



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Магнитогорский государственный технический университет им. Г.И.
Носова»



УТВЕРЖДАЮ
Директор ИЭиАС
В.Р. Храмшин

25.04.2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

ПРИКЛАДНОЙ АНАЛИЗ ДАННЫХ В PYTHON

Направление подготовки (специальность)
38.03.05 Бизнес-информатика

Направленность (профиль/специализация) программы
Аналитическая поддержка в управлении бизнес-процессами

Уровень высшего образования - бакалавриат

Форма обучения
очная

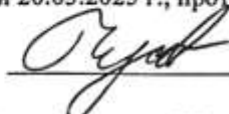
Институт/ факультет	Институт энергетики и автоматизированных систем
Кафедра	Бизнес-информатики и информационных технологий
Курс	4
Семестр	8

Магнитогорск
2025 год

Рабочая программа составлена на основе ФГОС ВО - бакалавриат по направлению подготовки 38.03.05 Бизнес-информатика (приказ Минобрнауки России от 29.07.2020 г. № 838)

Рабочая программа рассмотрена и одобрена на заседании кафедры Бизнес-информатики и информационных технологий 20.03.2025 г., протокол № 7

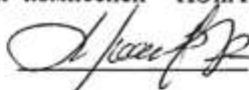
Зав. кафедрой



Г.Н. Чусавитина

Рабочая программа одобрена методической комиссией ИЭиАС 25.04.2025 г. протокол № 5

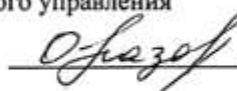
Председатель



В.Р. Храмшин

Согласовано:

Зав. кафедрой Менеджмента и государственного управления



О.Л. Назарова

Рабочая программа составлена:

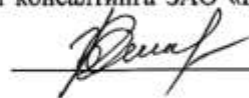
доцент кафедры БИиИТ, канд. пед. наук



Е.В. Карманова

Рецензент:

руководитель направления бизнес-анализа и консалтинга ЗАО «КонсОМ СКС»,
канд. техн. наук



В.А. Ошурков

Лист актуализации рабочей программы

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2026 - 2027 учебном году на заседании кафедры Бизнес-информатики и информационных

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ Г.Н. Чусавитина

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2027 - 2028 учебном году на заседании кафедры Бизнес-информатики и информационных

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ Г.Н. Чусавитина

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2028 - 2029 учебном году на заседании кафедры Бизнес-информатики и информационных

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ Г.Н. Чусавитина

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2029 - 2030 учебном году на заседании кафедры Бизнес-информатики и информационных

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ Г.Н. Чусавитина

1 Цели освоения дисциплины (модуля)

Целями освоения дисциплины являются: формирование компетенций в области обработки и анализа данных - освоение методов сбора, очистки и предварительной обработки данных с использованием Python; развитие навыков визуализации данных; освоение методов статистического и машинного обучения; автоматизация аналитических задач; применение аналитики для принятия управленческих решений; развитие навыков работы с реальными данными.

2 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина Прикладной анализ данных в Python входит в обязательную часть учебного плана образовательной программы.

Для изучения дисциплины необходимы знания (умения, владения), сформированные в результате изучения дисциплин/ практик:

Информационные процессы, системы и сети

Информатика

Основы алгоритмизации и программирования

Анализ больших данных

Знания (умения, владения), полученные при изучении данной дисциплины будут необходимы для изучения дисциплин/практик:

Подготовка к сдаче и сдача государственного экзамена

Подготовка к защите и защита выпускной квалификационной работы

Производственная – преддипломная практика

Прикладные информационные системы

3 Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля) и планируемые результаты обучения

В результате освоения дисциплины (модуля) «Прикладной анализ данных в Python» обучающийся должен обладать следующими компетенциями:

Код индикатора	Индикатор достижения компетенции
ОПК-3	Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации
ОПК-3.1	Участвует в управлении ИТ-проектами
ОПК-3.2	Принимает участие в разработке продуктов и услуг в сфере ИКТ (в том числе алгоритмов, программ, информационных ресурсов)

4. Структура, объём и содержание дисциплины (модуля)

Общая трудоемкость дисциплины составляет 4 зачетных единиц 144 акад. часов, в том числе:

