

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Заболотни Галина Ивановна

Должность: Директор филиала

Дата подписания: 28.02.2026 16:31:58

Уникальный программный ключ:

476db7d4acc6b30ef81301b72be235477473d63457266ce26b7e9e40f733b8b08

САМАРСКИЙ ПОЛИТЕХ  
Спортивный университет

**МИНОБРНАУКИ РОССИИ**

федеральное государственное бюджетное образовательное учреждение высшего образования

**«Самарский государственный технический университет»**

(ФГБОУ ВО «СамГТУ»)

УТВЕРЖДАЮ:

Директор филиала ФГБОУ ВО

«СамГТУ» в г. Новокуйбышевске

\_\_\_\_\_ / Г.И. Заболотни

« \_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ г.

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

### Б1.В.03 «Большие данные в прикладных информационных технологиях»

Код и направление подготовки (специальность)	09.04.02 Информационные системы и технологии
Направленность (профиль)	Прикладные информационные системы и технологии
Квалификация	Магистр
Форма обучения	Очная
Год начала подготовки	2026
Институт / факультет	Кафедры филиала ФГБОУ ВО «СамГТУ» в г. Новокуйбышевске
Выпускающая кафедра	Кафедра «Информатика и системы управления» (НФ-ИиСУ)
Кафедра-разработчик	Кафедра «Информатика и системы управления» (НФ-ИиСУ)
Объем дисциплины, ч. / з.е.	252 / 7
Форма контроля (промежуточная аттестация)	Экзамен

### **Б1.В.03 «Большие данные в прикладных информационных технологиях»**

Рабочая программа дисциплины разработана в соответствии с требованиями ФГОС ВО по направлению подготовки (специальности) **09.04.02 Информационные системы и технологии**, утвержденного приказом Министерства образования и науки РФ от № 917 от 19.09.2017 и соответствующего учебного плана.

Разработчик РПД:

Доцент,  
кандидат технических  
наук

\_\_\_\_\_  
(должность, степень, ученое звание)

А.Н. Лада

\_\_\_\_\_  
(ФИО)

\_\_\_\_\_  
Заведующий кафедрой

\_\_\_\_\_  
А.В. Волкодаева, кандидат  
экономических наук, доцент  
(ФИО, степень, ученое звание)

СОГЛАСОВАНО:

\_\_\_\_\_  
Заместитель директора

\_\_\_\_\_  
Е.Т. Демидова, кандидат  
юридических наук, доцент  
(ФИО, степень, ученое звание)

\_\_\_\_\_  
Руководитель образовательной  
программы

\_\_\_\_\_  
А.В. Волкодаева, кандидат  
экономических наук, доцент  
(ФИО, степень, ученое звание)

## Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы.....	4
2. Место дисциплины (модуля) в структуре образовательной программы.....	5
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся .....	6
4. Содержание дисциплины (модуля), структурированное по темам (разделам), с указанием отведенного на них количества академических часов и видов учебных занятий .....	6
4.1. Содержание лекционных занятий.....	6
4.2. Содержание лабораторных занятий.....	7
4.3. Содержание практических занятий.....	7
4.4. Содержание самостоятельной работы .....	8
5. Перечень учебной литературы и учебно-методического обеспечения по дисциплине (модулю).....	8
6. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения.....	9
7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», профессиональных баз данных, информационно-справочных систем.....	9
8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю) .....	9
9. Методические материалы .....	10
10. Фонд оценочных средств по дисциплине (модулю) .....	11

**1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы**

Наименование категории (группы) компетенций	Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения (знать, уметь, владеть, соотнесенные с индикаторами достижения компетенции)
Профессиональные компетенции			
	ПК-2. Способен разрабатывать, вводить в эксплуатацию и обслуживать базы данных; дополнять, модифицировать и совершенствовать базы данных и другие хранилища информации	ПК-2.1. Знать современные тенденции, технологии и регламенты интеграции БД на новые платформы и версии программного обеспечения (ПО).	<p>Знать современные модели данных, методологии проектирования и архитектуры систем хранения информации, а также комплекс методов обеспечения их целостности, безопасности, производительности и отказоустойчивости.</p> <p>Уметь анализировать требования предметной области и выбирать адекватную модель данных и технологию хранения на основе знания их архитектурных особенностей и ограничений.</p> <p>Владеть терминологией и систематизированными знаниями в области теории баз данных и систем хранения для анализа и сравнения различных технологических решений.</p>
		ПК-2.2. Проводить анализ системных проблем обработки информации на уровне БД, формировать предложения по перспективному развитию БД, осуществлять контроль обновлений БД.	<p>Знать последовательность этапов и практические методики жизненного цикла хранилища данных: от проектирования и разработки до развертывания, эксплуатации, оптимизации и модификации.</p> <p>Уметь проектировать, разрабатывать, внедрять и оптимизировать базы данных и сложные запросы в соответствии с поставленными задачами и требованиями.</p> <p>Владеть навыками практического применения методологий проектирования и инструментов для сквозной работы с хранилищами данных.</p>
		ПК-2.3. Владеть навыками внедрения в практику администрирования новых технологий с БД, осуществлять их обслуживание.	<p>Знать функциональные возможности и специфику конкретных инструментов и СУБД, необходимых для администрирования, мониторинга, миграции и обеспечения безопасности данных.</p> <p>Уметь применять инструментальные средства и языки для решения практических задач администрирования и управления данными.</p> <p>Владеть навыками администрирования СУБД, написания и оптимизации сложного SQL-кода, а также использования специализированных инструментов для проектирования, мониторинга и</p>

			защиты данных.
ПК-3 Способен выполнять разработку программного обеспечения общего и прикладного назначения, проводить оценку качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию	ПК-3.1 Знать современные тенденции развития информационных технологий		Знать современные методологии, архитектурные принципы и стандартные методы оценки качества на всех этапах жизненного цикла ПО.
			Уметь обосновывать выбор методологии разработки, архитектурных паттернов и инструментов контроля качества при планировании и реализации программных проектов.
			Владеть навыками в области инженерии ПО для критической оценки и выбора подходов к разработке, проектированию и обеспечению качества.
	ПК.3.2. Уметь решать задачи по разработке программного обеспечения общего и прикладного назначения, проведения оценки качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию.		Знать стандартные процессы, техники и инструменты, применяемые на каждом этапе работы с ПО: от анализа требований и проектирования до реализации, тестирования, отладки, интеграции и сопровождения.
			Уметь осуществлять полный цикл работ по созданию и поддержке программного обеспечения.
			Владеть методами и практиками сквозного процесса разработки ПО для преобразования требований в работоспособный, качественный и сопровождаемый программный продукт.
	ПК.3.3. Владеть навыками выбора и применения современных инструментальных средств для решения задач по разработке программного обеспечения общего и прикладного назначения, проведения оценки качества и работоспособности разработанного программного обеспечения, осуществления его сопровождения и модификации.		Знать синтаксис современных языков программирования, возможности промышленных фреймворков, а также принципы работы инструментов DevOps-стека.
			Уметь применять на практике языки программирования, фреймворки и инструменты коллективной разработки для создания, сборки, развертывания и поддержки программного кода.
			Владеть практическими навыками написания кода на современных языках с использованием фреймворков, а также уверенным использованием инструментов контроля.

## 2. Место дисциплины (модуля) в структуре образовательной программы

Место дисциплины (модуля) в структуре образовательной программы: **часть, формируемая участниками образовательных отношений**

Код компетенции	Предшествующие дисциплины	Параллельно осваиваемые дисциплины	Последующие дисциплины
ПК-2	Б1.В.01 Управление IT проектами		Б2.О.02(П) Производственная практика: технологическая (проектно-технологическая) практика

			Б2.В.01(Пд) Производственная практика: технологическая (проектно-технологическая) практика Б3.01 Выполнение и защита выпускной квалификационной работы.
ПК-3	Б1.В.01 Управление IT проектами	Б1.В.ДВ.01.01 Интеллектуальный анализ данных и машинное обучение Б1.В.ДВ.01.02 Интернет вещей	Б2.О.02(П) Производственная практика: технологическая (проектно-технологическая) практика Б2.В.01(Пд) Производственная практика: технологическая (проектно-технологическая) практика Б3.01 Выполнение и защита выпускной квалификационной работы.

**3. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся**

Вид учебной работы	Всего часов / часов в электронной форме	3 семестр часов / часов в электронной форме
<b>Аудиторная контактная работа (всего), в том числе:</b>	32	32
Лекции	8	8
Практические занятия	24	24
<b>Самостоятельная работа (всего), в том числе:</b>	184	184
Подготовка к практическим занятиям	184	184
<b>Контроль: экзамен</b>	36	36
<b>Итого: час</b>	<b>252</b>	<b>252</b>
<b>Итого: з.е.</b>	<b>7</b>	<b>7</b>

**4. Содержание дисциплины (модуля), структурированное по темам (разделам), с указанием отведенного на них количества академических часов и видов учебных занятий**

№ раздела	Наименование раздела дисциплины	Виды учебной нагрузки и их трудоемкость, часы				
		ЛЗ	ЛР	ПЗ	СРС	Всего часов
1	Основы данных, процесс и машинное обучение	4	0	8	72	84
2	Статистика, вероятность, гипотезы и визуализация	4	0	16	112	168
	<b>Итого</b>	<b>8</b>	<b>0</b>	<b>24</b>	<b>184</b>	<b>252</b>

**4.1. Содержание лекционных занятий**

№ занятия	Наименование раздела	Тема практического занятия	Содержание практического занятия (перечень дидактических единиц: рассматриваемых подтем, вопросов)	Количество часов / часов в электронной форме
<b>3 семестр</b>				
1	Основы данных, процесс и машинное обучение	Тема 1. Большие данные и экосистема больших данных	Ведение в анализ больших данных. Основные определения, термины, задачи анализа больших данных. Экосистема аналитики больших данных. Распределенные файловые системы.	2

2	Основы данных, процесс и машинное обучение	Тема 3. Машинное обучение и математические основы работы с данными	Машинное обучение на больших данных. Обзор источников информации для Big Data. Методики сбора данных.	2
3	Статистика, вероятность, гипотезы и визуализация	Тема 5. Теория вероятностей при работе с большими данными	Событие. Случайная величина. Центральная предельная теорема. Распределение вероятностей. Закон больших чисел. Применение при работе с большими данными.	2
4	Статистика, вероятность, гипотезы и визуализация	Тема 7. Визуализация больших данных	Создание дашбордов. Построение информационных панелей с использованием современных инструментов	2
<b>Итого за семестр:</b>				8
<b>Итого:</b>				8

#### 4.2. Содержание лабораторных занятий

Учебные занятия не реализуются.

