

Смоленский колледж телекоммуникаций (филиал)
федерального государственного бюджетного образовательного
учреждения высшего образования
«Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича»

УТВЕРЖДАЮ

Заместитель директора по УР



И.А. Овчинникова

« 14 » 05 2025 г.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТАМИ

по МДК 01.01. Разработка программных модулей в системах
искусственного интеллекта

специальность: 09.02.13 Интеграция решений с применением технологий
искусственного интеллекта

преподаватель: Котяткина Анастасия Николаевна

форма обучения – очная

Составлены в соответствии с рабочей программой дисциплины, утвержденной
«14» 05.2025 г.

Рассмотрены на заседании методической комиссии
гуманитарных и программно-вычислительных дисциплин

Протокол № 10 от «14» 05.2025 г.

Председатель МК



Т.Н. Строре

Методист



О.Г. Ряска

г. Смоленск, 2025

Содержание

1	Пояснительная записка	3
2	Особенности организации внеаудиторной самостоятельной работы студентов	3
3	Виды самостоятельной работы по МДК 01.01. Разработка программных модулей в системах искусственного интеллекта	4
	Приложения	30

1. Пояснительная записка

Самостоятельная работа – это планируемая работа студентов, выполняемая по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Она предназначена не только для овладения дисциплиной, но и для формирования навыков самостоятельной работы вообще, в учебной, научной, профессиональной деятельности, способности принимать на себя ответственность, самостоятельно решать проблемы, находить конструктивные решения, выход из кризисной ситуации и т.д. Таким образом, значимость самостоятельной работы студента выходит далеко за рамки отдельной дисциплины, играя существенную роль в развитии самостоятельности как черты характера, личностного качества, выраженного в способности мыслить, анализировать ситуации, вырабатывать собственное мнение, действовать по собственной инициативе, независимо от навязываемых взглядов.

Продумывая формы организации самостоятельной работы по дисциплине, преподаватель должен исходить из нескольких позиций:

- необходимые знания, умения и навыки, которые должен показать студент в результате выполнения всех заданий, выносимых на самостоятельное изучение (в соответствии с целью и задачами изучаемой дисциплины);
- формирование профессиональных компетентностей, которые должны проявиться через ЗУНы (знания, умения и навыки);
- формирование креативности студента в процессе изучения дисциплины и способности нестандартно мыслить при выполнении заданий для самостоятельной работы;
- развитие активной исследовательской позиции студента;
- воспитание чувства ответственности за своевременное выполнение задания.

Методические указания и рекомендации позволяют студенту выявить главное и второстепенное в изучаемой дисциплине, увидеть связь теории и практики, развивают способность к анализу полученных результатов, формируют способность формулировать тактические подходы к выполнению поставленных задач, например, подготовке к сдаче зачетов, экзаменов.

Таким образом, самостоятельная работа студентов способствует развитию у них творческой активности, повышению компетентности, совершенствованию мыслительных навыков, а также воспитывает личность будущего профессионала.

Студент, приступающий к изучению МДК 01.01 «Разработка программных модулей в системах искусственного интеллекта», получает информацию обо всех видах самостоятельной работы, об объеме и видах самостоятельной работы. Перед выполнением студентами самостоятельной внеаудиторной работы преподаватель проводит инструктаж по выполнению задания, который включает: цель задания, его содержание, сроки выполнения, ориентировочный объем работы, основные требования к результатам работы, критерии оценки.

2. Особенности организации внеаудиторной самостоятельной работы студентов

При предъявлении видов заданий на внеаудиторную самостоятельную работу рекомендуется использовать дифференцированный подход к студентам. Перед выполнением студентами внеаудиторной самостоятельной работы преподаватель проводит инструктаж по выполнению задания, который включает:

- цель задания,
- содержание,
- сроки выполнения,
- ориентировочный объем работы,
- основные требования к результатам работы,
- критерии оценки.

В процессе инструктажа преподаватель предупреждает студентов о возможных типичных ошибках, встречающихся при выполнении задания. Инструктаж проводится преподавателем за счет объема времени, отведенного на изучение дисциплины.

Во время выполнения студентами внеаудиторной самостоятельной работы и при необходимости преподаватель может проводить консультации за счет общего бюджета времени, отведенного на консультации.

Самостоятельная работа может осуществляться индивидуально или группами студентов в зависимости от цели, объема, конкретной тематики самостоятельной работы, уровня сложности, уровня умений студентов.

Контроль результатов внеаудиторной самостоятельной работы студентов может осуществляться в пределах времени, отведенного на обязательные учебные занятия по дисциплине и внеаудиторную самостоятельную работу студентов по дисциплине, может проходить в письменной, устной или смешанной форме, с представлением изделия или продукта творческой деятельности студента.

В качестве форм и методов контроля внеаудиторной самостоятельной работы студентов могут быть использованы: тестирование, защита практических и лабораторных занятий, письменная проверка и др.

3. Виды самостоятельной работы по МДК 01.01. Разработка программных модулей в системах искусственного интеллекта

На самостоятельную работу студента по МДК 01.01. Разработка программных модулей в системах искусственного интеллекта РУП отведено 26 часов

Тема 1.1. История и эволюция искусственного интеллекта (ИИ)

Цель ВСП: сформировать у студентов системное понимание ключевых этапов развития искусственного интеллекта, причинно-следственных связей между научными открытиями, технологическими возможностями и периодами подъема («зим») и спада («зим») в развитии дисциплины. Выработать способность анализировать исторический контекст для оценки современных трендов и прогнозирования будущих направлений развития ИИ.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание	Обязательный	2	1 неделя	Выписки из текста, электронная презентация	Устный опрос, демонстрация слайдов

4. Задание. Необходимо выполнить два этапа заданий.

Рекомендации по выполнению:

Распределите время: 45 минут на каждый блок. Не углубляйтесь в излишние детали в Блоке 1, ваша цель — выявить системные закономерности.

В Блоке 2 избегайте пространных рассуждений "вообще". Будьте конкретны, используйте термины и опирайтесь на факты из Блока 1.

Ценится ваше собственное, обоснованное мнение, а не пересказ чужих мыслей.

Этапы выполнения работы:

Этап 1. Аналитический хронологический конспект

Создайте структурированный конспект, отражающий эволюцию ИИ. Используйте таблицу или временную шкалу (в текстовом формате). В конспекте для каждого из перечисленных ниже периодов необходимо указать:

- хронологические рамки;
- ключевые идеи, технологии и ученые;
- яркие примеры и достижения.
- основные ограничения и причины следующего кризиса/сдвига парадигмы.

Периоды для анализа:

1. Предпосылки и рождение ИИ (1940-е – 1956): Работы Тьюринга, Маккаллока и Питтса, Дартмутская конференция.

2. Эпоха больших ожиданий и символьного подхода (1950-е – 1970-е): "Общий решатель задач", программы для игр в шахматы и шашки, LISP.
3. Первая "Зима ИИ" (1970-е – 1980-е): Критика Минского и Пейперта, Отчет Лайтхилла. Причины спада.
4. Расцвет экспертных систем (1980-е): MYCIN, XCON. Причины успеха и последующего заката.
5. Вторая "Зима ИИ" и rise of Machine Learning (конец 1980-х – 1990-е): Возвращение нейронных сетей (обратное распространение), статистические методы.
6. Эра больших данных и глубокого обучения (2000-е – 2010-е): Победа Deep Blue, AlphaGo, развитие сверточных сетей, доступ к большим данным и GPU.
7. Современный этап: Large Language Models и генеративный ИИ (2020-е – по н.в.): Трансформеры, GPT, Stable Diffusion и т.д.

Формат сдачи: Текстовый документ (Word/Google Doc) или таблица с четкой структурой.