- контактная работа – 49,3 акад. часов;
- аудиторная – 48 акад. часов;
- внеаудиторная – 1,3 акад. часов;
- самостоятельная работа – 94,7 акад. часов;
- в форме практической подготовки – 0 акад. час;

Форма аттестации - зачет

Раздел/ тема дисциплины	Семестр	Аудиторная контактная работа (в акад. часах)			Самостоятельная работа студента	Вид самостоятельной работы	Форма текущего контроля успеваемости и промежуточной аттестации	Код компетенции
		Лек.	лаб. зан.	практ. зан.				
1. Методы визуализации данных с использованием Python								
1.1 Особенности описательной (дескриптивной) аналитики	8	6		4	6,7	Выполнение практической работы	Отчет по практической работе	ОПК-3.1, ОПК-3.2
1.2 Методы и инструменты визуализации больших данных		4		4	10	Выполнение практической работы	Отчет по практической работе	ОПК-3.1, ОПК-3.2
1.3 Создание концептуальных карт		2		2	6	Выполнение практической работы	Отчет по практической работе	ОПК-3.2
Итого по разделу		12		10	22,7			
2. Методы и инструменты диагностической аналитики в образовании								
2.1 Методы кластерного анализа для выявления закономерностей в больших данных	8	2		2	14	Выполнение практической работы	Отчет по практической работе	ОПК-3.2
2.2 Методы классификации данных, их применение для анализа больших данных		4		4	16	Выполнение практической работы	Отчет по практической работе	ОПК-3.2
2.3 ROC-анализ		2		2	15	Выполнение практической работы	Отчет по практической работе	ОПК-3.2
Итого по разделу		8		8	45			
3. Методы и инструменты предиктивной аналитики в образовании								
3.1 Задачи регрессии при анализе больших данных	8	1		2	14	Выполнение практической	Отчет по практической	ОПК-3.2

						работы	работе	
3.2 Построение регрессионных деревьев для прогнозирования	8	3		4	13	Выполнение практической работы №14	Отчет по практической работе №14	ОПК-3.2
Итого по разделу		4		6	27			
Итого за семестр		24		24	94,7		зачёт	
Итого по дисциплине		24		24	94,7		зачет	

5 Образовательные технологии

Основными образовательными технологиями, положенными в основу преподавания дисциплины «Прикладной анализ данных в Python» являются:

- **активные технологии обучения:**

о метод ролевых игр - это разыгрывание участниками группы сценки с заранее распределенными ролями в интересах овладения определенной поведенческой или эмоциональной стороной жизненных ситуаций.

Ролевая игра проводится в небольших группах (3-5 участников);

о технологии кейс-стади - техника обучения, использующая описание ре-альных ситуаций. Обучающиеся должны проанализировать ситуацию, разобраться в сути проблем, предложить возможные решения и выбрать лучшее из них. Кейсы базируются на реальном фактическом материале, или же приближены к реальной ситуации;

о разработка проекта - это способ достижения дидактической цели через детальную разработку проблемы (технология), которая должна завершиться вполне реальным, осязаемым практическим результатом, оформленным тем или иным образом;

о работа в малых группах - это одна из самых популярных стратегий, так как она дает всем обучающимся (в том числе и стеснительным) возможность участвовать в работе, практиковать навыки сотрудничества, меж-личностного общения (в частности, умение активно слушать, вырабатывать общее мнение, разрешать возникающие разногласия);

- **интерактивные лекции:**

о лекций-дискуссий - преподаватель приводит отдельные примеры в виде ситуаций или кратко сформулированных проблем и предлагает студентам коротко обсудить, затем краткий анализ, выводы и лекция продолжается. Положительным в дискуссии является, то, что обучаемые согласятся с точкой зрения преподавателя с большой охотой, скорее в ходе дискуссии, нежели во время беседы, когда преподаватель лишь указывает на необходимость принять его позицию по обсуждаемому вопросу. Данный метод позволяет преподавателю видеть, насколько эффективно слушатели используют полученные знания в ходе дискуссии.

Активные технологии обучения преимущественно используются в рамках практических занятий, интерактивные лекции - в процессе изучения и закрепления нового учебного материала.

В качестве практико-ориентированного средства обучения выбран образовательный портал ФГБОУ ВО «МГТУ им. Г.И. Носова».

