

**ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ
«ИННОПРОГ»**

УТВЕРЖДАЮ

Генеральный директор ООО

«ИННОПРОГ»



подпись

инициалы, фамилия

01 февраля 2025 г.

Приказ об утверждении
программы

от 01 февраля 2025 г. № ОБР-3

**ДОПОЛНИТЕЛЬНАЯ ПРОФЕССИОНАЛЬНАЯ ПРОГРАММА
ПРОФЕССИОНАЛЬНОЙ ПЕРЕПОДГОТОВКИ**

«Data Science»

Вид программы:	дополнительная профессиональная программа профессиональной переподготовки
Новая квалификация:	специалист в области Data Science
Трудоемкость:	960 академических часов
Форма обучения:	очная, с применением электронного обучения и дистанционных образовательных технологий
Режим реализации:	12 месяцев (48 учебных недель), рекомендуемая нагрузка – 20 академических часов в неделю
Язык обучения:	русский

Иннополис, 2025

Содержание

1	Общая характеристика программы	5
1.1	Наименование программы	5
1.2	Нормативно-правовые основания разработки программы	5
1.3	Структура образовательной программы	6
1.4	Вид программы и целевая модель результата	7
1.5	Актуальность программы	8
1.6	Цель реализации программы	9
1.7	Задачи программы	9
1.8	Категория слушателей	10
1.9	Требования к уровню подготовки поступающего на обучение . .	10
1.10	Порядок приема, зачисления, отчисления, пересдачи и апелляции	11
1.11	Форма обучения, режим и язык обучения	11
1.12	Срок освоения и трудоемкость программы	12
1.13	Документ о квалификации	12
1.14	Характеристика новой квалификации и связанной с ней профес- сиональной деятельности	12
1.15	Связь Программы с профессиональными стандартами, трудовы- ми функциями и уровнями квалификации	13
1.16	Планируемые результаты освоения программы	13
1.16.1	Перечень формируемых профессиональных компетенций	13
1.16.2	Планируемые результаты обучения	16
2	Содержание программы	16
2.1	Учебный план	16
2.2	Календарный учебный график	18
2.3	Рабочие программы модулей	19
2.3.1	Модуль 1. Введение в профессию и организация обучения	22
2.3.2	Модуль 2. Python для анализа данных	24
2.3.3	Модуль 3. Продвинутый Python и ООП для Data Science	27
2.3.4	Модуль 4. Git, Linux и воспроизводимая среда	30
2.3.5	Модуль 5. SQL и PostgreSQL для аналитики	32
2.3.6	Модуль 6. Математика, статистика и продуктовые метрики	35

2.3.7	Модуль 7. Подготовка, очистка и исследовательский анализ данных	38
2.3.8	Модуль 8. Машинное обучение	41
2.3.9	Модуль 9. Глубокое обучение, NLP, компьютерное зрение и генеративный ИИ	45
2.3.10	Модуль 10. MLOps, API, деплой и мониторинг моделей .	48
2.3.11	Модуль 11. Проектная практика и портфолио	51
2.3.12	Модуль 12. Итоговая аттестация	54
2.4	Проектная практика и консультации	55
3	Организационно-педагогические условия реализации программы	56
3.1	Общие условия реализации	56
3.2	Материально-технические условия	57
3.3	Электронная информационно-образовательная среда	57
3.4	Кадровые условия	57
3.5	Учебно-методическое обеспечение	57
3.6	Применение электронного обучения и дистанционных образовательных технологий	58
3.7	Информационная безопасность, персональные данные и этика .	58
3.8	Локальные процедуры обучения	58
3.9	Обновление программы	59
4	Оценка качества освоения программы	59
4.1	Общие положения	59
4.2	Текущий контроль успеваемости	59
4.3	Промежуточная аттестация	60
4.4	Пересдача, апелляция и подтверждение самостоятельности . . .	61
4.5	Допуск к итоговой аттестации	61
4.6	Итоговая аттестация	61
4.7	Внутренняя и внешняя оценка качества реализации программы	62
5	Методические материалы	63
5.1	Методические указания для слушателей	63
5.2	Методические указания для преподавателей и проверяющих . .	63
5.3	Рекомендуемая литература и ресурсы	63

Приложение 1. Матрица формирования компетенций	64
Приложение 2. Оценочные материалы промежуточной аттестации .	67
Приложение 3. Оценочные материалы итоговой аттестации	69
Приложение 4. Перечень локальных нормативных актов, на которые должна опираться реализация программы	73
Приложение 5. Лист актуализации программы	74
Приложение 6. Требования к репозиторию, README и демонстрации	75
Приложение 7. Дорожная карта подготовки выпускного проекта . .	83
Приложение 8. Контрольный лист соответствия требованиям про- граммы	84

1. Общая характеристика программы

1.1. Наименование программы

Дополнительная профессиональная программа профессиональной переподготовки «Data Science» (далее – Программа) разработана для реализации обществом с ограниченной ответственностью «Иннопрог» в рамках дополнительного профессионального образования.

Программа относится к программам профессиональной переподготовки и направлена на получение слушателями компетенции, необходимой для выполнения нового вида профессиональной деятельности в сфере анализа данных, машинного обучения, искусственного интеллекта и разработки прикладных аналитических решений.

1.2. Нормативно-правовые основания разработки программы

Программа разработана в соответствии с:

- Федеральным законом от 29.12.2012 № 273-ФЗ «Об образовании в Российской Федерации»;
- приказом Министерства науки и высшего образования Российской Федерации от 24.03.2025 № 266 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам»;
- постановлением Правительства Российской Федерации от 11.10.2023 № 1678 «Об утверждении Правил применения организациями, осуществляющими образовательную деятельность, электронного обучения, дистанционных образовательных технологий при реализации образовательных программ»;
- постановлением Правительства Российской Федерации от 18.09.2020 № 1490 «О лицензировании образовательной деятельности»;
- постановлением Правительства Российской Федерации от 26.08.2013 № 729 «О федеральной информационной системе «Федеральный реестр сведений о документах об образовании и (или) о квалификации, документах об обучении»»;

- Федеральным законом от 27.07.2006 № 152-ФЗ «О персональных данных»;
- приказом Министерства труда и социальной защиты Российской Федерации от 06.07.2020 № 405н «Об утверждении профессионального стандарта «Специалист по большим данным»»;
- приказом Министерства труда и социальной защиты Российской Федерации от 20.07.2022 № 424н «Об утверждении профессионального стандарта «Программист»»;
- приказом Министерства труда и социальной защиты Российской Федерации от 18.01.2017 № 44н «Об утверждении профессионального стандарта «Разработчик Web и мультимедийных приложений»»;
- ГОСТ Р 7.0.97-2025 «Система стандартов по информации, библиотечному и издательскому делу. Организационно-распорядительная документация. Требования к оформлению документов»;
- локальными нормативными актами ООО «Иннопрог», регулирующими реализацию дополнительных профессиональных программ, применение электронного обучения и дистанционных образовательных технологий, порядок проведения аттестации, обработку персональных данных и выдачу документов о квалификации.

1.3. Структура образовательной программы

Настоящая Программа «Data Science» оформлена как комплекс основных характеристик образования, организационно-педагогических условий, оценочных и методических материалов, необходимых для реализации дополнительной профессиональной программы профессиональной переподготовки.

В состав Программы включены:

- общая характеристика Программы, включая цель, задачи, категорию слушателей, требования к поступающим, форму обучения, срок освоения, трудоемкость, документ о квалификации и характеристику новой квалификации;
- учебный план, определяющий перечень модулей, трудоемкость, распределение часов по видам учебной работы, формы промежуточной и итоговой аттестации;

- календарный учебный график, определяющий рекомендуемую последовательность освоения модулей и период итоговой аттестации;
- рабочие программы модулей, раскрывающие цели, результаты, тематическое содержание, самостоятельную работу, проектные задания и формы контроля;
- организационно-педагогические условия реализации, включая кадровые, материально-технические, информационно-методические условия, электронную информационно-образовательную среду и порядок применения электронного обучения и дистанционных образовательных технологий;
- оценочные материалы для текущего контроля, промежуточной и итоговой аттестации, критерии допуска, пересдачи, апелляции и подтверждения самостоятельности выполнения работ;
- методические материалы и приложения, обеспечивающие единообразное применение Программы, проверку проектных работ, оформление результатов и актуализацию документа.

Структура Программы обеспечивает проверяемость объема, содержания, планируемых результатов, условий реализации, форм аттестации и документов, выдаваемых по итогам обучения.

1.4. Вид программы и целевая модель результата

Программа является дополнительной профессиональной программой профессиональной переподготовки. Целевая модель результата освоения Программы состоит в подготовке специалиста, способного самостоятельно выполнять прикладной Data Science проект от постановки задачи и анализа данных до построения модели, интерпретации результата, оформления репозитория, подготовки демонстрационного прототипа и защиты решения перед заказчиком или аттестационной комиссией.

Итоговый результат освоения Программы выражается в сформированном портфолио слушателя, которое включает учебные notebooks, SQL-запросы, аналитические отчеты, модели машинного обучения, нейросетевые прототипы, воспроизводимый репозиторий, файл зависимостей, инструкцию запуска, пояснительную записку и материалы защиты. Такая модель результата позволяет

оценивать освоение программы не только по отдельным тестам, но и по способности слушателя выполнить полный цикл работы с данными.

Профессиональная деятельность выпускника ориентирована на прикладные задачи организаций, где требуется анализировать данные, формулировать гипотезы, рассчитывать метрики, строить прогнозные модели, проверять качество результатов, объяснять ограничения и готовить решения для дальнейшего внедрения. В рамках Программы Data Science рассматривается как междисциплинарная область на стыке программирования, аналитики, статистики, машинного обучения, инженерии данных и коммуникации с заказчиком.

1.5. Актуальность программы

Актуальность Программы определяется ростом спроса на специалистов, которые умеют переводить данные организаций в проверяемые аналитические выводы, прогнозные модели и прикладные цифровые решения. Для бизнеса Data Science важен в задачах прогнозирования спроса, сегментации клиентов, оценки рисков, персонализации предложений, анализа качества продуктов, автоматизации обработки текстов и изображений, поддержки управленческих решений.

Практическая ценность специалиста определяется комплексом навыков. Требуется корректно поставить задачу, проверить качество данных, выбрать метрики, исключить утечки, сравнить несколько решений, объяснить результат, обеспечить воспроизводимый запуск и показать, как модель может применяться в реальном процессе. Поэтому Программа построена как последовательная проектная траектория, где каждый модуль формирует часть будущего выпускного проекта.

Особенность направления Data Science состоит в необходимости соединять техническую реализацию с исследовательской дисциплиной. Слушатель должен понимать источник данных, границы применимости выборки, риск смещения, правила подготовки обучающей и тестовой частей, значение метрик, надежность статистических выводов и требования к воспроизводимости. Без этих элементов модель может показывать формально высокую метрику и одновременно давать слабый управленческий эффект. Поэтому в Программе отдельно выделены статистика, качество данных, интерпретация, инженерная подготовка модели, документация и защита проекта.