#### 4.3. Содержание практических занятий

№ занятия	Наименование раздела	Тема лекции	Содержание лекции (перечень дидактических единиц: рассматриваемых подтем, вопросов)	Количество часов / часов в электронной форме
<b>3 семестр</b>				
1	Основы данных, процесс и машинное обучение	Тема 1. Большие данные и экосистема больших данных	Экосистема аналитики больших данных. Распределенные файловые системы.	2
2	Основы данных, процесс и машинное обучение	Тема 2. Процесс исследования данных	Методы и методики процесса исследования данных, управление проектами в сфере аналитики данных.	2
3	Основы данных, процесс и машинное обучение	Тема 3. Машинное обучение и математические основы работы с данными	Машинное обучение на больших данных.	2
4	Основы данных, процесс и машинное обучение	Тема 3. Машинное обучение и математические основы работы с данными	Обзор источников информации для Big Data. Методики сбора данных.	2
5	Статистика, вероятность, гипотезы и визуализация	Тема 4. Описательная статистика	Модели данных. Подготовка исходных данных для анализа: первичная обработка и визуализация имеющихся данных.	2
6	Статистика, вероятность, гипотезы и визуализация	Тема 4. Описательная статистика	Описательная статистика имеющихся наборов данных и инструменты для этого.	2
7	Статистика, вероятность, гипотезы и визуализация	Тема 5. Теория вероятностей при работе с большими данными	Событие. Случайная величина. Центральная предельная теорема. Распределение вероятностей.	2
8	Статистика, вероятность,	Тема 5. Теория вероятностей при	Закон больших чисел. Применение при работе с	2

	гипотезы и визуализация	работе с большими данными	большими данными.	
9	Статистика, вероятность, гипотезы и визуализация	Тема 6. Проверка гипотез при анализе данных	Статистические тесты. Дисперсионный анализ.	2
10	Статистика, вероятность, гипотезы и визуализация	Тема 6. Проверка гипотез при анализе данных	Критерий Хи-квадрат. Технические детали.	2
11	Статистика, вероятность, гипотезы и визуализация	Тема 7. Визуализация больших данных	Создание дашбордов.	2
12	Статистика, вероятность, гипотезы и визуализация	Тема 7. Визуализация больших данных	Построение информационных панелей с использованием современных инструментов.	2
<b>Итого за семестр:</b>				24
<b>Итого:</b>				24

#### 4.4. Содержание самостоятельной работы

Наименование раздела	Вид самостоятельной работы	Содержание самостоятельной работы (перечень дидактических единиц: рассматриваемых подтем, вопросов)	Количество часов
<b>3 семестр</b>			
Основы данных, процесс и машинное обучение	Подготовка к практическим занятиям	Ведение в анализ больших данных. Основные определения, термины, задачи анализа больших данных. Экосистема аналитики больших данных. Распределенные файловые системы. Методы и методики процесса исследования данных, управление проектами в сфере аналитики данных. Машинное обучение на больших данных. Обзор источников информации для Big Data. Методики сбора данных.	72
Статистика, вероятность, гипотезы и визуализация	Подготовка к практическим занятиям	Модели данных. Подготовка исходных данных для анализа: первичная обработка и визуализация имеющихся данных. Описательная статистика имеющихся наборов данных и инструменты для этого. Событие. Случайная величина. Центральная предельная теорема. Распределение вероятностей. Закон больших чисел. Применение при работе с большими данными. Статистические тесты. Дисперсионный анализ. Критерий Хи-квадрат. Технические детали. Создание дашбордов. Построение информационных панелей с использованием современных инструментов.	116
<b>Итого за семестр:</b>			<b>184</b>
<b>Итого:</b>			<b>184</b>

#### 5. Перечень учебной литературы и учебно-методического обеспечения по дисциплине (модулю)

№ п/п	Библиографическое описание	Ресурс НТБ СамГТУ (ЭБС СамГТУ, IPRbooks и т.д.)
<b>Основная литература</b>		
1	Машинное обучение: Новый искусственный интеллект : пер. с англ. / Э. Алпайдин.- М., Точка, 2017Альпина Паблишер.- 191 с. - Режим доступа: <a href="https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 119813">https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 119813</a>	Электронный ресурс
2	<i>Data Mining : Учеб.пособие / И. А. Чубукова .- 2-е изд., испр..- М., Интернет-Ун-т Информ.Технологий, 2008М., БИНОМ.Лаб.знаний.- 382 с. - 303 с. - Режим доступа: <a href="https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 103322">https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 103322</a></i>	Электронный ресурс

3	Проектирование реляционных баз данных : учеб.пособие по курсовому проектированию / В. Н. Якимов; Самар.гос.техн.ун-т.- Самара, 2013.- 95 с. - Режим доступа: <a href="https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 112609">https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 112609</a>	Электронный ресурс
<b>Дополнительная литература</b>		
4	Общая и прикладная статистика : учеб. / П. Ф. Аскеров, Р. Н. Пахунова, А. В. Пахунов.- М., Инфра-М, 2014.- 271 с. - Режим доступа: <a href="https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 115654">https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 115654</a>	Электронный ресурс
5	Теория вероятностей : учеб. / А. И. Луценко.- Ростов н/Д, Феникс, 2009.- 251 с. - Режим доступа: <a href="https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 106132">https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 106132</a>	Электронный ресурс
6	Бизнес-аналитика : от данных к знаниям : учеб.пособие / Н. Б. Паклин, В. И. Орешков .- 2-е изд., испр. .- СПб., Питер, 2013.- 701 с. - Режим доступа: <a href="https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 111773">https://elib.samgtu.ru/getinfo?uid=els_samgtu catalog 111773</a>	Электронный ресурс

Доступ обучающихся к ЭР НТБ СамГТУ ([elib.samgtu.ru](http://elib.samgtu.ru)) осуществляется посредством электронной информационной образовательной среды университета и сайта НТБ СамГТУ по логину и паролю.

#### **6. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения**

При проведении лекционных занятий используется мультимедийное оборудование. Организовано взаимодействие обучающегося и преподавателя с использованием электронной информационной образовательной среды университета.

№ п/п	Наименование	Производитель	Способ распространения
1	Microsoft Windows 8.1 Professional операционная система	Microsoft (Зарубежный)	Лицензионное
2	Microsoft Office 2013	Microsoft (Зарубежный)	Лицензионное
3	Антивирус Kaspersky EndPoint Security	«Лаборатории Касперского» (Отечественный)	Лицензионное
4	Программное обеспечение «Антиплагиат.Эксперт»	АО «Антиплагиат» (Отечественный)	Лицензионное

#### **7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», профессиональных баз данных, информационно-справочных систем**

№ п/п	Наименование	Краткое описание	Режим доступа
1	Административно-управленческий портал	<a href="http://www.aup.ru/marketing">www.aup.ru/marketing</a>	Ресурсы открытого доступа
2	Консультант плюс	<a href="http://www.consultant.ru">http://www.consultant.ru</a>	Ресурсы открытого доступа
3	Электронная библиотека изданий СамГТУ	<a href="http://irbis.samgtu.local/cgi-bin/irbis64r_01/cgiirbis_64.exe">http://irbis.samgtu.local/cgi-bin/irbis64r_01/cgiirbis_64.exe</a>	Российские базы данных ограниченного доступа
4	Электронно-библиотечная система IPRbooks	<a href="http://www.iprbookshop.ru">http://www.iprbookshop.ru</a>	Российские базы данных ограниченного доступа
5	eLIBRARY.ru	<a href="http://www.eLIBRARY.ru">http://www.eLIBRARY.ru</a>	Российские базы данных ограниченного доступа

#### **8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)**

##### **Лекционные занятия**

Учебная аудитория для проведения занятий лекционного типа, текущего контроля и промежуточной аттестации.

Оборудование: набор демонстрационного оборудования (экран, проектор, переносной ноутбук), специализированная мебель.

### **Практические занятия**

Учебная аудитория для проведения практических и семинарских занятий, текущего контроля и промежуточной аттестации. Помещение оснащено специализированной мебелью, оборудованием и техническими средствами обучения.

### **Самостоятельная работа**

Аудитория для самостоятельной работы. Помещение оснащено специализированной мебелью, оборудованием и техническими средствами обучения, компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.

## **9. Методические материалы**

### **Методические рекомендации при работе на лекции**

До лекции студент должен просмотреть учебно-методическую и научную литературу по теме лекции с тем, чтобы иметь представление о проблемах, которые будут разбираться в лекции.

Перед началом лекции обучающимся сообщается тема лекции, план, вопросы, подлежащие рассмотрению, доводятся основные литературные источники. Весь учебный материал, сообщаемый преподавателем, должен не просто прослушиваться. Он должен быть активно воспринят, т.е. услышан, осмыслен, понят, зафиксирован на бумаге и закреплён в памяти. Приступая к слушанию нового учебного материала, полезно мысленно установить его связь с ранее изученным. Следя за техникой чтения лекции (акцент на существенном, повышение тона, изменение ритма, пауза и т.п.), необходимо вслед за преподавателем уметь выделять основные категории, законы и определять их содержание, проблемы, предполагать их возможные решения, доказательства и выводы. Осуществляя такую работу, можно значительно облегчить себе понимание учебного материала, его конспектирование и дальнейшее изучение.

Конспектирование лекции позволяет обработать, систематизировать и лучше сохранить полученную информацию с тем, чтобы в будущем можно было восстановить в памяти основные, содержательные моменты. Типичная ошибка, совершаемая обучающимся, дословное конспектирование речи преподавателя. Как правило, при записи «слово в слово» не остается времени на обдумывание, анализ и синтез информации. Отбирая нужную информацию, главные мысли, проблемы, решения и выводы, необходимо сокращать текст, строить его таким образом, чтобы потом можно было легко в нем разобраться. Желательно оставить в рабочих конспектах поля, на которых можно будет делать пометки из рекомендованной литературы, дополняющие материал прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений. С окончанием лекции работа над конспектом не может считаться завершённой. Нужно еще восстановить отдельные места, проверить, все ли понятно, уточнить что-то на консультации и т.п. с тем, чтобы конспект мог быть использован в процессе подготовки к практическим занятиям, зачету, экзамену. Конспект лекции – незаменимый учебный документ, необходимый для самостоятельной работы.

### **Методические рекомендации при подготовке и работе на практическом занятии**

Практические занятия по дисциплине проводятся в целях выработки практических умений и приобретения навыков в решении профессиональных задач.

Рекомендуется следующая схема подготовки к практическому занятию:

- 1) ознакомление с планом практического занятия, который отражает содержание предложенной темы;
- 2) проработка конспекта лекции;
- 3) чтение рекомендованной литературы;
- 4) подготовка ответов на вопросы плана практического занятия;
- 5) выполнение тестовых заданий, задач и др.

