

Документы
Информация о владельце: "Самарский государственный экономический университет"
ФИО: Кандрашина Елена Александровна
Должность: И.о. ректора ФГАОУ ВО «Самарский государственный экономический университет»
Дата подписания: 20.07.2026 15:22:11
Уникальный программный ключ:
2db64eb9605ce27edd3b8e8fdd32c70e0674ddd2

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ) «АНАЛИЗ БОЛЬШИХ ДАННЫХ»

Уровень высшего образования: бакалавриат

Направление подготовки: 27.03.02 Управление качеством

Направленность (профиль) подготовки: Экономика и управление качеством

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

Год набора (приема на обучение): 2026

Срок получения образования: 4 года

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

г. Самара, 2026

Разработчики:

Кандидат экономических наук Ермолина Л. В.

Рабочая программа дисциплины (модуля) составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 27.03.02 Управление качеством, утвержденного приказом Минобрнауки от 31.07.2020 № 869, с учетом трудовых функций профессиональных стандартов: "Специалист по техническому контролю качества продукции", утвержден приказом Минтруда России от 15.07.2021 № 480н; "Специалист по качеству", утвержден приказом Минтруда России от 22.04.2021 № 276н.

Согласование и утверждение

№	Подразделение или коллегиальный орган	Ответственное лицо	ФИО	Виза	Дата, протокол (при наличии)
1	Кафедра экономики, организации и стратегии развития предприятия	Заведующий кафедрой, руководитель подразделения, реализующего ОП	Шепелев А. В.	Рассмотрено	15.05.2026, № 10

1. Цель и задачи освоения дисциплины (модуля)

Цель освоения дисциплины - Сформировать у обучающихся комплекс теоретических знаний и практических навыков в области анализа больших данных, необходимых для управления качеством продукции, процессов и систем на основе данных, а также для решения прикладных задач управления качеством в цифровой экономике с использованием современных инструментов интеллектуального анализа данных

Задачи изучения дисциплины:

- Формирование у студентов теоретической базы по основным парадигмам машинного обучения (обучение с учителем, без учителя, с подкреплением), методам предобработки и визуализации данных, а также критериям выбора алгоритмов в зависимости от структуры данных и бизнес-задачи.

;

- Формирование у студентов практических навыков реализации алгоритмов регрессии, классификации, кластеризации и снижения размерности с использованием библиотек Python (scikit-learn, TensorFlow, PyTorch) и работы с большими распределёнными наборами данных через Spark MLlib для построения прогнозных моделей.;

- Формирование у студентов способности интегрировать модели машинного обучения в корпоративные информационные системы, автоматизировать процесс анализа больших данных (ETL-пайплайны), оценивать качество моделей метриками и интерпретировать результаты для поддержки управленческих решений в области маркетинга, логистики, финансов и риск-менеджмента.

.

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

Компетенции, индикаторы и результаты обучения

ОПК-7 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности

ОПК-7.2 Использует современные информационные технологии для поиска и обмена информацией при решении задач в области управления качеством

Знать:

ОПК-7.2/Зн1 Основные современные информационные технологии и платформы для машинного обучения и анализа больших данных и их функциональные возможности для решения профессиональных задач.

Уметь:

ОПК-7.2/Ум1 Выбирать и применять соответствующие программные средства для сбора, предобработки, анализа больших данных и построения моделей машинного обучения, адаптируя их под конкретные бизнес-задачи и требования к информационно-аналитической поддержке

Владеть:

ОПК-7.2/Нв1 Навыками работы с современными ИТ-инструментами для полного цикла работы с данными, методами интеграции и трансформации данных, а также приёмами интерпретации результатов и их использования для подготовки обоснованных управленческих решений.

ПК-7 Способен проводить внутренний аудит системы управления качеством организации, анализировать его результаты, разрабатывать мероприятия по предотвращению выпуска некачественной продукции

ПК-7.1 Анализирует и выбирает оптимальные методы и методики решения конкретных задач по управлению качеством продукции (работ, услуг)

Знать:

ПК-7.1/Зн1 Сущность информации, принципы кодирования и представления данных в различных системах счисления (двоичной, восьмеричной, шестнадцатеричной), а также основы цифрового представления данных, используемых в машинном обучении и при работе с большими объёмами данных.

Уметь:

ПК-7.1/Ум1 Выполнять преобразования между системами счисления, интерпретировать представление числовых данных в контексте задач машинного обучения и применять соответствующие методы обработки для подготовки данных к анализу.

