

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

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ) «МАШИННОЕ ОБУЧЕНИЕ НА БОЛЬШИХ ДАННЫХ»

Уровень высшего образования: магистратура

Направление подготовки: 09.04.03 Прикладная информатика

Направленность (профиль) подготовки: Искусственный интеллект и большие данные

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Год набора (приема на обучение): 2026

Срок получения образования: 2 года

Объем:
в зачетных единицах: 5 з.е.
в академических часах: 180 ак.ч.

г. Самара, 2026

Разработчики:

Не имеет Колотилина М. А.

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

Согласование и утверждение

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

1. Цель и задачи освоения дисциплины (модуля)

Цель освоения дисциплины - формирование результатов обучения, обеспечивающих достижение планируемых результатов освоения образовательной программы.

Задачи изучения дисциплины:

- Формирование научно-обоснованного подхода к управленческой деятельности в сложных производственных и социально-экономических системах с использованием математического аппарата. ;
- Развитие математической грамотности для работы с научной, методической и справочной литературой, для эффективного поиска и применения информации в профессиональной деятельности.;
- Приобретение навыков разработки инструментов и методов сбора исходных данных в рамках управления работами по сопровождению и проектами создания ИС ..

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

Компетенции, индикаторы и результаты обучения

УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий

УК-1.2 Определяет, интерпретирует и ранжирует информацию, требуемую для решения проблемных ситуаций

Знать:

УК-1.2/Зн1 Методы определения, интерпретации и ранжирования информации в задачах машинного обучения на больших данных (выбор признаков, оценка качества данных, интерпретируемость моделей); критерии значимости данных и признаков для построения моделей; подходы к ранжированию гипотез и альтернативных решений (метрики качества, вычислительная сложность, устойчивость к шумам); методы системного анализа для выделения ключевой информации в больших массивах.

Уметь:

УК-1.2/Ум1 Определять, интерпретировать и ранжировать информацию, требуемую для решения проблемных ситуаций в области ML на больших данных; проводить критический анализ признакового пространства и выбирать наиболее значимые признаки; интерпретировать результаты моделей (важность признаков, SHAP, LIME) и ранжировать их по влиянию на целевую переменную; ранжировать альтернативные алгоритмы по критериям точности, скорости обучения, масштабируемости и интерпретируемости; применять системный подход для выделения ключевых факторов, влияющих на качество модели.

Владеть:

УК-1.2/Вв1 Навыками определения, интерпретации и ранжирования информации при работе с большими данными и распределёнными ML-фреймворками (Spark MLlib, Dask); методами оценки важности признаков и отбора признаков (Feature Selection) для снижения размерности; техниками интерпретации моделей (SHAP, LIME, деревья решений) и ранжирования результатов; способностью ранжировать проблемные ситуации по степени влияния на бизнес-результат и предлагать приоритетные направления для улучшения модели; навыками документирования результатов анализа и интерпретации для принятия обоснованных управленческих решений.

ПК-3 Способен контролировать фактическое исполнение проекта малого и среднего уровня сложности в области ИТ

ПК-3.1 Изучает и анализирует информацию, необходимую для разработки в области ИТ малого и среднего уровня сложности

Знать:

ПК-3.1/Зн1 Методы и подходы к изучению и анализу информации, необходимой для разработки проектов в области ИТ малого и среднего уровня сложности (включая проекты по ML на больших данных); источники информации для планирования ML-проектов (требования заказчика, техническая документация, нормативные акты, бенчмарки, отраслевые стандарты); критерии оценки информации (достоверность, полнота, актуальность, релевантность) применительно к проектам малого и среднего уровня; принципы системного подхода для сбора и структурирования информации на этапах инициации и планирования проекта.

Уметь:

ПК-3.1/Ум1 Изучать и анализировать информацию, необходимую для разработки проекта по машинному обучению на больших данных (объём данных, характеристики данных, требования к производительности, инфраструктурные ограничения); собирать и систематизировать информацию о доступных ML-библиотеках, вычислительных ресурсах и инструментах управления проектами; проводить анализ требований заказчика и преобразовывать их в технические задачи; оценивать полноту и достоверность полученной информации для принятия решений по выбору алгоритмов, архитектуры и технологического стека; использовать результаты анализа информации для формирования плана проекта и технического задания.

Владеть:

ПК-3.1/Нв1 Навыками изучения и анализа информации для разработки проектов в области ML на больших данных; методами сбора и обработки информации из различных источников (научные статьи, документация, отраслевые отчёты, данные бенчмарков); способами критической оценки информации и выявления противоречий или пробелов; инструментами систематизации и визуализации информации (таблицы, диаграммы, ментальные карты) для подготовки отчётов и презентаций; навыками документирования результатов анализа информации и их использования для обоснования проектных решений перед заказчиком и руководством.

ПК-4 Способен управлять работами и проводить научные исследования и инструментарий в сфере ИТ

ПК-4.1 Изучает и анализирует современные научные методы исследований и инструментарий в области управления проектами

Знать:

ПК-4.1/Зн1 Современные научные методы исследований и инструментарий в области управления проектами в сфере ИТ (Agile, Scrum, Kanban, Waterfall); методы планирования и проведения научных исследований (эксперимент, сравнительный анализ, статистическая обработка, бенчмаркинг); инструменты управления проектами (Jira, Trello, MS Project, диаграммы Ганта, реестр рисков); подходы к интеграции научных методов в проектную деятельность (формулировка гипотез, сбор данных, эксперименты, анализ результатов); методы оценки эффективности проектных решений с использованием научного подхода.

Уметь:

ПК-4.1/Ум1 Изучать и анализировать современные научные методы исследований и инструментарий управления проектами для применения в проектах по машинному обучению на больших данных; выбирать и адаптировать научные методы (сравнительный анализ, А/В-тестирование, имитационное моделирование) для оценки эффективности ML-моделей и выбора оптимальных алгоритмов; применять инструменты управления проектами для планирования экспериментов, отслеживания результатов и управления рисками; анализировать современные подходы в области MLOps, Data Science и Big Data для повышения качества и скорости разработки; использовать результаты анализа научных методов для формирования стратегии исследований в проекте.

Владеть:

ПК-4.1/Нв1 Навыками изучения и анализа современных научных методов исследований и инструментария управления проектами; методами планирования и проведения экспериментов в ML-проектах (настройка гиперпараметров, кросс-валидация, сравнение моделей); инструментами управления проектами (Jira, Trello, диаграммы Ганта) для координации научной деятельности; способами интерпретации результатов исследований и их применения для принятия проектных решений; навыками документирования исследовательских работ и презентации результатов научному сообществу и руководству.

ПК-4.2 Управляет и контролирует проведение научных исследований и использование инструментария в области управления проектами в области ИТ малого и среднего уровня сложности

Знать:

ПК-4.2/Зн1 Методологию управления и контроля проведения научных исследований в области ИТ малого и среднего уровня сложности; современный инструментарий управления проектами для координации исследовательских и проектных работ (Jira, Trello, MS Project, диаграммы Ганта); методы планирования, мониторинга и оценки научных исследований (контрольные точки, метрики, вехи); формальные инструменты управления рисками и проблемами проекта при проведении исследований (реестр рисков, матрица вероятности и влияния, планы реагирования); подходы к интеграции научных исследований в проектный цикл ML-разработки.