Этап 2. Профессиональная рефлексия и прогноз

Ответьте письменно на следующие вопросы, аргументируя свою позицию на основе исторического анализа, проведенного на этапе 1.

Вопрос 1 (Анализ цикличности):

Основываясь на истории двух "Зим ИИ", проанализируйте, какие факторы (технологические, экономические, социальные) являются индикаторами возможного начала нового периода стагнации или, наоборот, стремительного роста для современных генеративных моделей? Считаете ли вы, что текущий бум устойчив? Почему?

Вопрос 2 (Смена парадигм):

История ИИ — это история смены парадигм: от символического подхода к коннекционизму и статистике. Какую следующую фундаментальную смену парадигмы, по вашему мнению, можно ожидать в области ИИ? (Например, гибридные модели, нейроморфные вычисления, ИИ, основанный на причинно-следственных связях). Обоснуйте свой ответ, опираясь на текущие ограничения глубокого обучения.

Вопрос 3 (Профессиональный контекст):

Как понимание истории "экспертных систем" может помочь современному специалисту по интеграции ИИ при внедрении, скажем, LLM в бизнес-процессы компании? Какие уроки можно извлечь из успехов и провалов прошлого?

Формат сдачи: Текстовый документ с четкими ответами на каждый вопрос (общий объем ответов ~1-1,5 страницы А4).

5. Критерии оценки

Критерий	Баллы	Описание
Блок 1: Полнота и точность конспекта	0-4	Учтены все указанные периоды, корректно указаны ключевые факты, имена, даты и технологические вехи.
Блок 1: Глубина анализа причин и следствий	0-3	Не просто перечислены факты, но и проанализированы связи между достижениями, ограничениями и сменами этапов.
Блок 2: Качество аргументации	0-3	Ответы на вопросы содержат четкие, логичные и убедительные аргументы, подкрепленные отсылками к историческим примерам.
Блок 2: Профессиональная глубина и прогноз	0-3	Ответы демонстрируют не просто знание истории, но и способность применять эти знания для анализа текущей ситуации и формирования прогнозов.
Структура и оформление	0-2	Работа хорошо структурирована, оформлена аккуратно,

Критерий	Баллы	Описание
		текст легко читается.
Итого	0-15	

Шкала перевода: 13-15 баллов = "отлично", 10-12 = "хорошо", 7-9 = "удовлетворительно", <7 = "неудовлетворительно".

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.
2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>
3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.2. Основные направления ИИ: машинное обучение, глубокое обучение, нейронные сети.

Цель ВСП: сформировать у студентов четкое, структурированное понимание трех ключевых направлений ИИ, их взаимосвязи и иерархии. Выработать способность проводить сравнительный анализ, понимать область применения каждой технологии и делать осознанный выбор подхода для решения конкретных бизнес-задач.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание 1	Обязательный	4	1 неделя	Текстовый документ (Word) или таблица с четкой структурой.	Устный опрос, демонстрация проекта

Задание 1. Вам предстоит выступить в роли специалиста по ИИ, которому нужно объяснить коллегам из нетехнического отдела (например, отдела маркетинга) разницу между ключевыми направлениями ИИ и обосновать выбор технологии для нового проекта.

Работа состоит из двух этапов.

Рекомендации по выполнению:

Начните с Этапа 1, чтобы создать для себя прочную концептуальную базу. Это основа для выполнения Этапа 2.

На Этапе 2 думайте как консультант. Ваша цель — не просто назвать технологию, а убедить "заказчика" в правильности вашего вывода.

Используйте простой и ясный язык, особенно при формулировке аналогий на Этапе 1.

Этапы выполнения работы:

Этап 1. Сравнительный концептуальный анализ.

Создайте сравнительную таблицу или инфографику в текстовом формате, которая раскрывает суть следующих направлений:

Машинное обучение (Machine Learning)

Нейронные сети (Neural Networks)

Глубокое обучение (Deep Learning)

Таблица должна включать следующие критерии для сравнения:

Ключевая идея/Определение: Краткая суть подхода (1-2 предложения).

Аналогия для нетехнических специалистов: Простая и понятная аналогия из жизни (например, "МО — это как ученик, который решает задачи по примерам...").

Архитектура/Принцип работы: Основные компоненты и как происходит процесс обучения (данные, признаки, модель).

Сильные стороны: Для каких задач подходит лучше всего.

Слабые стороны/Ограничения: Какие проблемы плохо решает, какие требования к данным предъявляет.

Примеры типичных задач: 2-3 конкретных примера применения.

Формат сдачи: Текстовый документ (Word/Google Doc) или таблица с четкой структурой.

Этап 2. Практический кейс: Выбор технологии для бизнес-задачи.

Ответьте письменно на следующие вопросы, аргументируя свой выбор на основе анализа, проведенного на этапе 1.

Сценарий:

Компания "Розничный Гигант" хочет внедрить ИИ-решения для улучшения бизнес-показателей. Вам, как специалисту по ИИ, поручили проанализировать три инициативы и предложить наиболее подходящую технологию (Машинное обучение, Глубокое обучение или Нейронные сети) для каждой из них.

Задачи компании:

Задача А: Прогнозирование оттока клиентов.

Описание: Есть структурированные данные о клиентах: длительность договора, частота покупок, средний чек, количество обращений в поддержку. Нужно определить клиентов с высокой вероятностью ухода к конкурентам в следующем месяце.

Задача Б: Система компьютерного зрения для проверки качества продукции.

Описание: На конвейере нужно автоматически обнаруживать брак (царапины, вмятины, неправильная сборка) с помощью камер. Точность должна быть максимально высокой, а дефекты могут быть разнообразными и не всегда очевидными.

Задача В: Автоматическая классификация отзывов на сайте.

Описание: Нужно анализировать тексты отзывов и автоматически относить их к категориям: "Положительный", "Негативный", "Нейтральный", а также определять тему: "доставка", "качество товара", "сервис".

Вопросы для анализа для каждой задачи (А, Б, В):

Какая технология является наиболее подходящей? (Обоснуйте, ссылаясь на ее сильные стороны из этапа 1).

Почему две другие технологии подходят меньше? (Укажите на их ограничения в контексте данной задачи).

Какого типа данные и в каком объеме потребуются для обучения модели под эту задачу?

Формат сдачи: Текстовый документ с четкими ответами на каждый вопрос по всем трем задачам.

Критерии оценки:

Критерий	Баллы	Описание
Блок 1: Полнота и точность сравнения	0-4	Таблица содержит все критерии для всех трех направлений. Определения четкие, сильные и слабые стороны отражены корректно.
Блок 1: Качество аналогий и примеров	0-3	Приведенные аналогии понятны нетехническому специалисту, примеры задач релевантны и конкретны.
Блок 2: Обоснованность	0-4	Для каждой бизнес-задачи выбор технологии

Критерий	Баллы	Описание
выбора технологии		аргументирован с ссылкой на сильные/слабые стороны технологий.
Блок 2: Глубина анализа альтернатив	0-3	Для каждой задачи корректно объяснено, почему другие технологии не подходят, показано понимание их ограничений.
Структура и оформление	0-2	Работа хорошо структурирована, оформлена аккуратно, текст грамотный и легкий для восприятия.
Итого	0-16	

Шкала перевода: 14-16 баллов = "отлично", 11-13 = "хорошо", 8-10 = "удовлетворительно", <8 = "неудовлетворительно".

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.
2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>
3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.3. Этические вопросы и вызовы, связанные с развитием ИИ.