Владеть:

ПК-7.1/Нв1 Навыками работы с числовыми типами данных в языках программирования и инструментах анализа, методами преобразования и агрегации данных из различных источников, а также приёмами интерпретации результатов вычислений в двоичных и десятичных форматах для подготовки управленческих решений

3. Место дисциплины в структуре ОП

Дисциплина (модуль) «Анализ больших данных» относится к обязательной 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
ОПК-7 - Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности		
ОПК-7.2 Использует современные информационные технологии для поиска и обмена информацией при решении задач в области управления качеством	Маркетинг, Технологические основы производства	Инвестиционный анализ, Маркетинг, Основы проектной деятельности, Подготовка к процедуре защиты и защита выпускной квалификационной работы, Практикум по управлению производственным предприятием, Производственная практика: организационно-управленческая практика
ПК-7 - Способен проводить внутренний аудит системы управления качеством организации, анализировать его результаты, разрабатывать мероприятия по предотвращению выпуска некачественной продукции		

ПК-7.1 Анализирует и выбирает оптимальные методы и методики решения конкретных задач по управлению качеством продукции (работ, услуг)		Бережливое производство, Деловые коммуникации и документооборот, Подготовка к процедуре защиты и защита выпускной квалификационной работы, Производственная практика: организационно-управленческая практика, Производственная практика: практика по профилю профессиональной деятельности, Производственная практика: преддипломная практика, Системы менеджмента качества, Стандарты качества в управленческой деятельности, Управление затратами на обеспечение качества, Управление рисками
---	--	---

4. Объем дисциплины (модуля) и виды учебной работы

Период обучения	Общая трудоемкость (часы)	Общая трудоемкость (ЗЕТ)	Контактная работа (часы, всего)	Лекционные занятия (часы)	Практические занятия (часы)	Групповая контактная работа (часы)	Индивидуальная контактная работа (часы)	Самостоятельная работа (часы)	Промежуточная аттестация
Третий семестр	144	4	54	36	18	2	0,3	53,7	Экзамен
Всего	144	4	54	36	18	2	0,3	53,7	34

5. Содержание дисциплины (модуля)

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

Наименование раздела, темы	Всего	Лекционные занятия	Практические занятия	Самостоятельная работа
Раздел 1. Введение в машинное обучение и обработку данных	62	18	8	33,7
Тема 1.1. Введение в машинное обучение и обработку данных. Постановка основных классов задач в машинном обучении.	9	2	1	4
Тема 1.2. Регрессия и классификация; кластеризация, снижение размерности	7	2	1	4

Тема 1.3. Регрессия и классификация; кластеризация, снижение размерности	7	2	1	4
Тема 1.4. Задача регрессии. Минимизация квадрата отклонения. Регрессионная функция: условное мат.ожидание	7	2	1	4
Тема 1.5. Линейная регрессия и метод k ближайших соседей. Переобучение и недообучение	7,3	2	1	4
Тема 1.6. Разложение ошибки на шум, смещение и разброс	7	2	1	4
Тема 1.7. Детектирование выбросов и аномалий. Что такое выбросы, типы выбросов	7	2	1	4
Тема 1.8. Очистка данных и технологии регуляризации. Основные виды регуляризации	10,7	4	1	5,7
Раздел 2. Технологии кластеризации и классификации. Нейронные сети. Генетические алгоритмы	48	18	10	20
Тема 2.1. Технологии кластеризации и классификации. Kmeans. EM-алгоритм	7	2	1	4
Тема 2.2. Другие методы кластеризации. Задачи классификации. Байесовский классификатор	5	2	1	2
Тема 2.3. Линейные методы для классификации. Логистическая регрессия, максимизация правдоподобия	5	2	1	2
Тема 2.4. Нейронные сети: общая архитектура. Многослойные сети. Обратное распространение ошибки	5	2	1	2
Тема 2.5. Стохастический градиентный спуск. Генетические алгоритмы	5	2	1	2
Тема 2.6. Извлечение признаков / выделение особенностей (Feature detection)	5	2	1	2
Тема 2.7. Преобразования признаков. Нормализация данных. Методы нормализации данных	5	2	1	2
Тема 2.8. Нормализация по методу минимакса. Нормализация по Z-показателю. Десятичное масштабирование	5	2	1	2

Тема 2.9. Нечеткие множества. Байесовы сети. Задачи байесовского вывода. Методика построения нечеткой байесовой сети	6	2	2	2
--	---	---	---	---

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	Введение в машинное обучение и обработку данных	тестирование	Экзамен
2	Технологии кластеризации и классификации. Нейронные сети. Генетические алгоритмы	тестирование	Экзамен