Подготовка обучающегося к практическому занятию производится по вопросам, разработанным для каждой темы практических занятий и (или) лекций. В процессе подготовки к практическим занятиям, необходимо обратить особое внимание на самостоятельное изучение рекомендованной литературы.

Работа студентов во время практического занятия осуществляется на основе заданий, которые выдаются обучающимся в начале или во время занятия. На практических занятиях приветствуется активное участие в обсуждении конкретных ситуаций, способность на основе полученных знаний находить наиболее эффективные решения поставленных проблем, уметь находить полезный дополнительный материал по тематике занятий. Обучающимся необходимо обращать внимание на основные понятия, алгоритмы, определять практическую значимость рассматриваемых вопросов. На практических занятиях обучающиеся должны уметь выполнить расчет по заданным параметрам или выработать определенные решения по обозначенной проблеме. Задания могут быть групповые и индивидуальные. В зависимости от сложности предлагаемых заданий, целей занятия, общей подготовки обучающихся преподаватель может подсказать обучающимся алгоритм решения или первое действие, или указать общее направление рассуждений. Полученные результаты обсуждаются с позиций их адекватности или эффективности в рассмотренной ситуации.

### **Методические рекомендации по выполнению самостоятельной работы**

Организация самостоятельной работы обучающихся ориентируется на активные методы овладения знаниями, развитие творческих способностей, переход от поточного к индивидуализированному обучению с учетом потребностей и возможностей обучающегося.

Самостоятельная работа с учебниками, учебными пособиями, научной, справочной литературой, материалами периодических изданий и Интернета является наиболее эффективным методом получения дополнительных знаний, позволяет значительно активизировать процесс овладения информацией, способствует более глубокому усвоению изучаемого материала. Все новые понятия по изучаемой теме необходимо выучить наизусть и внести в глоссарий, который целесообразно вести с самого начала изучения курса.

Самостоятельная работа реализуется:

- непосредственно в процессе аудиторных занятий;
- на лекциях, практических занятиях;
- в контакте с преподавателем вне рамок расписания;
- на консультациях по учебным вопросам, в ходе творческих контактов, при ликвидации задолженностей, при выполнении индивидуальных заданий и т.д.;
- в библиотеке, дома, на кафедре при выполнении обучающимся учебных и практических задач.

Эффективным средством осуществления обучающимся самостоятельной работы является электронная информационно-образовательная среда университета, которая обеспечивает доступ к учебным планам, рабочим программам дисциплин (модулей), практик, к изданиям электронных библиотечных систем.

### **10. Фонд оценочных средств по дисциплине (модулю)**

Фонд оценочных средств представлен в приложении № 1.

**Фонд оценочных средств  
по дисциплине**

**Б1.В.03 «Большие данные в прикладных информационных технологиях»**

<b>Код и направление подготовки (специальность)</b>	<u>09.04.02 Информационные системы и технологии</u>
<b>Направленность (профиль)</b>	<u>Прикладные информационные системы и технологии</u>
<b>Квалификация</b>	<u>Магистр</u>
<b>Форма обучения</b>	<u>Очная</u>
<b>Год начала подготовки</b>	<u>2026</u>
<b>Институт / факультет</b>	<u>Кафедры филиала ФГБОУ ВО «СамГТУ» в г. Новокуйбышевске</u>
<b>Выпускающая кафедра</b>	<u>Кафедра «Информатика и системы управления» (НФ-ИиСУ)</u>
<b>Кафедра-разработчик</b>	<u>Кафедра «Информатика и системы управления» (НФ-ИиСУ)</u>
<b>Объем дисциплины, ч. / з.е.</b>	<u>252 / 7</u>
<b>Форма контроля (промежуточная аттестация)</b>	<u>Экзамен</u>

**Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы**

Наименование категории (группы) компетенций	Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения (знать, уметь, владеть, соотнесенные с индикаторами достижения компетенции)
Профессиональные компетенции			
	ПК-2. Способен разрабатывать, вводить в эксплуатацию и обслуживать базы данных; дополнять, модифицировать и совершенствовать базы данных и другие хранилища информации	ПК-2.1. Знать современные тенденции, технологии и регламенты интеграции БД на новые платформы и версии программного обеспечения (ПО).	<p>Знать современные модели данных, методологии проектирования и архитектуры систем хранения информации, а также комплекс методов обеспечения их целостности, безопасности, производительности и отказоустойчивости.</p> <p>Уметь анализировать требования предметной области и выбирать адекватную модель данных и технологию хранения на основе знания их архитектурных особенностей и ограничений.</p> <p>Владеть терминологией и систематизированными знаниями в области теории баз данных и систем хранения для анализа и сравнения различных технологических решений.</p>
		ПК-2.2. Проводить анализ системных проблем обработки информации на уровне БД, формировать предложения по перспективному развитию БД, осуществлять контроль обновлений БД.	<p>Знать последовательность этапов и практические методики жизненного цикла хранилища данных: от проектирования и разработки до развертывания, эксплуатации, оптимизации и модификации.</p> <p>Уметь проектировать, разрабатывать, внедрять и оптимизировать базы данных и сложные запросы в соответствии с поставленными задачами и требованиями.</p> <p>Владеть навыками практического применения методологий проектирования и инструментов для сквозной работы с хранилищами данных.</p>
		ПК-2.3. Владеть навыками внедрения в практику администрирования новых технологий с БД, осуществлять их обслуживание.	<p>Знать функциональные возможности и специфику конкретных инструментов и СУБД, необходимых для администрирования, мониторинга, миграции и обеспечения безопасности данных.</p> <p>Уметь применять инструментальные средства и языки для решения практических задач администрирования и управления данными.</p> <p>Владеть навыками администрирования СУБД, написания и оптимизации сложного SQL-кода, а также использования специализированных инструментов для проектирования, мониторинга и</p>

			защиты данных.
ПК-3 Способен выполнять разработку программного обеспечения общего и прикладного назначения, проводить оценку качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию	ПК-3.1 Знать современные тенденции развития информационных технологий		Знать современные методологии, архитектурные принципы и стандартные методы оценки качества на всех этапах жизненного цикла ПО.
			Уметь обосновывать выбор методологии разработки, архитектурных паттернов и инструментов контроля качества при планировании и реализации программных проектов.
			Владеть навыками в области инженерии ПО для критической оценки и выбора подходов к разработке, проектированию и обеспечению качества.
	ПК.3.2. Уметь решать задачи по разработке программного обеспечения общего и прикладного назначения, проведения оценки качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию.		Знать стандартные процессы, техники и инструменты, применяемые на каждом этапе работы с ПО: от анализа требований и проектирования до реализации, тестирования, отладки, интеграции и сопровождения.
			Уметь осуществлять полный цикл работ по созданию и поддержке программного обеспечения.
			Владеть методами и практиками сквозного процесса разработки ПО для преобразования требований в работоспособный, качественный и сопровождаемый программный продукт.
	ПК.3.3. Владеть навыками выбора и применения современных инструментальных средств для решения задач по разработке программного обеспечения общего и прикладного назначения, проведения оценки качества и работоспособности разработанного программного обеспечения, осуществления его сопровождения и модификации.		Знать синтаксис современных языков программирования, возможности промышленных фреймворков, а также принципы работы инструментов DevOps-стека.
			Уметь применять на практике языки программирования, фреймворки и инструменты коллективной разработки для создания, сборки, развертывания и поддержки программного кода.
			Владеть практическими навыками написания кода на современных языках с использованием фреймворков, а также уверенным использованием инструментов контроля.

### Матрица соответствия оценочных средств запланированным результатам обучения

Код индикатора достижения компетенции	Результаты обучения	Оценочные средства	Текущий контроль успеваемости	Промежуточная аттестация
<b>Основы данных, процесс и машинное обучение</b>				
ПК-2.1. Знать современные тенденции, технологии и регламенты интеграции БД на новые платформы и версии	Знать современные модели данных, методологии проектирования и архитектуры систем хранения информации, а также комплекс методов обеспечения их	Тестовые задания	Да	Нет
		Зачет	Нет	Да

программного обеспечения (ПО).	целостности, безопасности, производительности и отказоустойчивости.			
	Уметь анализировать требования предметной области и выбирать адекватную модель данных и технологию хранения на основе знания их архитектурных особенностей и ограничений.	Практические задания	Да	Нет
		Зачет	Нет	Да
	Владеть терминологией и систематизированными знаниями в области теории баз данных и систем хранения для анализа и сравнения различных технологических решений.	Практические задания	Да	Нет
		Зачет	Нет	Да
	ПК-2.2. Проводить анализ системных проблем обработки информации на уровне БД, формировать предложения по перспективному развитию БД, осуществлять контроль обновлений БД.	Знать последовательность этапов и практические методики жизненного цикла хранилища данных: от проектирования и разработки до развертывания, эксплуатации, оптимизации и модификации.	Тестовые задания	Да
Зачет			Нет	Да
Уметь проектировать, разрабатывать, внедрять и оптимизировать базы данных и сложные запросы в соответствии с поставленными задачами и требованиями.		Практические задания	Да	Нет
		Зачет	Нет	Да
Владеть навыками практического применения методологий проектирования и инструментов для сквозной работы с хранилищами данных.		Практические задания	Да	Нет
		Зачет	Нет	Да
ПК-2.3. Владеть навыками внедрения в практику администрирования новых технологий с БД, осуществлять их обслуживание.	Знать функциональные возможности и специфику конкретных инструментов и СУБД, необходимых для администрирования, мониторинга, миграции и обеспечения безопасности данных.	Тестовые задания	Да	Нет
		Зачет	Нет	Да
	Уметь применять инструментальные средства и языки для решения практических задач администрирования и управления данными.	Практические задания	Да	Нет
		Зачет	Нет	Да
	Владеть навыками администрирования СУБД, написания и оптимизации сложного SQL-кода, а также использования специализированных инструментов для проектирования, мониторинга и защиты данных.	Практические задания	Да	Нет
		Зачет	Нет	Да