6 Учебно-методическое обеспечение самостоятельной работы обучающихся

Представлено в приложении 1.

7 Оценочные средства для проведения промежуточной аттестации

Представлены в приложении 2.

8 Учебно-методическое и информационное обеспечение дисциплины

а) Основная литература:

1. Ланских, Ю. В. Интеллектуальный анализ данных : учебное пособие / Ю. В. Ланских, В. Г. Ланских. — Киров : ВятГУ, 2023. — 240 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/408569> (дата обращения: 27.03.2025). — Режим доступа: для авториз. пользователей.

2. Демидова, Л. А. Кластерный анализ. Python : учебное пособие / Л. А.

Демидова. — Москва : РТУ МИРЭА, 2022. — 103 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/240092> (дата обращения: 13.03.2025). — Режим доступа: для авториз. пользователей.

б) Дополнительная литература:

Демидова, Л. А. Разведочный анализ данных. Python : учебно-методическое пособие / Л. А. Демидова. — Москва : РТУ МИРЭА, 2022 — Часть 1 — 2022. — 107 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/310970> (дата обращения: 13.03.2025). — Режим доступа: для авториз. пользователей.

Демидова, Л. А. Разведочный анализ данных. Python : учебно-методическое пособие / Л. А. Демидова. — Москва : РТУ МИРЭА, 2023 — Часть 2 — 2023. — 92 с. — ISBN 978-5-7339-1933-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/382691> (дата обращения: 13.03.2025). — Режим доступа: для авториз. пользователей.

Олейник, Т. А. Статистический анализ данных с использованием языка программирования Python: практикум : учебное пособие / Т. А. Олейник. — Москва : МИЭТ, 2023. — 144 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/461597> (дата обращения: 13.03.2025). — Режим доступа: для авториз. пользователей.

в) Методические указания:

1. Курзаева, Л. В. Анализ данных в задачах исследования социально-экономических систем и процессов : учебное пособие / Л. В. Курзаева ; МГТУ. - Магнитогорск : МГТУ, 2016. - 1 электрон. опт. диск (CD-ROM). - Загл. с титул. экрана. - URL: <https://host.megaprolib.net/MP0109/Download/MObject/20637> (дата обращения: 15.01.2025). - Макрообъект. - Текст : электронный. - Сведения доступны также на CD-ROM.

2. Курзаева, Л. В. Введение в теорию систем и системный анализ : учебное пособие / Л. В. Курзаева ; МГТУ. - [2-е изд., подгот. по печ. изд. 2013 г.]. - Магнитогорск : МГТУ, 2015. - 1 электрон. опт. диск (CD-ROM). - Загл. с титул. экрана. - URL: <https://host.megaprolib.net/MP0109/Download/MObject/1796> (дата обращения: 15.01.2025). - Макрообъект. - Текст : электронный. - Сведения доступны также на CD-ROM.

г) Программное обеспечение и Интернет-ресурсы:

Программное обеспечение

Наименование ПО	№ договора	Срок действия лицензии
7Zip	свободно распространяемое ПО	бессрочно
Deductor Studio Academic	Соглашение о сотрудничестве №06-2901\08 от 29.01.2008	бессрочно
Anaconda Python	свободно распространяемое ПО	бессрочно
Adobe Reader	свободно распространяемое ПО	бессрочно
Браузер Yandex	свободно распространяемое ПО	бессрочно

Профессиональные базы данных и информационные справочные системы

Название курса	Ссылка
Электронные ресурсы библиотеки МГТУ им. Г.И. Носова	https://host.megaprolib.net/MP0109/Web
Национальная информационно-аналитическая система – Российский индекс научного цитирования (РИНЦ)	URL: https://elibrary.ru/project_risc.asp

9 Материально-техническое обеспечение дисциплины (модуля)

Материально-техническое обеспечение дисциплины включает:

Учебные аудитории для проведения занятий лекционного типа Персональный компьютер (или ноутбук) с пакетом Office, с выходом в Интернет и с доступом в электронную информационно-образовательную среду университета. Доска, мультимедийный проектор, экран. Мультимедийные презентации к лекциям, учебно-наглядные пособия

Учебные аудитории для проведения лабораторных занятий Персональные компьютеры с пакетом Office; с выходом в Интернет и с доступом в электронную информационно-образовательную среду университета. Браузер MozillaFirefox.