Уметь:

ПК-4.2/Ум1 Управлять и контролировать проведение научных исследований в проектах по машинному обучению на больших данных малого и среднего уровня сложности; применять инструментарий управления проектами для планирования этапов исследований, распределения ресурсов и отслеживания прогресса; координировать работу команды при выполнении исследовательских задач (эксперименты, анализ данных, валидация моделей); оценивать качество и достоверность полученных научных результатов; управлять рисками, связанными с проведением исследований (неполнота данных, ошибки моделирования, дрейф данных, срыв сроков); формировать отчёты о ходе и результатах исследований для руководства и заказчика.

Владеть:

ПК-4.2/Нв1 Навыками управления и контроля проведения научных исследований в ML-проектах; методами применения инструментария управления проектами (Agile/Scrum, планирование вех, управление изменениями) для исследовательских работ; способами координации научной деятельности и представления результатов руководству и заказчику; навыками выявления и управления исследовательскими рисками; навыками интеграции результатов научных исследований в практические решения проектов; способностью документировать ход и результаты исследований, формировать отчёты и презентации.

3. Место дисциплины в структуре ОП

Дисциплина (модуль) «Машинное обучение на больших данных» относится к формируемой участниками образовательных отношений части образовательной программы и изучается в семестре(ах): 3.

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

Компетенция	Предшествующие дисциплины	Последующие дисциплины
ПК-3 - Способен контролировать фактическое исполнение проекта малого и среднего уровня сложности в области ИТ		
ПК-3.1 Изучает и анализирует информацию, необходимую для разработки в области ИТ малого и среднего уровня сложности	Анализ прикладных систем, Интеллектуальные информационные системы, Методы оптимизации, Производственная практика: (проектно-технологическая) практика	Анализ прикладных систем, Выполнение и защита выпускной квалификационной работы, Интеллектуальные информационные системы, Производственная практика: (проектно-технологическая) практика, Производственная практика: преддипломная, Тестирование искусственного интеллекта
ПК-4 - Способен управлять работами и проводить научные исследования и инструментарий в сфере ИТ		
ПК-4.1 Изучает и анализирует современные научные методы исследований и инструментарий в области управления проектами	Методы оптимизации, Производственная практика: (проектно-технологическая) практика	Выполнение и защита выпускной квалификационной работы, Производственная практика: (проектно-технологическая) практика, Производственная практика: преддипломная
ПК-4.2 Управляет и контролирует проведение научных исследований и использование инструментария в области управления проектами в области ИТ малого и среднего уровня сложности	Массово параллельные вычисления для ускорения машинного обучения, Производственная практика: (проектно-технологическая) практика	Выполнение и защита выпускной квалификационной работы, Массово параллельные вычисления для ускорения машинного обучения, Производственная практика: (проектно-технологическая) практика, Производственная практика: преддипломная
УК-1 - Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий		
УК-1.2 Определяет, интерпретирует и ранжирует информацию, требуемую для решения проблемных ситуаций	Методы машинного обучения, Русский язык и культура речи, Русский язык как иностранный	Выполнение и защита выпускной квалификационной работы

4. Объем дисциплины (модуля) и виды учебной работы

Период обучения	Общая трудоемкость (часы)	Общая трудоемкость (ЗЕТ)	Контактная работа (часы, всего)	Практические занятия (часы)	Групповая контактная работа (часы)	Индивидуальная контактная работа (часы)	Самостоятельная работа (часы)	Промежуточная аттестация
Третий семестр	180	5	20	20	2	0,3	123,7	Экзамен
Всего	180	5	20	20	2	0,3	123,7	34

5. Содержание дисциплины (модуля)

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

Наименование раздела, темы	Всего	Практические занятия	Самостоятельная работа
Раздел 1. Математические основы машинного обучения. Задачи классификации	70	8	59,7
Тема 1.1. Выполнение упражнений на kaggle.com в интерактивном датасете	21,7	2	19,7
Тема 1.2. Выполнение упражнений на kaggle.com в интерактивном датасете	23	2	20
Тема 1.3. Выполнение упражнений на kaggle.com в интерактивном датасете	25,3	4	20
Раздел 2. Деревья в машинном обучении. Случайный лес. Линейные классификаторы: основные понятия	76	12	64
Тема 2.1. Выполнение упражнений на kaggle.com в интерактивном датасете	24	4	20
Тема 2.2. Выполнение упражнений на kaggle.com в интерактивном датасете	26	4	22
Тема 2.3. Выполнение упражнений на kaggle.com в интерактивном датасете	26	4	22

5.2. Контрольные мероприятия по дисциплине

Вид контроля	Форма контроля/Оценочное средство
Текущий контроль	Тестирование
Промежуточная аттестация	Экзамен

№ п/п	Наименование раздела	Вид контроля/ используемые оценочные материалы	
		Текущий	Промежут. аттестация
1	Математические основы машинного обучения. Задачи классификации	Тестирование	Экзамен
2	Деревья в машинном обучении. Случайный лес. Линейные классификаторы: основные понятия	Тестирование	Экзамен

6. Оценочные материалы текущего контроля

1. Математические основы машинного обучения. Задачи классификации Тестирование

№ п/п	Содержание вопроса		Компетенция
		Правильный ответ (ключ ответа)	
1	В рамках критического анализа проблемной ситуации, связанной с высокой размерностью данных (10 000 признаков), определите, интерпретируйте и ранжируйте методы снижения размерности. Какой метод обеспечивает сохранение максимальной дисперсии данных? a) Lasso-регрессия b) PCA (Principal Component Analysis) c) t-SNE d) Отбор признаков по корреляции Ответ: b		УК-1
2	При системном подходе к выбору алгоритма классификации для больших данных (миллионы записей) ранжируйте алгоритмы по скорости обучения на распределённом кластере. Какой алгоритм обучается быстрее всего? a) Random Forest (на Spark) b) XGBoost c) Линейная логистическая регрессия (на Spark MLlib) d) Нейронная сеть (на TensorFlow) Ответ: c		УК-1
3	В рамках определения и интерпретации информации о качестве модели на больших данных вы должны выбрать метрику для несбалансированного набора данных (99% отрицательных, 1% положительных). Какая метрика является наиболее информативной? a) Accuracy b) F1-score c) MAE d) R ² Ответ: b		УК-1
4	При ранжировании методов обработки пропусков в данных для больших массивов (10 ТБ) выберите наиболее масштабируемый метод для распределённой среды: a) Заполнение средним (на одном узле) b) Удаление строк с пропусками (на Spark) c) Imputation с использованием медианы (в Dask) d) Imputation с использованием KNN (на локальной машине) Ответ: c		УК-1
5	В рамках выработки стратегии действий при обнаружении дрейфа данных (data drift) в потоковом ML-приложении ранжируйте варианты реагирования. Какое действие следует выполнить в первую очередь? a) Переобучить модель заново на всех данных b) Выполнить мониторинг и определить степень дрейфа, затем принять решение о переобучении c) Увеличить размер батча d) Сменить алгоритм Ответ: b		УК-1
6	Соотнесите метод ранжирования признаков с его характеристикой при определении и интерпретации информации о важности признаков:		УК-1