ПК-3.1 Знать современные тенденции развития информационные технологий	Знать современные методологии, архитектурные принципы и стандартные методы оценки качества на всех этапах жизненного цикла ПО.	Тестовые задания	Да	Нет
		Зачет	Нет	Да
	Уметь обосновывать выбор методологии разработки, архитектурных паттернов и инструментов контроля качества при планировании и реализации программных проектов.	Практические задания	Да	Нет
		Зачет	Нет	Да
	Владеть навыками в области инженерии ПО для критической оценки и выбора подходов к разработке, проектированию и обеспечению качества.	Практические задания	Да	Нет
		Зачет	Нет	Да
ПК.3.2. Уметь решать задачи по разработке программного обеспечения общего и прикладного назначения, проведения оценки качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию.	Знать стандартные процессы, техники и инструменты, применяемые на каждом этапе работы с ПО: от анализа требований и проектирования до реализации, тестирования, отладки, интеграции и сопровождения.	Тестовые задания	Да	Нет
		Зачет	Нет	Да
	Уметь осуществлять полный цикл работ по созданию и поддержке программного обеспечения.	Практические задания	Да	Нет
		Зачет	Нет	Да
	Владеть методами и практиками сквозного процесса разработки ПО для преобразования требований в работоспособный, качественный и сопровождаемый программный продукт.	Практические задания	Да	Нет
		Зачет	Нет	Да
ПК.3.3. Владеть навыками выбора и применения современных инструментальных средств для решения задач по разработке программного обеспечения общего и прикладного назначения, проведения оценки качества и работоспособности разработанного программного обеспечения, осуществления его сопровождения и модификации.	Знать синтаксис современных языков программирования, возможности промышленных фреймворков, а также принципы работы инструментов DevOps-стека.	Тестовые задания	Да	Нет
		Зачет	Нет	Да
	Уметь применять на практике языки программирования, фреймворки и инструменты коллективной разработки для создания, сборки, развертывания и поддержки программного кода.	Практические задания	Да	Нет
		Зачет	Нет	Да
	Владеть практическими навыками написания кода на современных языках с использованием фреймворков, а также	Практические задания	Да	Нет
		Зачет	Нет	Да

	уверенным использованием инструментов контроля.				
<b>Статистика, вероятность, гипотезы и визуализация</b>					
ПК-2.1. Знать современные тенденции, технологии и регламенты интеграции БД на новые платформы и версии программного обеспечения (ПО).	Знать современные модели данных, методологии проектирования и архитектуры систем хранения информации, а также комплекс методов обеспечения их целостности, безопасности, производительности и отказоустойчивости.	Тестовые задания	Да	Нет	
		Экзамен	Нет	Да	
	Уметь анализировать требования предметной области и выбирать адекватную модель данных и технологию хранения на основе знания их архитектурных особенностей и ограничений.	Практические задания	Да	Нет	
		Экзамен	Нет	Да	
	Владеть терминологией и систематизированными знаниями в области теории баз данных и систем хранения для анализа и сравнения различных технологических решений.	Практические задания	Да	Нет	
		Экзамен	Нет	Да	
	ПК-2.2. Проводить анализ системных проблем обработки информации на уровне БД, формировать предложения по перспективному развитию БД, осуществлять контроль обновлений БД.	Знать последовательность этапов и практические методики жизненного цикла хранилища данных: от проектирования и разработки до развертывания, эксплуатации, оптимизации и модификации.	Тестовые задания	Да	Нет
			Экзамен	Нет	Да
Уметь проектировать, разрабатывать, внедрять и оптимизировать базы данных и сложные запросы в соответствии с поставленными задачами и требованиями.		Практические задания	Да	Нет	
		Экзамен	Нет	Да	
Владеть навыками практического применения методологий проектирования и инструментов для сквозной работы с хранилищами данных.		Практические задания	Да	Нет	
		Экзамен	Нет	Да	
ПК-2.3. Владеть навыками внедрения в практику администрирования новых технологий с БД, осуществлять их обслуживание.		Знать функциональные возможности и специфику конкретных инструментов и СУБД, необходимых для администрирования, мониторинга, миграции и обеспечения безопасности данных.	Тестовые задания	Да	Нет
			Экзамен	Нет	Да
	Уметь применять инструментальные средства и языки для решения практических	Практические задания	Да	Нет	
		Экзамен	Нет	Да	

	задач администрирования и управления данными.				
	Владеть навыками администрирования СУБД, написания и оптимизации сложного SQL-кода, а также использования специализированных инструментов для проектирования, мониторинга и защиты данных.	Практические задания	Да	Нет	
		Экзамен	Нет	Да	
ПК-3.1 Знать современные тенденции развития информационных технологий	Знать современные методологии, архитектурные принципы и стандартные методы оценки качества на всех этапах жизненного цикла ПО.	Тестовые задания	Да	Нет	
		Экзамен	Нет	Да	
	Уметь обосновывать выбор методологии разработки, архитектурных паттернов и инструментов контроля качества при планировании и реализации программных проектов.	Практические задания	Да	Нет	
		Экзамен	Нет	Да	
	Владеть навыками в области инженерии ПО для критической оценки и выбора подходов к разработке, проектированию и обеспечению качества.	Практические задания	Да	Нет	
		Экзамен	Нет	Да	
	ПК-3.2. Уметь решать задачи по разработке программного обеспечения общего и прикладного назначения, проведения оценки качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию.	Знать стандартные процессы, техники и инструменты, применяемые на каждом этапе работы с ПО: от анализа требований и проектирования до реализации, тестирования, отладки, интеграции и сопровождения.	Тестовые задания	Да	Нет
			Экзамен	Нет	Да
Уметь осуществлять полный цикл работ по созданию и поддержке программного обеспечения.		Практические задания	Да	Нет	
		Экзамен	Нет	Да	
Владеть методами и практиками сквозного процесса разработки ПО для преобразования требований в работоспособный, качественный и сопровождаемый программный продукт.		Практические задания	Да	Нет	
		Экзамен	Нет	Да	
ПК-3.3. Владеть навыками выбора и применения современных инструментальных средств для решения задач по разработке программного обеспечения общего и прикладного назначения,	Знать синтаксис современных языков программирования, возможности промышленных фреймворков, а также принципы работы инструментов DevOps-стека.	Тестовые задания	Да	Нет	
		Экзамен	Нет	Да	
	Уметь применять на практике языки	Практические задания	Да	Нет	

проведения оценки качества и работоспособности разработанного программного обеспечения, осуществления его сопровождения и модификации.	программирования, фреймворки и инструменты коллективной разработки для создания, сборки, развертывания и поддержки программного кода.	Экзамен	Нет	Да
	Владеть практическими навыками написания кода на современных языках с использованием фреймворков, а также уверенным использованием инструментов контроля.	Практические задания	Да	Нет
		Экзамен	Нет	Да

**Типовые задания для промежуточной аттестации по дисциплине  
Б1.В.03 «Большие данные в прикладных информационных технологиях»**

(шифр и наименование дисциплины)

**для направления подготовки 09.04.02 Информационные системы и технологии**  
(шифр и наименование направления подготовки, специальности)

**2026 ГОД ПРИЕМА**

(год приема на образовательную программу)

**Контролируемая (ые) компетенция(и):**

**ПК-2** Способен разрабатывать, вводить в эксплуатацию и обслуживать базы данных; дополнять, модифицировать и совершенствовать базы данных и другие хранилища информации

**ПК-3** Способен выполнять разработку программного обеспечения общего и прикладного назначения, проводить оценку качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию  
(шифр и наименование компетенции(й))

**Спецификация тестовых заданий**

Содержание дисциплины (разделы / темы)	Число заданий									
	закрытые			открытые				комбинированные		всего
	однозначный выбор варианта ответа	многозначный выбор варианта ответа	задание на сопоставление	задание на установление правильной	задания на дополнение	задания с развернутым ответом	практико-ориентированные задания	Задания с выбором одного ответа и обоснованием выбора ответа	Задания с выбором нескольких ответов и обоснованием выбора ответов	
<b>Раздел 1. Основы данных, процесс и машинное обучение</b>	8	6	6	6	10	6	6	6	6	60
Тема 1. Большие данные и экосистема больших данных	2	2	2	2	4	2	2	2	2	20
Тема 2. Процесс исследования данных	2	2	2	2	4	2	2	2	2	20
Тема 3. Машинное обучение и математические основы работы с данными	4	2	2	2	2	2	2	2	2	20
<b>Раздел 2. Статистика, вероятность, гипотезы и визуализация</b>	11	9	9	8	11	8	8	8	8	80
Тема 4. Описательная статистика	3	3	2	2	2	2	2	2	2	20
Тема 5. Теория вероятностей при работе с большими данными	2	2	3	2	3	2	2	2	2	20
Тема 6. Проверка гипотез при анализе данных	4	2	2	2	2	2	2	2	2	20
Тема 7. Визуализация больших данных	2	2	2	2	4	2	2	2	2	20

**Количество заданий в комплекте оценочных материалов**

Код компетенции	Наименование компетенции	Количество заданий
ПК-2	Способен разрабатывать, вводить в эксплуатацию и обслуживать базы данных; дополнять, модифицировать и совершенствовать базы данных и	<b>70</b>

	другие хранилища информации	
ПК-3	Способен выполнять разработку программного обеспечения общего и прикладного назначения, проводить оценку качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию	<b>70</b>

### Сценарии выполнения диагностических заданий

Тип задания	Последовательность действий при выполнении задания
Задание закрытого типа с однозначным выбором варианта ответа	1. Внимательно прочитать текст задания. 2. Выбрать единственный вариант ответа из предложенных.
Задание закрытого типа с многозначным выбором вариантов ответа	1. Внимательно прочитать текст задания. 2. Выбрать несколько вариантов ответа из предложенных.
Задание закрытого типа на установление соответствия	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидаются пары элементов. 2. Внимательно прочитать оба списка: список 1 - вопросы, утверждения, факты, понятия и т.д.; список 2 - утверждения, свойства объектов и т.д. 3. Сопоставить элементы списка 1 с элементами списка 2, сформировать пары элементов. 4. Записать буквы вариантов ответа (например, АБВГ)
Задание закрытого типа на установление последовательности	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается последовательность элементов. 2. Внимательно прочитать предложенные варианты ответа. 3. Построить верную последовательность из предложенных элементов. 4. Записать буквы вариантов ответа в нужной последовательности без пробелов и знаков препинания (например, БВА)
Задание открытого типа на дополнение	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается недостающее дополнение. 2. Определить какой информации не хватает. 3. Внесение пропущенного слова. 4. Записать в ответ только дополнение.
Задание открытого типа с развернутым ответом	1. Внимательно прочитать текст задания и понять суть вопроса. 2. Продумать логику и полноту ответа. 3. Записать ответ, используя четкие компактные формулировки. 4. В случае расчетной задачи записать решение и ответ.
Задание комбинированного типа: практико-ориентированные задания	1. Внимательно прочитать текст задания. 2. Выполните указанные в задания действия
Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов. 2. Внимательно прочитать предложенные варианты ответа. 3. Выбрать один ответ, наиболее верный. 4. Записать только букву выбранного варианта ответа. 5. Записать аргументы, обосновывающие выбор ответа
Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается несколько из предложенных вариантов. 2. Внимательно прочитать предложенные варианты ответа. 3. Выбрать несколько верных вариантов ответов. 4. Записать последовательно буквы выбранных вариантов без пробелов и знаков препинания (например, АБВ). 5. Записать аргументы, обосновывающие выбор каждого из ответов