Цель ВСР: сформировать у студентов способность идентифицировать, анализировать и предлагать решения для ключевых этических проблем, возникающих на всех этапах жизненного цикла ИИ-систем (от сбора данных до развертывания и мониторинга). Развить навык критической оценки социальных последствий технологий и познакомить с существующими принципами и рамками этичного ИИ.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание	Обязательный	2	1 неделя	Текстовый документ (Word) или таблица с четкой структурой.	Устный опрос, демонстрация проекта

Задание. Вам предстоит выступить в роли Этика-консультанта в ИИ-компании. Руководство поручило вам подготовить аналитический обзор и дать рекомендации по трем наиболее актуальным этическим вызовам.

Работа состоит из двух этапов.

Рекомендации по выполнению:

На Этапе 1 думайте как аналитик: ваша задача — не просто перечислить, а показать причинно-следственные связи между технологией и ее социальным impact.

На Этапе 2 думайте как инженер и юрист: предлагайте не абстрактные пожелания ("нужно быть честным"), а конкретные инженерные или процессные решения ("внедрить аудит модели по метрике X раз в квартал").

Используйте ресурсы: ознакомление с реальными хартиями и кейсами придаст вашим рассуждениям вес и профессиональную глубину.

Этапы выполнения работы:

Этап 1. Карта этических рисков.

Создайте таблицу-классификатор, в которой систематизированы ключевые этические проблемы ИИ. Таблица должна содержать следующие столбцы:

- категория проблемы (например, Справедливость и предвзятость, Прозрачность и объяснимость, Конфиденциальность и данные, Безопасность и контроль, Социально-экономические последствия);

- суть проблемы (краткое описание, в чем заключается этический вызов)

Технические причины (что на уровне данных, алгоритмов или архитектуры системы приводит к этой проблеме?);

- потенциальные негативные последствия (для личности, группы людей, общества в целом);

- пример-иллюстрация (конкретный кейс или гипотетический сценарий)

***Проанализируйте следующие проблемы (минимум 5, включая обязательные):**

1. Algorithmic Bias (Обязательная)

2. "Черный ящик" (Black Box Problem) и объяснимость (Обязательная)

3. Конфиденциальность и слежка

4. Автономные системы и вопрос ответственности (например, беспилотные автомобили)

5. Влияние ИИ на рынок труда и социальное неравенство

6. Манипулятивное поведение (микротаргетинг, deepfakes)

Формат сдачи: Текстовый документ (Word/Google Doc) или электронная таблица с четкой структурой.

Этап 2. Разработка этического меморандума для проекта.

Выберите ОДИН из предложенных ниже сценариев внедрения ИИ. На его основе составьте Этический меморандум — документ, который идентифицирует риски и предлагает практические шаги по их mitigation (смягчению).

Сценарии на выбор:

Сценарий А: Система отбора кандидатов.

Компания разрабатывает ИИ-систему для автоматического анализа резюме и проведения первичных видео-интервью для отбора кандидатов на позиции в отдел продаж. Система обучается на данных о прошлых успешных сотрудниках.

Сценарий Б: Система прогнозирования преступности.

Муниципалитет хочет внедрить систему, которая использует данные о прошлых преступлениях, демографическую статистику и видео с камер наблюдения для прогнозирования "зон повышенного риска" и оптимизации маршрутов патрулирования полиции.

Сценарий В: Генеративный ИИ для новостного агрегатора.

Стартап создает сервис, который с помощью LLM (большой языковой модели) автоматически генерирует краткие новостные сводки на основе тысяч источников. Цель — максимальная персонализация контента для каждого пользователя.

Структура Этического меморандума:

1. **Описание проекта:** Кратко опишите выбранный сценарий.

2. **Идентификация рисков:** Перечислите 3 наиболее критичных этических риска для данного проекта, основываясь на своей "Карте рисков" из Блока 1.

3. **Рекомендации по смягчению рисков:** Для каждого выявленного риска предложите 1-2 конкретных, практических действия.

Пример: Риск: "Смещенность алгоритма против определенных demographic групп".

Рекомендация: "Внедрить регулярный аудит модели на fairness с помощью метрик (например, demographic parity, equalized odds). Разнообразить набор данных для обучения. Создать комитет по этике с привлечением внешних экспертов".

4. Принципы этичного ИИ: Назовите 2-3 ключевых принципа из существующих хартий (например, принципы OECD по ИИ, рекомендации ЕС), которые являются основой для ваших рекомендаций.

Формат сдачи: Текстовый документ с четким разделением на пункты.

Критерии оценки:

Критерий	Баллы	Описание
Блок 1: Систематизация и глубина анализа	0-4	В таблице представлены все требуемые категории, проблемы проанализированы глубоко, с указанием технических причин и последствий. Примеры релевантны.
Блок 1: Понимание взаимосвязей	0-3	Студент демонстрирует понимание, как технические решения порождают этические проблемы (связь "столбцов" в таблице).
Блок 2: Качество идентификации рисков	0-3	Для выбранного сценария выявлены наиболее релевантные и нетривиальные этические риски.
Блок 2: Практическая ценность рекомендаций	0-4	Рекомендации являются конкретными, выполнимыми и напрямую направлены на смягчение выявленных рисков.
Структура, оформление и аргументация	0-2	Работа логично структурирована, оформлена аккуратно, все выводы и рекомендации хорошо аргументированы.
Итого	0-16	

Шкала перевода: 14-16 баллов = "отлично", 11-13 = "хорошо", 8-10 = "удовлетворительно", <8 = "неудовлетворительно".

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.
2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>
3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.4. Методы сбора данных: веб-скрапинг, API, базы данных.

Цель ВСП: сформировать у студентов системное понимание преимуществ, ограничений и юридических аспектов различных методов сбора данных. Выработать практические навыки выбора оптимального метода для конкретной бизнес-задачи и проектирования процесса сбора данных.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание 1	Обязательный	2	1 неделя	Текстовый документ (Word) или	Устный опрос, демонстрация проекта

				таблица с четкой структурой.	
--	--	--	--	------------------------------	--

Задание 1. Вам предстоит выступить в роли Data Engineer в команде по разработке ИИ-решений. Вам поручено проанализировать три проекта и предложить стратегию сбора данных для каждого.

Работа состоит из двух этапов.

Рекомендации по выполнению:

На Этапе 1 сфокусируйтесь на выявлении ключевых различий, которые влияют на принятие решения в реальной жизни (например, "скорость" vs "юридическая чистота").

На Этапе 2 думайте как архитектор решения. Не останавливайтесь на выборе метода, а сразу продумывайте, как это будет работать на практике и что может пойти не так.

Помните: не существует единственно правильного ответа. Ценится ваша способность взвесить "за" и "против" и предложить решение, оптимальное для конкретных условий задачи.

Этапы выполнения работы:

Этап 1. Сравнительный анализ методов сбора данных.

Создайте **сравнительную таблицу**, в которой детально проанализируйте три основных метода сбора данных:

1. Веб-скрапинг (Web Scraping)
2. Работа с API (Application Programming Interface)
3. Запросы к базам данных (SQL/NoSQL)

Таблица должна включать следующие критерии для сравнения:

- Принцип работы: Краткое техническое описание процесса получения данных.

- Скорость и объем: Потенциал для сбора больших данных, скорость получения.

Структурированность данных: Уровень организации данных на выходе (сырые, полуструктурированные, структурированные).

- Техническая сложность: Уровень требуемых навыков для реализации (простой/средний/сложный).

- Юридические и этические ограничения: Основные риски и правила использования (лицензии, robots.txt, условия использования, GDPR).

- Стабильность и надежность: Частота изменения интерфейса/структуры данных, устойчивость метода.

- Типичные кейсы применения: 2-3 примера задач, для которых метод лучше всего подходит.

Формат сдачи: Текстовый документ (Word/Google Doc) или электронная таблица с четкой структурой.