	<table><tr><th>Метод</th><th>Характеристика</th></tr><tr><td>1. Важность на основе случайного леса (Gini Importance)</td><td>А. Интерпретация отдельных предсказаний модели</td></tr><tr><td>2. SHAP (SHapley Additive exPlanations)</td><td>Б. Анализ вклада каждого признака в общую вариацию</td></tr><tr><td>3. Коэффициенты линейной регрессии</td><td>В. Интерпретация как сила и направление влияния признака на целевую переменную</td></tr></table>	Метод	Характеристика	1. Важность на основе случайного леса (Gini Importance)	А. Интерпретация отдельных предсказаний модели	2. SHAP (SHapley Additive exPlanations)	Б. Анализ вклада каждого признака в общую вариацию	3. Коэффициенты линейной регрессии	В. Интерпретация как сила и направление влияния признака на целевую переменную	
Метод	Характеристика									
1. Важность на основе случайного леса (Gini Importance)	А. Интерпретация отдельных предсказаний модели									
2. SHAP (SHapley Additive exPlanations)	Б. Анализ вклада каждого признака в общую вариацию									
3. Коэффициенты линейной регрессии	В. Интерпретация как сила и направление влияния признака на целевую переменную									
	Ответ: 1–Б, 2–А, 3–В									
7	Соотнесите проблемную ситуацию с соответствующей стратегией действий при критическом анализе ML-проекта: <table><tr><th>Проблемная ситуация</th><th>Стратегия действий</th></tr><tr><td>1. Качество модели низкое (переобучение)</td><td>А. Внедрить мониторинг дрейфа и автоматическое переобучение</td></tr><tr><td>2. Модель хорошо работает на обучении, но плохо на тесте</td><td>Б. Использовать кросс-валидацию, регуляризацию, добавить данные</td></tr><tr><td>3. Данные изменяются со временем (дрейф)</td><td>В. Уменьшить сложность модели, применить регуляризацию, сократить признаки</td></tr></table>	Проблемная ситуация	Стратегия действий	1. Качество модели низкое (переобучение)	А. Внедрить мониторинг дрейфа и автоматическое переобучение	2. Модель хорошо работает на обучении, но плохо на тесте	Б. Использовать кросс-валидацию, регуляризацию, добавить данные	3. Данные изменяются со временем (дрейф)	В. Уменьшить сложность модели, применить регуляризацию, сократить признаки	УК-1
Проблемная ситуация	Стратегия действий									
1. Качество модели низкое (переобучение)	А. Внедрить мониторинг дрейфа и автоматическое переобучение									
2. Модель хорошо работает на обучении, но плохо на тесте	Б. Использовать кросс-валидацию, регуляризацию, добавить данные									
3. Данные изменяются со временем (дрейф)	В. Уменьшить сложность модели, применить регуляризацию, сократить признаки									
	Ответ: 1–В, 2–Б, 3–А									
8	Соотнесите критерий ранжирования информации с его описанием при системном подходе к анализу данных: <table><tr><th>Критерий</th><th>Описание</th></tr><tr><td>1. Полнота данных</td><td>А. Соответствие данных требуемому формату и диапазонам</td></tr><tr><td>2. Актуальность данных</td><td>Б. Доля отсутствующих значений в наборе данных</td></tr><tr><td>3. Достоверность данных</td><td>В. Степень соответствия данных текущему состоянию предметной области</td></tr></table>	Критерий	Описание	1. Полнота данных	А. Соответствие данных требуемому формату и диапазонам	2. Актуальность данных	Б. Доля отсутствующих значений в наборе данных	3. Достоверность данных	В. Степень соответствия данных текущему состоянию предметной области	УК-1
Критерий	Описание									
1. Полнота данных	А. Соответствие данных требуемому формату и диапазонам									
2. Актуальность данных	Б. Доля отсутствующих значений в наборе данных									
3. Достоверность данных	В. Степень соответствия данных текущему состоянию предметной области									
	Ответ: 1–Б, 2–В, 3–А									
9	Расположите в правильной последовательности этапы критического анализа проблемной ситуации при выборе алгоритма машинного обучения для больших данных: 1. Определение требований к модели (точность, скорость, интерпретируемость) 2. Ранжирование алгоритмов по критериям (точность, скорость, масштабируемость) 3. Интерпретация результатов тестирования и выбор лучшего алгоритма 4. Системный анализ данных и постановка задачи (классификация/регрессия/кластеризация) 5. Проведение экспериментов с отобранными алгоритмами на тестовой выборке 6. Выработка стратегии по улучшению модели (если необходимо)	УК-1								
	Ответ: 4 → 1 → 2 → 5 → 3 → 6									
10	Расположите в правильной последовательности этапы критического анализа проблемной ситуации при выборе алгоритма машинного обучения для больших данных: 1. Определение типа дрейфа (дрейф признаков, дрейф целевой переменной) 2. Интерпретация влияния дрейфа на качество модели (метрики) 3. Ранжирование признаков по степени дрейфа (с использованием PSI, KS-теста) 4. Системный анализ причин дрейфа (изменение источника данных, внешние факторы) 5. Выработка стратегии реагирования (переобучение, корректировка, мониторинг) 6. Реализация стратегии и повторный мониторинг	УК-1								
	Ответ: 1 → 2 → 3 → 4 → 5 → 6									

11	Дайте развернутый ответ В рамках критического анализа проблемной ситуации в ML-проекте по прогнозированию оттока клиентов определите, интерпретируйте и ранжируйте факторы, влияющие на качество модели (данные, признаки, алгоритм, инфраструктура), и предложите стратегию действий по улучшению качества.		УК-1
	Ответ:	Ранжирование: качество данных → отбор признаков → выбор алгоритма → инфраструктура; стратегия: очистка данных, Feature Engineering, тестирование нескольких алгоритмов, оптимизация гиперпараметров, мониторинг дрейфа.	
12	Дайте развернутый ответ При системном подходе к выбору фреймворка для распределённого ML (Spark MLlib, Dask, TensorFlow) определите и ранжируйте критерии выбора для проекта с данными 50 ТБ и требованием к интерпретируемости моделей. Предложите стратегию выбора.		УК-1
	Ответ:	Критерии: масштабируемость, скорость, интерпретируемость, стоимость; ранжирование: Spark MLlib (масштабируемость, интерпретируемость), Dask (гибкость, Python-экосистема), TensorFlow (глубокое обучение, сложнее интерпретация); стратегия: начать с Spark MLlib для линейных моделей и случайного леса.	
13	Дайте развернутый ответ В рамках ранжирования информации о качестве модели рассчитайте F1-score, если Precision = 0.8, Recall = 0.6. Ответ дайте с точностью до двух знаков.		УК-1
	Ответ:	$F1 = 2 * (0.8 * 0.6) / (0.8 + 0.6) = 0.96 / 1.4 \approx 0.69$.	
14	Дайте развернутый ответ Опишите стратегию действий по обработке дисбаланса классов в задаче обнаружения мошеннических транзакций (99.9% – не мошеннические, 0.1% – мошеннические). Определите и ранжируйте методы борьбы с дисбалансом.		УК-1
	Ответ:	Стратегия: 1) Сбор дополнительных данных (если возможно); 2) Использование методов сэмплирования (SMOTE для увеличения миноритарного класса, оверсемплинг); 3) Применение алгоритмов, устойчивых к дисбалансу (XGBoost с параметром scale_pos_weight); 4) Использование метрик F1, ROC-AUC вместо Accuracy; 5) Аугментация данных.	
15	Дайте развернутый ответ В рамках критического анализа проблемной ситуации выяснилось, что модель, обученная на данных за 2024 год, показывает низкое качество на данных за 2025 год. Определите, интерпретируйте и ранжируйте возможные причины (дрейф данных, изменение бизнес-процессов, переобучение) и предложите стратегию по устранению проблемы.		УК-1
	Ответ:	Причины: дрейф данных (наиболее вероятно), изменение бизнес-процессов, переобучение. Стратегия: мониторинг дрейфа (PSI), сбор новых данных, регулярное переобучение (retraining), использование онлайн-обучения или обучение с подкреплением.	