Программа также учитывает изменение практики работы с данными. В современных проектах Data Science специалисту требуется не только строить модель в notebook, но и уметь подготовить данные из SQL-источников, оформить эксперименты, сохранить артефакты, написать минимальный API для инференса, контейнеризовать решение, описать ограничения и представить результат в форме, понятной бизнес-заказчику. Эти требования отражены в учебном плане, проектной практике и итоговой аттестации.

1.6. Цель реализации программы

Цель Программы – сформировать у слушателей профессиональные компетенции, необходимые для выполнения нового вида профессиональной деятельности в области Data Science, включая анализ и подготовку данных, статистическую проверку гипотез, построение и оценку моделей машинного обучения, применение нейронных сетей и генеративного ИИ, а также инженерное оформление аналитического решения.

1.7. Задачи программы

Для достижения цели Программы решаются следующие задачи:

- сформировать базовые и продвинутые навыки программирования на Python для анализа данных;
- обучить работе с табличными данными, файлами, базами данных, SQL и PostgreSQL;
- сформировать навыки математической статистики, проверки гипотез, A/B-тестирования и расчета продуктовых метрик;
- обучить методам очистки, профилирования, визуализации и исследовательского анализа данных;
- сформировать навыки построения, сравнения и интерпретации моделей машинного обучения;
- дать практику применения глубокого обучения, NLP, компьютерного зрения и генеративного ИИ;
- обучить воспроизводимому оформлению проекта с использованием Git, Linux, Docker, тестирования и документации;

- сформировать навыки подготовки ML-сервиса, API, базового мониторинга и модели к демонстрационному применению;
- подготовить слушателя к выполнению и защите комплексного выпускного проекта.

1.8. Категория слушателей

К освоению Программы допускаются лица, имеющие среднее профессиональное и (или) высшее образование, а также лица, получающие среднее профессиональное и (или) высшее образование. Программа ориентирована на начинающих специалистов, аналитиков, разработчиков, студентов, предпринимателей и сотрудников организаций, которым необходимо освоить инструменты анализа данных и машинного обучения для решения прикладных задач.

Программа подходит слушателям, которые планируют перейти в область Data Science с позиций аналитика данных, Python-разработчика, BI-специалиста, инженера, экономиста, продуктового менеджера или начинающего исследователя данных. Допускается разный стартовый уровень подготовки, поскольку начальные модули включают Python, работу с файлами, таблицами, Git, Linux и SQL. При этом слушатель должен быть готов регулярно выполнять самостоятельные задания, читать техническую документацию, оформлять выводы в письменном виде и дорабатывать проект по замечаниям преподавателя.

1.9. Требования к уровню подготовки поступающего на обучение

Для успешного освоения Программы рекомендуется уверенное владение компьютером, готовность работать с математическими понятиями, таблицами, графиками и текстовыми отчетами, наличие персонального компьютера или ноутбука с доступом к сети Интернет, возможность устанавливать программное обеспечение, работать с браузером, терминалом и средой разработки. Начальный опыт программирования полезен, но не является обязательным.

Рекомендуемый минимальный уровень включает базовую математическую грамотность, понимание процентов и средних величин, навык работы с электронными таблицами, готовность использовать командную строку, устанавливать Python-пакеты и выполнять задания в репозитории. При зачислении может проводиться входная диагностика. Ее результат используется для реко-

мендации дополнительных материалов, индивидуальной траектории повторения и выбора сложности первых проектных заданий.

1.10. Порядок приема, зачисления, отчисления, пересдачи и апелляции

Прием слушателей осуществляется на основании заявления, договора об образовании и документов, подтверждающих наличие или получение среднего профессионального и (или) высшего образования. Конкретный порядок приема, зачисления, перевода, отчисления и восстановления определяется локальными нормативными актами ООО «Иннопрог» и условиями договора. Порядок пересдачи промежуточной и итоговой аттестации, сроки повторной проверки, требования к доработке проекта и порядок подачи апелляции фиксируются в локальном акте о текущем контроле, промежуточной и итоговой аттестации.

Зачисление оформляется распорядительным актом организации. До начала обучения слушателю предоставляется доступ к электронной информационно-образовательной среде, учебному плану, графику, правилам выполнения заданий и каналам коммуникации. Отчисление может производиться по инициативе слушателя, по соглашению сторон, при нарушении условий договора или локальных актов, а также при систематическом невыполнении учебного графика после уведомления слушателя и предоставления возможности устранить задолженность.

При неудовлетворительном результате промежуточной аттестации слушателю предоставляется возможность пересдачи или доработки задания в порядке, установленном локальным актом. Апелляция подается в письменной или электронной форме с указанием конкретных критериев, с которыми слушатель не согласен. Рассмотрение апелляции осуществляется комиссией или уполномоченным лицом с оформлением решения в электронной информационно-образовательной среде или ином установленном организацией порядке.

1.11. Форма обучения, режим и язык обучения

Обучение по Программе осуществляется в очной форме с применением электронного обучения и дистанционных образовательных технологий. Электронное обучение и дистанционные образовательные технологии применяются при проведении учебных занятий, консультаций, самостоятельной работы,

текущего контроля, промежуточной аттестации, проектной практики и итоговой аттестации при условии обеспечения доступа слушателя к электронной информационно-образовательной среде, учебным материалам, заданиям, средствам коммуникации и процедурам идентификации слушателя.

1.12. Срок освоения и трудоемкость программы

Общая трудоемкость Программы составляет 960 академических часов. Один академический час принимается равным 45 минутам. Срок освоения при рекомендуемой нагрузке составляет 12 месяцев (48 учебных недель). Трудоемкость включает лекционные занятия, практические занятия, самостоятельную работу, промежуточную аттестацию, проектную практику и итоговую аттестацию.

1.13. Документ о квалификации

Лицам, успешно освоившим Программу и прошедшим итоговую аттестацию, выдается диплом о профессиональной переподготовке установленного организацией образца. Лицам, осваивающим среднее профессиональное или высшее образование, диплом о профессиональной переподготовке выдается одновременно с получением соответствующего документа об образовании и о квалификации. Сведения о выданных документах о квалификации подлежат внесению в федеральную информационную систему «Федеральный реестр сведений о документах об образовании и (или) о квалификации, документах об обучении».

1.14. Характеристика новой квалификации и связанной с ней профессиональной деятельности

Выпускник, освоивший Программу, может выполнять трудовые действия, связанные с анализом данных, подготовкой аналитических витрин, построением моделей машинного обучения, разработкой прототипов Data Science решений, подготовкой отчетов, визуализаций, API для модели и воспроизводимых проектных материалов.

1.15. Связь Программы с профессиональными стандартами, трудовыми функциями и уровнями квалификации

Таблица 1 – Связь программы с профессиональными стандартами и трудовыми функциями

Профессиональный ориентир	Связанные трудовые действия и функции	Модули
Профессиональный стандарт «Специалист по большим данным», утвержденный приказом Минтруда России от 06.07.2020 № 405н	анализ больших данных, подготовка данных, проведение аналитического исследования, применение технологий больших данных	модули 5-12
Профессиональный стандарт «Программист», утвержденный приказом Минтруда России от 20.07.2022 № 424н	разработка программных компонентов, отладка, тестирование, сопровождение программного кода	модули 2-4, 8-11
Профессиональный стандарт «Разработчик Web и мультимедийных приложений», утвержденный приказом Минтруда России от 18.01.2017 № 44н	создание простых web-компонентов, интерфейсов демонстрации и взаимодействия с web-сервисами	модули 10-12

1.16. Планируемые результаты освоения программы

1.16.1. Перечень формируемых профессиональных компетенций

Таблица 2 – Перечень формируемых профессиональных компетенций

Код	Компетенция	Модули формирования	Средства оценки
ПК-1	Анализировать бизнес-задачу, формулировать аналитическую или ML-постановку, выбирать критерии успеха и ограничения проекта.	все модули, проектная практика, итоговая аттестация	практические задания, зачеты, проект, защита

Код	Компетенция	Модули формирования	Средства оценки
ПК-2	Собирать, описывать, профилировать и проверять качество данных из файлов, баз данных, API и иных источников.	модули 2, 5, 7, 11	практические задания, зачеты, проект, защита
ПК-3	Использовать Python, NumPy, pandas и инженерные практики для подготовки данных и автоматизации аналитических пайплайнов.	модули 2, 3, 7, 11	практические задания, зачеты, проект, защита
ПК-4	Работать с SQL и PostgreSQL, строить аналитические витрины, агрегировать данные и готовить выборки для анализа и моделирования.	модули 5, 7, 11	практические задания, зачеты, проект, защита
ПК-5	Применять математическую статистику, проверку гипотез, A/B-тестирование и продуктовые метрики для принятия решений на основе данных.	модуль 6, модули 7, 11	практические задания, зачеты, проект, защита
ПК-6	Строить, оценивать, сравнивать и интерпретировать модели машинного обучения для задач классификации, регрессии, кластеризации и прогнозирования.	модули 8, 11, 12	практические задания, зачеты, проект, защита

Код	Компетенция	Модули формирования	Средства оценки
ПК-7	Применять методы глубокого обучения, NLP, компьютерного зрения и генеративного ИИ в прикладных задачах с учетом ограничений данных и моделей.	модули 9, 11, 12	практические задания, зачеты, проект, защита
ПК-8	Готовить модель к использованию: сохранять артефакты, создавать API, контейнеризовать сервис и описывать воспроизводимый запуск.	модули 4, 10, 11, 12	практические задания, зачеты, проект, защита
ПК-9	Визуализировать данные, готовить аналитические отчеты, дашборды и аргументированные выводы для заказчика.	модули 6, 7, 11, 12	практические задания, зачеты, проект, защита
ПК-10	Обеспечивать воспроизводимость, тестирование, контроль версий, документирование и качество аналитического кода.	модули 3, 4, 10, 11	практические задания, зачеты, проект, защита
ПК-11	Соблюдать требования информационной безопасности, защиты персональных данных, этики и ответственного использования моделей.	модули 1, 7, 9, 10, 11	практические задания, зачеты, проект, защита
ПК-12	Представлять результат Data Science проекта, защищать решения, отвечать на вопросы по данным, методам, метрикам и бизнес-эффекту.	модули 11, 12	практические задания, зачеты, проект, защита

1.16.2. Планируемые результаты обучения

Таблица 3 – Планируемые результаты обучения

Группа результата	Содержание результата
Знать	этапы жизненного цикла Data Science проекта, виды данных, требования к качеству данных, основы статистики, методы ML и DL, принципы интерпретации моделей, требования к воспроизводимости, безопасности и документации.
Уметь	формулировать постановку задачи, собирать и очищать данные, писать Python и SQL код, проводить EDA, строить и сравнивать модели, проверять гипотезы, разрабатывать API для модели, оформлять отчет и презентацию.
Владеть	инструментами Python, Jupyter, pandas, NumPy, scikit-learn, PyTorch или TensorFlow, PostgreSQL, Git, Linux, Docker, MLflow или DVC, средствами визуализации и практиками командной работы.
Иметь опыт	выполнения end-to-end проектов: от описания задачи и подготовки данных до модели, интерпретации, воспроизводимого запуска и защиты результата перед комиссией.