### Система оценивания заданий

Указания по оцениванию	Результат оценивания (баллы, полученные за выполнение задания / характеристика правильности ответа)
Задание закрытого типа с однозначным выбором варианта ответа считается верным, если правильно определен вариант ответа	За правильный вариант ответа начисляется 1 балл
Задание закрытого типа с многозначным выбором вариантов ответа считается верным, если правильно определены все варианты ответа	За правильный вариант ответа начисляется 1 балл
Задание закрытого типа на установление соответствия считается верным, если правильно установлены все соответствия (позиции из одного столбца верно сопоставлены с позициями другого)	Количество баллов определяется числом пар для сопоставления. За каждое правильно установленное соответствие начисляется 1 балл.
Задание закрытого типа на установление последовательности считается верным, если	Максимальный балл определяется количеством элементов в последовательности. В случае ошибки

правильно указана вся последовательность цифр	в одном месте - снижение на один балл. За каждое правильно указанное место элемента в последовательности начисляется 1 балл.
Задание открытого типа на дополнение, где предоставляется предложение или фрагмент текста, в котором пропущено одно или несколько слов или фраз. Задача состоит в том, чтобы заполнить пропуски, восстановив тем самым исходный смысл предложения.	2 балла засчитывается, если студент вписал правильный ответ в соответствии с ключом. 1 балл может быть засчитан за близкий к правильному ответ, если он демонстрирует частичное понимание.
Задание открытого типа с развернутым ответом считается верным, если ответ совпадает с эталонным по содержанию и полноте	Максимальный балл - 4. Студент может получить 4 балла за полный и правильный ответ, логично изложенный и с корректной терминологией, или меньше за неполные или неточно сформулированные ответы. Полнота (1 балл), Правильность (1 балл), Логичность (1 балл), Терминология (1 балл).
Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа считается верным, если правильно указана цифра и приведены корректные аргументы, используемые при выборе ответа	За правильный выбор ответа начисляется 1 балл. За качественное обоснование - еще 2-3 балла. Критерии оценивания обоснования должны быть четко определены (например, логичность, полнота, использование фактов). Неправильный выбор ответа - 0 баллов, даже если обоснование частично верное.
Задание комбинированного типа с выбором нескольких вариантов ответа и обоснованием выбора ответа считается верным, если правильно указана цифра и приведены корректные аргументы, используемые при выборе ответа	За правильный выбор ответа начисляется 1 балл. За качественное обоснование - еще 2-3 балла. Критерии оценивания обоснования должны быть четко определены (например, логичность, полнота, использование фактов). Неправильный выбор ответа - 0 баллов, даже если обоснование частично верное.

### Тестовые задания с ключами ответов

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
<b><u>ПК-2 Способен разрабатывать, вводить в эксплуатацию и обслуживать базы данных; дополнять, модифицировать и совершенствовать базы данных и другие хранилища информации</u></b>											
1.	<b>Выберите правильный вариант ответа:</b> Ключевой характеристикой больших данных, описывающей их высокую скорость поступления и обработки, является: А) Объем (Volume) Б) Разнообразие (Variety) В) Скорость (Velocity) Г) Достоверность (Veracity)	В	Задание закрытого типа с однозначным выбором варианта ответа	1	1						
2.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных систем являются компонентами экосистемы Hadoop для обработки больших данных? А) HDFS Б) Apache Spark В) Microsoft SQL Server Г) Apache Hive	АБГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	1						
3.	<b>Установите соответствие между типами данных в контексте больших данных и их описанием:</b> <u>Тип данных:</u> 1. Структурированные данные 2. Неструктурированные данные 3. Полуструктурированные данные <u>Описание:</u> А) Данные, не имеющие предопределенной модели (тексты, изображения) Б) Данные, организованные в строго заданном формате (таблицы реляционных	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>А</td> <td>В</td> </tr> </table>	1	2	3	Б	А	В	Задание закрытого типа на установление соответствия	2	1
1	2	3									
Б	А	В									

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	БД) В) Данные с нечеткой структурой, но содержащие метки (JSON, XML) Запишите выбранные буквы под соответствующими цифрами. <table border="1" style="margin-left: 40px;"> <tr> <td style="width: 30px; text-align: center;">1</td> <td style="width: 30px; text-align: center;">2</td> <td style="width: 30px; text-align: center;">3</td> </tr> <tr> <td style="height: 15px;"></td> <td style="height: 15px;"></td> <td style="height: 15px;"></td> </tr> </table>	1	2	3							
1	2	3									
4.	<b>Укажите правильную последовательность этапов жизненного цикла данных в парадигме Data Lake:</b> А) Хранение сырых данных Б) Очистка и преобразование данных В) Анализ и извлечение инсайтов Г) Постановка задачи и сбор данных Ответ запишите в виде последовательности букв слева направо.	ГАБВ	Задание закрытого типа на установление последовательности	2	1						
5.	<b>Прочитайте и дополните фразу:</b> Распределенная файловая система, лежащая в основе экосистемы Hadoop и обеспечивающая отказоустойчивое хранение больших объемов данных на кластере стандартного оборудования, называется .	HDFS (Hadoop Distributed File System)	Задание открытого типа на дополнение	1	1						
6.	<b>Дайте развернутый ответ:</b> Опишите принцип работы архитектуры Massively Parallel Processing (MPP) в системах хранения и обработки больших данных.	Архитектура MPP предполагает разделение данных и вычислительной нагрузки между множеством независимых узлов (серверов), которые работают параллельно. Каждый узел обрабатывает свою часть данных.	Задание открытого типа с развернутым ответом	3	1						
7.	<b>Практико-ориентированное задание:</b> Предложите архитектурное решение для хранения и первичной обработки потока данных с IoT-датчиков нефтехимического предприятия, генерирующего 10 ТБ неструктурированных данных в сутки.	Для приема потока данных использовать Apache Kafka. Для долговременного хранения сырых данных – Data Lake на основе HDFS или облачного объектного хранилища (S3). Для пакетной обработки и очистки – Apache Spark.	Задание комбинированного типа: практико-ориентированные задания	3	1						
8.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Какой подход к хранению данных в первую очередь обеспечивает гибкость для будущего неопределенного анализа в проекте Big Data? А) Создание детализированной реляционной схемы данных Б) Использование Data Lake для хранения	Б Data Lake позволяет хранить данные в исходном формате, что обеспечивает максимальную гибкость для	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	1						

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	данных в сыром виде В) Нормализация всех таблиц до третьей нормальной формы	последующего анализа по различным, еще не известным на момент сохранения, сценариям, в отличие от жестких схем реляционных баз данных.									
9.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных задач эффективно решаются с помощью технологии Apache Spark в контексте больших данных? А) Поточковая обработка данных в реальном времени Б) Интерактивный SQL-анализ больших наборов данных В) Долговременное архивирование холодных данных на ленточные накопители Г) Итеративные алгоритмы машинного обучения	АБГ Spark предоставляет библиотеки для потоковой обработки (Structured Streaming), выполнения SQL-запросов (Spark SQL) и машинного обучения (MLlib), эффективно используя распределенные вычисления в памяти.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	1						
10.	<b>Прочитайте и дополните фразу:</b> Принцип _____ в архитектуре больших данных предполагает, что отказ отдельного аппаратного компонента (ноды) считается нормальным событием, и система должна автоматически продолжать работу.	Отказоустойчивости	Задание открытого типа на дополнение	1	1						
11.	<b>Выберите правильный вариант ответа:</b> Первый этап методологии CRISP-DM, на котором определяются цели проекта с точки зрения бизнеса, называется: А) Понимание данных (Data Understanding) Б) Понимание бизнеса (Business Understanding) В) Моделирование (Modeling) Г) Подготовка данных (Data Preparation)	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	2						
12.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных задач относятся к этапу «Подготовка данных» (Data Preparation) в процессе исследования данных? А) Формулирование бизнес-требований Б) Очистка данных от выбросов В) Обучение модели машинного обучения Г) Преобразование форматов данных	БГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	2						
13.	<b>Установите соответствие между этапами методологии CRISP-DM и их ключевыми результатами:</b> Этапы: 1. Понимание данных (Data Understanding) 2. Моделирование (Modeling) 3. Внедрение (Deployment) <b>Результаты:</b> А) Отчет о качестве данных, описательные статистики Б) Обученная и валидированная прогнозная модель	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>А</td> <td>Б</td> <td>В</td> </tr> </table>	1	2	3	А	Б	В	Задание закрытого типа на установление соответствия	2	2
1	2	3									
А	Б	В									

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	В) Интегрированное в производственную среду аналитическое решение Запишите выбранные буквы под соответствующими цифрами. <table border="1" data-bbox="352 344 839 405"> <tr> <td data-bbox="352 344 515 376">1</td> <td data-bbox="515 344 678 376">2</td> <td data-bbox="678 344 839 376">3</td> </tr> <tr> <td data-bbox="352 376 515 405"></td> <td data-bbox="515 376 678 405"></td> <td data-bbox="678 376 839 405"></td> </tr> </table>	1	2	3							
1	2	3									
14.	<b>Укажите правильную последовательность основных этапов типового ETL-процесса (Extract, Transform, Load):</b> А) Загрузка (Load) Б) Извлечение (Extract) В) Преобразование (Transform) Ответ запишите в виде последовательности букв слева направо.	БВА	Задание закрытого типа на установление последовательности	1	2						
15.	<b>Прочитайте и дополните фразу:</b> Процесс выявления, анализа и исправления записей, содержащих ошибки, несоответствия формату или аномальные значения в наборе данных, называется _____ данных.	очистка	Задание открытого типа на дополнение	1	2						
16.	<b>Дайте развернутый ответ:</b> Опишите цели и основные техники этапа «Разведка данных» (Data Exploration) в процессе аналитического проекта.	Цель этапа – первичное понимание данных, выявление закономерностей, аномалий, проверка гипотез и формирование идей для дальнейшего моделирования. Основные техники включают расчет описательных статистик (среднее, медиана, дисперсия), построение визуализаций (гистограммы, диаграммы рассеяния, box plot), анализ корреляций и проверку распределений признаков.	Задание открытого типа с развернутым ответом	3	2						
17.	<b>Практико-ориентированное задание:</b> Разработайте план подготовки данных (Data Preparation) для анализа логов веб-сервера с целью выявления бот-атаки. Включите в план не менее трех ключевых шагов.	1) Очистка: удаление записей с пропущенными обязательными полями (IP, timestamp). 2) Преобразование: парсинг строки user-agent и извлечение признаков (тип браузера, ОС). 3) Агрегация: подсчет	Задание комбинированного типа: практико-ориентированные задания	3	2						