6. Оценочные материалы текущего контроля

1. Введение в машинное обучение и обработку данных тестирование

№ п/п	Содержание вопроса		Компетенция
		Правильный ответ (ключ ответа)	
1	Выберите один правильный ответ Понимая принципы работы информационных технологий и используя методы сбора и анализа данных для поддержки управленческих решений, выберите метод машинного обучения, который используется для прогнозирования непрерывной целевой переменной: a) Классификация b) Регрессия c) Кластеризация d) Поиск ассоциативных правил Ответ: b		ПК-7, ОПК-7
2	Выберите один правильный ответ Для информационно-аналитической поддержки принятия решений в задачах анализа больших данных применяется метод уменьшения размерности. Какой из перечисленных методов является методом уменьшения размерности? a) PCA (Principal Component Analysis) b) Линейная регрессия c) Метод k-ближайших соседей d) Деревья решений Ответ: a		ПК-7, ОПК-7
3	Выберите один правильный ответ Используя программные средства обработки больших данных для сбора и анализа информации, выберите инструмент, который относится к распределённой обработке данных и использует модель MapReduce: a) Apache Spark b) Apache Hadoop c) Apache Kafka d) Redis Ответ: b		ПК-7, ОПК-7

4	Выберите один правильный ответ Принимая управленческие решения на основе анализа больших данных, важно понимать метрики оценки моделей машинного обучения. Для задачи классификации с несбалансированными классами более информативной метрикой является: a) Accuracy (точность) b) F1-score c) Mean Absolute Error d) R-squared Ответ: b	ПК-7, ОПК-7
5	Выберите один правильный ответ Для сбора и анализа больших данных в рамках информационно-аналитической поддержки управленческих решений используется распределённая файловая система. Какая из перечисленных систем является распределённой файловой системой в экосистеме Hadoop? a) HDFS b) HBase c) Hive d) ZooKeeper Ответ: a	ПК-7, ОПК-7
6	Прочитайте кейс и дайте развернутый ответ Ситуация: На предприятии по производству хлебобулочных изделий внедряется система сбора данных с каждого этапа технологической линии (температура печи, влажность, скорость конвейера, качество муки). Объем данных составляет 2 ТБ в смену. Руководство требует немедленно внедрить «нейросеть для предсказания брака», не понимая разницы между статистическим контролем (SPC) и ML. Вопрос: Дайте аргументированное заключение для руководства. Почему в данном случае для первичного анализа лучше подходят методы описательной статистики и визуализации, а не сложное машинное обучение? Сформулируйте критерии перехода от «Больших данных» к «Умному использованию данных» в системе менеджмента качества. Ответ: 1. Проблема качества данных: Большой объем (2 ТБ) ≠ информация. Данные могут быть зашумлены, иметь пропуски и содержать мультиколлинеарность (температура и влажность коррелируют). 2. Эффект «Гарбаджа»: ML-модель на неочищенных данных даст ложные выводы, что в управлении качеством критично (риски выпуска небезопасной продукции). 3. Приоритет SPC: Контрольные карты Шухарта позволяют отследить момент разладки процесса без обучения сложных моделей и требуют меньших вычислительных затрат. 4. Критерии перехода: Переход к ML оправдан, когда: а) тренды нелинейны; б) факторов > 20 и они сильно пересекаются; в) требуется прогноз на шаг вперед для компенсации инерционности процессов.	ПК-7, ОПК-7
7	Прочитайте ситуацию и дайте развернутый ответ Ситуация: На заводе СГЭУ внедряют цифровой двойник производства. Необходимо автоматизировать следующие процессы: 1. Прогноз остаточного ресурса оборудования (износ режущего инструмента). 2. Кластеризация поставщиков по уровню стабильности партий сырья. 3. Выявление аномалий в показателях датчиков (превышение вибрации). Вопрос: Соотнесите каждую из трех задач с конкретным типом машинного обучения (Регрессия, Кластеризация, Обнаружение аномалий). Объясните, на каком этапе цикла Деминга (Plan, Do, Check, Act) применяется каждая из моделей и как результат влияет на корректирующие действия. Ответ: 1. Прогноз ресурса — Регрессия (зависит от времени). Этап Plan (Планирование ТО). Результат: график замены инструмента. 2. Кластеризация поставщиков — Обучение без учителя (Кластеризация). Этап Act (Действие/Улучшение). Результат: пересмотр договоров с поставщиками из «рискованного» кластера. 3. Аномалии вибрации — Обнаружение выбросов (Outlier Detection). Этап Check (Проверка). Результат: мгновенная остановка линии (Act), если аномалия критична.	ПК-7, ОПК-7
8	Раскройте термин «Инженерия признаков» (Feature Engineering). В управлении качеством принят термин «критические параметры процесса». Как связаны эти два понятия? Приведите пример создания нового составного признака (например, «индекс стабильности работы экструдера») из сырых данных датчиков и поясните, почему качество признаков важнее, чем сложность самой модели ML. Ответ: 1. Определение: Преобразование сырых данных в признаки, которые лучше выявляют закономерности (масштабирование, кодирование категорий, создание отношений признаков). 2. Связь с УК: Критические параметры процесса (по ГОСТ Р ИСО 9001) — это и есть целевые признаки. 3. Пример: Есть сырые данные: «Давление на входе» и «Давление на выходе». Создаем новый признак «Перепад давления» (вход - выход). Именно перепад коррелирует с разрывом пленки, а не абсолютные значения. 4. Вывод: Хороший признак позволяет использовать простую линейную регрессию вместо «черного ящика» (нейросети), что критично для интерпретируемости результатов аудита. 5.	ПК-7, ОПК-7