2. Содержание программы

2.1. Учебный план

Таблица 4 – Учебный план

№	Наименование модуля	Всего	Лекц.	Практ.	Сам. раб.	ПА	ИА	Форма контроля
1	Введение в профессию и организация обучения	12	6	3	3	0	0	входной контроль
2	Python для анализа данных	84	24	39	14	7	0	зачет
3	Продвинутый Python и ООП для Data Science	64	18	28	11	7	0	зачет
4	Git, Linux и воспроизводимая среда	40	12	18	7	3	0	зачет
5	SQL и PostgreSQL для аналитики	76	21	34	14	7	0	зачет
6	Математика, статистика и продуктовые метрики	88	27	34	20	7	0	зачет
7	Подготовка, очистка и исследовательский анализ данных	96	24	44	21	7	0	зачет
8	Машинное обучение	136	37	61	31	7	0	зачет
9	Глубокое обучение, NLP, компьютерное зрение и генеративный ИИ	124	34	59	24	7	0	зачет

Продолжение таблицы 4

№	Наименование модуля	Всего	Лекц.	Практ.	Сам. раб.	ПА	ИА	Форма контроля
10	MLOps, API, деплой и мониторинг моделей	64	18	28	14	4	0	зачет
11	Проектная практика и портфолио	136	0	99	37	0	0	допуск к ИА
12	Итоговая аттестация	40	0	0	0	0	40	защита проекта
Итого		960	221	447	196	56	40	–

Промежуточная аттестация проводится по завершении модулей, определенных учебным планом. Итоговая аттестация проводится отдельно от промежуточной аттестации в форме защиты итогового аттестационного проекта.

2.2. Календарный учебный график

Начало обучения определяется по мере комплектования учебной группы. Ниже представлен рекомендуемый график при сроке обучения 12 месяцев (48 недель).

Таблица 5 – Календарный учебный график

№	Наименование модуля	Период обучения	Часы
1	Введение в профессию и организация обучения	1 неделя	12
2	Python для анализа данных	2-5 недели	84
3	Продвинутый Python и ООП для Data Science	6-8 недели	64
4	Git, Linux и воспроизводимая среда	9-10 недели	40
5	SQL и PostgreSQL для аналитики	11-14 недели	76
6	Математика, статистика и продуктовые метрики	15-18 недели	88
7	Подготовка, очистка и исследовательский анализ данных	19-23 недели	96

№	Наименование модуля	Период обучения	Часы
8	Машинное обучение	24-30 недели	136
9	Глубокое обучение, NLP, компьютерное зрение и генеративный ИИ	31-36 недели	124
10	MLOps, API, деплой и мониторинг моделей	37-39 недели	64
11	Проектная практика и портфолио	40-46 недели	136
12	Итоговая аттестация	47-48 недели	40
Итого			960

2.3. Рабочие программы модулей

В рабочих программах модулей определяются цели, результаты, тематическое содержание, формы организации занятий, самостоятельной работы и контроля. Перечень тем соотносится с фактической структурой учебных материалов направления «Data Science».

Проектная траектория программы. Практическая часть Программы организована через последовательную разработку проектных работ. Сложность работ повышается от анализа таблиц и SQL-отчетов к статистическим исследованиям, ML-моделям, нейросетевым прототипам, API и воспроизводимому выпускному проекту.

Таблица 6 – Проектная траектория программы

№	Проект	Основные инструменты	Проверяемый результат
1	Аналитический отчет по продажам или заказам	Python, pandas, визуализация	очистка данных, расчет метрик, выводы
2	SQL-витрина клиентских транзакций	SQL, PostgreSQL	запросы, объединение таблиц, оконные функции
3	EDA качества данных	pandas, визуализация, статистика	профилирование, пропуски, выбросы, карта рисков
4	A/B-тест продуктовой гипотезы	статистика, метрики	гипотеза, расчет показателей, интерпретация

№	Проект	Основные инструменты	Проверяемый результат
5	Дашборд бизнес-метрик	визуализация, BI-подход	метрики, сегменты, графики, аналитический комментарий
6	Модель прогнозирования спроса или цены	регрессия, feature engineering	baseline, метрики, анализ ошибок
7	Модель классификации оттока или заявок	классификация, интерпретация	ROC AUC, F1, настройка порога, SHAP
8	Кластеризация клиентских сегментов	unsupervised learning	сегменты, профили, бизнес-интерпретация
9	Прогноз временного ряда	time series	лаги, сезонность, backtesting
10	NLP-классификатор отзывов или обращений	NLP, transformers	токенизация, модель, метрики, ошибки
11	CV-прототип классификации изображений	нейронные сети, transfer learning	обучение, валидация, ограничения
12	ML-сервис с API и Docker	FastAPI, Docker, логирование	инференс, контейнер, тестовый запрос, README
13	Итоговый Data Science проект	end-to-end DS	данные, EDA, модель, интерпретация, демонстрация, защита

Сквозные требования к проектным работам. Каждая проектная работа, начиная со второго модуля, должна иметь воспроизводимую структуру. Минимальный комплект материалов включает исходные данные или инструкцию их получения, код, результаты расчета, краткий аналитический вывод и описание ограничений. Для проектов, где используются персональные данные или чувствительная информация, применяются обезличенные, учебные или открытые наборы данных.

Таблица 7 – Сквозные требования к оформлению проектных работ

№	Компонент	Обязательное содержание	Проверяемый результат
1	Постановка задачи	описание бизнес-вопроса, целевого показателя, ограничений, ожидаемого результата	задача понятна без устного пояснения
2	Данные	источник, состав полей, объем, период, ограничения, правила использования	данные можно проверить и повторно загрузить
3	Качество данных	пропуски, дубликаты, выбросы, типы данных, риск смещения и утечек	риски данных выявлены до моделирования
4	Код	структура, функции, notebook или модули, комментарии, единый стиль	код читается и запускается проверяющим
5	Метрики	выбор метрик, baseline, критерии успешности и способ сравнения	качество результата измерено корректно
6	Выводы	интерпретация результата, ограничения, рекомендации и дальнейшие шаги	выводы основаны на расчетах и графиках
7	Репозиторий	README, зависимости, структура папок, команды запуска, сведения об авторе	проект воспроизводится по инструкции
8	Самостоятельность	история коммитов, индивидуальные решения, ответы на вопросы преподавателя	подтверждена авторская работа слушателя

2.3.1. Модуль 1. Введение в профессию и организация обучения

Трудоемкость модуля – 12 академических часов. В том числе лекционные занятия – 6 ч., практические занятия – 3 ч., самостоятельная работа – 3 ч., промежуточная аттестация – 0 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать представление о профессии, правилах обучения, проектной траектории и требованиях к результатам.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 8 – Тематический план модуля 1. Введение в профессию и организация обучения

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Профессия Data Science и жизненный цикл аналитического проекта	4	2	1	1	практическое задание
2	Организация обучения, ЭИОС, расписание, правила коммуникации	4	2	1	1	практическое задание
3	Требования к проектам, репозиторию, данным, отчетам и защите	3	1	1	1	проектное задание
4	Входная диагностика: Python, математика, аналитическое мышление	1	1	0	0	практическое задание

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. входной контроль. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «карта профессиональной траектории Data Science специалиста и индивидуальный план проектного результата». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- определить целевую роль и типы задач Data Science;
- создать структуру учебного репозитория;
- оформить README с правилами запуска материалов;
- пройти входную диагностику по Python, математике и аналитическому мышлению;

Чек-лист результата модуля:

- репозиторий создан и содержит разделы для модулей;
- слушатель понимает требования к портфолио;
- оформлен индивидуальный план итогового проекта;
- зафиксированы пробелы для самостоятельного повторения;

Типовые ошибки, подлежащие исправлению до зачета:

- отсутствует структура репозитория;
- не указаны цель обучения и ожидаемый проект;
- нет инструкции по запуску материалов;
- результаты диагностики не разобраны;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры

репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель понимает, из каких этапов состоит Data Science проект, какие материалы входят в портфолио, какие требования предъявляются к данным, коду, отчетам и защите.

2.3.2. Модуль 2. Python для анализа данных

Трудоемкость модуля – 84 академических часов. В том числе лекционные занятия – 24 ч., практические занятия – 39 ч., самостоятельная работа – 14 ч., промежуточная аттестация – 7 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать базовые навыки Python для анализа данных, обработки файлов, работы с таблицами и подготовки аналитического отчета.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 9 – Тематический план модуля 2. Python для анализа данных

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Рабочая среда: Python, Jupyter, VS Code, виртуальные окружения	9	3	4	2	практическое задание
2	Переменные, типы данных, ввод и вывод, f-строки	9	3	4	2	практическое задание
3	Условия, циклы, функции, декомпозиция аналитической задачи	8	2	4	2	практическое задание

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
4	Списки, словари, множества, кортежи и вложенные структуры	7	2	4	1	практическое задание
5	Работа с файлами CSV, JSON, TXT, кодировки и базовая валидация	7	2	4	1	практическое задание
6	Исключения, обработка ошибок и диагностические сообщения	8	2	4	1	практическое задание
7	NumPy: массивы, индексация, векторные операции	6	2	3	1	практическое задание
8	pandas: DataFrame, Series, загрузка, фильтрация и сортировка	6	2	3	1	практическое задание
9	Группировки, агрегации, объединение таблиц и простые отчеты	6	2	3	1	проектное задание
10	Ноутбук как аналитический документ: структура, выводы, графики	6	2	3	1	практическое задание
11	Мини-проект: отчет по продажам, заказам или складским остаткам	6	2	3	1	проектное задание
12	Промежуточная аттестация по модулю	6	0	0	0	зачет

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «аналитический отчет по продажам, заказам, посещениям или складским остаткам». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- загрузить данные из CSV, JSON или Excel;
- очистить пропуски и дубликаты;
- рассчитать агрегированные показатели;
- построить базовые графики и сформулировать выводы;

Чек-лист результата модуля:

- код разделен на логические блоки;
- таблицы загружаются без ручного изменения исходника;
- расчеты воспроизводимы;
- выводы связаны с полученными числами;

Типовые ошибки, подлежащие исправлению до зачета:

- результат зависит от локального абсолютного пути;
- нет проверки структуры данных;
- выводы не подтверждаются расчетами;
- исходные данные изменены вручную без описания;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель умеет использовать Python, NumPy и pandas для первичной обработки данных, строит понятный notebook и оформляет аналитический вывод.

2.3.3. Модуль 3. Продвинутый Python и ООП для Data Science

Трудоемкость модуля – 64 академических часов. В том числе лекционные занятия – 18 ч., практические занятия – 28 ч., самостоятельная работа – 11 ч., промежуточная аттестация – 7 ч., итоговая аттестация – 0 ч.