Аудитории для самостоятельной работы: компьютерные классы; читальные залы библиотеки Персональные компьютеры с пакетом Office; с выходом в Интернет и с доступом в электронную информационно-образовательную среду университета. Браузер MozillaFirefox.

Аудитории для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации Персональные компьютеры с пакетом Office и выходом в Интернет и с доступом в электронную информационно-образовательную среду университета. Браузер MozillaFirefox.

Аудитория для хранения и профилактического обслуживания учебного оборудования Мебель для хранения и обслуживания оборудования (шкафы, столы), учебно-методические материалы, компьютеры, ноутбуки, принтеры.

Учебно-методическое обеспечение самостоятельной работы обучающихся

Перечень тем для подготовки к практическим занятиям:

№ занятия	Краткое содержание практического занятия
1	Основы предобработки данных. Очистка данных. Устранение "выбросов". Анализ атрибутов и описательных статистик для образовательных данных. Анализ и обнаружение взаимосвязей в образовательных данных. Современные метрики образовательных проектов. Задание: дан датасет государственных школ Чикаго - Табели успеваемости (2011-2012). Необходимо проанализировать датасет на предмет выбросов, какие поля содержат нулевые значения, некорректные типы данных. Оцените имеется ли корреляция между отдельными показателями датасета. Какие прогнозируемые метрики можно предложить к данному датасету? <i>Полезные источники:</i> Три способа обнаружения выбросов - История данных Колина Горри (англ.) - http://colingorrie.github.io/outlier-detection.html Макшанов, А. В. Большие данные. Big Data : учебник для вузов / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. — Санкт-Петербург : Лань, 2021. — 188 с. — ISBN 978-5-8114-6810-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/165835 (дата обращения: 15.06.2022). — Режим доступа: для авториз. пользователей.
2	Факторный анализ. Задание: дан набор данных по продаже недвижимости в г. Магнитогорске. Необходимо провести факторный анализ, выделив, какие факторы влияют на стоимость недвижимости. <i>Полезные источники:</i> Талипов, Н. Г. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 308 с. — ISBN 978-5-7579-2488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193530 (дата обращения: 11.06.2022). — Режим доступа: для авториз. пользователей.
3	Особенности описательной (дескриптивной) аналитики. Задание: дан набор данных по продаже недвижимости в г. Магнитогорске. Определить: 1. Сколько квартир представлено в датасете? 2. Какова максимальная стоимость квартир? 3. В каком районе города самые дорогие и самые дешевые квартиры? 4. Существует ли корреляция между площадью квартиры и ее стоимостью? 5. Имеются ли аномальные квартиры в наборе данных? <i>Полезные источники:</i> Талипов, Н. Г. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 308 с. — ISBN 978-5-7579-2488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193530 (дата обращения: 11.06.2022). — Режим доступа: для авториз. пользователей.
4	Методы и инструменты визуализации данных. Задание: дан набор данных по продаже недвижимости в г. Магнитогорске. Построить на основе предложенных данных следующие виды графиков: линейный, “ящик с усами”, круговая диаграмма, столбиковая диаграмма, гистограмма, график рассеяния, пузырьковая диаграмма, географическая диаграмма. <i>Полезные источники:</i> Сервис для создания интерактивных графиков - https://chart-studio.plotly.com/