9	Классическое и статистическое vs Машинное обучение: эволюция определений В чем принципиальное терминологическое различие между «Классическим статистическим моделированием» (например, регрессия по МНК) и «Машинным обучением» в контексте анализа больших данных? В своем развернутом ответе используйте понятия «генерализующая способность» и «вычислительная итеративность». Поясните, почему для малых выборок (до 100 наблюдений) классическая статистика надежнее, а для больших данных ML эффективнее, и как это влияет на выбор инструментария в магистерской работе по экономике. Ответ: 1. Классическая статистика: Дедуктивный подход. Сначала выдвигается гипотеза о распределении данных, затем оцениваются параметры. Цель — интерпретация коэффициентов. 2. Машинное обучение: Индуктивный подход. Алгоритм сам ищет функцию, итеративно минимизируя ошибку (градиентный спуск). Цель — точность предсказания на новых данных (генерализация). 3. Ключевой нюанс: Статистика неэффективна при 10 000 признаков (проклятие размерности) и «сырых» данных. ML эффективнее за счет регуляризации и валидации. 4. Вывод: Если у вас данные опроса 50 человек — используйте SPSS и классическую регрессию. Если у вас 50 000 чеков магазина — используйте ML. Смешивание терминов («нейросеть для 30 наблюдений») — это методологическая ошибка.	ПК-7, ОПК-7
10	"Переобучение" и "недообучение" как угрозы валидации процессов Дайте определения терминам «Переобучение» и «Недообучение» модели машинного обучения. Переведите эти понятия на язык аудитора системы качества: как бы вы назвали эти явления в терминах риск-менеджмента? Объясните, как признаки переобучения выявляются на тестовой выборке (принцип кросс-валидации) и почему внедрение переобученной модели в реальный процесс Quality Management крайне опасно. Ответ: 1. Недообучение: модель не уловила зависимость (высокая систематическая ошибка). В УК — «грубый контроль», пропуск брака. 2. Переобучение: модель запомнила шумы, а не закономерность. В УК — «ложная тревога», остановка станка из-за разового скачка напряжения, что ведет к потерям времени. 3. Кросс-валидация — разбивка данных на части для проверки. Если на учебных данных точность 99%, а на тестовых — 70% — это переобучение. 4. Опасность: при переобучении система качества теряет робастность; при малейшем изменении сырья модель перестает работать, и отдел качества теряет доверие к цифровым инструментам.	ПК-7, ОПК-7
11	Соответствие между типом задач ML и их применением в управлении качеством Установите соответствие между типом машинного обучения и конкретной задачей в системе менеджмента качества предприятия: Тип задачи ML 1. Регрессия (Предсказание числа) 2. Классификация 3. Кластеризация 4. Обнаружение аномалий Пример задачи в УК А. Сегментация складов по уровню дефектности хранения без эталонных меток Б. Выявление нестандартных всплесков энергопотребления станка (сигнал о поломке) В. Прогноз точного значения износа режущей фрезы (в часах до замены) Г. Распознавание на фото, является ли дефект сварного шва «трещиной» или «порой» Ответ: 1 — В 2 — Г 3 — А 4 — Б	ПК-7, ОПК-7
12	Соответствие характеристик больших данных этапам жизненного цикла контроля качества Установите соответствие между классической характеристикой Big Data (3V + V) и проблемой, возникающей на конкретном этапе управления качеством: Характеристика Big Data 1. Volume (Объем) 2. Velocity (Скорость) 3. Variety (Разнообразие) 4. Veracity (Достоверность) Описание проблемы в Управлении качеством А. Датчики температуры отправляют показания с разной частотой, а фото брака не оцифрованы (JSON + JPEG + PDF) Б. 30% данных с линии розлива содержат ошибки из-за сбоев калибровки, доверять им нельзя В. Накоплено 15 ТБ логов за 5 лет, но серверное хранилище предприятия переполнено Г. Системе нужно принять решение об отбраковке партии за 0,5 секунды, пока бутылка движется по конвейеру Ответ: 1 — В 2 — Г 3 — А 4 — Б	ПК-7, ОПК-7