Цель модуля. развить навыки инженерного оформления аналитического кода, модульной структуры, тестирования, логирования и повторного использования компонентов.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 10 – Тематический план модуля 3. Продвинутый Python и ООП для Data Science

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Функции высшего порядка, итераторы, генераторы и comprehensions	9	3	4	2	практическое задание
2	Модули, пакеты, импорт, структура аналитического проекта	9	3	4	2	практическое задание
3	ООП: классы, методы, инкапсуляция и модели предметной области	8	2	4	2	практическое задание
4	dataclass, типизация, аннотации и читаемость кода	7	2	4	1	практическое задание

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
5	Логирование, конфигурации, переменные окружения и секреты	6	2	3	1	практическое задание
6	Тестирование аналитического кода с использованием pytest	6	2	3	1	практическое задание
7	Мини-пайплайн подготовки данных как Python-пакет	6	2	3	1	практическое задание
8	Код-ревью, стиль, форматирование и статический анализ	6	2	3	1	практическое задание
9	Промежуточная аттестация по модулю	7	0	0	0	зачет

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «мини-пакет для загрузки, проверки и подготовки набора данных». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- выделить функции загрузки и проверки данных;
- описать конфигурации проекта;

- добавить логирование ключевых шагов;
- покрыть критические функции тестами;

Чек-лист результата модуля:

- код вынесен из notebook в модули;
- есть файл зависимостей;
- тесты запускаются одной командой;
- ошибки данных обрабатываются явно;

Типовые ошибки, подлежащие исправлению до зачета:

- вся логика остается в одном notebook;
- нет тестов на ключевые преобразования;
- исключения подавляются без сообщения;
- секреты или локальные пути записаны в коде;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель оформляет аналитический код как поддерживаемый Python-проект, применяет типизацию, тестирование, логирование и повторно используемые функции.

2.3.4. Модуль 4. Git, Linux и воспроизводимая среда

Трудоемкость модуля – 40 академических часов. В том числе лекционные занятия – 12 ч., практические занятия – 18 ч., самостоятельная работа – 7 ч., промежуточная аттестация – 3 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать навыки работы с Git, Linux, Docker, зависимостями и воспроизводимым запуском проекта.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 11 – Тематический план модуля 4. Git, Linux и воспроизводимая среда

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Git: репозиторий, коммиты, ветки, merge, pull request	9	3	4	2	практическое задание
2	Работа с удаленным репозиторием, README, .gitignore и история проекта	9	3	4	2	практическое задание
3	Linux CLI: файловая система, права, процессы, архивы	7	2	4	1	практическое задание
4	Виртуальные окружения, зависимости, requirements и pyproject	6	2	3	1	практическое задание
5	Docker для аналитика: образ, контейнер, volume, запуск ноутбука	6	2	3	1	практическое задание
6	Промежуточная аттестация по модулю	3	0	0	0	зачет

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «воспроизводимая среда Data Science проекта с Git, Linux, Docker и инструкцией запуска». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- создать ветки и осмысленную историю коммитов;
- подготовить окружение и файл зависимостей;
- описать команды запуска;
- собрать контейнер или воспроизводимую локальную среду;

Чек-лист результата модуля:

- репозиторий не содержит временных файлов и секретов;
- зависимости устанавливаются по инструкции;
- команды запуска проверены;
- README описывает структуру проекта;

Типовые ошибки, подлежащие исправлению до зачета:

- нет коммитов по этапам работы;
- зависимости устанавливаются вручную без фиксации;
- данные и артефакты расположены хаотично;
- инструкция запуска непроверяема;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель умеет готовить среду, которая позволяет преподавателю или заказчику запустить проект и проверить результат без устных пояснений.

2.3.5. Модуль 5. SQL и PostgreSQL для аналитики

Трудоемкость модуля – 76 академических часов. В том числе лекционные занятия – 21 ч., практические занятия – 34 ч., самостоятельная работа – 14 ч., промежуточная аттестация – 7 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать умение извлекать, преобразовывать и анализировать данные в PostgreSQL, строить аналитические витрины и готовить выборки для моделей.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 12 – Тематический план модуля 5. SQL и PostgreSQL для аналитики

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Реляционная модель, таблицы, ключи, типы данных	9	3	4	2	практическое задание
2	SELECT, WHERE, ORDER BY, LIMIT и базовая фильтрация	8	2	4	2	практическое задание

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
3	JOIN, UNION, под-запросы и работа с несколькими таблицами	8	2	4	2	практическое задание
4	GROUP BY, HAVING, агрегаты и аналитические витрины	8	2	4	2	практическое задание
5	Оконные функции, ранжирование, скользящие показатели	6	2	3	1	практическое задание
6	PostgreSQL: индексы, explain plan и базовая оптимизация	6	2	3	1	практическое задание
7	Проектирование схемы для аналитического кейса	6	2	3	1	практическое задание
8	Подключение PostgreSQL к Python, выгрузки и загрузки	6	2	3	1	практическое задание
9	Мини-проект: SQL-отчет по транзакциям или событиям	6	2	3	1	проектное задание
10	Контроль качества данных в базе	6	2	3	1	практическое задание
11	Промежуточная аттестация по модулю	7	0	0	0	зачет

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «SQL-витрина данных для аналитики клиентов, заказов, платежей или событий продукта». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- спроектировать простую схему таблиц;
- написать запросы с объединениями и агрегациями;
- применить оконные функции;
- выгрузить витрину для анализа в Python;

Чек-лист результата модуля:

- таблицы имеют понятные ключи;
- запросы читаемы и отформатированы;
- агрегации соответствуют бизнес-вопросу;
- витрина пригодна для EDA и моделирования;

Типовые ошибки, подлежащие исправлению до зачета:

- неверные соединения дают дубли;
- фильтры меняют смысл метрик;
- нет проверки количества строк после join;
- SQL-запросы не сохранены в репозитории;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель работает с PostgreSQL как с источником данных, строит аналитические выборки и понимает влияние SQL-логики на качество дальнейшей модели.

2.3.6. Модуль 6. Математика, статистика и продуктовые метрики

Трудоемкость модуля – 88 академических часов. В том числе лекционные занятия – 27 ч., практические занятия – 34 ч., самостоятельная работа – 20 ч., промежуточная аттестация – 7 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать статистическую основу для проверки гипотез, анализа экспериментов, расчета метрик и интерпретации результатов.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 13 – Тематический план модуля 6. Математика, статистика и продуктовые метрики

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Линейная алгебра для Data Science: векторы, матрицы, расстояния	8	3	3	2	практическое задание
2	Основы анализа: функции, производные, градиент и оптимизация	8	3	3	2	практическое задание
3	Вероятность, случайные величины и распределения	8	3	3	2	практическое задание
4	Описательная статистика, выбросы, квантили и распределения	7	2	3	2	практическое задание

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
5	Доверительные интервалы и оценивание параметров	7	2	3	2	практическое задание
6	Проверка гипотез, p-value, ошибки первого и второго рода	7	2	3	2	практическое задание
7	A/B-тестирование, размер выборки и мощность теста	7	2	3	2	практическое задание
8	Корреляция, регрессия и ограничения причинных выводов	7	2	3	2	практическое задание
9	Продуктовые метрики: конверсия, удержание, когорты, LTV	6	2	3	1	практическое задание
10	Интерпретация результатов для бизнес-заказчика	6	2	3	1	практическое задание
11	Мини-проект: проверка гипотезы по пользовательским данным	5	2	2	1	проектное задание
12	Промежуточная аттестация по модулю	6	0	0	0	зачет
13	Самостоятельная подготовка к зачету и доработка отчета	6	2	2	1	проектное задание

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «статистическое исследование продуктовой гипотезы с расчетом метрик и интерпретацией результата». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- сформулировать нулевую и альтернативную гипотезы;
- выбрать статистический критерий;
- рассчитать доверительные интервалы;
- подготовить вывод для бизнес-заказчика;

Чек-лист результата модуля:

- метрика определена до расчета;
- выбранный критерий соответствует типу данных;
- интерпретация учитывает размер эффекта;
- отдельно описаны ограничения выборки;

Типовые ошибки, подлежащие исправлению до зачета:

- p -value трактуется как вероятность правоты гипотезы;
- выбор критерия не объяснен;
- игнорируется множественная проверка;
- вывод сделан без учета практической значимости;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель использует статистику для проверки гипотез, различает статистическую и практическую значимость, корректно объясняет выводы.

2.3.7. Модуль 7. Подготовка, очистка и исследовательский анализ данных

Трудоемкость модуля – 96 академических часов. В том числе лекционные занятия – 24 ч., практические занятия – 44 ч., самостоятельная работа – 21 ч., промежуточная аттестация – 7 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать навыки профилирования, очистки, объединения источников, построения признаков и исследовательского анализа данных.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 14 – Тематический план модуля 7. Подготовка, очистка и исследовательский анализ данных

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Профилирование данных, типы признаков и карта источников	8	2	4	2	практическое задание
2	Пропуски, дубликаты, аномалии и правила очистки	8	2	4	2	практическое задание
3	Объединение источников, ключи, гранулярность и утечки данных	8	2	4	2	практическое задание
4	Работа с датами, временем, категориями и текстовыми полями	8	2	4	2	практическое задание

Продолжение таблицы 14

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
5	Feature engineering для табличных данных	8	2	4	2	практическое задание
6	Масштабирование, кодирование категорий и подготовка пайплайнов	7	2	3	2	практическое задание
7	Визуальный EDA: распределения, связи, сегменты и выбросы	7	2	3	2	практическое задание
8	Документирование предположений, ограничений и рисков данных	7	2	3	2	практическое задание
9	Автоматизация подготовки данных с pandas и sklearn Pipeline	6	2	3	1	практическое задание
10	Контроль качества данных: схемы, проверки и отчеты	6	2	3	1	проектное задание
11	Мини-проект: EDA и подготовка датасета для модели	6	2	3	1	проектное задание
12	Промежуточная аттестация по модулю	7	0	0	0	зачет
13	Доработка проекта по результатам ревью	5	1	3	1	проектное задание
14	Подготовка итогового EDA-отчета	5	1	3	1	проектное задание

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «полный EDA-отчет с профилированием качества данных, визуализациями и планом подготовки признаков». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- построить паспорт набора данных;
- оценить пропуски, выбросы и дубликаты;
- исследовать связи между признаками;
- описать стратегию подготовки признаков;

Чек-лист результата модуля:

- каждое преобразование данных объяснено;
- пропуски и выбросы обработаны осознанно;
- визуализации имеют подписи и выводы;
- сформирован список рисков данных;

Типовые ошибки, подлежащие исправлению до зачета:

- удаляются строки без оценки потерь;
- категориальные признаки кодируются без проверки;
- визуализации не отвечают на вопрос;
- есть утечка целевой переменной в признаки;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной ин-

терпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель проводит исследовательский анализ данных, готовит признаки и выявляет ограничения, которые влияют на качество моделей.