Этап 2. Разработка стратегии сбора данных для бизнес-кейсов (45 минут)

Проанализируйте **ТРИ** предложенных ниже бизнес-сценария. Для каждого ответьте на следующие вопросы:

Вопросы для каждого сценария:

1. Какой метод сбора данных является оптимальным? Можно предложить комбинацию методов.

2. Обоснуйте свой выбор, ссылаясь на сравнительную таблицу из Блока 1 (скорость, юридические аспекты, сложность и т.д.).

3. Опишите пошаговый план действий для реализации выбранного метода (3-4 ключевых шага).

4. Какие основные риски (технические, юридические, операционные) необходимо учесть в этом сценарии?

Бизнес-сценарии:

Сценарий А: Анализ цен конкурентов.

Задача: Ежедневно отслеживать изменения цен на 100 товаров-аналогов в 5 крупнейших интернет-магазинах для гибкого ценообразования.

Сценарий Б: Обогащение данных о клиентах.

Задача: В имеющейся CRM-системе есть база данных клиентов с их названиями и индустрией. Необходимо автоматически получать данные о их финансовых показателях (выручка, количество сотрудников) для сегментации.

Сценарий В: Создание датасета для обучения чат-бота.

Задача: Для обучения NLP-модели чат-бота, специализирующегося на кулинарии, необходимо собрать большую коллекцию (100 000+) рецептов, ингредиентов и пошаговых инструкций.

Формат сдачи: Текстовый документ с четкими ответами на вопросы для каждого из трех сценариев.

Критерии оценки:

Критерий	Баллы	Описание
Блок 1: Полнота и глубина сравнения	0-4	Таблица содержит все критерии для всех трех методов. Анализ демонстрирует глубокое понимание особенностей каждого метода.
Блок 1: Качество примеров и оценка рисков	0-3	Приведенные кейсы применения релевантны, юридические и технические ограничения описаны корректно.
Блок 2: Обоснованность выбора метода	0-4	Для каждого сценария выбор метода (или комбинации) логически обоснован с ссылкой на сравнительный анализ.
Блок 2: Практическая реализуемость плана и оценка рисков	0-3	Предложенный пошаговый план является конкретным и выполнимым. Выявлены наиболее существенные риски.
Структура и оформление	0-2	Работа хорошо структурирована, оформлена аккуратно, текст грамотный и легкий для восприятия.
Итого	0-16	

Шкала перевода: 14-16 баллов = "отлично", 11-13 = "хорошо", 8-10 = "удовлетворительно", <8 = "неудовлетворительно".

Информационное обеспечение:

- Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.
- Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROОбразование : [сайт]. — URL: <https://profspo.ru/books/98551>
- Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.5. Подготовка данных для обучения моделей ИИ.

Цель ВСР: сформировать у студентов системное понимание полного цикла подготовки данных (Data Preprocessing Pipeline). Выработать практические навыки анализа качества данных, выбора методов их обработки и обоснования принятых решений для конкретных типов данных и бизнес-задач.

Трудоемкость

Количество заданий	Характер задачи	Норма времени (в	Срок выполнения	Форма представления	Форма контроля
--------------------	-----------------	------------------	-----------------	---------------------	----------------

(задач, упражнений)	(обязательный/рекомендательный)	часах по рабочей программе)	(в неделях)	материала	каждого задания
Задание 1	Обязательный	2	1 неделя	Текстовый документ с четкими ответами на вопросы для каждого из двух сценариев.	Устный опрос, демонстрация проекта

Задание 1. Вам предстоит выступить в роли **Data Scientist / Data Engineer** в проекте по разработке ИИ-модели. Вам предоставили "сырые" датасеты и поставили задачу спроектировать процесс их подготовки к обучению.

Работа состоит из двух этапов.

Рекомендации по выполнению:

На Этапе 1 думайте как **архитектор процесса**. Ваша цель — создать универсальную и логичную схему, а не просто перечислить этапы.

На Этапе 2 будьте **конкретны**. Вместо "обработаем пропуски" напишите "пропуски в столбце support_calls заполним медианным значением, так как распределение скошено, и это уменьшит влияние выбросов".

Для генерации новых признаков (фичеринжиниринг) подумайте, какая **дополнительная информация** может быть полезна модели, но не явно заложена в исходных данных (например, "отношение количества звонков в поддержку к общему времени пользования услугой").

Этап 1. Проектирование конвейера предобработки данных.

Создайте **схему/блок-схему** (в текстовом или графическом формате) универсального конвейера подготовки данных. Для каждого этапа укажите:

- Название этапа
- Основные задачи, решаемые на этом этапе
- Ключевые методы и технологии для решения этих задач
- Критерии успешного завершения этапа

Основные этапы для отражения в конвейере:

1. Анализ и Понимание Данных (Data Understanding)
2. Обработка Пропущенных Значений (Handling Missing Values)
3. Обработка Выбросов (Handling Outliers)
4. Кодирование Категориальных Признаков (Encoding Categorical Features)
5. Масштабирование и Нормализация Чисел (Scaling/Normalization)
6. Работа с Текстом (Text Preprocessing)
7. Работа с Изображениями (Image Preprocessing)
8. Разбиение на Выборки (Train/Validation/Test Split)

Формат сдачи: Текстовый документ с четко структурированным списком или созданная в любом редакторе блок-схема, вставленная в документ.

Этап 2. Разработка стратегии предобработки для конкретных кейсов.

Проанализируйте **ДВА** предложенных ниже сценария. Для каждого ответьте на следующие вопросы:

Вопросы для каждого сценария:

- 1. Качество данных:** Какие **3 основные проблемы** с качеством данных вы ожидаете встретить в этом датасете? (Например, пропуски, выбросы, несбалансированность, шум и т.д.)
- 2. План предобработки:** Для каждой выявленной проблемы предложите **конкретный метод обработки** и обоснуйте его выбор.
- 3. Фичеринжиниринг:** Какие **2 новых признака (feature)** вы могли бы создать на основе имеющихся данных, чтобы улучшить качество будущей модели?
- 4. Валидация:** Как вы будете проверять, что процесс предобработки прошел корректно и не внес искажений в данные?

Сценарии:**Сценарий А: Прогнозирование оттока клиентов телеком-компаний.**

Данные: Таблица с клиентами. Столбцы: duration_days (дней с нами), monthly_charge (ежемесячный платеж), support_calls (обращения в поддержку), contract_type [Monthly, Yearly], payment_method [Card, Electronic, Cash], churned [0, 1] - целевая переменная.

Сценарий Б: Классификация тональности отзывов о ресторанах.

Данные: Таблица с отзывами. Столбцы: review_id (id отзыва), review_text (текст отзыва), reviewer_rating (оценка от 1 до 5), sentiment [positive, negative, neutral] - целевая переменная.

Формат сдачи: Текстовый документ с четкими ответами на вопросы для каждого из двух сценариев.

Критерии оценки:

Критерий	Баллы	Описание
Блок 1: Полнота и логичность конвейера	0-4	Конвейер отражает все ключевые этапы, их последовательность логична, методы и критерии подобраны корректно.
Блок 1: Глубина понимания методов	0-3	Продемонстрировано понимание нюансов методов (например, разница между One-Hot Encoding и Label Encoding).
Блок 2: Качество анализа проблем данных	0-3	Для каждого сценария выявлены релевантные и нетривиальные проблемы качества данных.
Блок 2: Обоснованность и креативность решений	0-4	Предложенные методы предобработки и новые признаки логично обоснованы и направлены на улучшение качества модели.
Структура и оформление	0-2	Работа хорошо структурирована, оформлена аккуратно, ответы изложены ясно и по делу.
Итого	0-16	

Шкала перевода: 14-16 баллов = "отлично", 11-13 = "хорошо", 8-10 = "удовлетворительно", <8 = "неудовлетворительно".