13	Соответствие методов снижения размерности целям отдела качества. Установите соответствие между методом обработки данных и конечной управленческой целью в отделе качества Метод (Столбец А)Цель применения в УК (Столбец Б) 1. Отбор признаков (Feature Selection) А. Сжать 50 коррелирующих датчиков в 3 главных компоненты для построения визуального дашборда для директора 2. Извлечение признаков (PCA / t-SNE) Б. Исключить из модели признак «цвет упаковки», который никак не влияет на герметичность шва 3. Регуляризация (L1 / Lasso) В. В автоматическом режиме «занулить» веса незначимых параметров, чтобы модель не переобучалась на шумах 4. Агрегация данных (Средние/Медианы) Г. Сгладить высокочастотные колебания датчика вибрации, чтобы оценить общий тренд износа подшипника Установите соответствие между методом обработки данных и конечной управленческой целью в отделе качества: Метод 1. Отбор признаков 2. Извлечение признаков (PCA / t-SNE) 3. Регуляризация (L1 / Lasso) 4. Агрегация данных (Средние/Медианы) Цель применения в УК А. Сжать 50 коррелирующих датчиков в 3 главных компоненты для построения визуального дашборда для директора Б. Исключить из модели признак «цвет упаковки», который никак не влияет на герметичность шва В. В автоматическом режиме «занулить» веса незначимых параметров, чтобы модель не переобучалась на шумах Г. Сгладить высокочастотные колебания датчика вибрации, чтобы оценить общий тренд износа подшипника	ПК-7, ОПК-7
14	Этапы подготовки данных (ETL) перед построением модели контроля качества Расположите в правильном порядке этапы трансформации сырых данных с производственной линии перед их использованием в алгоритме машинного обучения для прогнозирования брака. Расположите в правильном порядке этапы трансформации сырых данных с производственной линии перед их использованием в алгоритме машинного обучения для прогнозирования брака. 1. Трансформация — приведение данных датчиков к единому масштабу. 2. Очистка — удаление выбросов и заполнение пропусков. 3. Сбор — получение сырых логов с PLC-контроллеров станка. 4. Извлечение признаков — создание составного признака «перепад температур». 5. Загрузка — сохранение подготовленного датасета в витрину данных для обучения.	ПК-7, ОПК-7
15	Жизненный цикл данных в системе менеджмента качества (от датчика до корректирующего действия) Установите правильную последовательность преобразования данных в управленческое решение в рамках цикла непрерывного улучшения качества (на основе PDCA). Установите правильную последовательность преобразования данных в управленческое решение в рамках цикла непрерывного улучшения качества (на основе PDCA). Варианты этапов (в произвольном порядке): 1. Визуализация на дашборде оператора (отображение тренда выхода параметра за границы поля допуска). 2. Сбор сырых данных с датчиков вибрации и температуры (Big Data). 3. Применение предиктивной модели (Machine Learning) для оценки остаточного ресурса узла. 4. Формирование отчета о необходимости внепланового технического обслуживания (Act). 5. Агрегация и фильтрация данных за последнюю рабочую смену.	ПК-7, ОПК-7

2. Технологии кластеризации и классификации. Нейронные сети. Генетические алгоритмы тестирования

№ п/п	Содержание вопроса		Компетенция
		Правильный ответ (ключ ответа)	