2.3.8. Модуль 8. Машинное обучение

Трудоемкость модуля – 136 академических часов. В том числе лекционные занятия – 37 ч., практические занятия – 61 ч., самостоятельная работа – 31 ч., промежуточная аттестация – 7 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать умение строить, оценивать, сравнивать и интерпретировать модели машинного обучения.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 15 – Тематический план модуля 8. Машинное обучение

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Жизненный цикл ML-проекта и постановка задачи	8	2	4	2	практическое задание
2	Разделение выборки, утечки данных и baseline-модель	8	2	4	2	практическое задание
3	Метрики классификации: accuracy, precision, recall, F1, ROC AUC	8	2	4	2	практическое задание
4	Метрики регрессии: MAE, RMSE, MAPE, R2	8	2	4	2	практическое задание

Продолжение таблицы 15

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
5	Линейные модели, регуляризация и интерпретация коэффициентов	7	2	3	2	практическое задание
6	Логистическая регрессия и настройка порога решения	7	2	3	2	практическое задание
7	Деревья решений, случайный лес и важность признаков	7	2	3	2	практическое задание
8	Градиентный бустинг, XGBoost, CatBoost и LightGBM	7	2	3	2	практическое задание
9	Кросс-валидация, подбор гиперпараметров и валидационные схемы	7	2	3	2	практическое задание
10	Кластеризация: k-means, DBSCAN, иерархические методы	7	2	3	2	практическое задание
11	Снижение размерности: PCA, t-SNE, UMAP и визуализация сегментов	7	2	3	2	практическое задание
12	Работа с несбалансированными классами	7	2	3	2	практическое задание
13	Интерпретация моделей: permutation importance, SHAP, partial dependence	6	2	3	1	практическое задание
14	Временные ряды: лаги, сезонность, backtesting и прогнозирование	6	2	3	1	практическое задание

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
15	Рекомендательные системы: похожесть, ranking и простые рекомендации	6	2	3	1	практическое задание
16	Анализ ошибок, устойчивость и ограничения модели	6	2	3	1	практическое задание
17	Мини-проект: end-to-end ML для бизнес-задачи	6	2	3	1	проектное задание
18	Модельная карта: данные, метрики, риски, интерпретация	6	2	3	1	проектное задание
19	Промежуточная аттестация по модулю	7	0	0	0	зачет
20	Доработка модели и подготовка к защите результата	5	1	3	1	проектное задание

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «модель машинного обучения для классификации, регрессии, сегментации или прогноза временного ряда». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- сформировать baseline;
- разделить данные на обучающую и тестовую части;
- обучить несколько моделей;
- оценить метрики и объяснить ошибки;

Чек-лист результата модуля:

- выбрана подходящая метрика;
- исключена утечка данных;
- есть сравнение baseline и улучшенной модели;
- показан анализ ошибок и важности признаков;

Типовые ошибки, подлежащие исправлению до зачета:

- модель оценивается на обучающих данных;
- нет baseline;
- метрика выбрана без связи с задачей;
- результат не интерпретирован для бизнеса;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель строит ML-модели, сравнивает подходы, оценивает качество и формулирует практические рекомендации по результатам модели.

2.3.9. Модуль 9. Глубокое обучение, NLP, компьютерное зрение и генеративный ИИ

Трудоемкость модуля – 124 академических часов. В том числе лекционные занятия – 34 ч., практические занятия – 59 ч., самостоятельная работа – 24 ч., промежуточная аттестация – 7 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать прикладные навыки работы с нейронными сетями, NLP, компьютерным зрением, предобученными моделями и генеративным ИИ.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 16 – Тематический план модуля 9. Глубокое обучение, NLP, компьютерное зрение и генеративный ИИ

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Нейронные сети: тензоры, слои, функции потерь и оптимизация	8	2	4	2	практическое задание
2	PyTorch: Dataset, DataLoader, модель, обучение и валидация	8	2	4	2	практическое задание
3	TensorFlow или Keras: быстрый прототип нейросетевой модели	8	2	4	2	практическое задание
4	Регуляризация, dropout, batch normalization и ранняя остановка	8	2	4	2	практическое задание
5	Компьютерное зрение: сверточные сети и классификация изображений	8	2	4	2	практическое задание

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
6	Transfer learning для изображений	8	2	4	2	практическое задание
7	NLP: токенизация, векторизация, TF-IDF и эмбединги	8	2	4	2	практическое задание
8	Трансформеры и модели для классификации текстов	7	2	4	1	практическое задание
9	Практика с Hugging Face и предобученными моделями	6	2	3	1	практическое задание
10	Генеративный ИИ: промпты, RAG-логика, ограничения и проверка фактов	6	2	3	1	практическое задание
11	Оценка качества генеративных ответов и риски галлюцинаций	6	2	3	1	практическое задание
12	Этика, персональные данные и безопасное применение моделей	6	2	3	1	практическое задание
13	Мини-проект NLP: классификация обращений или отзывов	6	2	3	1	проектное задание
14	Мини-проект CV: классификация изображений или дефектов	6	2	3	1	проектное задание
15	Интеграция предобученной модели в аналитический прототип	6	2	3	1	практическое задание
16	Промежуточная аттестация по модулю	7	0	0	0	зачет

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
17	Доработка нейросетевого проекта по результатам проверки	6	2	3	1	проектное задание
18	Подготовка модельного отчета и демонстрации	6	2	3	1	проектное задание

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «нейросетевой прототип для текста, изображения или генеративного сценария с анализом качества и ограничений». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- подготовить данные для NLP или CV;
- использовать transfer learning или предобученную модель;
- оценить качество на валидации;
- описать риски и ограничения использования;

Чек-лист результата модуля:

- данные подготовлены в корректном формате;
- модель обучается или донастраивается воспроизводимо;
- качество измеряется понятными метриками;
- описаны ошибки и этические ограничения;

Типовые ошибки, подлежащие исправлению до зачета:

- используется готовая модель без проверки качества;
- нет разделения данных;
- выводы основаны только на отдельных примерах;
- игнорируются персональные данные и ограничения генеративного ИИ;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель понимает базовые сценарии глубокого обучения, NLP, компьютерного зрения и генеративного ИИ, умеет оценивать качество и границы применимости.

2.3.10. Модуль 10. MLOps, API, деплой и мониторинг моделей

Трудоемкость модуля – 64 академических часов. В том числе лекционные занятия – 18 ч., практические занятия – 28 ч., самостоятельная работа – 14 ч., промежуточная аттестация – 4 ч., итоговая аттестация – 0 ч.

Цель модуля. сформировать навыки подготовки модели к использованию: сохранение артефактов, API, контейнеризация, мониторинг и управление экспериментами.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 17 – Тематический план модуля 10. MLOps, API, деплой и мониторинг моделей

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Артефакты ML-проекта: модель, пре-процессор, метаданные, версии	9	3	4	2	практическое задание
2	MLflow или DVC: эксперименты, параметры, метрики и артефакты	9	3	4	2	практическое задание
3	FastAPI для инференса модели	8	2	4	2	практическое задание
4	Валидация входных данных, обработка ошибок и логирование	8	2	4	2	практическое задание
5	Docker Compose для ML-сервиса и зависимостей	7	2	3	2	практическое задание
6	Базовый CI/CD: проверки, тесты и сборка контейнера	7	2	3	2	практическое задание
7	Мониторинг качества, data drift, model drift и журнал прогнозов	6	2	3	1	практическое задание
8	Безопасность API, секреты, персональные данные и ограничения доступа	6	2	3	1	практическое задание
9	Промежуточная аттестация по модулю	4	0	0	0	зачет

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. зачет. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «минимальный ML-сервис с API, Docker, логированием, тестовым запросом и описанием мониторинга». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- сохранить модель и препроцессинг;
- создать endpoint инференса;
- добавить проверку входных данных;
- подготовить Dockerfile и инструкцию запуска;

Чек-лист результата модуля:

- API возвращает прогноз в заданном формате;
- модель и preprocessing загружаются корректно;
- ошибки входных данных обрабатываются;
- README содержит команды запуска и тестовый запрос;

Типовые ошибки, подлежащие исправлению до зачета:

- модель работает только внутри notebook;
- API не проверяет входные данные;
- контейнер не собирается;
- нет логирования и описания мониторинга качества;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной ин-

терпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель готовит модель к демонстрационному применению, понимает базовые требования к инференсу, контейнеризации, тестированию и сопровождению.

2.3.11. Модуль 11. Проектная практика и портфолио

Трудоемкость модуля – 136 академических часов. В том числе лекционные занятия – 0 ч., практические занятия – 99 ч., самостоятельная работа – 37 ч., промежуточная аттестация – 0 ч., итоговая аттестация – 0 ч.

Цель модуля. обеспечить разработку комплексного выпускного проекта с данными, анализом, моделью, интерпретацией, демонстрацией и защитой.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 18 – Тематический план модуля 11. Проектная практика и портфолио

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Выбор темы, постановка бизнес-задачи и критерии успеха	11	0	7	4	проектная консультация
2	Поиск, описание и правовая проверка источников данных	11	0	7	4	проектная консультация
3	Проектирование структуры репозитория и окружения	11	0	7	4	проектная консультация
4	Сбор, загрузка и первичная валидация данных	11	0	7	4	проектная консультация

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
5	EDA, статистический анализ и формирование гипотез	13	0	10	3	проектная консультация
6	Feature engineering и подготовка обучающей выборки	13	0	10	3	проектная консультация
7	Baseline, выбор метрик и первая версия модели	13	0	10	3	проектная консультация
8	Улучшение модели, подбор параметров и анализ ошибок	13	0	10	3	проектная консультация
9	Интерпретация, риски, ограничения и модельная карта	10	0	7	3	проектная консультация
10	API, демонстрационный интерфейс или аналитический дашборд	13	0	10	3	проектная консультация
11	Подготовка отчета, презентации, README и инструкции	10	0	7	3	проектная консультация
12	Предзащита и доработка проекта	7	0	7	0	проектная консультация

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. допуск к ИА. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «комплексный портфолио-проект Data Science с данными, EDA, моделью, интерпретацией и

воспроизводимым запуском». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- утвердить тему и паспорт проекта;
- собрать и описать данные;
- построить baseline и улучшенную модель;
- подготовить репозиторий, отчет, презентацию и демонстрацию;

Чек-лист результата модуля:

- проект имеет ясную бизнес-задачу;
- данные описаны и проверены;
- модель сравнивается с baseline;
- результат можно воспроизвести по инструкции;

Типовые ошибки, подлежащие исправлению до зачета:

- проект сводится к запуску готового notebook;
- отсутствует самостоятельный анализ данных;
- нет связи метрик с задачей;
- репозиторий не позволяет проверить работу;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель собирает все освоенные элементы в цельный Data Science проект и готовит материалы для допуска к итоговой аттестации.

2.3.12. Модуль 12. Итоговая аттестация

Трудоемкость модуля – 40 академических часов. В том числе лекционные занятия – 0 ч., практические занятия – 0 ч., самостоятельная работа – 0 ч., промежуточная аттестация – 0 ч., итоговая аттестация – 40 ч.

Цель модуля. проверить готовность выпускника самостоятельно решать прикладные задачи Data Science и представлять результат комиссии.

Планируемые результаты освоения модуля. По завершении модуля слушатель:

- применять инструменты модуля при решении практических задач;
- оформлять код, данные и выводы в воспроизводимом виде;
- объяснять принятые решения, ограничения и качество результата;

Содержание модуля и виды учебной работы.

Таблица 19 – Тематический план модуля 12. Итоговая аттестация

№	Тема	Всего	Лекц.	Практ.	Сам. раб.	Контроль
1	Проверка комплекта материалов итогового проекта	0	0	0	0	итоговая аттестация
2	Демонстрация воспроизводимого запуска и результата	0	0	0	0	итоговая аттестация
3	Публичная защита, вопросы комиссии и разбор решений	40	0	0	0	итоговая аттестация
4	Оформление решения комиссии и итоговой ведомости	0	0	0	0	итоговая аттестация

Практический результат модуля. Слушатель выполняет практическое или проектное задание, размещает результат в репозитории, оформляет инструкцию запуска, краткий отчет и выводы по качеству результата.