2. Деревья в машинном обучении. Случайный лес. Линейные классификаторы: основные понятия Тестирование

№ п/п	Содержание вопроса		Компетенция
		Правильный ответ (ключ ответа)	
1	В рамках мониторинга реализации проекта по разработке ML-модели на больших данных вы обнаружили, что фактическое время обучения модели на 30% превышает плановое. Какой инструмент контроля фактического исполнения следует применить для выявления причины? а) Диаграмма Ганта б) Профилирование распределённого кластера (Spark UI, Ganglia) с) SWOT-анализ д) Матрица ответственности RACI		ПК-3
	Ответ:	б	
2	При контроле фактического исполнения проекта по внедрению пайплайна ML выяснилось, что использование GPU составляет 95%, но обучение модели замедлилось. Какое управляющее действие следует предпринять для мониторинга реализации? а) Увеличить количество узлов кластера б) Проверить эффективность использования GPU (утилизацию, узкие места) и оптимизировать код с) Заменить GPU на более новые модели д) Перейти с PyTorch на TensorFlow		ПК-3
	Ответ:	б	
3	В рамках мониторинга реализации проекта малого уровня сложности по разработке ML-решения вы должны выбрать метрику для оценки фактического исполнения по срокам и объёму работ: а) Использование GPU/CPU б) Процент завершённых задач от общего объёма работ (Earned Value) с) Количество строк кода д) Частота коммитов в Git		ПК-3
	Ответ:	б	

4	При контроле фактического исполнения проекта по созданию ML-пайплайна вы обнаружили риск превышения бюджета из-за необходимости дополнительного оборудования для обработки данных. Какой инструмент управления рисками следует применить? a) Реестр рисков и план реагирования b) Диаграмма Ганта c) Доска Канбан d) Устав проекта	ПК-3								
	Ответ: a									
5	В рамках мониторинга реализации проекта по масштабированию ML-модели на кластере вы фиксируете отставание по срокам на 20% на этапе настройки гиперпараметров. Какое управляющее действие является приоритетным для контроля фактического исполнения? a) Уволить руководителя проекта b) Пересмотреть план-график, сократить количество экспериментов и использовать автоматизированный подбор гиперпараметров (Hureport, Optuna) c) Увеличить бюджет на 40% d) Прекратить проект	ПК-3								
	Ответ: b									
6	Соотнесите этап мониторинга ML-проекта с его содержанием при контроле фактического исполнения: <table><thead><tr><th>Этап мониторинга</th><th>Содержание</th></tr></thead><tbody><tr><td>1. Сбор данных о ходе работ</td><td>А. Оценка отклонений от плана по срокам, бюджету и качеству модели (метрики)</td></tr><tr><td>2. Анализ отклонений</td><td>Б. Регулярное обновление статуса задач в Jira, фиксация метрик экспериментов (MLflow)</td></tr><tr><td>3. Корректирующие действия</td><td>В. Перераспределение ресурсов, оптимизация пайплайна, изменение плана экспериментов</td></tr></tbody></table>	Этап мониторинга	Содержание	1. Сбор данных о ходе работ	А. Оценка отклонений от плана по срокам, бюджету и качеству модели (метрики)	2. Анализ отклонений	Б. Регулярное обновление статуса задач в Jira, фиксация метрик экспериментов (MLflow)	3. Корректирующие действия	В. Перераспределение ресурсов, оптимизация пайплайна, изменение плана экспериментов	ПК-3
Этап мониторинга	Содержание									
1. Сбор данных о ходе работ	А. Оценка отклонений от плана по срокам, бюджету и качеству модели (метрики)									
2. Анализ отклонений	Б. Регулярное обновление статуса задач в Jira, фиксация метрик экспериментов (MLflow)									
3. Корректирующие действия	В. Перераспределение ресурсов, оптимизация пайплайна, изменение плана экспериментов									
	Ответ: 1–Б, 2–А, 3–В									
7	Соотнесите инструмент управления проектами с его применением для мониторинга реализации и контроля фактического исполнения в ML-проекте: <table><thead><tr><th>Инструмент</th><th>Применение в проекте</th></tr></thead><tbody><tr><td>1. Jira + BigGantt</td><td>А. Отслеживание задач, спринтов, бэклога и визуализация сроков (диаграмма Ганта)</td></tr><tr><td>2. MLflow / Weights & Biases</td><td>Б. Логирование экспериментов, сравнение метрик моделей, отслеживание версий моделей</td></tr><tr><td>3. Prometheus + Grafana</td><td>В. Мониторинг производительности кластера (CPU, GPU, память, сеть) в реальном времени</td></tr></tbody></table>	Инструмент	Применение в проекте	1. Jira + BigGantt	А. Отслеживание задач, спринтов, бэклога и визуализация сроков (диаграмма Ганта)	2. MLflow / Weights & Biases	Б. Логирование экспериментов, сравнение метрик моделей, отслеживание версий моделей	3. Prometheus + Grafana	В. Мониторинг производительности кластера (CPU, GPU, память, сеть) в реальном времени	ПК-3
Инструмент	Применение в проекте									
1. Jira + BigGantt	А. Отслеживание задач, спринтов, бэклога и визуализация сроков (диаграмма Ганта)									
2. MLflow / Weights & Biases	Б. Логирование экспериментов, сравнение метрик моделей, отслеживание версий моделей									
3. Prometheus + Grafana	В. Мониторинг производительности кластера (CPU, GPU, память, сеть) в реальном времени									
	Ответ: 1–А, 2–Б, 3–В									
8	Соотнесите тип проблемы с соответствующим управляющим действием при контроле фактического исполнения ML-проекта: <table><thead><tr><th>Проблема</th><th>Управляющее действие</th></tr></thead><tbody><tr><td>1. Отставание по срокам на этапе сбора данных</td><td>А. Оптимизация ETL-пайплайна, параллелизация, использование Spark/Dask</td></tr><tr><td>2. Снижение качества модели на тестовых данных</td><td>Б. Переобучение с новыми данными, анализ дрейфа, корректировка признакового пространства</td></tr><tr><td>3. Превышение бюджета на GPU-время</td><td>В. Оптимизация кода, уменьшение объёма данных, переход на более дешёвый инстанс</td></tr></tbody></table>	Проблема	Управляющее действие	1. Отставание по срокам на этапе сбора данных	А. Оптимизация ETL-пайплайна, параллелизация, использование Spark/Dask	2. Снижение качества модели на тестовых данных	Б. Переобучение с новыми данными, анализ дрейфа, корректировка признакового пространства	3. Превышение бюджета на GPU-время	В. Оптимизация кода, уменьшение объёма данных, переход на более дешёвый инстанс	ПК-3
Проблема	Управляющее действие									
1. Отставание по срокам на этапе сбора данных	А. Оптимизация ETL-пайплайна, параллелизация, использование Spark/Dask									
2. Снижение качества модели на тестовых данных	Б. Переобучение с новыми данными, анализ дрейфа, корректировка признакового пространства									
3. Превышение бюджета на GPU-время	В. Оптимизация кода, уменьшение объёма данных, переход на более дешёвый инстанс									
	Ответ: 1–А, 2–Б, 3–В									
9	Расположите в правильной последовательности шаги мониторинга реализации ML-проекта по внедрению модели в продуктивную среду: 1 Выявление отклонений от плана и их анализ (по срокам, метрикам, бюджету) 2 Разработка плана корректирующих действий (оптимизация, переобучение, дополнительные ресурсы) 4 Сбор фактических данных о ходе работ (статусы задач, метрики моделей, использование ресурсов) 5 Сравнение плановых и фактических показателей (Earned Value, метрики модели) 6 Реализация корректирующих действий и обновление плана 7 Документирование изменений и отчётность перед руководством	ПК-3								