1	Установите соответствие Понимая принципы работы информационных технологий и используя методы анализа данных для поддержки управленческих решений, установите соответствие между типом задачи машинного обучения и примером применения: <table><tr><td>Тип задачи</td><td>Пример</td></tr><tr><td>1. Регрессия</td><td>А. Определение типа цветка по характеристикам</td></tr><tr><td>2. Классификация</td><td>Б. Группировка клиентов по поведению</td></tr><tr><td>3. Кластеризация</td><td>В. Прогноз цен на недвижимость</td></tr><tr><td>4. Снижение размерности</td><td>Г. Визуализация многомерных данных в 2D</td></tr></table> Ответ: 1 – В, 2 – А, 3 – Б, 4 – Г.	Тип задачи	Пример	1. Регрессия	А. Определение типа цветка по характеристикам	2. Классификация	Б. Группировка клиентов по поведению	3. Кластеризация	В. Прогноз цен на недвижимость	4. Снижение размерности	Г. Визуализация многомерных данных в 2D	ПК-7, ОПК-7
Тип задачи	Пример											
1. Регрессия	А. Определение типа цветка по характеристикам											
2. Классификация	Б. Группировка клиентов по поведению											
3. Кластеризация	В. Прогноз цен на недвижимость											
4. Снижение размерности	Г. Визуализация многомерных данных в 2D											
2	Установите соответствие Для информационно-аналитической поддержки принятия решений необходимо выбирать правильные метрики. Установите соответствие между метрикой и её характеристикой: <table><tr><td>Метрика</td><td>Характеристика</td></tr><tr><td>1. Точность (Precision)</td><td>А. Доля верно предсказанных объектов среди всех</td></tr><tr><td>2. Полнота (Recall)</td><td>Б. Доля истинно положительных среди предсказанных положительных</td></tr><tr><td>3. F-мера</td><td>В. Доля истинно положительных среди всех реальных положительных</td></tr><tr><td>4. Ассигасу</td><td>Г. Гармоническое среднее точности и полноты</td></tr></table> Ответ: 1 – Б, 2 – В, 3 – Г, 4 – А.	Метрика	Характеристика	1. Точность (Precision)	А. Доля верно предсказанных объектов среди всех	2. Полнота (Recall)	Б. Доля истинно положительных среди предсказанных положительных	3. F-мера	В. Доля истинно положительных среди всех реальных положительных	4. Ассигасу	Г. Гармоническое среднее точности и полноты	ПК-7, ОПК-7
Метрика	Характеристика											
1. Точность (Precision)	А. Доля верно предсказанных объектов среди всех											
2. Полнота (Recall)	Б. Доля истинно положительных среди предсказанных положительных											
3. F-мера	В. Доля истинно положительных среди всех реальных положительных											
4. Ассигасу	Г. Гармоническое среднее точности и полноты											
3	Установите соответствие Используя методы и программные средства сбора и обработки информации, установите соответствие между этапом обработки данных и его описанием: <table><tr><td>Этап</td><td>Описание</td></tr><tr><td>1. Сбор данных</td><td>А. Удаление дубликатов и заполнение пропусков</td></tr><tr><td>2. Предобработка</td><td>Б. Извлечение данных из различных источников</td></tr><tr><td>3. Выделение признаков</td><td>В. Преобразование данных в числовой формат</td></tr><tr><td>4. Масштабирование</td><td>Г. Нормализация данных для единого диапазона</td></tr></table> Ответ: 1 – Б, 2 – А, 3 – В, 4 – Г.	Этап	Описание	1. Сбор данных	А. Удаление дубликатов и заполнение пропусков	2. Предобработка	Б. Извлечение данных из различных источников	3. Выделение признаков	В. Преобразование данных в числовой формат	4. Масштабирование	Г. Нормализация данных для единого диапазона	ПК-7, ОПК-7
Этап	Описание											
1. Сбор данных	А. Удаление дубликатов и заполнение пропусков											
2. Предобработка	Б. Извлечение данных из различных источников											
3. Выделение признаков	В. Преобразование данных в числовой формат											
4. Масштабирование	Г. Нормализация данных для единого диапазона											
4	Установите последовательность В рамках информационно-аналитической поддержки управленческих решений выберите правильную последовательность шагов при построении модели машинного обучения (из предложенных вариантов): Варианты (буквы): А. 1→2→3→4→5 Б. 2→1→3→4→5 В. 1→3→2→4→5 Г. 1→2→4→3→5 Шаги: 1. Сбор и подготовка данных 2. Выбор модели 3. Обучение на обучающей выборке 4. Оценка на тестовой выборке 5. Настройка гиперпараметров и повторное обучение Ответ: А (1→2→3→4→5).	ПК-7, ОПК-7										

5	Установите последовательность Понимая принципы обработки информации для анализа, выберите правильную последовательность этапов предобработки данных перед обучением модели (из предложенных вариантов): Варианты (буквы): А. 1→2→3→4→5 Б. 2→1→3→4→5 В. 2→3→1→4→5 Г. 1→3→2→4→5 Шаги: - Удаление выбросов - Заполнение пропусков - Нормализация/стандартизация - Кодирование категориальных признаков - Разделение на обучающую и тестовую выборки Ответ: А (1→2→3→4→5)	ПК-7, ОПК-7
6	Дайте развернутый ответ Используя методы анализа данных для информационной поддержки решений, рассчитайте среднюю абсолютную ошибку (MAE) для модели регрессии, если истинные значения: [10, 15, 20], предсказанные: [12, 14, 18]. Ответ: $MAE = (10-12 + 15-14 + 20-18) / 3 = (2+1+2)/3 = 5/3 \approx 1,67$.	ПК-7, ОПК-7
7	Дайте развернутый ответ Оценивая качество модели классификации для принятия управленческих решений, определите точность (accuracy) по матрице ошибок: TP=80, TN=50, FP=10, FN=20 Ответ: $Accuracy = (TP+TN)/(TP+TN+FP+FN) = (80+50)/(80+50+10+20) = 130/160 = 0,8125$ (81,25%)	ПК-7, ОПК-7
8	Дайте развернутый ответ В процессе подготовки данных для машинного обучения примените нормализацию min-max для значения 7, если минимальное значение признака = 2, максимальное = 12. Ответ: Нормализованное значение = $(7-2)/(12-2) = 5/10 = 0,5$.	ПК-7, ОПК-7
9	Дайте развернутый ответ Анализируя сходство объектов с использованием методов обработки данных, рассчитайте евклидово расстояние между точками A(1, 2) и B(4, 6). Ответ: Расстояние = $\sqrt{((4-1)^2 + (6-2)^2)} = \sqrt{(3^2 + 4^2)} = \sqrt{(9+16)} = \sqrt{25} = 5$.	ПК-7, ОПК-7
10	Дайте развернутый ответ Для оценки вычислительной сложности модели в рамках информационно-аналитической поддержки, рассчитайте количество весов в полносвязном слое нейронной сети, если на вход подаётся 100 признаков, а слой содержит 50 нейронов (без смещений). Ответ: Количество весов = $100 \times 50 = 5000$.	ПК-7, ОПК-7
11	Выберите один из предложенных вариантов ответа Какой тип обучения нейронной сети используется, когда на вход подается только вектор данных, а требуемый выходной сигнал отсутствует, и сеть должна самостоятельно выявить скрытые закономерности и структуру в данных? А) Обучение с учителем Б) Обучение без учителя В) Обучение с подкреплением Г) Обучение с частичным привлечением учителя Ответ: Б) Обучение без учителя	ПК-7, , ОПК-7, ПК-7
12	Выберите один верный ответ из предложенных вариантов Генетические алгоритмы, применяемые в анализе данных, относятся к классу эвристических алгоритмов поиска. Какая идея лежит в их основе? А) Моделирование процессов работы мозга человека для распознавания образов Б) Формализация знаний экспертов в виде продукционных правил В) Моделирование естественных эволюционных процессов для решения оптимизационных задач Г) Использование нечеткой логики для работы с неопределенностью Ответ: В) Моделирование естественных эволюционных процессов для решения оптимизационных задач	ПК-7, ОПК-7
13	Выберите один верный ответ из предложенных вариантов В задаче классификации алгоритмы машинного обучения решают задачу ... А) Отнесения объектов к одной из заранее определенных категорий Б) Группировки объектов по схожим признакам без наличия меток классов В) Прогнозирования непрерывного числового значения на основе исторических данных Г) Снижения размерности данных для упрощения вычислений Ответ: А) Отнесения объектов к одной из заранее определенных категорий	ПК-7, ОПК-7