Форма контроля. защита проекта. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость запуска, качество данных, обоснованность выводов и самостоятельность выполнения.

2.4. Проектная практика и консультации

Проектная практика является сквозным элементом Программы. Ее назначение состоит в том, чтобы объединить знания и навыки, полученные в отдельных модулях, в один завершенный Data Science проект. В ходе проектной практики слушатель выбирает предметную область, описывает источники данных, выполняет очистку и исследовательский анализ, формирует признаки, строит baseline и улучшенную модель, анализирует ошибки, готовит модельную карту, отчет, репозиторий, презентацию и демонстрацию.

Итоговая аттестация. Итоговая аттестация проводится в форме защиты итогового аттестационного проекта. К защите допускается слушатель, освоивший учебный план, выполнивший обязательные практические задания, прошедший промежуточную аттестацию и представивший комплект материалов выпускного проекта.

Комплект материалов включает репозиторий проекта, исходный код, ноутбуки, скрипты подготовки данных и обучения модели, описание источников данных, отчет, сохраненные артефакты модели или инструкции по их получению, README, файл зависимостей или контейнерную конфигурацию, результаты тестирования, презентацию и демонстрацию в формате ноутбука, дашборда, API или web-прототипа.

Детализация практической подготовки по модулю. В рамках модуля слушатель выполняет практические задания, связанные с темой «защита итогового аттестационного проекта перед комиссией». Задания выполняются в репозитории, сопровождаются краткими выводами, комментариями к коду и материалами для проверки преподавателем. Практическая подготовка направлена на то, чтобы результат модуля можно было использовать как часть итогового портфолио.

Основные действия слушателя при выполнении заданий модуля:

- представить проблему и данные;
- показать EDA, модель и метрики;

- продемонстрировать воспроизводимый запуск;
- ответить на вопросы по ограничениям и практической ценности;

Чек-лист результата модуля:

- доклад структурирован и укладывается в регламент;
- демонстрация подтверждает работоспособность;
- ответы раскрывают самостоятельность выполнения;
- пояснительная записка соответствует требованиям;

Типовые ошибки, подлежащие исправлению до зачета:

- невозможно воспроизвести результат;
- слушатель не объясняет выбранные метрики;
- есть признаки несамостоятельного выполнения;
- выводы не связаны с полученными данными;

Минимальный уровень освоения модуля. Минимальный уровень считается достигнутым, если слушатель выполнил обязательное задание, представил воспроизводимый результат, объяснил ход решения, исправил критические замечания преподавателя и подтвердил самостоятельность выполнения работы.

Повышенный уровень освоения модуля. Повышенный уровень фиксируется при наличии дополнительной проверки качества, аккуратной структуры репозитория, самостоятельного улучшения базового решения, подробной интерпретации результата и готовности объяснить ограничения выбранного подхода.

Итоговый образовательный результат модуля. Слушатель подтверждает сформированность компетенций через публичную защиту проекта, демонстрацию материалов и ответы на вопросы комиссии.

3. Организационно-педагогические условия реализации программы

3.1. Общие условия реализации

Реализация Программы обеспечивается совокупностью организационных, кадровых, информационных, методических и технических условий. Образовательный процесс организуется с использованием электронной

информационно-образовательной среды, учебных материалов, практических заданий, консультаций, проверочных мероприятий и проектной работы.

3.2. Материально-технические условия

Для освоения Программы слушателю требуется персональный компьютер или ноутбук, стабильный доступ к сети Интернет, браузер, среда разработки, доступ к электронной платформе обучения и возможность установки программного обеспечения. Рекомендуется использовать Python, JupyterLab, VS Code, Git, Docker, PostgreSQL, библиотеки pandas, NumPy, scikit-learn, Matplotlib, PyTorch или TensorFlow, FastAPI, MLflow или DVC.

3.3. Электронная информационно-образовательная среда

Электронная информационно-образовательная среда обеспечивает размещение учебных материалов, доступ к расписанию, прием практических работ, фиксацию результатов проверки, идентификацию и аутентификацию слушателя, коммуникацию с преподавателями и хранение результатов обучения. Порядок идентификации, фиксации результатов контрольных мероприятий, применения прокторинга при необходимости и доступа к материалам определяется локальным актом организации об электронном обучении и дистанционных образовательных технологиях.

3.4. Кадровые условия

Реализация Программы обеспечивается преподавателями, наставниками, методистами и специалистами сопровождения, имеющими профильное образование, профессиональный опыт в области программирования, анализа данных, машинного обучения, искусственного интеллекта, разработки программных решений или реализации образовательных программ.

3.5. Учебно-методическое обеспечение

Учебно-методическое обеспечение включает рабочие программы модулей, видеолекции, конспекты, презентации, демонстрационные ноутбуки, практические задания, тесты, чек-листы, критерии проверки, шаблоны репозитори-

ев, README, отчетов, модельных карт и презентаций, наборы учебных данных и фонд оценочных средств.

3.6. Применение электронного обучения и дистанционных образовательных технологий

Программа реализуется с применением электронного обучения и дистанционных образовательных технологий. Электронное обучение используется для предоставления учебных материалов, заданий, записей, методических рекомендаций, средств коммуникации, фиксации результатов контроля и обратной связи. Дистанционные образовательные технологии используются для организации взаимодействия слушателя и педагогических работников на расстоянии, проведения консультаций, проверки работ, промежуточной и итоговой аттестации.

При реализации Программы организация обеспечивает идентификацию слушателя в электронной информационно-образовательной среде, доступ к материалам в течение периода обучения, возможность сдачи работ, получение обратной связи и фиксацию результатов освоения Программы. Технические требования к устройству, программному обеспечению, каналам связи и учетным данным доводятся до слушателя до начала обучения или при предоставлении доступа к образовательной среде.

3.7. Информационная безопасность, персональные данные и этика

При выполнении заданий и проектов слушатели обязаны соблюдать требования законодательства о персональных данных, правила работы с учебными и открытыми наборами данных, условия лицензий, требования к конфиденциальности и правила академической добросовестности. Использование генеративных моделей допускается как вспомогательный инструмент при соблюдении требований самостоятельности, проверки результатов и корректного указания способа использования.

3.8. Локальные процедуры обучения

Организация реализации Программы обеспечивается локальными нормативными актами ООО «Иннопрог», регулирующими прием и зачисление слу-

шателей, режим занятий, применение электронного обучения и дистанционных образовательных технологий, текущий контроль, промежуточную и итоговую аттестацию, передачи, апелляции, отчисление, восстановление, выдачу документов о квалификации, внесение сведений в ФИС ФРДО и обработку персональных данных.

Локальные процедуры применяются в части, не противоречащей законодательству Российской Федерации, условиям договора об образовании и настоящей Программе. Слушатель информируется о значимых правилах реализации Программы до начала обучения или при предоставлении доступа к образовательной среде.

3.9. Обновление программы

Программа актуализируется при изменении законодательства, профессиональных стандартов, образовательных технологий, программного обеспечения, требований рынка труда, содержания учебных материалов, результатов внутреннего мониторинга качества и обратной связи от слушателей, преподавателей и работодателей.

4. Оценка качества освоения программы

4.1. Общие положения

Оценка качества освоения Программы включает текущий контроль, промежуточную аттестацию, проектную практику, предзащиту, итоговую аттестацию, внутренний мониторинг качества и возможность внешней независимой оценки качества реализации Программы.

4.2. Текущий контроль успеваемости

Текущий контроль проводится в форме практических заданий, тестов, мини-проектов, проверки репозиторий, устных разборов, автоматизированной проверки кода, ревью ноутбуков и анализа отчетов. Проверяется полнота выполнения задания, корректность кода и расчетов, воспроизводимость результата, качество обработки данных, корректность выбора метрик, качество визуализаций и выводов, соблюдение сроков и самостоятельность выполнения.

4.3. Промежуточная аттестация

Промежуточная аттестация проводится по завершении модулей в форме зачета. Зачет может включать практическую работу, мини-проект, тест, устное собеседование, проверку репозитория или комбинированное задание.

Таблица 20 – Оценочные материалы промежуточной аттестации

№	Модуль	Форма задания	Критерии зачета
2	Python для анализа данных	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
3	Продвинутый Python и ООП для Data Science	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
4	Git, Linux и воспроизводимая среда	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
5	SQL и PostgreSQL для аналитики	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
6	Математика, статистика и продуктовые метрики	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
7	Подготовка, очистка и исследовательский анализ данных	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
8	Машинное обучение	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
9	Глубокое обучение, NLP, компьютерное зрение и генеративный ИИ	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
10	MLOps, API, деплой и мониторинг моделей	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления

4.4. Пересдача, апелляция и подтверждение самостоятельности

Пересдача промежуточной аттестации проводится после устранения замечаний. Критическими основаниями для неудовлетворительного результата являются отсутствие воспроизводимого результата, невозможность запустить код, грубые ошибки в расчетах, утечка целевой переменной, отсутствие описания источника данных, признаки несамостоятельного выполнения, несоответствие проекта заявленной теме и отсутствие пояснений к выводам.

4.5. Допуск к итоговой аттестации

К итоговой аттестации допускается слушатель, который освоил учебный план, выполнил обязательные практические задания, прошел промежуточную аттестацию по модулям, подготовил итоговый проект, прошел предзащиту или внутреннюю проверку готовности проекта и устранил критические замечания.

4.6. Итоговая аттестация

Итоговая аттестация проводится в форме защиты итогового аттестационного проекта. Защита включает краткое выступление слушателя, демонстрацию результата, ответы на вопросы комиссии и проверку воспроизводимости ключевых частей решения.

Таблица 21 – Критерии оценивания итогового проекта

№	Критерий	Баллы	Что оценивается
1	Постановка задачи и связь с бизнес-целью	10	ясная проблема, целевая метрика, критерии успеха, ограничения
2	Данные и качество подготовки	15	источники, описание, очистка, проверки качества, отсутствие утечек
3	EDA, статистика и гипотезы	15	исследовательский анализ, визуализации, статистические выводы
4	Моделирование и метрики	20	baseline, выбор алгоритмов, валидация, сравнение моделей, анализ ошибок
5	Интерпретация и применимость результата	10	объяснение факторов, ограничения, риски, практический сценарий использования

№	Критерий	Баллы	Что оценивается
6	Инженерное оформление	10	репозиторий, структура, README, зависимости, тесты, логирование
7	Демонстрация, API, дашборд или сервисный прототип	10	воспроизводимый запуск, удобная демонстрация, понятный сценарий проверки
8	Защита и ответы на вопросы	10	аргументация решений, владение материалом, самостоятельность выполнения

Таблица 22 – Шкала перевода баллов итоговой аттестации

Баллы	Результат	Описание
90-100	отлично	проект полностью воспроизводим, данные и модель обоснованы, метрики достигнуты, выводы применимы, защита уверенная
70-89	зачтено	проект решает задачу, основные материалы представлены, есть отдельные замечания по оформлению, модели или глубине анализа
50-69	требуется доработка	есть рабочие фрагменты, но проект требует существенной доработки по данным, модели, запуску или выводам
0-49	неудовлетворительно	отсутствует воспроизводимый результат, нарушена самостоятельность, данные или модель не подтверждают заявленные выводы

4.7. Внутренняя и внешняя оценка качества реализации программы

Внутренняя оценка качества реализации Программы включает анализ результатов текущего контроля, промежуточной аттестации, итоговой аттеста-

ции, обратной связи слушателей, отзывов преподавателей и качества выпускных проектов. Внешняя независимая оценка качества может проводиться с участием работодателей, профильных специалистов, представителей профессионального сообщества или иных внешних экспертов.