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.
2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>
3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.6. Виды обучения: обучение с учителем, обучение без учителя, обучение с подкреплением

Цель ВСР: сформировать у студентов четкое концептуальное понимание трех основных парадигм машинного обучения, их фундаментальных различий, областей применения и требований к данным. Развить навык осознанного выбора типа обучения в зависимости от поставленной бизнес-задачи и имеющихся данных.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание 1	Обязательный	2	1 неделя	Текстовый документ с четкими ответами на вопросы для каждого из двух сценариев.	Устный опрос, демонстрация проекта

Задание 1. Вам предстоит выступить в роли AI Solution Architect. Ваша задача — проанализировать различные сценарии и определить, какой тип обучения наиболее применим в каждом случае, а также объяснить свою позицию. Работа состоит из двух этапов.

Рекомендации по выполнению:

Ключ к успеху — в понимании фундаментальных различий в целях и данных для каждого типа обучения. Сфокусируйтесь на этом на Этапе 1.

На Этапе 2 самым важным является не выбор, а его обоснование. Даже если выбор правильный, но обоснование слабое, балл будет снижен. Аргументируйте, почему выбранный метод подходит, а другие — нет.

Используйте простой и ясный язык, особенно при формулировке аналогий на Этапе 1.

Этап 1. Сравнительный концептуальный анализ.

Создайте развернутую сравнительную таблицу, которая раскрывает суть трех парадигм обучения:

1. Обучение с учителем (Supervised Learning)
2. Обучение без учителя (Unsupervised Learning)
3. Обучение с подкреплением (Reinforcement Learning)

Таблица должна включать следующие критерии для сравнения:

- Ключевая идея/Аналогия: Краткая суть подхода и простая аналогия из жизни (например, "Обучение с учителем — как ученик, который решает задачи с готовыми ответами в конце учебника").
- Наличие и тип данных: Что представляет собой набор данных? Есть ли размеченные (labeled) данные? Если да, то какого типа (метки, ответы)?
- Цель алгоритма: Какую задачу решает модель в процессе обучения?
- Основные типы задач: Конкретные примеры задач, характерных для этого типа обучения (например, для обучения с учителем: классификация, регрессия).
- Процесс обучения: Как происходит взаимодействие модели с данными/средой?
- Критерий успеха: По чему мы понимаем, что модель научилась?
- Плюсы и минусы: Сильные и слабые стороны подхода.

Формат сдачи: Текстовый документ (Word/Google Doc) или электронная таблица с четкой структурой.

Этап 2. Классификация бизнес-задач и обоснование выбора.

Проанализируйте шесть предложенных ниже бизнес-сценариев. Для каждого сценария выполните два действия:

Задание для каждого сценария:

1. Определите тип обучения: Выберите один из трех основных типов (С учителем / Без учителя / С подкреплением), который является наиболее подходящим для решения данной задачи.

2. Развернутое обоснование: Объясните свой выбор, отвечая на следующие вопросы:
Какие данные доступны? Соответствуют ли они требованиям выбранного типа обучения?
Какова конечная цель модели? Согласуется ли она с целью выбранного типа обучения?
Почему два других типа обучения подходят меньше?

Бизнес-сценарии:

1. Прогнозирование спроса на товары: Магазин хочет предсказать продажи мороженого на следующую неделю на основе исторических данных о продажах, погоде и днях недели.
2. Сегментация клиентов: Банк хочет разбить свою клиентскую базу на однородные группы на основе их транзакционного поведения (суммы, частоты, категории трат) для разработки персональных предложений. У банка нет заранее определенных шаблонов групп.
3. Автономный складской робот: Робот должен научиться оптимально перемещаться по складу, собирая товары. Он получает "вознаграждение" за быструю и правильную сборку заказа и "штраф" за столкновения или ошибки.
4. Обнаружение мошеннических операций: Кредитная организация хочет автоматически помечать подозрительные транзакции. В их распоряжении есть большая база данных операций, каждая из которых вручную размечена как "мошенническая" или "легитимная".
5. Рекомендательная система для стримингового сервиса: Сервис хочет предлагать пользователям фильмы, которые им, вероятно, понравятся, на основе истории их просмотров и оценок, которые они ставили другим фильмам.
6. Снижение размерности для визуализации: Исследовательская компания имеет датасет с 500 характеристиками пациентов. Чтобы выявить скрытые закономерности и визуализировать данные в 2D- или 3D-пространстве, необходимо уменьшить количество признаков, сохранив максимальное количество информации.

Формат сдачи: Текстовый документ с таблицей, где для каждого сценария указан выбранный тип обучения и приведено развернутое обоснование.

Критерии оценки:

Критерий	Баллы	Описание
Блок 1: Полнота и точность сравнения	0-4	Таблица содержит все критерии для всех трех парадигм. Определения четкие, цели и процессы описаны корректно.
Блок 1: Качество аналогий и понимание сути	0-3	Приведенные аналогии точно отражают суть парадигм. Продемонстрировано глубокое понимание фундаментальных различий.
Блок 2: Правильность классификации сценариев	0-3	Для всех шести сценариев выбран корректный тип обучения.
Блок 2: Глубина и качество обоснования выбора	0-4	Для каждого сценария приведено логичное, развернутое и убедительное обоснование, демонстрирующее понимание сильных и слабых сторон каждого типа обучения.
Структура и оформление	0-2	Работа хорошо структурирована, оформлена аккуратно, текст грамотный и легкий для восприятия.
Итого	0-16	

Шкала перевода: 14-16 баллов = "отлично", 11-13 = "хорошо", 8-10 = "удовлетворительно", <8 = "неудовлетворительно".

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.
2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>
3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.7. Кластеризация.

Цель ВСП: сформировать у студентов практическое понимание методов кластеризации как инструмента для обнаружения скрытых структур в данных. Развить навыки выбора подходящего алгоритма кластеризации, оценки качества результатов и бизнес-интерпретации полученных кластеров.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание 1	Обязательный	2	1 неделя	Работа выполняется в электронном виде. Формат документа: PDF или DOC/DOCX	Устный опрос, демонстрация проекта

Задание 1. Вам предстоит выступить в роли Data Analyst в компании, где стоит задача сегментировать клиентов для повышения эффективности маркетинга. Проведите анализ данных и предложите решение на основе методов кластеризации. Работа состоит из двух этапов.

Рекомендации по выполнению:

Начните с изучения теоретических основ каждого алгоритма

Уделите внимание практическим аспектам применения методов

Продумайте, как результаты кластеризации можно использовать для бизнеса

Используйте визуализацию для лучшего понимания работы алгоритмов

Не забывайте о важности предобработки данных для качества кластеризации

Примечание: Для выполнения задания достаточно теоретического анализа, написание кода не требуется.

Этап 1. Сравнительный анализ алгоритмов кластеризации.

Создайте **сравнительную таблицу** основных алгоритмов кластеризации:

1. K-means (K-средних)
2. DBSCAN
3. Иерархическая кластеризация

Таблица должна включать следующие критерии:

Критерий	K-means	DBSCAN	Иерархическая
Основной принцип работы			
Требуется ли указания числа кластеров			

Критерий	K-means	DBSCAN	Иерархическая
Типы кластеров, которые может находить (сферические, произвольной формы)			
Устойчивость к выбросам			
Метрики для оценки качества			
Преимущества			
Недостатки			
Лучшие сценарии применения			

Формат сдачи: Текстовый документ или электронная таблица с заполненной таблицей.

Этап 2. Практический кейс: Сегментация пользователей интернет-магазина.