5	Создание концептуальных карт. <i>Полезные источники:</i> Загорулько, Ю. А. Искусственный интеллект. Инженерия знаний : учебное пособие для вузов / Ю. А. Загорулько, Г. Б. Загорулько. — Москва : Издательство Юрайт, 2020. — 93 с. — (Высшее образование). — ISBN 978-5-534-07198-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/455500 (дата обращения: 15.06.2022).
6	Методы кластерного анализа для выявления закономерностей в данных. Задание. дан набор данных по продаже недвижимости в г. Магнитогорске. Необходимо разделить множество квартир на равномошные группы, используя методы кластеризации. Охарактеризуйте результат. <i>Полезные источники:</i> Талипов, Н. Г. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 308 с. — ISBN 978-5-7579-2488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193530 (дата обращения: 11.06.2022). — Режим доступа: для авториз. пользователей.
7	Методы классификации данных, их применение для анализа данных. Задание. дан набор данных по продаже недвижимости в г. Магнитогорске. Необходимо разбить множество на классы, используя метод бинарной классификации. Критерии разбиения подобрать самостоятельно. <i>Полезные источники:</i> Талипов, Н. Г. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 308 с. — ISBN 978-5-7579-2488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193530 (дата обращения: 11.06.2022). — Режим доступа: для авториз. пользователей.
8	ROC-анализ для оценки динамики образовательного (учебного процесса) Задание. дан набор данных по продаже недвижимости в г. Магнитогорске. На основе построенной в предыдущем задании классификации построить ROC-кривую и написать пояснение к ней. <i>Полезные источники:</i> Талипов, Н. Г. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 308 с. — ISBN 978-5-7579-2488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193530 (дата обращения: 11.06.2022). — Режим доступа: для авториз. пользователей.
9	<i>Полезные источники:</i> Задачи регрессии при анализе данных. Задание. дан набор данных по продаже недвижимости в г. Магнитогорске. Необходимо построить прогноз стоимости квартир на основе предложенного признакового пространства. <i>Полезные источники:</i> Талипов, Н. Г. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 308 с. — ISBN 978-5-7579-2488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193530 (дата обращения: 11.06.2022). — Режим доступа: для авториз. пользователей.
10	Построение регрессионных деревьев для прогнозирования результата. Задание. дан набор данных по продаже недвижимости в г. Магнитогорске. Необходимо построить прогноз стоимости квартир на основе предложенного признакового пространства с использованием методов деревьев с подбором лучших

	гиперпараметров модели. <i>Полезные источники:</i> Талипов, Н. Г. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 308 с. — ISBN 978-5-7579-2488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193530 (дата обращения: 11.06.2022). — Режим доступа: для авториз. пользователей.
11	Построение наивного байесовского классификатора. Задание. Необходимо построить прогноз стоимости квартир на основе предложенного признакового пространства. Оценить возможность построения на данном множестве наивного байесовского классификатора. Обосновать ответ. <i>Полезные источники:</i> Талипов, Н. Г. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 308 с. — ISBN 978-5-7579-2488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193530 (дата обращения: 11.06.2022). — Режим доступа: для авториз. пользователей.

7 Оценочные средства для проведения промежуточной аттестации по дисциплине «Прикладной анализ данных в Python»

а) Планируемые результаты обучения и оценочные средства для проведения промежуточной аттестации:

Код индикатора	Индикатор достижения компетенции	Оценочные средства
ОПК-3: Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации		
ОПК-3.1	Участствует в управлении ИТ-проектами	Теоретические вопросы 1. Методы, инструменты анализа данных. 2. Методы и инструменты визуализации образовательных данных. 3. Приемы агрегирования данных. Построение дашбордов на основе данных. 4. Методы структурирования знаний. 5. Подходы и инструменты построения Интеллект-карт, деревьев решений. 6. Концептуальная модель базы данных. 7. Инструменты построения концептуальных карт Практические задания: 1. Построить с использованием сервиса Yandex DataLens инфопанель (дашборд) для визуализации параметров качества теста по полученным результатам тестирования. 2. Используя он-лайн сервис постройте Интеллект-карту по следующим понятиям «Большие данные», «Технологии машинного обучения», «Интеллектуальные сервисы для маркетинга». 3. Используя он-лайн сервис постройте дерево решений по предметной области – получение зачета по дисциплине; публикация статьи в журнале, индексируемом SCOPUS/Web of Science с квартилем Q2 – Q3. 4. Реализуйте следующие стадии концептуального анализа для предметной области «Управление бизнес-процессам»: определение входных и выходных данных, словарь терминов, выявление объектов, понятий и их атрибутов, выявление связей между понятиями. Комплексное задание:

Разработать рекомендательную систему, которая поможет бизнесу эффективно проводить A/B тестирование элементов сайта (например, дизайна, текстов, кнопок СТА, ценовых стратегий) для повышения конверсии и других ключевых метрик.

Требования к выполнению:

1. Анализ данных и подготовка

- Сбор данных:
 - Определить источники данных (Google Analytics, Яндекс.Метрика, внутренние логи, CSV/Excel).
 - Сформировать датасет с историческими данными о поведении пользователей (просмотры, клики, конверсии).
- Предобработка данных:
 - Очистка от пропусков, аномалий, дубликатов.
 - Нормализация и категоризация данных (например, сегментация пользователей).