5. Методические материалы

5.1. Методические указания для слушателей

Для успешного освоения Программы слушателю рекомендуется регулярно изучать материалы модулей, выполнять задания в установленные сроки, вести рабочий репозиторий, оформлять README к каждому проекту, фиксировать источники данных, проверять воспроизводимость запуска, использовать чек-листы самопроверки и документировать применение внешних инструментов.

5.2. Методические указания для преподавателей и проверяющих

Преподавателям и проверяющим рекомендуется проверять ход рассуждения, код, структуру проекта и воспроизводимость результата, давать конкретную обратную связь по данным, признакам, метрикам, модели, визуализациям и выводам, фиксировать критические ошибки отдельно от рекомендательных замечаний и использовать единые критерии оценивания.

5.3. Рекомендуемая литература и ресурсы

Рекомендуются официальная документация Python, pandas, NumPy, Matplotlib, scikit-learn, PostgreSQL, Git, Docker, PyTorch, TensorFlow, Hugging Face, FastAPI, MLflow и DVC, учебные материалы ООО «Иннопрог», открытые наборы данных для учебных проектов и методические рекомендации преподавателей.

Приложение 1. Матрица формирования компетенций

Таблица 23 – Перечень формируемых профессиональных компетенций

Код	Компетенция	Модули формирования	Средства оценки
ПК-1	Анализировать бизнес-задачу, формулировать аналитическую или ML-постановку, выбирать критерии успеха и ограничения проекта.	все модули, проектная практика, итоговая аттестация	практические задания, зачеты, проект, защита
ПК-2	Собирать, описывать, профилировать и проверять качество данных из файлов, баз данных, API и иных источников.	модули 2, 5, 7, 11	практические задания, зачеты, проект, защита
ПК-3	Использовать Python, NumPy, pandas и инженерные практики для подготовки данных и автоматизации аналитических пайплайнов.	модули 2, 3, 7, 11	практические задания, зачеты, проект, защита
ПК-4	Работать с SQL и PostgreSQL, строить аналитические витрины, агрегировать данные и готовить выборки для анализа и моделирования.	модули 5, 7, 11	практические задания, зачеты, проект, защита
ПК-5	Применять математическую статистику, проверку гипотез, A/B-тестирование и продуктовые метрики для принятия решений на основе данных.	модуль 6, модули 7, 11	практические задания, зачеты, проект, защита

Код	Компетенция	Модули формирования	Средства оценки
ПК-6	Строить, оценивать, сравнивать и интерпретировать модели машинного обучения для задач классификации, регрессии, кластеризации и прогнозирования.	модули 8, 11, 12	практические задания, зачеты, проект, защита
ПК-7	Применять методы глубокого обучения, NLP, компьютерного зрения и генеративного ИИ в прикладных задачах с учетом ограничений данных и моделей.	модули 9, 11, 12	практические задания, зачеты, проект, защита
ПК-8	Готовить модель к использованию: сохранять артефакты, создавать API, контейнеризовать сервис и описывать воспроизводимый запуск.	модули 4, 10, 11, 12	практические задания, зачеты, проект, защита
ПК-9	Визуализировать данные, готовить аналитические отчеты, дашборды и аргументированные выводы для заказчика.	модули 6, 7, 11, 12	практические задания, зачеты, проект, защита
ПК-10	Обеспечивать воспроизводимость, тестирование, контроль версий, документирование и качество аналитического кода.	модули 3, 4, 10, 11	практические задания, зачеты, проект, защита
ПК-11	Соблюдать требования информационной безопасности, защиты персональных данных, этики и ответственного использования моделей.	модули 1, 7, 9, 10, 11	практические задания, зачеты, проект, защита

Код	Компетенция	Модули формирования	Средства оценки
ПК-12	Представлять результат Data Science проекта, защищать решения, отвечать на вопросы по данным, методам, метрикам и бизнес-эффекту.	модули 11, 12	практические задания, зачеты, проект, защита

Приложение 2. Оценочные материалы промежуточной аттестации

Таблица 24 – Оценочные материалы промежуточной аттестации

№	Модуль	Форма задания	Критерии зачета
2	Python для анализа данных	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
3	Продвинутый Python и ООП для Data Science	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
4	Git, Linux и воспроизводимая среда	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
5	SQL и PostgreSQL для аналитики	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
6	Математика, статистика и продуктовые метрики	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
7	Подготовка, очистка и исследовательский анализ данных	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
8	Машинное обучение	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
9	Глубокое обучение, NLP, компьютерное зрение и генеративный ИИ	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления
10	MLOps, API, деплой и мониторинг моделей	практическая работа, мини-проект или проверка репозитория	корректность, воспроизводимость, выводы, качество оформления

Пример структуры практического задания: изучить описание задачи и набор данных, загрузить данные, проверить типы столбцов, выявить пропуски и

аномалии, выполнить преобразования, рассчитать метрики, построить визуализации, сформулировать выводы и оформить репозиторий.

Приложение 3. Оценочные материалы итоговой аттестации

Итоговый аттестационный проект должен быть выполнен в форме завершеного Data Science решения. Тема проекта утверждается преподавателем или комиссией с учетом содержания Программы, доступности данных, реализуемости задачи и возможности проверки результата.

Таблица 25 – Критерии оценивания итогового проекта

№	Критерий	Баллы	Что оценивается
1	Постановка задачи и связь с бизнес-целью	10	ясная проблема, целевая метрика, критерии успеха, ограничения
2	Данные и качество подготовки	15	источники, описание, очистка, проверки качества, отсутствие утечек
3	EDA, статистика и гипотезы	15	исследовательский анализ, визуализации, статистические выводы
4	Моделирование и метрики	20	baseline, выбор алгоритмов, валидация, сравнение моделей, анализ ошибок
5	Интерпретация и применимость результата	10	объяснение факторов, ограничения, риски, практический сценарий использования
6	Инженерное оформление	10	репозиторий, структура, README, зависимости, тесты, логирование
7	Демонстрация, API, дашборд или сервисный прототип	10	воспроизводимый запуск, удобная демонстрация, понятный сценарий проверки
8	Защита и ответы на вопросы	10	аргументация решений, владение материалом, самостоятельность выполнения

Таблица 26 – Шкала перевода баллов итоговой аттестации

Баллы	Результат	Описание
90-100	отлично	проект полностью воспроизводим, данные и модель обоснованы, метрики достигнуты, выводы применимы, защита уверенная
70-89	зачтено	проект решает задачу, основные материалы представлены, есть отдельные замечания по оформлению, модели или глубине анализа
50-69	требуется доработка	есть рабочие фрагменты, но проект требует существенной доработки по данным, модели, запуску или выводам
0-49	неудовлетворительно	отсутствует воспроизводимый результат, нарушена самостоятельность, данные или модель не подтверждают заявленные выводы

Приложение 5. Примерная тематика итоговых проектов.

Таблица 27 – Проектная траектория программы

№	Проект	Основные инструменты	Проверяемый результат
1	Аналитический отчет по продажам или заказам	Python, pandas, визуализация	очистка данных, расчет метрик, выводы
2	SQL-витрина клиентских транзакций	SQL, PostgreSQL	запросы, объединение таблиц, оконные функции
3	EDA качества данных	pandas, визуализация, статистика	профилирование, пропуски, выбросы, карта рисков
4	A/B-тест продуктовой гипотезы	статистика, метрики	гипотеза, расчет показателей, интерпретация
5	Дашборд бизнес-метрик	визуализация, BI-подход	метрики, сегменты, графики, аналитический комментарий
6	Модель прогнозирования спроса или цены	регрессия, feature engineering	baseline, метрики, анализ ошибок
7	Модель классификации оттока или заявок	классификация, интерпретация	ROC AUC, F1, настройка порога, SHAP
8	Кластеризация клиентских сегментов	unsupervised learning	сегменты, профили, бизнес-интерпретация
9	Прогноз временного ряда	time series	лаги, сезонность, backtesting
10	NLP-классификатор отзывов или обращений	NLP, transformers	токенизация, модель, метрики, ошибки
11	CV-прототип классификации изображений	нейронные сети, transfer learning	обучение, валидация, ограничения
12	ML-сервис с API и Docker	FastAPI, Docker, логирование	инференс, контейнер, тестовый запрос, README
13	Итоговый Data Science проект	end-to-end DS	данные, EDA, модель, интерпретация, демонстрация, защита

Допускается утверждение иной темы итогового проекта, если она соответствует цели Программы, позволяет проверить планируемые результаты обу-

чения и содержит практическую часть с данными, анализом, моделью, воспроизводимым запуском и защитой результата.

**Приложение 4. Перечень локальных нормативных актов, на которые
должна опираться реализация программы**

Таблица 28 – Перечень локальных нормативных актов

№	Локальный акт	Назначение
1	Положение о реализации дополнительных профессиональных программ	порядок разработки, утверждения, актуализации и реализации ДПП
2	Положение об электронном обучении и дистанционных образовательных технологиях	ЭИОС, идентификация слушателя, фиксация результатов и прокторинг при необходимости
3	Положение о текущем контроле, промежуточной и итоговой аттестации	формы контроля, критерии оценивания, пересдачи и апелляции
4	Положение о защите персональных данных	правила обработки персональных данных слушателей и доступа к учебным данным
5	Порядок выдачи, учета и хранения документов о квалификации	выдача дипломов и передача сведений во ФРДО
6	Положение о самостоятельности выполнения работ и академической добросовестности	правила использования внешней помощи и генеративных моделей

Приложение 5. Лист актуализации программы

Приложение 6. Лист актуализации программы.

Таблица 29 – Лист актуализации программы

№	Дата	Основание изменения	Содержание изменения	Ответственный
1	01 февраля 2025 г.	Утверждение программы	Первичная редакция Программы	ООО «Иннопрог»

Приложение 6. Требования к репозиторию, README и демонстрации

Приложение 7. Детализированные чек-листы и оценочные материалы по модулям. Настоящее приложение конкретизирует проверяемые результаты по модулям, требования к портфолио, примерные задания и критерии качества. Материалы могут использоваться преподавателем при разработке конкретных вариантов практических заданий и при подготовке индивидуальной обратной связи слушателю.

7.1. Чек-лист промежуточной аттестации по модулям.