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
		количества запросов с одного IP за короткие временные интервалы.			
18.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Какой метод обработки пропущенных значений (missing values) в числовом признаке является наиболее предпочтительным, если его распределение имеет значительные выбросы? А) Замена средним арифметическим значением Б) Замена медианным значением В) Удаление всех строк с пропусками	Б Медиана является робастной мерой центральной тенденции и не чувствительна к выбросам, в отличие от среднего арифметического. Удаление строк может привести к значительной потере данных.	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	2
19.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных инструментов являются подходящими для выполнения задач ETL (Extract, Transform, Load) в контексте больших данных? А) Apache NiFi Б) Apache Spark В) Microsoft Excel Г) Apache Airflow	АБГ Apache NiFi предназначен для построения потоков данных, Spark – для распределенной обработки и трансформации больших объемов данных, Airflow – для оркестрации сложных ETL-пайплайнов.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	2
20.	<b>Прочитайте и дополните фразу:</b> Техника _____ применяется для приведения числовых признаков с разными диапазонами значений к единому масштабу, например, в интервал [0, 1].	нормализация	Задание открытого типа на дополнение	1	2
21.	<b>Выберите правильный вариант ответа:</b> Алгоритм машинного обучения, который обучается на неразмеченных данных для поиска скрытых структур или кластеров, относится к классу: А) Обучение с учителем (Supervised Learning) Б) Обучение с подкреплением (Reinforcement Learning) В) Обучение без учителя (Unsupervised Learning)	В	Задание закрытого типа с однозначным выбором варианта ответа	1	3
22.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных шагов являются типичными для предобработки данных перед обучением модели машинного обучения? А) Кодирование категориальных признаков (One-Hot Encoding) Б) Разделение данных на обучающую и тестовую выборки В) Подбор гиперпараметров модели Г) Удаление признаков с высокой корреляцией	АБГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	3
23.	<b>Установите соответствие между типом</b>	1 2 3	Задание	2	3

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы									
	<p><b>задачи машинного обучения и ее примером:</b>  <u>Тип задачи:</u>            1. Классификация (Classification)            2. Регрессия (Regression)            3. Кластеризация (Clustering)  <u>Пример:</u>            А) Прогнозирование стоимости недвижимости на основе характеристик            Б) Сегментация клиентов по поведенческим паттернам            В) Определение типа растения по изображению листа            Запишите выбранные буквы под соответствующими цифрами.</p> <table border="1" data-bbox="352 645 839 707"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td></td> <td></td> <td></td> </tr> </table>	1	2	3				<table border="1" data-bbox="868 241 1062 264"> <tr> <td>В</td> <td>А</td> <td>Б</td> </tr> </table>	В	А	Б	закрытого типа на установление соответствия		
1	2	3												
В	А	Б												
24.	<p><b>Укажите правильную последовательность этапов в типовом потоке работы машинного обучения:</b>            А) Оценка модели на тестовой выборке            Б) Подготовка и предобработка данных            В) Обучение модели на обучающей выборке            Г) Постановка задачи и сбор данных            Ответ запишите в виде последовательности букв слева направо.</p>	ГБВА	Задание закрытого типа на установление последовательности	2	3									
25.	<p><b>Прочитайте и дополните фразу:</b>            Метрика _____, рассчитываемая как отношение истинно положительных предсказаний к сумме истинно положительных и ложноотрицательных, особенно важна для оценки моделей классификации на несбалансированных данных.</p>	Полнота	Задание открытого типа на дополнение	1	3									
26.	<p><b>Дайте развернутый ответ:</b>            Опишите принципиальную разницу между задачами регрессии и классификации в машинном обучении.</p>	В задаче регрессии целевая переменная (что мы предсказываем) является непрерывной числовой величиной (например, цена, температура, время). В задаче классификации целевая переменная является категориальной, то есть модель относит объект к одному из заранее определенных классов (например, «спам»/«не спам», тип заболевания).	Задание открытого типа с развернутым ответом	2	3									
27.	<p><b>Практико-ориентированное задание:</b>            Для прогнозирования вероятности сбоя насосного оборудования на</p>	Следует использовать модель бинарной	Задание комбинированного типа:	2	3									

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
	нефтехимическом заводе по данным вибрационных датчиков предложите тип модели машинного обучения и обоснуйте выбор.	классификации (например, Gradient Boosting или Logistic Regression). Целевая переменная – «сбой» / «норма». Задача сводится к предсказанию дискретного класса (сбой или нет) на основе непрерывных признаков от датчиков, что является классической задачей классификации.	практико-ориентированные задания		
28.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> При работе с очень большим набором данных для обучения модели какой подход к валидации является наиболее вычислительно эффективным? А) Перекрестная проверка на 10 фолдах (10-fold CV) Б) Разделение на одну обучающую и одну тестовую выборку (Hold-out) В) Перекрестная проверка по отдельным наблюдениям (Leave-One-Out)	Б Простое разделение на обучающую/тестовую выборку требует обучения модели всего один раз, что значительно менее затратно по вычислительным ресурсам и времени для больших данных, по сравнению с многократным обучением при k-fold CV.	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	3
29.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных библиотек Python являются де-факто стандартом для реализации алгоритмов машинного обучения и работы с данными? А) NumPy Б) Scikit-learn В) TensorFlow Г) Microsoft Power BI	АБВ NumPy – основа для численных операций, Scikit-learn – основная библиотека для классических алгоритмов ML, TensorFlow – фреймворк для глубокого обучения.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	3
30.	<b>Выберите правильный вариант ответа:</b> Автоматизированный процесс подбора комбинации гиперпараметров модели, максимизирующей ее качество на валидационной выборке, называется: А) Обучение модели (Model Training) Б) Тюнинг гиперпараметров (Hyperparameter Tuning) В) Кросс-валидация (Cross-Validation)	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	3
31.	<b>Выберите правильный вариант ответа:</b> Мера центральной тенденции, которая равна значению, делящему упорядоченную выборку пополам, называется: А) Среднее арифметическое (Mean)	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	4

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	Б) Медиана (Median) В) Мода (Mode)										
32.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных метрик являются мерами разброса (вариативности) данных? А) Дисперсия (Variance) Б) Среднее арифметическое (Mean) В) Стандартное отклонение (Standard Deviation) Г) Медиана (Median)	АВ	Задание закрытого типа с многозначным выбором вариантов ответа	2	4						
33.	Установите соответствие между статистической мерой и ее кратким описанием: <u>Мера:</u> 1. Квартиль (Quartile) 2. Дисперсия (Variance) 3. Коэффициент вариации (Coefficient of Variation) <u>Описание:</u> А) Средний квадрат отклонений значений от среднего Б) Значение, ниже которого расположен определенный процент данных В) Относительная мера разброса, выраженная в процентах Запишите выбранные буквы под соответствующими цифрами.	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>А</td> <td>В</td> </tr> </table>	1	2	3	Б	А	В	Задание закрытого типа на установление соответствия	2	4
1	2	3									
Б	А	В									
	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td></td> <td></td> <td></td> </tr> </table>	1	2	3							
1	2	3									
34.	<b>Укажите правильную последовательность расчета трех квартилей (Q1, Q2, Q3) для упорядоченного набора данных:</b> А) Найти медиану всей выборки (Q2) Б) Найти медиану верхней половины данных (Q3) В) Найти медиану нижней половины данных (Q1) Ответ запишите в виде последовательности букв слева направо.	АВБ	Задание закрытого типа на установление последовательности	2	4						
35.	<b>Прочитайте и дополните фразу:</b> График, который отображает распределение данных через пять характеристик: минимум, первый квартиль (Q1), медиану (Q2), третий квартиль (Q3) и максимум, называется _____.	боксплот	Задание открытого типа на дополнение	1	4						
36.	<b>Дайте развернутый ответ:</b> Объясните, в чем заключается практическая польза от расчета коэффициента вариации (CV) по сравнению со стандартным отклонением.	Коэффициент вариации (CV) позволяет сравнивать степень разброса данных между выборками с разными единицами измерения или сильно отличающимися средними значениями, так как он представляет собой	Задание открытого типа с развернутым ответом	2	4						

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
		относительную меру (стандартное отклонение / среднее).			
37.	<b>Практико-ориентированное задание:</b> Для набора данных о времени отклика веб-сервера (в мс): [105, 110, 115, 120, 125, 130, 200] рассчитайте и интерпретируйте медиану и стандартное отклонение.	Медиана = 120 мс. Это означает, что половина запросов обрабатывается быстрее 120 мс, а половина – медленнее. Стандартное отклонение ≈ 31.6 мс. Это говорит о среднем разбросе времени отклика относительно среднего значения. Высокое значение вызвано выбросом (200 мс).	Задание комбинированного типа: практико-ориентированные задания	3	4
38.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Для анализа асимметрии распределения набора данных следует использовать: А) Дисперсию Б) Коэффициент вариации В) Коэффициент асимметрии	В Коэффициент асимметрии количественно измеряет степень и направление асимметрии распределения (сдвиг влево или вправо). Дисперсия и коэффициент вариации измеряют разброс, но не форму распределения.	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	4
39.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных визуализаций являются наиболее эффективными для первичного описательного анализа одного числового признака? А) Круговая диаграмма Б) Гистограмма В) Линейный график Г) Боксплот	БГ Гистограмма показывает форму распределения и частоту значений. Box plot наглядно отображает медиану, квартили, размах и выбросы.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	4
40.	<b>Выберите правильные варианты ответов:</b> Какие утверждения о нормальном распределении являются верными? А) Медиана, мода и среднее арифметическое совпадают Б) Оно полностью описывается двумя параметрами: средним и стандартным отклонением	АБ	Задание закрытого типа с многозначным выбором вариантов ответа	2	4