Исходные данные:

Интернет-магазин собрал данные о покупателях за последние 6 месяцев:

- customer_id - идентификатор покупателя
- total_purchases - общая сумма покупок (руб.)
- purchase_frequency - частота покупок (раз в месяц)
- avg_cart_value - средний чек (руб.)
- last_purchase_days - дней с последней покупки

Задание:

1. Предобработка данных

Какие шаги предобработки необходимы перед кластеризацией?

Почему это важно?

2. Выбор алгоритма

Какой алгоритм кластеризации вы выберете и почему?

Какие гиперпараметры нужно будет настроить?

3. Интерпретация результатов

Как вы будете оценивать качество кластеризации?

Предположите, какие типы клиентов могут быть выделены в кластеры

Какие маркетинговые стратегии можно предложить для каждого кластера?

Пример ожидаемых результатов:

Кластер	Характеристики	Рекомендации по работе
1	Высокие чеки, редкие покупки	Премиум-обслуживание, персональные предложения
2	Частые покупки, низкие чеки	Программа лояльности, рассылки о новинках
3	Неактивные клиенты	Реактивация, специальные предложения

Формат сдачи: Текстовый документ с ответами на все вопросы кейса.

Критерии оценки:

Критерий	Баллы	Описание
----------	-------	----------

Критерий	Баллы	Описание
Полнота сравнительного анализа	0-4	Все алгоритмы полноценно сравнены по ключевым критериям
Понимание принципов работы алгоритмов	0-3	Продemonстрировано глубокое понимание различий в работе алгоритмов
Обоснованность выбора алгоритма для кейса	0-4	Выбор алгоритма аргументирован с учетом особенностей данных
Качество бизнес-интерпретации результатов	0-3	Кластеры охарактеризованы осмысленно, предложены релевантные маркетинговые стратегии
Структура и оформление работы	0-2	Работа хорошо структурирована, ответы изложены четко и грамотно
Итого	0-16	

Шкала оценки:

14-16: Отлично

11-13: Хорошо

8-10: Удовлетворительно

0-7: Неудовлетворительно

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.

2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>

3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.8. Валидация моделей: кросс-валидация, разделение данных на тренировочные и тестовые.

Цель ВСП: сформировать у студентов понимание важности корректной валидации моделей машинного обучения, освоить методы разделения данных и оценки качества моделей на независимых выборках. Развить навыки выбора стратегии валидации в зависимости от характеристик данных и решаемой задачи.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание 1	Обязательный	2	1 неделя	Текстовый документ с четкими	Устный опрос, демонстрация проекта

				ответами на вопросы для каждого из двух сценариев.	
--	--	--	--	--	--

Задание 1. Вам предстоит выступить в роли **Data Scientist** в компании, разрабатывающей ML-модели для прогнозирования. Проведите анализ различных методов валидации и предложите оптимальные стратегии для конкретных бизнес-кейсов. Работа состоит из двух этапов.

Рекомендации по выполнению:

Начните с изучения теоретических основ каждого метода валидации.

Уделите внимание особенностям разных типов данных

(сбалансированные/несбалансированные, временные ряды)

Продумайте, как выбранная стратегия валидации повлияет на надежность оценки модели

Используйте конкретные примеры и аргументы при обосновании выбора

Не забывайте о практических аспектах (вычислительная сложность, объем данных)

Примечание: для выполнения задания достаточно теоретического анализа, написание кода не требуется.

Этап 1. Сравнительный анализ методов валидации.

Создайте **сравнительную таблицу** методов валидации моделей:

Методы для анализа:

1. Простое разделение (Hold-out)
2. K-Fold кросс-валидация
3. Stratified K-Fold кросс-валидация
4. Временные ряды (Time Series Split)

Таблица должна включать следующие критерии:

Критерий	Hold-out	K-Fold	Stratified K-Fold	Time Series Split
Принцип работы				
Оптимальное соотношение train/test				
Устойчивость к переобучению				
Требования к данным				
Вычислительная сложность				
Лучшие сценарии применения				
Основные преимущества				
Основные недостатки				

Дополнительное задание:

Объясните, в чем заключается проблема **утечки данных (data leakage)** при неправильной валидации и как ее избежать.

Этап 2. Практический кейс: Выбор стратегии валидации (45 минут)

Проанализируйте три бизнес-сценария и для каждого ответьте на вопросы:

Вопросы для каждого сценария:

1. Какой метод валидации вы рекомендуете и почему?
2. Какие параметры метода вы установите (размер тестовой выборки, количество фолдов и т.д.)?
3. Как будете оценивать качество модели?

4. Какие риски неправильной валидации существуют в данном кейсе?

Сценарии:

Сценарий 1: Классификация заболеваний

Задача: Предсказание наличия редкого заболевания по медицинским показателям

Данные: 10,000 пациентов, распределение классов: 98% - здоровы, 2% - больны

Особенности: Несбалансированные данные, важна точность предсказания меньшинственного класса

Сценарий 2: Прогнозирование продаж

Задача: Прогноз ежедневных продаж сети магазинов на следующий месяц

Данные: Ежедневные продажи за 3 года (1,095 записей)

Особенности: Временной ряд с сезонностью, трендом

Сценарий 3: Классификация изображений

Задача: Определение пород собак по фотографиям

Данные: 50,000 изображений, равномерно распределенных по 100 породам

Особенности: Сбалансированные данные, большая выборка

Критерии оценки:

Критерий	Баллы	Описание
Полнота сравнительного анализа	0-4	Все методы полноценно сравнены по ключевым критериям
Понимание принципов валидации	0-3	Продемонстрировано понимание целей и методов валидации
Обоснованность выбора методов для кейсов	0-4	Выбор стратегий валидации аргументирован особенностями данных и задач
Качество анализа рисков	0-3	Корректно идентифицированы риски неправильной валидации
Структура и оформление работы	0-2	Работа хорошо структурирована, ответы изложены четко
Итого	0-16	

Шкала оценки:

14-16: Отлично

11-13: Хорошо

8-10: Удовлетворительно

0-7: Неудовлетворительно

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.

2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>

3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.9. Принципы проектирования архитектуры ИИ-систем: модульность, масштабируемость, эффективность.

Цель ВСП: сформировать у студентов системное понимание ключевых принципов проектирования архитектуры ИИ-систем. Развить навыки анализа, проектирования и оценки архитектурных решений с позиций модульности, масштабируемости и эффективности для создания надежных и поддерживаемых ИИ-решений.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание 1	Обязательный	4	1 неделя	Текстовый документ с четкими ответами на вопросы для каждого из двух сценариев.	Устный опрос, демонстрация проекта

Задание 1. Вам предстоит выступить в роли **Архитектора ИИ-систем**. Ваша задача - проанализировать принципы проектирования, оценить существующие решения и разработать архитектурную концепцию для конкретного бизнес-кейса. Работа состоит из трех этапов.

Рекомендации по выполнению:

Распределите время равномерно между тремя блоками.

В Block 1 фокусируйтесь на практической применимости принципов.

В Block 2 используйте конкретные примеры и аналогии.

В Block 3 уделите внимание не только схеме, но и обоснованию решений.

Думайте о компромиссах - идеальной архитектуры не существует.

Учитывайте операционные аспекты (мониторинг, деплой, обновления).

Результат работы: Готовность к роли архитектора ИИ-систем с пониманием ключевых принципов проектирования и способностью их применять на практике.

Этап 1. Анализ принципов проектирования.