	Ответ:	3 → 4 → 1 → 2 → 5 → 6	
10	Расположите в правильной последовательности этапы контроля фактического исполнения проекта по разработке распределённого ML-решения: 1 Сбор требований к производительности и ресурсам (время обучения, стоимость, точность) 2 Разработка плана-графика и распределение задач между членами команды 3 Регулярные встречи для обновления статуса и выявления проблем (Scrum, Kanban) 4 Оценка фактически достигнутых результатов (метрики: время обучения, Accuracy, использование GPU) 5 Анализ отклонений и принятие управленческих решений 6 Закрытие проекта и фиксация итоговых результатов		ПК-3
	Ответ:	1 → 2 → 3 → 4 → 5 → 6	
11	Дайте развернутый ответ В рамках мониторинга реализации проекта по разработке ML-модели для предсказания спроса выявили, что фактическое время обучения модели на 25% превышает плановое, а точность на 5% ниже ожидаемой. Предложите план корректирующих действий с использованием формальных инструментов управления рисками.		ПК-3
	Ответ:	1. Провести профилирование кластера (Spark UI, GPU utilisation); 2. Оптимизировать код (уменьшить данные, использовать кэширование, выборку); 3. Проверить качество данных (дрейф, пропуски); 4. Обновить реестр рисков; 5. Пересмотреть график и выделить дополнительные вычислительные ресурсы (по возможности).	
12	Дайте развернутый ответ В рамках контроля фактического исполнения проекта рассчитайте показатель освоенного объёма (Earned Value) на текущую дату, если плановая стоимость выполненных работ (PV) составляет 500 000 руб., фактическая стоимость (AC) – 550 000 руб., а процент выполнения работ – 80% от общего объёма стоимостью 800 000 руб. Ответ дайте в рублях.		ПК-3
	Ответ:	$EV = 800\,000 \cdot 0.8 = 640\,000$ руб.	
13	Дайте развернутый ответ Опишите, как вы будете осуществлять мониторинг использования ресурсов (GPU, память, время) при выполнении проекта по распределённому обучению ML-модели на больших данных. Какие инструменты и метрики примените для контроля фактического исполнения?		ПК-3
	Ответ:	Использовать NVIDIA DCGM, Prometheus + Grafana для сбора метрик; ключевые метрики: GPU Utilisation, Memory Usage, Power Draw, Temperature, Throughput (samples/sec), Model Loss; отслеживать отклонения в реальном времени, оповещения при превышении порогов; логировать метрики в MLflow.	
14	Дайте развернутый ответ В рамках мониторинга реализации проекта по внедрению ML-модели в продуктивную среду вы обнаружили, что модель начала давать ошибки на 10% данных из-за дрейфа. Предложите стратегию действий по восстановлению качества модели с применением формальных инструментов управления проблемами.		ПК-3
	Ответ:	1. Зафиксировать проблему в реестре проблем; 2. Провести анализ дрейфа (PSI, распределения); 3. Переобучить модель на новых данных (с учётом дрейфа); 4. Внедрить автоматический мониторинг дрейфа; 5. Разработать план регулярного переобучения (retraining schedule); 6. Обновить документацию и уведомить заказчика.	
15	Дайте развернутый ответ Опишите, какие риски наиболее характерны для проектов по разработке ML-решений на больших данных, и как вы будете их мониторить и контролировать в рамках проекта малого уровня сложности.		ПК-3
	Ответ:	Риски: 1) дрейф данных – мониторинг PSI; 2) недоступность вычислительных ресурсов – резервирование, облачные решения; 3) неполнота/некачественные данные – автоматические проверки качества; 4) переобучение – кросс-валидация, регуляризация; 5) превышение бюджета – контроль через реестр рисков, ежемесячный пересмотр затрат. Управление: регулярные ретроспективы, план реагирования на риски.	
16	В рамках управления проведением научных исследований по оценке эффективности распределённых алгоритмов ML для больших данных вы должны выбрать метод, позволяющий объективно сравнить производительность различных реализаций: а) Экспертная оценка б) Бенчмаркинг на стандартизированных наборах данных с) Анализ научной литературы д) Опрос разработчиков		ПК-4
	Ответ:	б	