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы												
	В) Имеет прямоугольную форму																
41.	<b>Выберите правильный вариант ответа:</b> Закон больших чисел утверждает, что с ростом объема выборки: А) Распределение выборочных средних стремится к нормальному Б) Выборочное среднее стремится к математическому ожиданию генеральной совокупности В) Стандартное отклонение выборки уменьшается	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	5												
42.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных распределений вероятностей являются дискретными? А) Нормальное распределение Б) Распределение Пуассона В) Равномерное распределение (непрерывное) Г) Биномиальное распределение	БГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	5												
43.	<b>Установите соответствие между распределением вероятностей и типичной сферой его применения:</b> <u>Распределение:</u> 1. Нормальное распределение 2. Распределение Пуассона 3. Экспоненциальное распределение <u>Сфера применения:</u> А) Моделирование времени между независимыми событиями (например, отказы оборудования) Б) Описание ошибок измерений, антропометрических данных В) Моделирование числа событий в фиксированном интервале времени (например, количество заявок в кол-центр) Запишите выбранные буквы под соответствующими цифрами. <table border="1" data-bbox="352 1283 839 1341"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td></td> <td></td> <td></td> </tr> </table>	1	2	3				<table border="1" data-bbox="868 790 1070 848"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>В</td> <td>А</td> </tr> </table>	1	2	3	Б	В	А	Задание закрытого типа на установление соответствия	2	5
1	2	3															
1	2	3															
Б	В	А															
44.	<b>Укажите правильную последовательность шагов при проверке гипотезы о законе распределения с помощью критерия согласия (например, хи-квадрат):</b> А) Расчет статистики критерия по данным Б) Формулировка нулевой и альтернативной гипотез В) Сравнение расчетного значения с критическим Г) Принятие или отклонение нулевой гипотезы Ответ запишите в виде последовательности букв слева направо.	БАВГ	Задание закрытого типа на установление последовательности	2	5												
45.	<b>Прочитайте и дополните фразу:</b> _____ вероятность — это вероятность наступления события А при условии, что событие В уже произошло. Обозначается как $P(A B)$ .	Условная	Задание открытого типа на дополнение	1	5												
46.	<b>Дайте развернутый ответ:</b> объясните значение центральной предельной теоремы (ЦПТ) для анализа больших данных.	Она обосновывает использование методов, основанных на нормальности, для выборочных	Задание открытого типа с развернутым ответом	3	5												

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
		статистик даже тогда, когда исходные данные не нормальны, если выборка велика.									
47.	<b>Практико-ориентированное задание:</b> Известно, что среднее время обработки транзакции в распределенной системе составляет 50 мс, стандартное отклонение – 5 мс (распределение близко к нормальному). Оцените вероятность того, что среднее время, рассчитанное по выборке из 100 транзакций, будет лежать в интервале от 49 до 51 мс.	Используя ЦПТ, стандартная ошибка среднего $= 5 / \sqrt{100} = 0.5$ мс. Z-оценки для 49 и 51: $(49-50)/0.5 = -2$ и $(51-50)/0.5 = 2$ . По таблице Z, $P(-2 < Z < 2) \approx 0.9544$ . Вероятность $\approx 95.4\%$ .	Задание комбинированного типа: практико-ориентированные задания	3	5						
48.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Для моделирования количества отказов сложного оборудования на определенном временном интервале при условии, что отказы независимы и редки, наиболее подходящим распределением является: А) Нормальное распределение Б) Распределение Пуассона В) Равномерное распределение	Б Распределение Пуассона специально моделирует количество событий (отказов) в фиксированном интервале времени при условии их независимости и постоянной средней интенсивности, что характерно для редких событий.	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	5						
49.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> В каких задачах анализа больших данных принципиально важно использование байесовского подхода к вероятности? А) А/В-тестирование с последовательным анализом Б) Обновление вероятности гипотезы по мере поступления новых данных В) Расчет корреляции Пирсона между двумя признаками	АБ Байесовский подход позволяет непрерывно обновлять апостериорную вероятность гипотезы (например, «вариант А лучше В») по мере накопления данных, что эффективно для последовательного А/В-тестирования и динамических систем.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	5						
50.	Установите соответствие между основным понятием теории вероятностей и его определением: <u>Понятие:</u> 1. Независимость событий 2. Математическое ожидание 3. Дисперсия случайной величины <u>Определение:</u> А) Мера разброса значений вокруг среднего	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>В</td> <td>А</td> </tr> </table>	1	2	3	Б	В	А	Задание закрытого типа на установление соответствия	2	5
1	2	3									
Б	В	А									

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы												
	<p>Б) Наступление одного события не влияет на вероятность другого</p> <p>В) Средневзвешенное значение случайной величины</p> <p>Запишите выбранные буквы под соответствующими цифрами.</p> <table border="1" style="width: 100%; text-align: center;"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td> </td> <td> </td> <td> </td> </tr> </table>	1	2	3													
1	2	3															
51.	<p><b>Выберите правильный вариант ответа:</b></p> <p>Нулевая гипотеза (H0) в статистическом тесте обычно представляет собой утверждение о:</p> <p>А) Наличии ожидаемого эффекта или различия</p> <p>Б) Отсутствии эффекта, различия или статус-кво</p> <p>В) Достоверности полученных результатов</p>	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	6												
52.	<p><b>Выберите правильные варианты ответов:</b></p> <p>Какие из перечисленных факторов влияют на мощность (power) статистического критерия?</p> <p>А) Уровень значимости (alpha)</p> <p>Б) Объем выборки (sample size)</p> <p>В) Величина ожидаемого эффекта (effect size)</p> <p>Г) Количество знаков после запятой в данных</p>	АБВ	Задание закрытого типа с многозначным выбором вариантов ответа	2	6												
53.	<p><b>Установите соответствие между статистическим критерием и типом проверяемой гипотезы:</b></p> <p><u>Критерий:</u></p> <p>t-критерий Стьюдента для независимых выборок</p> <p>Критерий хи-квадрат Пирсона</p> <p>Критерий Манна-Уитни</p> <p><u>Тип:</u></p> <p>А) Сравнение медиан двух независимых выборок (непараметрический аналог t-критерия)</p> <p>Б) Сравнение средних значений двух независимых выборок, данных с нормальным распределением</p> <p>В) Проверка связи между двумя категориальными переменными</p> <p>Запишите выбранные буквы под соответствующими цифрами.</p> <table border="1" style="width: 100%; text-align: center;"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td> </td> <td> </td> <td> </td> </tr> </table>	1	2	3				<table border="1" style="width: 100%; text-align: center;"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>В</td> <td>А</td> </tr> </table>	1	2	3	Б	В	А	Задание закрытого типа на установление соответствия	2	6
1	2	3															
1	2	3															
Б	В	А															
54.	<p><b>Укажите правильную последовательность шагов процедуры проверки статистической гипотезы:</b></p> <p>А) Выбор уровня значимости (alpha) и расчет p-value</p> <p>Б) Формулировка нулевой (H0) и альтернативной (H1) гипотез</p> <p>В) Принятие решения: отвергнуть или не отвергнуть H0</p> <p>Г) Выбор и расчет статистики критерия</p> <p>Ответ запишите в виде последовательности букв слева направо.</p>	БГАВ	Задание закрытого типа на установление последовательности	2	6												
55.	<p><b>Прочитайте и дополните фразу:</b></p> <p>Максимально допустимая вероятность совершить ошибку первого рода, то есть отвергнуть верную нулевую гипотезу, называется уровнем</p>	значимости	Задание открытого типа на дополнение	1	6												

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
56.	<p>Дайте развернутый ответ: Объясните, в чем состоит разница между ошибками первого и второго рода при проверке статистических гипотез.</p>	<p>Ошибка первого рода – это отвержение верной нулевой гипотезы. Ошибочно считается, что эффект есть. Ошибка второго рода – это принятие ложной нулевой гипотезы. Ошибочно не замечается существующий эффект.</p>	<p>Задание открытого типа с развернутым ответом</p>	2	6
57.	<p><b>Практико-ориентированное задание:</b> Сформулируйте нулевую (H0) и альтернативную (H1) гипотезы для проверки утверждения: «После внедрения новой системы кэширования среднее время загрузки страницы сайта уменьшилось».</p>	<p>H0: <math>\mu_{\text{новая}} \geq \mu_{\text{старая}}</math> (среднее время после внедрения больше или равно времени до внедрения). H1: <math>\mu_{\text{новая}} &lt; \mu_{\text{старая}}</math> (среднее время после внедрения меньше времени до внедрения). Это левосторонний тест.</p>	<p>Задание комбинированного типа: практико-ориентированные задания</p>	2	6
58.	<p><b>Выберите правильный вариант ответа и дайте обоснование:</b> Для сравнения средних значений двух независимых выборок, распределения которых существенно отличаются от нормального, следует использовать: А) Двухвыборочный t-критерий Стьюдента Б) Критерий Манна-Уитни (U-критерий) В) Критерий хи-квадрат</p>	<p>Б Критерий Манна-Уитни является непараметрическим и не требует предположения о нормальности распределения данных. Он проверяет гипотезу о равенстве медиан или сдвиге распределений. t-критерий чувствителен к нарушению нормальности.</p>	<p>Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа</p>	2	6
59.	<p><b>Выберите правильные варианты ответов и дайте обоснование:</b> Почему при планировании A/B-теста важно заранее определить необходимый размер выборки (sample size calculation)? А) Чтобы избежать ложного обнаружения эффекта из-за множественных проверок Б) Чтобы обеспечить достаточную мощность теста для обнаружения значимого эффекта В) Чтобы гарантировать нормальность распределения данных Г) Чтобы минимизировать риск ошибки второго рода (не обнаружить реальный эффект)</p>	<p>БГ Расчет размера выборки позволяет убедиться, что тест будет обладать достаточной статистической мощностью (1 - beta) для обнаружения эффекта заданной величины.</p>	<p>Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов</p>	3	6

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы												
60.	<p><b>Укажите правильную последовательность этапов расчета p-value для двустороннего критерия (two-tailed test):</b></p> <p>А) Найти площадь под кривой распределения за пределами полученного значения статистики с обеих сторон  Б) Рассчитать наблюдаемое значение статистики критерия по данным  В) Определить теоретическое распределение статистики при верной H0</p> <p>Ответ запишите в виде последовательности букв слева направо.</p>	ВБА	Задание закрытого типа на установление последовательности	2	6												
61.	<p><b>Выберите правильный вариант ответа:</b></p> <p>Для визуализации корреляционной матрицы, показывающей взаимосвязи между множеством числовых признаков, наиболее эффективна:</p> <p>А) Гистограмма  Б) Тепловая карта (Heatmap)  В) Линейный график</p>	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	7												
62.	<p><b>Выберите правильные варианты ответов:</b></p> <p>Какие из перечисленных принципов, сформулированных Эдвардом Тафти, являются основополагающими для создания эффективных дашбордов и визуализаций?</p> <p>А) Максимальное насыщение графиков декоративными элементами  Б) Устранение «визуального шума» (ненужных линий, теней, 3D-эффектов)  В) Представление максимально детализированных данных без агрегации  Г) Максимальное соотношение данных и чернил (Data-Ink Ratio)</p>	БГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	7												
63.	<p><b>Установите соответствие между типом данных/аналитической задачей и рекомендуемым видом графика:</b></p> <p><u>Тип данных:</u></p> <ol style="list-style-type: none"> <li>Анализ тренда одного показателя во времени</li> <li>Сравнение долей частей целого</li> <li>Исследование взаимосвязи между двумя числовыми переменными</li> </ol> <p><u>Вид графика:</u></p> <p>А) Круговая диаграмма или кольцевая диаграмма  Б) Линейный график (Line Chart)  В) Диаграмма рассеяния (Scatter Plot)</p> <p>Запишите выбранные буквы под соответствующими цифрами.</p> <table border="1" style="margin-left: 20px;"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>А</td> <td>В</td> </tr> </table>	1	2	3	Б	А	В	<table border="1" style="margin-left: 20px;"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>А</td> <td>В</td> </tr> </table>	1	2	3	Б	А	В	Задание закрытого типа на установление соответствия	2	7
1	2	3															
Б	А	В															
1	2	3															
Б	А	В															
64.	<p><b>Укажите правильную последовательность шагов при проектировании интерактивного дашборда для бизнес-аналитики:</b></p> <p>А) Определение ключевых показателей (KPI) и целевой аудитории  Б) Выбор инструмента визуализации и создание прототипа  В) Сбор требований и постановка задач от стейкхолдеров  Г) Тестирование удобства использования (usability testing) и финальная доработка</p>	ВАБГ	Задание закрытого типа на установление последовательности	2	7												