14	Выберите один верный ответ из предложенных вариантов Одним из ключевых свойств нейронных сетей, основанном на параллельной обработке информации, является способность ... А) выполнять вычисления строго последовательно по алгоритму Б) воспроизводить полный образ при предъявлении его искаженного или неполного фрагмента В) хранить информацию в виде четких правил «Если-То» Г) использовать адресную память для быстрого доступа к данным		ПК-7, ОПК-7
	Ответ:	Б) воспроизводить полный образ при предъявлении его искаженного или неполного фрагмента	
15	Выберите один верный ответ из предложенных вариантов Для предотвращения переобучения модели (overfitting) в нейронных сетях и других методах машинного обучения используется регуляризация. Ее основная цель — ... А) снизить сложность модели и повысить ее способность к обобщению на новых данных Б) увеличить количество скрытых слоев в нейронной сети В) повысить точность модели на обучающей выборке Г) ускорить процесс обучения за счет параллельных вычислений		ПК-7, ОПК-7
	Ответ:	А) снизить сложность модели и повысить ее способность к обобщению на новых данных	

7. Оценочные материалы промежуточной аттестации

Экзамен третий семестр

№ п/п	Содержание вопроса		Компетенция
	Правильный ответ (ключ ответа)		
1	Дайте развернутый ответ Понимая принципы работы информационных технологий и используя программный инструментарий для анализа данных, объясните, чем задача регрессии отличается от задачи классификации, и приведите пример бизнес-задачи для каждого типа.		ПК-7, ОПК-7
	Ответ:	Регрессия предсказывает непрерывную величину (например, прогноз продаж), а классификация – категориальную метку (например, определение, является ли клиент платёжеспособным); обе задачи решаются методами машинного обучения с использованием соответствующих библиотек.	
2	Дайте развернутый ответ Для информационно-аналитической поддержки принятия управленческих решений при прогнозировании продаж применяется линейная регрессия. Опишите, как интерпретируется коэффициент при независимой переменной в модели регрессии.		ПК-7, ОПК-7
	Ответ:	Коэффициент показывает, на сколько единиц изменится зависимая переменная (например, объём продаж) при изменении независимой переменной (например, цены) на одну единицу при прочих равных условиях.	
3	Дайте развернутый ответ Используя методы и программные средства сбора и обработки данных, опишите цель сжатия данных в машинном обучении, и назовите один из методов сжатия (например, PCA).		ПК-7, ОПК-7
	Ответ:	Сжатие данных позволяет уменьшить размерность признакового пространства, снижая вычислительные затраты и устраняя шум; одним из методов является главных компонент (PCA), который проецирует данные на оси максимальной дисперсии.	
4	Дайте развернутый ответ Применяя методы анализа для выявления нетипичных наблюдений, объясните, что такое выбросы, и как их наличие влияет на качество модели, и назовите один метод обнаружения выбросов.		ПК-7, ОПК-7
	Ответ:	Выбросы – это точки, сильно отклоняющиеся от основной совокупности, которые могут искажать оценки модели, снижая её точность; для их обнаружения часто используют метод межквартильного размаха (IQR) или алгоритмы, основанные на расстоянии (например, Z-оценка).	
5	Дайте развернутый ответ Для подготовки данных к обучению и предотвращения переобучения опишите, как регуляризация (например, L1 или L2) помогает улучшить обобщающую способность модели, и в чём разница между этими видами регуляризации.		ПК-7, ОПК-7
	Ответ:	Регуляризация добавляет штраф за большие веса в функцию потерь; L1-регуляризация приводит к разреженным решениям (обнуляет неважные признаки), а L2 – уменьшает веса, не обнуляя их.	
6	Дайте развернутый ответ Понимая принципы работы информационных технологий, сравните задачи кластеризации и классификации, указав, какая из них использует размеченные данные, а какая нет.		ПК-7, ОПК-7
	Ответ:	Классификация относится к обучению с учителем (требует размеченных меток), а кластеризация – к обучению без учителя (группирует объекты по схожести без эталонных ответов).	
7	Дайте развернутый ответ Для решения сложных задач прогнозирования и классификации применяются нейронные сети. Объясните назначение функции активации в нейроне и приведите пример одной из функций активации.		ПК-7, ОПК-7
	Ответ:	Функция активации вводит нелинейность в выход нейрона, позволяя сети аппроксимировать сложные зависимости; примеры: сигмоида, ReLU (Rectified Linear Unit) или гиперболический тангенс.	
8	Дайте развернутый ответ Используя методы оптимизации для поиска эффективных решений, опишите основной принцип работы генетического алгоритма, перечислив его ключевые этапы (селекция, кроссовер, мутация).		ПК-7, ОПК-7