17	При контроле проведения научных исследований в проекте по сравнению алгоритмов классификации на больших данных вы обнаружили, что результаты экспериментов плохо воспроизводятся. Какой инструментальный следует применить для повышения воспроизводимости? a) Jira (для отслеживания задач) b) Git + DVC (контроль версий кода и данных) + фиксация параметров эксперимента (MLflow) c) Диаграмма Ганта d) SWOT-анализ	ПК-4								
	Ответ: b									
18	В рамках управления научными исследованиями по масштабированию ML-моделей вы должны выбрать метрику для оценки эффективности использования кластера при проведении экспериментов: a) Количество публикаций b) Scaling Efficiency (эффективность масштабирования) c) Количество строк кода d) Частота коммитов	ПК-4								
	Ответ: b									
19	При контроле проведения научных исследований вы обнаружили, что гипотеза о преимуществе нового параллельного алгоритма не подтвердилась. Какое управленческое решение следует принять в рамках управления проектом? a) Скрыть отрицательные результаты b) Пересмотреть гипотезу, провести дополнительный анализ и скорректировать план исследований c) Прекратить проект d) Уволить исследователя	ПК-4								
	Ответ: b									
20	В рамках управления научными исследованиями в проекте малого уровня сложности по ускорению обучения нейросети на больших данных вы должны выбрать инструментальный управления проектами для планирования этапов экспериментов: a) Trello / Jira b) Adobe Photoshop c) Microsoft Word d) Notepad++	ПК-4								
	Ответ: a									
21	Соотнесите этап научного исследования с его содержанием при управлении и контроле проведения исследований в области ML на больших данных: <table><thead><tr><th>Этап исследования</th><th>Содержание</th></tr></thead><tbody><tr><td>1. Постановка задачи и выдвижение гипотезы</td><td>А. Сравнение производительности алгоритмов на бенчмарках, статистическая обработка</td></tr><tr><td>2. Проведение экспериментов</td><td>Б. Формулировка цели, выбор метрик, определение плана эксперимента (дизайн эксперимента)</td></tr><tr><td>3. Анализ результатов и выводы</td><td>В. Запуск кода на различных конфигурациях (количество узлов, GPU, параметры модели), сбор метрик</td></tr></tbody></table>	Этап исследования	Содержание	1. Постановка задачи и выдвижение гипотезы	А. Сравнение производительности алгоритмов на бенчмарках, статистическая обработка	2. Проведение экспериментов	Б. Формулировка цели, выбор метрик, определение плана эксперимента (дизайн эксперимента)	3. Анализ результатов и выводы	В. Запуск кода на различных конфигурациях (количество узлов, GPU, параметры модели), сбор метрик	ПК-4
Этап исследования	Содержание									
1. Постановка задачи и выдвижение гипотезы	А. Сравнение производительности алгоритмов на бенчмарках, статистическая обработка									
2. Проведение экспериментов	Б. Формулировка цели, выбор метрик, определение плана эксперимента (дизайн эксперимента)									
3. Анализ результатов и выводы	В. Запуск кода на различных конфигурациях (количество узлов, GPU, параметры модели), сбор метрик									
	Ответ: 1–Б, 2–В, 3–А									
22	Соотнесите инструмент управления проектами с его применением при проведении научных исследований в ML-проекте: <table><thead><tr><th>Инструмент</th><th>Применение в исследовании</th></tr></thead><tbody><tr><td>1. <u>Git</u> + DVC</td><td>А. Управление требованиями и отслеживание задач исследования</td></tr><tr><td>2. <u>Jira</u> / <u>Trello</u></td><td>Б. <u>Версионирование</u> кода, данных и воспроизводимость экспериментов</td></tr><tr><td>3. <u>MLflow</u> / <u>Weights & Biases</u></td><td>В. Логирование метрик, параметров и визуализация результатов экспериментов</td></tr></tbody></table>	Инструмент	Применение в исследовании	1. <u>Git</u> + DVC	А. Управление требованиями и отслеживание задач исследования	2. <u>Jira</u> / <u>Trello</u>	Б. <u>Версионирование</u> кода, данных и воспроизводимость экспериментов	3. <u>MLflow</u> / <u>Weights & Biases</u>	В. Логирование метрик, параметров и визуализация результатов экспериментов	ПК-4
Инструмент	Применение в исследовании									
1. <u>Git</u> + DVC	А. Управление требованиями и отслеживание задач исследования									
2. <u>Jira</u> / <u>Trello</u>	Б. <u>Версионирование</u> кода, данных и воспроизводимость экспериментов									
3. <u>MLflow</u> / <u>Weights & Biases</u>	В. Логирование метрик, параметров и визуализация результатов экспериментов									
	Ответ: 1–Б, 2–А, 3–В									
23	Соотнесите тип исследования с соответствующим инструментарием управления для контроля проведения исследований: <table><thead><tr><th>Тип исследования</th><th>Инструментарий управления</th></tr></thead><tbody><tr><td>1. Сравнительный анализ алгоритмов</td><td>А. План эксперимента, метрики (время, точность), реестр результатов</td></tr><tr><td>2. Оптимизация <u>гиперпараметров</u></td><td>Б. <u>GridSearch/RandomSearch/Bayesian</u>, логирование всех параметров в <u>MLflow</u></td></tr><tr><td>3. Масштабирование на кластер</td><td>В. Планирование ресурсов, мониторинг узлов, логирование времени выполнения</td></tr></tbody></table>	Тип исследования	Инструментарий управления	1. Сравнительный анализ алгоритмов	А. План эксперимента, метрики (время, точность), реестр результатов	2. Оптимизация <u>гиперпараметров</u>	Б. <u>GridSearch/RandomSearch/Bayesian</u> , логирование всех параметров в <u>MLflow</u>	3. Масштабирование на кластер	В. Планирование ресурсов, мониторинг узлов, логирование времени выполнения	ПК-4
Тип исследования	Инструментарий управления									
1. Сравнительный анализ алгоритмов	А. План эксперимента, метрики (время, точность), реестр результатов									
2. Оптимизация <u>гиперпараметров</u>	Б. <u>GridSearch/RandomSearch/Bayesian</u> , логирование всех параметров в <u>MLflow</u>									
3. Масштабирование на кластер	В. Планирование ресурсов, мониторинг узлов, логирование времени выполнения									

	Ответ: 1–А, 2–Б, 3–В	
24	Расположите в правильной последовательности этапы управления научным исследованием по оценке эффективности распределённых ML-алгоритмов: 1 Формулировка гипотезы и целей исследования 2 Выбор инструментария и планирование экспериментов 3 Проведение экспериментов и сбор данных 4 Анализ полученных данных и проверка гипотезы 5 Корректировка плана исследований при необходимости 6 Формулировка выводов и рекомендаций Ответ: 1 → 2 → 3 → 4 → 5 → 6	ПК-4
25	Расположите в правильной последовательности шаги контроля проведения научных исследований в проекте по разработке нового ML-алгоритма для больших данных: 1 Оценка достоверности и воспроизводимости результатов 2 Регулярная проверка выполнения плана экспериментов 3 Определение критериев качества исследования (метрики, пороговые значения) 4 Сравнение полученных результатов с литературными данными и предшествующими работами 5 Принятие решения о продолжении или корректировке исследований 6 Документирование результатов и подготовка отчёта Ответ: 3 → 2 → 1 → 4 → 5 → 6	ПК-4
26	Дайте развернутый ответ В рамках управления научными исследованиями по сравнению производительности трёх алгоритмов классификации (XGBoost, Random Forest, логистическая регрессия) на больших данных опишите план эксперимента. Какие инструменты управления проектами вы примените для обеспечения воспроизводимости? Ответ: План: 1) фиксация версий кода и данных (Git + DVC); 2) определение метрик (Accuracy, F1, время обучения); 3) проведение замеров на разных объёмах данных; 4) статистическая обработка; 5) документирование. Инструменты: Jira для задач, MLflow для логирования, DVC для данных.	ПК-4
27	Дайте развернутый ответ В рамках контроля проведения исследований вы получили следующие результаты: ускорение на GPU составило 12 раз по сравнению с CPU при использовании 1 GPU и 20 раз при использовании 2 GPU. Рассчитайте эффективность масштабирования (Scaling Efficiency) для 2 GPU (в процентах). Ответ дайте числом. Ответ: $(20 / 12) / 2 * 100\% \approx 83.3\%$	ПК-4
28	Дайте развернутый ответ Опишите, как вы будете управлять процессом научных исследований по оптимизации гиперпараметров модели глубокого обучения в условиях неопределённости (изменение требований, ошибки в данных). Какие инструменты управления рисками примените? Ответ: Ввести итеративный план с контрольными точками, реестр рисков, резервирование времени; использовать прототипирование и раннее тестирование; применять Git + DVC для воспроизводимости; регулярные ретроспективы; автоматизировать GridSearch/Bayesian с валидацией.	ПК-4
29	Дайте развернутый ответ В рамках контроля проведения научного исследования вы обнаружили, что экспериментальные данные противоречат теоретическим ожиданиям (модель показывает низкую точность). Опишите план действий по проверке и корректировке исследования. Ответ: 1) Проверить корректность кода и данных; 2) Повторить эксперименты с другими параметрами; 3) Провести статистический анализ ошибок; 4) Пересмотреть гипотезу или признаковое пространство; 5) При необходимости скорректировать план исследования; 6) Задokumentировать отклонения и выводы.	ПК-4
30	Дайте развернутый ответ Опишите, какой инструментарий управления проектами вы будете использовать для координации работы распределённой команды при проведении исследований по масштабируемости ML-моделей на кластере (5 исследователей). Обоснуйте выбор. Ответ: Jira – для планирования и отслеживания задач; Slack/Teams – для коммуникации; Git + DVC – для управления версиями кода, данных и экспериментов; MLflow – для логирования и сравнения результатов; Confluence – для общей документации; обоснование: прозрачность, воспроизводимость, контроль версий, единое хранилище результатов, снижение рисков потери данных.	ПК-4