2. Проектирование системы A/B тестирования

- Определение гипотез:
 - Выбрать элементы для тестирования (например, цвет кнопки, заголовок, расположение формы).
 - Сформулировать нулевую и альтернативную гипотезы.
- Разделение трафика:
 - Разработать алгоритм рандомизации пользователей в группы A и B (например, на основе user_id или cookies).
 - Обеспечить корректное распределение (50/50 или иное, в зависимости от задачи).

3. Разработка рекомендательного механизма

- Метрики для оценки:
 - Определить ключевые метрики (CTR, конверсия, средний чек, время на странице).
 - Рассчитать статистическую значимость (t-тест, U-критерий Манна-Уитни, хи-квадрат).
- Автоматизация анализа:
 - Написать скрипт на Python, который:
 - Сравнивает результаты групп A и B.
 - Определяет победителя (или отсутствие значимой разницы).
 - Формирует отчет с визуализацией (графики, таблицы).

4. Визуализация и интерпретация

- Дашборд для мониторинга:
 - Использовать matplotlib/seaborn или Plotly/Dash для наглядного представления результатов.
 - Включить динамику метрик, p-value, доверительные интервалы.
- Рекомендации для бизнеса:
 - На основе анализа предложить лучший вариант (A/B) или дополнительные тесты.
 - Указать возможные ограничения (время теста, сезонность, выборка).

5. Техническая реализация (Python)

- Используемые библиотеки:

ОПК-3.2	Принимает участие в разработке продуктов и услуг в сфере ИКТ (в том числе алгоритмов, программ, информационных ресурсов)	Теоретические вопросы: 1. Технология факторного анализа данных. 2. Методы классификации данных для анализа данных. 3. Задачи регрессии при анализе данных. Прогнозирование на основе регрессионных деревьев. 4. ROC-анализ данных 5. Построение наивного байесовского классификатора. Практические задания: 1. На данных по стоимости машин применить метод Линейной регрессии без нормализации данных. Получить коэффициент детерминации, оценки MAE, MSE, вывести коэффициенты полученной формулы регрессии. Представить анализ полученных оценок и коэффициентов, выводы оформить в отдельную ячейку ноутбука. 2. Для датасета CityStar решите задачу классификации планировки квартир. Реализуйте множественную классификацию (не менее 3 классов с планировками квартир). Обосновать выделение соответствующих классов датасета. 3. Выполните задачу кластеризации для датасета CityStar для любого k, выполните визуализацию для лучшего результата кластеризации. Присвойте выделенным кластерам метки, обоснуйте данные метки описательной статистикой. Комплексное задание: Реализовать приложение для увеличения конверсии рекламной кампании. Данные: История кампаний, демография, клики. Требования: 1. Определить и описать методы машинного обучения (нельзя использовать глубокое обучение): • Реализация нескольких подходов (например, разные алгоритмы машинного обучения). • Выбор оптимальной модели на основе метрик качества. 2. Создать интеллектуальное приложение, демонстрирующее работу модели (например, веб-сервис, настольное приложение или скрипт с интерфейсом).
---------	--	---

б) Порядок проведения промежуточной аттестации, показатели и критерии оценивания:

Промежуточная аттестация по дисциплине «Прикладной анализ данных в Python» включает теоретические вопросы, позволяющие оценить уровень усвоения обучающимися знаний, и практические задания, выявляющие степень сформированности умений и владений, проводится в форме зачета.

Показатели и критерии оценивания на зачет (в соответствии с формируемыми компетенциями и планируемыми результатами обучения):

– на оценку **«зачтено»** – обучающийся показывает высокий уровень сформированности компетенций, т.е. выполняет тренировочные, практические и практические работы в установленные сроки; разрабатывает проектные задания по дисциплине с учетом заявленных требований, владеет терминологическим аппаратом, демонстрирует глубокое теоретическое знание вопроса в области программирования, грамотно определяет логико-структурные связи, обосновывает свое решение и формулирует необходимые выводы.

– на оценку **«не зачтено»** – результат обучения не достигнут, обучающийся не может показать знания на уровне воспроизведения и объяснения информации, не может показать интеллектуальные навыки решения простых задач в области программирования.