Создайте развернутую аналитическую таблицу трех ключевых принципов проектирования ИИ-систем:

Принципы для анализа:

1. Модульность

2. Масштабируемость

3. Эффективность

Таблица должна включать следующие аспекты:

Критерий	Модульность	Масштабируемость	Эффективность
Определение и ключевые характеристики			
Технические методы реализации			
Метрики и показатели оценки			

Критерий	Модульность	Масштабируемость	Эффективность
Паттерны проектирования			
Типичные антипаттерны			
Влияние на стоимость разработки и поддержки			
Примеры из практики			

Дополнительное задание:

Проанализируйте взаимосвязь между тремя принципами. Как улучшение одного параметра может повлиять на другие? Приведите конкретные примеры синергии и компромиссов.

Этап 2. Критический анализ существующих решений.

Проанализируйте две архитектурные концепции и выполните сравнительный анализ:

Архитектура 1: Монолитная система рекомендаций

- Единое приложение для обработки данных, обучения моделей и генерации рекомендаций
- Общие базы данных и кэши
- Пакетное обновление моделей раз в сутки

Архитектура 2: Микросервисная система компьютерного зрения

- Отдельные сервисы для предобработки изображений, инференса моделей, постобработки
- Сервис обнаружения аномалий в работе системы
- Автоматическое масштабирование под нагрузку

Вопросы для анализа:

1. Оцените каждую архитектуру по принципам модульности, масштабируемости и эффективности (по 10-балльной шкале)
2. Выявите 2-3 ключевых преимущества и недостатка каждой архитектуры
3. Предложите по 2 улучшения для каждой архитектуры
4. Для каких бизнес-сценариев подходит каждая архитектура?

Блок 3. Проектирование архитектуры ИИ-системы.

Разработайте архитектурную концепцию для системы умного анализа отзывов:

Бизнес-требования:

- Обработка 10,000+ отзывов в день в реальном времени
- Поддержка 5+ языков
- Анализ тональности, извлечение ключевых тем, обнаружение жалоб
- Интеграция с CRM и системами поддержки клиентов
- Возможность быстрого добавления новых языков и типов анализа

Технические требования:

Время отклика < 2 секунд

Доступность 99.9%

Возможность обработки пиковых нагрузок (до 50,000 отзывов/день)

Простота обновления и дообучения моделей

Задание:

1. Разработайте блок-схему архитектуры системы (компоненты и связи)
2. Опишите принципы взаимодействия между компонентами
3. Обоснуйте, как ваше решение обеспечивает:
 - Модульность (разделение ответственности, переиспользование компонентов)
 - Масштабируемость (горизонтальное/вертикальное масштабирование)
 - Эффективность (оптимизация ресурсов, производительность)
4. Предложите план внедрения с указанием приоритетов и этапов

Критерии оценки:

Критерий	Баллы	Описание
Глубина анализа принципов	0-6	Полнота раскрытия принципов, качество примеров и метрик
Качество сравнительного анализа	0-5	Объективность оценки, глубина выявленных преимуществ и недостатков
Практическая ценность архитектурного решения	0-8	Соответствие требованиям, реализуемость, баланс принципов
Обоснованность технических решений	0-5	Логичность выбора архитектурных паттернов и технологий
Структура и оформление работы	0-4	Четкость изложения, качество схем и документации
Итого	0-28	

Шкала оценки:

24-28: Отлично

19-23: Хорошо

14-18: Удовлетворительно

0-13: Неудовлетворительно

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.

2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>

3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.10. Контейнеризация ИИ-систем.

Цель ВСР: сформировать у студентов практическое понимание принципов контейнеризации ИИ-систем, освоить инструменты контейнеризации и оркестрации, развить навыки проектирования отказоустойчивых и масштабируемых развертываний ML-моделей в продакшн-среде.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание 1	Обязательный	2	1 неделя	Формат: PDF + архив с конфигурационными файлами	Устный опрос, демонстрация проекта

Задание 1. Вам предстоит выступить в роли **MLOps инженера**. Ваша задача - спроектировать контейнеризованное решение для развертывания и масштабирования ИИ-системы, учитывая все этапы ML-жизненного цикла. Работа состоит из трех этапов.

Рекомендации по выполнению:

Начинайте с проектирования архитектуры, затем переходите к деталям реализации

Учитывайте специфику разных типов ML-нагрузок (training/inference)

Продумайте стратегии мониторинга и логирования

Обращайте внимание на security аспекты контейнеризации

Тестируйте предложенные docker-конфигурации на практике

Документируйте принятые архитектурные решения

Ожидаемый результат:

Полноценный проект контейнеризованной ML-системы, готовый к прототипированию и демонстрации принципов MLOps.

Этап 1. Анализ инструментов и практик контейнеризации.

Часть 1.1: Сравнительный анализ инструментов.

Создайте сравнительную таблицу инструментов контейнеризации и оркестрации:

Критерий	Docker	Kubernetes	Docker Compose	Kubeflow
Основное назначение				
Уровень абстракции				
Сложность настройки				
Масштабируемость				
Интеграция с ML-инструментами				
Лучшие сценарии использования в ML				

Часть 1.2: Анализ практик

Ответьте на вопросы:

1. Какие преимущества дает контейнеризация на разных этапах ML-жизненного цикла?
2. В чем разница между контейнеризацией тренировки и инференса моделей?
3. Как решаются проблемы управления версиями моделей в контейнерах?
4. Какие специфические требования к контейнеризации GPU-нагрузок?

Этап 2. Проектирование контейнеризованной ML-системы.

Техническое задание:

Разработайте архитектуру системы прогнозирования спроса для сети розничных магазинов.

Требования:

Ежедневное переобучение моделей на новых данных

Обслуживание 1000+ запросов в минуту в пиковые часы

Поддержка A/B тестирования моделей

Мониторинг дрейфа данных и концепта

Откат на предыдущую версию модели при проблемах

Задание:

1. Спроектируйте схему контейнеризованной архитектуры системы
2. Определите набор необходимых Docker-образов и их взаимосвязи
3. Разработайте docker-compose.yml для локального развертывания
4. Опишите манифесты Kubernetes для продакшн-развертывания
5. Предложите стратегию управления версиями моделей и контейнеров

Этап 3. Решение практических кейсов.

Кейс 1: Миграция монолитной ML-системы в контейнеры

Существующая система:

- Python-скрипты для обработки данных и обучения
- Flask API для обслуживания моделей
- Ручное развертывание и обновление
- Проблемы с воспроизводимостью и масштабированием

Задачи:

1. Разработайте план миграции на контейнеры
2. Определите этапы и их приоритетность
3. Предложите решение для бесшовного перехода

Кейс 2: Оптимизация costs в облачном Kubernetes

Ситуация:

- Высокие затраты на GPU-ноды в GKE/AKS/EKS
- Неравномерная нагрузка на инференс
- Простой ресурсов в ночное время

Задачи:

1. Предложите стратегию автоматического масштабирования
2. Разработайте подход к mixed-precision инференсу
3. Предложите решение для batch-обработки в периоды низкой нагрузки

Кейс 3: Multi-cloud развертывание

Требования:

- Развертывание системы в разных облачных провайдерах
- Единая точка входа для инференса
- Синхронизация моделей между регионами

Задачи:

1. Предложите архитектуру решения
2. Опишите инструменты для кроссплатформенного развертывания
3. Разработайте стратегию геораспределения нагрузки

Критерии оценки:

Критерий	Баллы	Описание
Глубина анализа инструментов	0-6	Полнота сравнения, понимание нюансов применения
Качество архитектурного решения	0-8	Соответствие требованиям, масштабируемость, отказоустойчивость
Практическая применимость решений	0-8	Реализуемость, оптимизация ресурсов, управление версиями
Оформление технической документации	0-4	Качество схем, docker-compose, манифестов
Структура и полнота работы	0-2	Логичность изложения, завершенность
Итого	0-28	

Шкала оценки:

- 24-28: Отлично
19-23: Хорошо
14-18: Удовлетворительно
0-13: Неудовлетворительно

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.
2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>
3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

Тема 1.11. Обеспечение безопасности и надежности ИИ-систем

Цель ВСР: сформировать у студентов системное понимание угроз безопасности и требований к надежности ИИ-систем. Развить навыки идентификации рисков, проектирования защищенных архитектур и разработки стратегий обеспечения отказоустойчивости AI-решений.