	Ответ: Генетический алгоритм моделирует естественную эволюцию: отбирает лучшие особи (селекция), скрещивает их для создания потомства (кроссовер) и вносит случайные изменения (мутация), итерационно улучшая популяцию.	
9	Дайте развернутый ответ В рамках информационно-аналитической поддержки решений опишите, что такое выделение признаков и зачем необходимо нормализовать данные перед обучением моделей. Ответ: Выделение признаков – это процесс создания новых значимых признаков из исходных данных для улучшения качества модели; нормализация (масштабирование) приводит все признаки к единому диапазону, что необходимо для корректной работы алгоритмов, чувствительных к масштабу (например, градиентный спуск).	ПК-7, ОПК-7
10	Дайте развернутый ответ Понимая современные методы анализа неопределённости, объясните, как нечеткие множества или байесовы сети помогают в принятии решений при неполной информации, и приведите пример их применения. Ответ: Нечеткие множества позволяют моделировать нечёткие условия (например, «высокий риск»), а байесовы сети представляют вероятностные зависимости между переменными; они используются в системах поддержки принятия решений, например, для диагностики заболеваний или оценки кредитных рисков.	ПК-7, ОПК-7

7.1. Уровни овладения

Компетенция: ОПК-7 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности.

Индикатор достижения компетенции: ОПК-7.2 Использует современные информационные технологии для поиска и обмена информацией при решении задач в области управления качеством.

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-7 Способен проводить внутренний аудит системы управления качеством организации, анализировать его результаты, разрабатывать мероприятия по предотвращению выпуска некачественной продукции.

Индикатор достижения компетенции: ПК-7.1 Анализирует и выбирает оптимальные методы и методики решения конкретных задач по управлению качеством продукции (работ, услуг).

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. Чуваков, А. Б. Подготовка технологических процессов и программирование обработки деталей на станках с ЧПУ: учебник для вузов / А. Б. Чуваков. - Москва: Юрайт, 2026. - 152 с - 978-5-534-08647-8. - Текст: электронный // ИКО Юрайт: [сайт]. - URL: <https://urait.ru/bcode/600201> (дата обращения: 21.05.2026). - Режим доступа: по подписке
2. Платонов, А. В. Машинное обучение: учебное пособие для вузов / А. В. Платонов. - 2-е изд. - Москва: Юрайт, 2026. - 89 с - 978-5-534-20732-3. - Текст: электронный // ИКО Юрайт: [сайт]. - URL: <https://urait.ru/bcode/589132> (дата обращения: 21.05.2026). - Режим доступа: по подписке

Дополнительная литература

1. Зыков, С. В. Программирование. Функциональный подход: учебник и практикум для вузов / С. В. Зыков. - 2-е изд. - Москва: Юрайт, 2026. - 150 с - 978-5-534-16942-3. - Текст: электронный // ИКО Юрайт: [сайт]. - URL: <https://urait.ru/bcode/584399> (дата обращения: 21.05.2026). - Режим доступа: по подписке

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

Профессиональные базы данных

Не используются.

Ресурсы «Интернет»

1. <http://www.pravo.gov.ru/ips/> - Информационно-правовая система «Законодательство России»
2. <https://rosstat.gov.ru/> - Федеральная служба государственной статистики (Росстат)

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. "Astra Linux Special Edition" РУСБ.10015-01;
2. Консультант Плюс;

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