7. Оценочные материалы промежуточной аттестации

Экзамен третий семестр

№ п/п	Содержание вопроса		Компетенция
		Правильный ответ (ключ ответа)	
1	Дайте развернутый ответ В рамках критического анализа проблемной ситуации в ML-проекте по прогнозированию оттока клиентов определите, интерпретируйте и ранжируйте факторы, влияющие на качество модели (данные, признаки, алгоритм, инфраструктура), и предложите стратегию действий по улучшению качества.		УК-1
	Ответ:	Ранжирование: качество данных → отбор признаков → выбор алгоритма → инфраструктура; стратегия: очистка данных, Feature Engineering, тестирование нескольких алгоритмов, оптимизация гиперпараметров, мониторинг дрейфа.	
2	Дайте развернутый ответ При системном подходе к выбору фреймворка для распределённого ML (Spark MLlib, Dask, TensorFlow) определите и ранжируйте критерии выбора для проекта с данными 50 ТБ и требованием к интерпретируемости моделей. Предложите стратегию выбора.		УК-1
	Ответ:	Критерии: масштабируемость, скорость, интерпретируемость, стоимость; ранжирование: Spark MLlib (масштабируемость, интерпретируемость), Dask (гибкость, Python-экосистема), TensorFlow (глубокое обучение, сложнее интерпретация); стратегия: начать с Spark MLlib для линейных моделей и случайного леса.	
3	Дайте развернутый ответ В рамках ранжирования информации о качестве модели рассчитайте F1-score, если Precision = 0.85, Recall = 0.7. Ответ дайте с точностью до двух знаков.		УК-1
	Ответ:	$F1 = 2 * (0.85 * 0.7) / (0.85 + 0.7) = 1.19 / 1.55 \approx 0.77$.	
4	Дайте развернутый ответ Опишите стратегию действий по обработке дисбаланса классов в задаче обнаружения мошеннических транзакций (99.9% – не мошеннические, 0.1% – мошеннические). Определите и ранжируйте методы борьбы с дисбалансом.		УК-1
	Ответ:	Стратегия: 1) Сбор дополнительных данных (если возможно); 2) Использование методов сэмплирования (SMOTE для увеличения миноритарного класса, оверсэмплинг); 3) Применение алгоритмов, устойчивых к дисбалансу (XGBoost с параметром scale_pos_weight); 4) Использование метрик F1, ROC-AUC вместо Accuracy; 5) Аугментация данных.	
5	Дайте развернутый ответ В рамках критического анализа проблемной ситуации выяснилось, что модель, обученная на данных за 2024 год, показывает низкое качество на данных за 2025 год. Определите, интерпретируйте и ранжируйте возможные причины (дрейф данных, изменение бизнес-процессов, переобучение) и предложите стратегию по устранению проблемы.		УК-1
	Ответ:	Причины: дрейф данных (наиболее вероятно), изменение бизнес-процессов, переобучение. Стратегия: мониторинг дрейфа (PSI), сбор новых данных, регулярное переобучение (retraining), использование онлайн-обучения или обучение с подкреплением.	
6	Дайте развернутый ответ В рамках мониторинга реализации проекта по разработке ML-модели для предсказания спроса выявили, что фактическое время обучения модели на 25% превышает плановое, а точность на 5% ниже ожидаемой. Предложите план корректирующих действий с использованием формальных инструментов управления рисками.		ПК-3
	Ответ:	1. Провести профилирование кластера (Spark UI, GPU utilisation); 2. Оптимизировать код (уменьшить данные, использовать кэширование, выборку); 3. Проверить качество данных (дрейф, пропуски); 4. Обновить реестр рисков; 5. Пересмотреть график и выделить дополнительные вычислительные ресурсы (по возможности).	
7	Дайте развернутый ответ В рамках контроля фактического исполнения проекта рассчитайте показатель освоенного объёма (Earned Value) на текущую дату, если плановая стоимость выполненных работ (PV) составляет 600 000 руб., фактическая стоимость (AC) – 650 000 руб., а процент выполнения работ – 75% от общего объёма стоимостью 1 000 000 руб. Ответ дайте в рублях.		ПК-3
	Ответ:	$EV = 1\,000\,000 * 0.75 = 750\,000$ руб.	
8	Дайте развернутый ответ Опишите, как вы будете осуществлять мониторинг использования ресурсов (GPU, память, время) при выполнении проекта по распределённому обучению ML-модели на больших данных. Какие инструменты и метрики примените для контроля фактического исполнения?		ПК-3
	Ответ:	Использовать NVIDIA DCGM, Prometheus + Grafana для сбора метрик; ключевые метрики: GPU Utilisation, Memory Usage, Power Draw, Temperature, Throughput (samples/sec), Model Loss; отслеживать отклонения в реальном времени, оповещения при превышении порогов; логировать метрики в MLflow.	
9	Дайте развернутый ответ В рамках мониторинга реализации проекта по внедрению ML-модели в продуктивную среду вы обнаружили, что модель начала давать ошибки на 10% данных из-за дрейфа. Предложите стратегию действий по восстановлению качества модели с применением формальных инструментов управления проблемами.		ПК-3

