговый	Достигнуто овладение минимально необходимыми знаниями, умениями и навыками. Студент владеет основной терминологией, умеет применять теоретические знания для решения поставленных задач в стандартных ситуациях.	41-60
Ниже порогового	Компетенция не освоена	0-40

Индикатор достижения компетенции: УК-6.2 Анализирует и использует инструменты непрерывного образования с учетом требований рынка труда и способы ее совершенствования на основе самооценки.

Уровень	Характеристика	Оценка в баллах
Повышенный	Достигнуто полное овладение знаниями, умениями и навыками. Студент свободно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	81-100
Базовый	Достигнуто достаточное овладение знаниями, умениями и навыками. Студент уверенно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	61-80
Пороговый	Достигнуто овладение минимально необходимыми знаниями, умениями и навыками. Студент владеет основной терминологией, умеет применять теоретические знания для решения поставленных задач в стандартных ситуациях.	41-60
Ниже порогового	Компетенция не освоена	0-40

Компетенция: ПК-2 Способен планировать и использовать методы научных исследований и инструментарий в области управления проектами в сфере ИТ малого и среднего уровня сложности.

Индикатор достижения компетенции: ПК-2.1 Изучает и планирует использование методов научных исследований и инструментария в области управления проектами.

Уровень	Характеристика	Оценка в баллах
Повышенный	Достигнуто полное овладение знаниями, умениями и навыками. Студент свободно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	81-100
Базовый	Достигнуто достаточное овладение знаниями, умениями и навыками. Студент уверенно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	61-80
Пороговый	Достигнуто овладение минимально необходимыми знаниями, умениями и навыками. Студент владеет основной терминологией, умеет применять теоретические знания для решения поставленных задач в стандартных ситуациях.	41-60
Ниже порогового	Компетенция не освоена	0-40

Индикатор достижения компетенции: ПК-2.2 Использует методы научных исследований и инструментарий в области управления проектами в сфере ИТ малого и среднего уровня сложности.

Уровень	Характеристика	Оценка в баллах
Повышенный	Достигнуто полное овладение знаниями, умениями и навыками. Студент свободно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	81-100
Базовый	Достигнуто достаточное овладение знаниями, умениями и навыками. Студент уверенно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	61-80
Пороговый	Достигнуто овладение минимально необходимыми знаниями, умениями и навыками. Студент владеет основной терминологией, умеет применять теоретические знания для решения поставленных задач в стандартных ситуациях.	41-60
Ниже порогового	Компетенция не освоена	0-40

8. Материально-техническое и учебно-методическое обеспечение дисциплины

8.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. Терехов, А. М. Методы статистического анализа в индустрии туризма. Практический курс: учебник для вузов / А. М. Терехов, П. Г. Николенко, М. В. Ефремова. - Москва: Юрайт, 2026. - 154 с - 978-5-534-18626-0. - Текст: электронный // ИКО Юрайт: [сайт]. - URL: <https://urait.ru/bcode/589618> (дата обращения: 21.05.2026). - Режим доступа: по подписке

Дополнительная литература

1. Платонов, А. В. Машинное обучение: учебное пособие для вузов / А. В. Платонов. - 2-е изд. - Москва: Юрайт, 2026. - 89 с - 978-5-534-20732-3. - Текст: электронный // ИКО Юрайт: [сайт]. - URL: <https://urait.ru/bcode/589132> (дата обращения: 21.05.2026). - Режим доступа: по подписке

2. Федоров, Д. Ю. Программирование на Python: учебное пособие для вузов / Д. Ю. Федоров. - 6-е изд. - Москва: Юрайт, 2026. - 187 с - 978-5-534-19666-5. - Текст: электронный // ИКО Юрайт: [сайт]. - URL: <https://urait.ru/bcode/585806> (дата обращения: 21.05.2026). - Режим доступа: по подписке

8.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся

Профессиональные базы данных

Не используются.

Ресурсы «Интернет»

1. <https://edu.1cfresh.com/> - 1С:Предприятие 8 через Интернет для учебных заведений
2. <https://stepik.org> - Платформа с онлайн-курсами от авторов-практиков
3. <https://www.rusprofile.ru/> - Система комплексного анализа рынков и компаний (СКАР) / Rusprofile

8.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине

Перечень программного обеспечения

(обновление производится по мере появления новых версий программы)

1. MySQL Server 8.0.45;
2. MySQL Workbench 8.0.45;
3. Python версии 3.14.4 ;
4. BpWin и ErWin Обновление;
5. Консультант Плюс;
6. Мой офис;

Перечень информационно-справочных систем

(обновление выполняется еженедельно)

Не используется.

8.4. Специальные помещения, лаборатории и лабораторное оборудование

Учебные аудитории для проведения занятий лекционного типа	Комплекты ученической мебели Мультимедийный проектор Доска Экран
---	---

Учебные аудитории для проведения практических занятий (занятий семинарского типа)	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СПб
Учебные аудитории для групповых и индивидуальных консультаций	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СПб
Учебные аудитории для текущего контроля и промежуточной аттестации	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СПб
Помещения для самостоятельной работы	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СПб
Помещения для хранения и профилактического обслуживания оборудования	Комплекты специализированной мебели для хранения