Трудоемкость

Количество заданий (задач, упражнений)	Характер задачи (обязательный/рекомендательный)	Норма времени (в часах по рабочей программе)	Срок выполнения (в неделях)	Форма представления материала	Форма контроля каждого задания
Задание 1	Обязательный	2	1 неделя	Формат: PDF + архив с конфигурационными файлами	Устный опрос, демонстрация проекта

Задание 1. Вам предстоит выступить в роли Security AI Engineer. Проведите анализ угроз безопасности и надежности для реальных ИИ-систем и разработайте практические рекомендации по их устранению. Работа состоит из двух этапов.

Рекомендации по выполнению:

Начните с систематизации знаний об угрозах ИИ-систем

Для каждого кейса учитывайте domain-specific требования

Предлагайте многоуровневую защиту (prevention, detection, response)

Уделяйте внимание как техническим, так и процессным аспектам безопасности

Думайте о балансе между безопасностью и производительностью

Используйте принцип "security by design"

Этап 1. Анализ угроз и уязвимостей ИИ-систем.

Часть 1.1: Классификация угроз

Создайте таблицу классификации угроз безопасности ИИ-систем:

Категория угроз	Конкретные примеры	Воздействие на систему	Методы защиты
Adversarial атаки			
Инъекции в ML-модели			
Кража моделей			
Отравление данных			
Нарушения конфиденциальности			
Атаки на инфраструктуру			

Часть 1.2: Анализ надежности

Ответьте на вопросы:

1. Какие метрики надежности наиболее важны для продакшн ИИ-систем?
2. Как обеспечить отказоустойчивость при деградации качества модели?
3. Какие стратегии мониторинга дрейфа данных наиболее эффективны?
4. Как организовать graceful degradation при сбоях в работе ИИ-компонентов?

Этап 2. Практические кейсы.

Кейс 1: Система распознавания лиц для физического доступа

Контекст:

- Система контроля доступа в secure facility
- CNN-модель для верификации сотрудников
- Real-time инференс на edge-устройствах

Задание:

1. Определите 3 наиболее критичные угрозы безопасности
2. Предложите конкретные контрмеры для каждой угрозы
3. Разработайте план реагирования на инциденты
4. Предложите архитектурные решения для повышения надежности

Кейс 2: FinTech система оценки кредитного скоринга

Контекст:

- ML-модель для кредитного скоринга в банке
- Обработка персональных данных клиентов
- Требования регуляторов (ЦБ РФ, GDPR)

Задание:

1. Выявите риски, связанные с конфиденциальностью данных
2. Предложите методы обеспечения explainability и fairness модели
3. Разработайте стратегию защиты от adversarial атак
4. Спроектируйте систему мониторинга аномалий предсказаний

Кейс 3: Автономная система управления складом

Контекст:

- RL-агенты для управления роботами-погрузчиками
- Real-time принятие решений
- Интеграция с системами безопасности склада

Задание:

1. Проанализируйте риски для физической безопасности
2. Предложите механизмы безопасного прерывания работы системы
3. Разработайте протоколы валидации решений ИИ
4. Создайте концепцию redundant систем контроля

Критерии оценки:

Критерий	Баллы	Описание
Полнота анализа угроз	0-4	Системный подход, охват различных категорий угроз
Практическая ценность рекомендаций	0-5	Конкретность и реализуемость предложенных мер защиты
Глубина проработки кейсов	0-4	Учет специфики домена, комплексность решений
Понимание метрик надежности	0-3	Корректное использование профессиональной терминологии

Критерий	Баллы	Описание
Структура и оформление	0-2	Логичность изложения, качество таблиц и схем
Итого	0-18	

Шкала оценки:

16-18: Отлично

13-15: Хорошо

10-12: Удовлетворительно

0-9: Неудовлетворительно

Информационное обеспечение:

1. Иванов В. В. Машинное обучение: Практическое руководство. / В.В Иванов — СПб.: Питер, 2021, — 380 с.

2. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 128 с. — ISBN 978-5-00101-908-4. — Текст : электронный // ЭБС PROФобразование : [сайт]. — URL: <https://profspo.ru/books/98551>

3. Жданов А.А. Автономный искусственный интеллект [Электронный ресурс] / Жданов А.А.— Электрон. текстовые данные.— М.: Лаборатория знаний, 2024.— 360 с.— Режим доступа: <https://profspo.ru/books/135845>

В оформлении презентаций выделяют два блока: оформление слайдов и представление информации на них. Для создания качественной презентации необходимо соблюдать ряд требований, предъявляемых к оформлению данных блоков.

Оформление слайдов:

Стиль	<ul style="list-style-type: none"> · Соблюдайте единый стиль оформления · Избегайте стилей, которые будут отвлекать от самой презентации. · Вспомогательная информация (управляющие кнопки) не должны преобладать над основной информацией (текстом, иллюстрациями).
Фон	Для фона предпочтительны холодные тона
Использование цвета	<ul style="list-style-type: none"> · На одном слайде рекомендуется использовать не более трех цветов: один для фона, один для заголовка, один для текста. · Для фона и текста используйте контрастные цвета. · Обратите внимание на цвет гиперссылок (до и после использования).
Анимационные эффекты	<ul style="list-style-type: none"> · Используйте возможности компьютерной анимации для представления информации на слайде. · Не стоит злоупотреблять различными анимационными эффектами, они не должны отвлекать внимание от содержания информации на слайде.

Представление информации:

Содержание информации	<ul style="list-style-type: none"> · Используйте короткие слова и предложения. · Минимизируйте количество предлогов, наречий, прилагательных. · Заголовки должны привлекать внимание аудитории.
Расположение информации на странице	<ul style="list-style-type: none"> · Предпочтительно горизонтальное расположение информации. · Наиболее важная информация должна располагаться в центре экрана. · Если на слайде располагается картинка(фото), надпись должна располагаться под ней.
Шрифты	<ul style="list-style-type: none"> · Для заголовков – не менее 24. · Для информации не менее 18. · Шрифты без засечек легче читать с большого расстояния. · Нельзя смешивать разные типы шрифтов в одной презентации. · Для выделения информации следует использовать жирный шрифт, курсив или подчеркивание. · Нельзя злоупотреблять прописными буквами (они читаются хуже строчных).
Способы выделения информации	· Следует использовать: рамки; границы, заливку; штриховку, стрелки; рисунки, диаграммы, фото.
Объем информации	<ul style="list-style-type: none"> · Не стоит заполнять один слайд слишком большим объемом информации: люди могут одновременно запомнить не более трех фактов, выводов, определений. · Наибольшая эффективность достигается тогда, когда ключевые пункты отображаются по одному на каждом отдельном слайде.
Виды слайдов	Для обеспечения разнообразия следует использовать разные виды слайдов: с текстом; с таблицами; с диаграммами, иллюстрациями, фото и т.д.

Основные критерии оценки презентации:

1. Структура. Структура презентации соответствует общепринятой структуре (Наличие заголовка, фамилии авторов).

2. Содержание.

3. Оформление. Вставка иллюстраций, фото (по необходимости), использование эффектов анимации, звукового сопровождения. Отсутствие орфографических и пунктуационных ошибок. Текст легко читается. Презентация не перегружена анимацией и картинками.

4. Коллективная работа. Слаженная работа в группе.

5. Понятность. Презентация не содержит логических ошибок и понятна практически без комментариев.