	Ответ: 1. Зафиксировать проблему в реестре проблем; 2. Провести анализ дрейфа (PSI, распределения); 3. Переобучить модель на новых данных (с учётом дрейфа); 4. Внедрить автоматический мониторинг дрейфа; 5. Разработать план регулярного переобучения (retraining schedule); 6. Обновить документацию и уведомить заказчика.	
10	Дайте развернутый ответ Опишите, какие риски наиболее характерны для проектов по разработке ML-решений на больших данных, и как вы будете их мониторить и контролировать в рамках проекта малого уровня сложности. Ответ: Риски: 1) дрейф данных – мониторинг PSI; 2) недоступность вычислительных ресурсов – резервирование, облачные решения; 3) неполнота/некачественные данные – автоматические проверки качества; 4) переобучение – кросс-валидация, регуляризация; 5) превышение бюджета – контроль через реестр рисков, ежемесячный пересмотр затрат. Управление: регулярные ретроспективы, план реагирования на риски.	ПК-3
11	Дайте развернутый ответ В рамках управления научными исследованиями по сравнению производительности трёх алгоритмов классификации (XGBoost, Random Forest, логистическая регрессия) на больших данных опишите план эксперимента. Какие инструменты управления проектами вы примените для обеспечения воспроизводимости? Ответ: План: 1) фиксация версий кода и данных (Git + DVC); 2) определение метрик (Accuracy, F1, время обучения); 3) проведение замеров на разных объёмах данных; 4) статистическая обработка; 5) документирование. Инструменты: Jira для задач, MLflow для логирования, DVC для данных.	ПК-4
12	Дайте развернутый ответ В рамках контроля проведения исследований вы получили следующие результаты: ускорение на GPU составило 12 раз по сравнению с CPU при использовании 1 GPU и 20 раз при использовании 2 GPU. Рассчитайте эффективность масштабирования (Scaling Efficiency) для 2 GPU (в процентах). Ответ дайте числом. Ответ: $(20 / 12) / 2 * 100\% \approx 83.3\%$.	ПК-4
13	Дайте развернутый ответ Опишите, как вы будете управлять процессом научных исследований по оптимизации гиперпараметров модели глубокого обучения в условиях неопределённости (изменение требований, ошибки в данных). Какие инструменты управления рисками примените? Ответ: Ввести итеративный план с контрольными точками, реестр рисков, резервирование времени; использовать прототипирование и раннее тестирование; применять Git + DVC для воспроизводимости; регулярные ретроспективы; автоматизировать GridSearch/Bayesian с валидацией.	ПК-4
14	Дайте развернутый ответ В рамках контроля проведения научного исследования вы обнаружили, что экспериментальные данные противоречат теоретическим ожиданиям (модель показывает низкую точность). Опишите план действий по проверке и корректировке исследования. Ответ: 1) Проверить корректность кода и данных; 2) Повторить эксперименты с другими параметрами; 3) Провести статистический анализ ошибок; 4) Пересмотреть гипотезу или признаковое пространство; 5) При необходимости скорректировать план исследования; 6) Задokumentировать отклонения и выводы.	ПК-4
15	Дайте развернутый ответ Опишите, какой instrumentarium управления проектами вы будете использовать для координации работы распределённой команды при проведении исследований по масштабируемости ML-моделей на кластере (5 исследователей). Обоснуйте выбор. Ответ: Jira – для планирования и отслеживания задач; Slack/Teams – для коммуникации; Git + DVC – для управления версиями кода, данных и экспериментов; MLflow – для логирования и сравнения результатов; Confluence – для общей документации; обоснование: прозрачность, воспроизводимость, контроль версий, единое хранилище результатов, снижение рисков потери данных.	ПК-4

7.1. Уровни овладения

Компетенция: УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий.

Индикатор достижения компетенции: УК-1.2 Определяет, интерпретирует и ранжирует информацию, требуемую для решения проблемных ситуаций.

Уровень	Характеристика	Оценка в баллах
Повышенный	Достигнуто полное овладение знаниями, умениями и навыками. Студент свободно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	81-100

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

Компетенция: ПК-3 Способен контролировать фактическое исполнение проекта малого и среднего уровня сложности в области ИТ.

Индикатор достижения компетенции: ПК-3.1 Изучает и анализирует информацию, необходимую для разработки в области ИТ малого и среднего уровня сложности.

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

Компетенция: ПК-4 Способен управлять работами и проводить научные исследования и инструментарий в сфере ИТ.

Индикатор достижения компетенции: ПК-4.1 Изучает и анализирует современные научные методы исследований и инструментарий в области управления проектами.

Уровень	Характеристика	Оценка в баллах
Повышенный	Достигнуто полное овладение знаниями, умениями и навыками. Студент свободно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	81-100
Базовый	Достигнуто достаточное овладение знаниями, умениями и навыками. Студент уверенно владеет терминологией, умеет применять теоретические знания в различных ситуациях для решения поставленных задач.	61-80

Пороговый	Достигнуто овладение минимально необходимыми знаниями, умениями и навыками. Студент владеет основной терминологией, умеет применять теоретические знания для решения поставленных задач в стандартных ситуациях.	41-60
Ниже порогового	Компетенция не освоена	0-40

Индикатор достижения компетенции: ПК-4.2 Управляет и контролирует проведение научных исследований и использование инструментария в области управления проектами в области ИТ малого и среднего уровня сложности.

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

8. Материально-техническое и учебно-методическое обеспечение дисциплины

8.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. Королев, А. В. Экономико-математические методы и моделирование: учебник и практикум для вузов / А. В. Королев. - Москва: Юрайт, 2026. - 280 с - 978-5-534-00883-8. - Текст: электронный // ИКО Юрайт: [сайт]. - URL: <https://urait.ru/bcode/584034> (дата обращения: 21.05.2026). - Режим доступа: по подписке

2. Методы оптимальных решений (Экономико-математические методы и модели). Задачник: Учебно-практическое пособие / С.И. Макаров, С.А. Севастьянова, Р.И. Горбунова [и др.]; С.И. Макаров, С.А. Севастьянова, Р.И. Горбунова [и др.]; под. ред. С.И. Макаров, С.А. Севастьянова. - Москва: КноРус, 2020. - 202 с. - 978-5-406-07701-6. - Текст: электронный // book_ru: [сайт]. - URL: <https://book.ru/book/933559> (дата обращения: 21.05.2026). - Режим доступа: по подписке

Дополнительная литература

1. Макаров, С.И. Математика для экономистов: Учебное пособие / С.И. Макаров. - Москва: КноРус, 2020. - 263 с. - 978-5-406-07840-2. - Текст: электронный // book_ru: [сайт]. - URL: <https://book.ru/book/934068> (дата обращения: 21.05.2026). - Режим доступа: по подписке

2. Математика для экономистов (математический анализ и линейная алгебра). Задачник: Учебно-практическое пособие / С.И. Макаров, М.В. Мищенко, Л.И. Уфимцева [и др.]; С.И. Макаров, М.В. Мищенко, Л.И. Уфимцева [и др.]; под. ред. С.И. Макаров, М.В. Мищенко. - Москва: КноРус, 2018. - 358 с. - 978-5-406-06423-8. - Текст: электронный // book_ru: [сайт]. - URL: <https://book.ru/book/930056> (дата обращения: 21.05.2026). - Режим доступа: по подписке

3. Зенков, А. В. Методы оптимальных решений: учебник для вузов / А. В. Зенков. - Москва: Юрайт, 2026. - 201 с - 978-5-534-05377-7. - Текст: электронный // ИКО Юрайт: [сайт]. - URL: <https://urait.ru/bcode/586112> (дата обращения: 21.05.2026). - Режим доступа: по подписке

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

Профессиональные базы данных

1. <http://www.gks.ru/> - Профессиональная база данных «Официальная статистика» (Официальный сайт Федеральной службы государственной статистики)
2. <http://www.gov.ru/> - Профессиональная база данных «Информационные системы Министерства экономического развития Российской Федерации в сети Интернет» (Портал «Официальная Россия»)
3. <https://www.minfin.ru/ru/> - Профессиональная база данных «Финансово-экономические показатели Российской Федерации» (Официальный сайт Министерства финансов РФ)
4. <http://pravo.gov.ru/> - Государственная система правовой информации «Официальный интернет-портал правовой информации»
5. <https://rosinfra.ru/> - Платформа «Росинфра». «Национальный Центр развития государственно-частного партнерства»

Ресурсы «Интернет»

Не используются.

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

Перечень программного обеспечения

(обновление производится по мере появления новых версий программы)

1. Microsoft Windows XP;
2. Anaconda Distribution;
3. Mathcad 13 CLASSROOM;
4. Excel Compare;
5. Консультант Плюс;
6. Мой офис;

Перечень информационно-справочных систем

(обновление выполняется еженедельно)

1. Справочно-правовая система "Гарант-Максимум";

8.4. Специальные помещения, лаборатории и лабораторное оборудование

Учебные аудитории для проведения занятий лекционного типа	Комплекты ученической мебели Мультимедийный проектор Доска Экран
Учебные аудитории для проведения практических занятий (занятий семинарского типа)	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СИ

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