Таблица 30 – Чек-лист промежуточной аттестации по модулям

№	Модуль	Минимальный проверяемый результат	Признаки высокого уровня выполнения
1	Введение в профессию	структура репозитория, входная диагностика, индивидуальная траектория проекта	выбран реалистичный итоговый проект, описаны данные и критерии успеха
2	Python для анализа данных	notebook с загрузкой данных, обработкой, расчетами и выводами	аккуратная декомпозиция, проверка качества данных, пояснения к графикам
3	Продвинутый Python и ООП	мини-пакет, функции подготовки данных, тесты и логирование	типизация, конфигурации, обработка ошибок, читаемая архитектура
4	Git, Linux и среда	репозиторий, история коммитов, файл зависимостей, инструкция запуска	Docker, проверенная команда запуска, отсутствие секретов и временных файлов
5	SQL и PostgreSQL	SQL-запросы, витрина данных, выгрузка для анализа	оконные функции, контроль дублей, объяснение логики join и фильтров
6	Статистика и метрики	проверка гипотезы, расчет метрик, доверительный интервал, вывод	анализ размера эффекта, ограничения выборки, корректная бизнес-интерпретация

№	Модуль	Минимальный проверяемый результат	Признаки высокого уровня выполнения
7	Подготовка и EDA	паспорт данных, пропуски, выбросы, визуализации, признаки	карта рисков данных, план feature engineering, проверка утечек
8	Машинное обучение	baseline, несколько моделей, метрики, анализ ошибок	подбор гиперпараметров, интерпретация, выбор порога, сравнение сценариев
9	Deep Learning и ИИ	прототип NLP, CV или генеративного сценария, метрики качества	transfer learning, анализ ошибок, описание этических и правовых ограничений
10	MLOps и API	сохраненная модель, API, Docker, README, тестовый запрос	логирование, валидация входа, тесты, описание мониторинга и деградации качества
11	Проектная практика	паспорт итогового проекта, EDA, модель, отчет, презентация	воспроизводимый end-to-end проект с clear baseline и бизнес-рекомендациями
12	Итоговая аттестация	защита проекта, демонстрация запуска, ответы на вопросы комиссии	уверенное объяснение выбора методов, метрик, ограничений и практической ценности

7.2. Примерные практические задания по модулям.

Таблица 31 – Примерные практические задания по модулям

№	Задание	Состав результата	Критические условия приемки
1	Описать роль Data Science специалиста в выбранной предметной области	краткое описание задачи, схема жизненного цикла, список данных и метрик	задача связана с реальным процессом и имеет измеримый результат
2	Подготовить аналитический отчет по открытым данным	notebook, графики, расчет показателей, текстовые выводы	данные загружаются автоматически, выводы подтверждены расчетами

№	Задание	Состав результата	Критические условия приемки
3	Оформить пайплайн подготовки данных как Python-проект	модули, тесты, логирование, файл зависимостей	пайплайн запускается по инструкции, тесты проходят
4	Настроить репозиторий и среду выполнения	README, структура папок, Docker или виртуальная среда, история коммитов	проект воспроизводится на другом компьютере по инструкции
5	Построить SQL-витрину для анализа заказов	схема таблиц, SQL-запросы, выгрузка, описание логики	нет дублей из-за ошибочного объединения таблиц
6	Проверить продуктовую гипотезу	формулировка гипотез, расчет критерия, интерпретация, ограничения	вывод учитывает статистическую и практическую значимость
7	Провести EDA и подготовку признаков	паспорт данных, визуализации, обработка пропусков, список признаков	исключены явные утечки и описаны риски данных
8	Обучить модель машинного обучения	baseline, модель, метрики, анализ ошибок, важность признаков	качество оценено на отложенной выборке или через кросс-валидацию
9	Создать нейросетевой прототип	данные, модель, обучение или донастройка, метрики, выводы	результат оценен количественно и на примерах ошибок
10	Обернуть модель в API	сохранение артефактов, API, Docker, тестовый запрос, логирование	сервис запускается и возвращает корректный прогноз
11	Подготовить итоговый проект	паспорт, EDA, модель, отчет, презентация, репозиторий	тема утверждена, материалы проверяемы, запуск воспроизводим
12	Защитить итоговый проект	доклад, демонстрация, ответы, оценочный лист	слушатель объясняет данные, методы, метрики и ограничения

7.3. Требования к итоговому Data Science проекту. Итоговый проект должен показывать способность слушателя самостоятельно пройти полный цикл Data Science работы. Допускаются проекты по прогнозированию спроса, классификации обращений, сегментации клиентов, анализу оттока, оценке риска, прогнозу временного ряда, обработке текстов, обработке изображений, рекомендательным системам, генеративным сценариям, аналитике образовательных или коммерческих данных.

Таблица 32 – Паспорт итогового Data Science проекта

№	Раздел паспорта проекта	Обязательное содержание
1	Название и область применения	краткое название, предметная область, пользователь результата, сценарий применения
2	Проблема и цель	описание управленческой или продуктовой проблемы, цель анализа или моделирования
3	Данные	источник, состав, объем, период, ограничения, правила использования, риск смещения
4	Целевая переменная и метрики	целевой показатель, бизнес-метрика, ML-метрика, критерий приемки результата
5	EDA	структура данных, пропуски, выбросы, распределения, корреляции, сегменты, первые выводы
6	Подготовка признаков	кодирование, масштабирование, обработка категорий, текстов, дат, временных признаков
7	Baseline	простая базовая модель или эвристика, с которой сравнивается улучшенное решение
8	Основная модель	выбранный алгоритм, настройка, метрики, анализ ошибок, устойчивость результата
9	Интерпретация	важность признаков, объяснение результата, ограничения и риски применения

№	Раздел паспорта проекта	Обязательное содержание
10	Демонстрация	notebook, отчет, дашборд, API или минимальный интерфейс для проверки результата
11	Репозиторий	README, зависимости, структура папок, инструкция запуска, данные или инструкция получения данных
12	Вывод и развитие	итоговая рекомендация, практическая ценность, ограничения и план дальнейшего улучшения

7.4. Оценочный лист итоговой защиты.

Таблица 33 – Оценочный лист итоговой защиты Data Science проекта

№	Критерий	Баллы	Пояснение к оцениванию
1	Постановка задачи и бизнес-контекст	10	ясность проблемы, пользователь результата, критерии успеха, ограничения
2	Данные и качество данных	15	источник, описание полей, проверка пропусков, выбросов, дублей и рисков
3	EDA и визуализация	10	графики, сегменты, закономерности, выводы, связь анализа с моделью
4	Подготовка признаков	10	корректность преобразований, отсутствие утечек, воспроизводимость preprocessing
5	Модель и метрики	20	baseline, выбор алгоритма, метрики, валидация, анализ ошибок
6	Интерпретация и ограничения	10	объяснимость, практическая значимость, ограничения данных и модели

№	Критерий	Баллы	Пояснение к оцениванию
7	Инженерное оформление	10	репозиторий, README, зависимости, API или демонстрационный запуск
8	Доклад и ответы	10	логика выступления, владение материалом, ответы на вопросы комиссии
9	Самостоятельность и академическая добросовестность	5	авторская работа, история изменений, отсутствие некорректных заимствований
Итого		100	минимальный положительный результат – 70 баллов

7.5. Примерные вопросы для собеседования и защиты.

1. Как сформулирована задача проекта и почему выбранная метрика подходит для ее оценки?
2. Какой источник данных использовался, какие ограничения и риски качества данных выявлены?
3. Какие пропуски, выбросы или дубликаты обнаружены и как они обработаны?
4. Какие признаки оказались наиболее значимыми и как это влияет на интерпретацию результата?
5. Как был построен baseline и насколько основная модель улучшила результат?
6. Как исключалась утечка данных между обучающей и тестовой выборками?
7. Почему выбран именно этот алгоритм или класс моделей?
8. Какие ошибки модели наиболее существенны с точки зрения бизнеса?
9. Как можно проверить устойчивость результата на новых данных?
10. Как запустить проект, какие зависимости нужны и где описана инструкция запуска?

11. Какие персональные данные или чувствительные признаки могли присутствовать в данных и как обеспечена безопасность?
12. Какие следующие шаги нужны для промышленного внедрения решения?

7.6. Критические основания для снижения оценки или недопуска. Критическими основаниями являются отсутствие воспроизводимого запуска, отсутствие данных или инструкции их получения, отсутствие самостоятельного анализа, невозможность объяснить ключевые решения, использование модели без оценки качества, подмена проекта демонстрацией чужого готового решения, некорректная работа с персональными данными, публикация секретов, существенное расхождение между заявленным результатом и фактическими материалами репозитория.

При наличии критического основания преподаватель или аттестационная комиссия вправе направить работу на обязательную доработку, назначить пересдачу или выставить неудовлетворительный результат в порядке, предусмотренном локальными актами организации.

Приложение 8. Требования к репозиторию, README и демонстрации. Для подтверждения освоения Программы и выполнения итогового проекта слушатель представляет проектные материалы в форме, позволяющей проверить самостоятельность, воспроизводимость и качество результата.

Минимальный состав проектных материалов:

- репозиторий или архив проекта с исходным кодом и структурой каталогов;
- файл README.md или иной документ с описанием цели, функций, структуры, установки и запуска проекта;
- описание используемых технологий, зависимостей, версии языка/фреймворка/инструментов;
- инструкции по подготовке окружения, настройке переменных, запуску приложения, тестов или демонстрационного сценария;
- демонстрационные данные, скриншоты, ссылка на развернутую версию или видеодемонстрация при наличии такой возможности;
- краткое описание ограничений, известных проблем и возможных направлений развития проекта.

Материалы должны быть достаточны для повторной проверки проекта преподавателем, экспертом или членом аттестационной комиссии без дополнительных устных пояснений, за исключением вопросов по защите.

Приложение 7. Дорожная карта подготовки выпускного проекта

Приложение 9. Дорожная карта подготовки выпускного проекта. Дорожная карта используется как методический ориентир для поэтапной подготовки итогового аттестационного проекта по квалификации «специалист в области Data Science».

1. Выбор темы и согласование прикладной задачи с преподавателем или куратором.
2. Формулирование цели проекта, целевой аудитории, пользовательских сценариев и критериев приемки.
3. Подготовка структуры проекта, репозитория, базовой документации и плана реализации.
4. Реализация минимально работоспособной версии и фиксация результата в репозитории.
5. Расширение функциональности, обработка ошибок, подготовка данных, интерфейса или демонстрационного сценария в соответствии со спецификой направления.
6. Проверка качества, тестирование, рефакторинг, оформление инструкций по запуску и использованию.
7. Предзащита или консультационная проверка, устранение замечаний, подготовка презентации и демонстрации.
8. Итоговая защита проекта с ответами на вопросы комиссии и фиксацией результата итоговой аттестации.

Приложение 8. Контрольный лист соответствия требованиям программы

Приложение 10. Контрольный лист соответствия требованиям программы. Контрольный лист применяется для внутренней проверки полноты Программы и готовности ее реализации.

Таблица 34 – Контрольный лист соответствия требованиям программы

№	Проверяемый элемент	Статус	Комментарий
1	Указаны вид программы, квалификация, трудоемкость, форма обучения, язык и срок освоения.	выполнено	
2	Представлены учебный план, календарный учебный график и рабочие программы модулей.	выполнено	
3	Описаны планируемые результаты, профессиональные компетенции и связь с квалификационными ориентирами.	выполнено	
4	Определены организационно-педагогические условия, ЭИОС, ЭО и ДОТ, кадровые и методические условия.	выполнено	
5	Описаны текущий контроль, промежуточная аттестация, итоговая аттестация, передача и апелляция.	выполнено	

№	Проверяемый элемент	Статус	Комментарий
6	Определены документы по итогам обучения и перечень локальных актов, обеспечивающих реализацию.	выполнено	