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
	Ответ запишите в виде последовательности букв слева направо.				
65.	<b>Прочитайте и дополните фразу:</b> Техника _____ (например, t-SNE, UMAP) используется для визуализации многомерных данных на двумерной плоскости, пытаясь сохранить локальные структуры и сходства между объектами.	уменьшение размерности	Задание открытого типа на дополнение	1	7
66.	<b>Дайте развернутый ответ:</b> Объясните, почему для визуализации больших данных часто используются агрегированные представления (агрегаты, сводные таблицы), а не отображение всех «сырых» записей.	Отображение миллионов «сырых» точек на графике приводит к перекрытию, визуальному шуму и делает график нечитаемым. Агрегация позволяет сжать информацию, выявить общие паттерны, тренды, сохранив при этом ясность и информативность визуализации.	Задание открытого типа с развернутым ответом	2	7
67.	<b>Практико-ориентированное задание:</b> Спроектируйте макет дашборда для мониторинга ключевых показателей (KPI) производственной линии в реальном времени. Перечислите не менее трех основных виджетов (типов графиков/таблиц).	1) Спидометр/большая цифра: для отображения текущей скорости производства (шт./час). 2) Линейный график: для отображения тренда скорости производства за последние 8 часов. 3) Тепловая карта состояния оборудования: цветом обозначается состояние каждого агрегата (норма, предупреждение, авария).	Задание комбинированного типа: практико-ориентированные задания	3	7
68.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Для визуализации временного ряда, демонстрирующего ярко выраженную сезонность (например, еженедельные циклы), лучше всего подходит: А) Гистограмма Б) Линейный график с достаточно длинным горизонтальным масштабом В) Круговая диаграмма	Б Линейный график идеально передает изменения показателя во времени. Достаточно длинный горизонтальный масштаб (оси X) позволяет визуально идентифицирова	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	7

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
		ть повторяющиеся сезонные паттерны, циклы и тренды.									
69.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных современных инструментов наиболее подходят для создания интерактивной визуализации и дашбордов на основе больших данных? А) Tableau Б) Microsoft Power BI В) Grafana Г) Adobe Photoshop	АБВ Photoshop – графический редактор, не инструмент для аналитической визуализации данных.	Задание комбинированного типа с выбором нескольких ответов и обоснованием	3	7						
70.	Прочитайте и дополните фразу: _____ график (или диаграмма) позволяет наглядно сравнить распределение одного количественного признака across несколько категорий, отображая медиану, квартили и выбросы для каждой группы.	боксплот	Задание открытого типа на дополнение	1	7						
<b><u>ПК-3 Способен выполнять разработку программного обеспечения общего и прикладного назначения, проводить оценку качества и работоспособности разработанного программного обеспечения, осуществлять его сопровождение и модификацию</u></b>											
71.	<b>Выберите правильный вариант ответа:</b> Какой принцип разработки ПО для больших данных гласит, что обработка должна происходить на узлах хранения, а не на центральном сервере? А) Принцип максимальной централизации Б) Принцип перемещения вычислений к данным В) Принцип перемещения данных к вычислениям	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	1						
72.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных инструментов являются частью стека технологий для оркестрации и управления пайплайнами обработки больших данных? А) Apache Airflow Б) Apache Kafka В) Luigi Г) Git	АВ	Задание закрытого типа с многозначным выбором вариантов ответа	2	1						
73.	Установите соответствие между паттерном проектирования распределенных систем и его описанием: <b>Паттерн:</b> 1. Пайплайн 2. Фангут 3. Агрегатор <b>Описание:</b> А) Распределение одной задачи между множеством параллельных исполнителей Б) Последовательная обработка данных через цепочку отдельных этапов В) Объединение результатов нескольких параллельных задач в единый ответ Запишите выбранные буквы под соответствующими цифрами.	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>А</td> <td>В</td> </tr> </table>	1	2	3	Б	А	В	Задание закрытого типа на установление соответствия	2	1
1	2	3									
Б	А	В									
74.	<b>Укажите правильную последовательность этапов разработки при использовании</b>	БВГА	Задание закрытого типа на	2	1						

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
	<b>методологии CI/CD:</b> А) Автоматическое развертывание в production Б) Написание кода и запуск юнит-тестов В) Интеграция изменений в общую ветку и запуск сборки Г) Автоматическое тестирование сборки Ответ запишите в виде последовательности букв слева направо.		установление последовательности		
75.	Прочитайте и дополните фразу: _____ подход в разработке ПО подразумевает, что инфраструктура для работы с большими данными (серверы, кластеры) описывается кодом, который можно версионировать и автоматически разворачивать.	IaC	Задание открытого типа на дополнение	1	1
76.	<b>Дайте развернутый ответ:</b> Опишите, как контейнеризация (например, с помощью Docker) упрощает разработку и развертывание приложений для экосистемы больших данных.	Контейнеризация упаковывает приложение со всеми зависимостями в единый изолированный образ. Это гарантирует идентичность окружения на этапах разработки, тестирования и эксплуатации, упрощает оркестрацию сложных распределенных приложений и обеспечивает быструю масштабируемость компонентов.	Задание открытого типа с развернутым ответом	3	1
77.	<b>Практико-ориентированное задание:</b> Разработайте схему мониторинга для распределенного Spark-приложения, обрабатывающего данные в реальном времени. Укажите не менее трех ключевых метрик, которые необходимо отслеживать, и обоснуйте их важность.	1) Задержка обработки критична для SLA систем реального времени. 2) Количество обработанных записей в секунду – показатель производительности и нагрузки. 3) Количество сбойных задач – индикатор стабильности работы приложения и проблем с данными или ресурсами.	Задание комбинированного типа: практико-ориентированные задания	3	1
78.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Для обеспечения отказоустойчивости распределенного приложения обработки данных при сбое одного из узлов кластера	А Репликация данных и вычислений позволяет	Задание комбинированного типа с выбором одного ответа	2	1

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	следует использовать: А) Репликацию данных и вычислений на нескольких узлах Б) Хранение всех данных только на одном надежном сервере В) Полный отказ от резервирования для экономии ресурсов	другому узлу немедленно взять на себя работу, обеспечивая непрерывность обработки и выполнение требований к доступности системы.	и обоснованием выбора ответа								
79.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных практик непосредственно направлены на повышение качества кода в проектах по обработке больших данных? А) Проведение ревьюирования кода Б) Написание модульных и интеграционных тестов для Spark-заданий В) Использование глобальных переменных для ускорения разработки Г) Ведение технической документации по архитектуре и пайплайнам	АБГ Ревьюирование кода выявляет ошибки и улучшает читаемость. Тестирование гарантирует корректность логики обработки данных. Документация облегчает сопровождение и модификацию.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	1						
80.	<b>Прочитайте и дополните фразу:</b> requirements.txt файл используется в проектах на _____ для управления зависимостями, что обеспечивает воспроизводимость окружения на разных этапах жизненного цикла ПО.	Python	Задание открытого типа на дополнение	1	1						
81.	<b>Выберите правильный вариант ответа:</b> Какой инструмент чаще всего используется для версионирования кода, конфигураций и скриптов в аналитических проектах? А) JIRA Б) Git В) Confluence	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	2						
82.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных артефактов следует версионировать в репозитории проекта по исследованию данных? А) Исходный код скриптов предобработки и анализа Б) Исходные сырые файлы данных объемом в терабайты В) Конфигурационные файлы (например, docker-compose.yml, .env.example) Г) Отчеты в формате PDF	АВГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	2						
83.	<b>Установите соответствие между практикой разработки и ее целью в контексте исследовательских скриптов:</b> <u>Практика:</u> Разделение кода на функции и модули Использование ячеек Markdown для комментариев Автоматизация запуска пайплайна через скрипт <u>Цель:</u> А) Повышение читаемости и объяснимости анализа Б) Обеспечение возможности повторного использования кода и его тестирования	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>А</td> <td>В</td> </tr> </table>	1	2	3	Б	А	В	Задание закрытого типа на установление соответствия	2	2
1	2	3									
Б	А	В									

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	<p>В) Обеспечение воспроизводимости результатов без ручных действий Запишите выбранные буквы под соответствующими цифрами.</p> <table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td></td> <td></td> <td></td> </tr> </table>	1	2	3							
1	2	3									
84.	<p><b>Укажите правильную последовательность шагов при документировании результатов исследовательского анализа для передачи в работу:</b>            А) Описание логики преобразования данных и выбранных моделей            Б) Фиксация версий всех используемых библиотек и данных            В) Формулировка бизнес-инсайтов и рекомендаций            Г) Описание постановки задачи и исходных данных            Ответ запишите в виде последовательности букв слева направо.</p>	ГАБВ	Задание закрытого типа на установление последовательности	2	2						
85.	<p><b>Прочитайте и дополните фразу:</b>            Процесс автоматического преобразования исследовательского кода в готовый к развертыванию модуль или пайплайн называется _____.</p>	рефакторинг	Задание открытого типа на дополнение	1	2						
86.	<p><b>Дайте развернутый ответ:</b>            Опишите, почему важно выделять конфигурационные параметры (пути к данным, настройки модели) в отдельные файлы или переменные окружения при разработке скриптов для анализа данных.</p>	Это отделяет логику приложения от его настроек, что делает код более гибким, безопасным и легко конфигурируемым для разных сред без изменения самого исходного кода.	Задание открытого типа с развернутым ответом	2	2						
87.	<p><b>Практико-ориентированное задание:</b>            Предложите структуру Python-проекта для исследования данных, которая облегчит его последующее преобразование.</p>	1) /src или /scripts 2) /configs 3) /tests 4) requirements.txt/pyproject.toml 5) Dockerfile	Задание комбинированного типа: практико-ориентированные задания	3	2						
88.	<p><b>Выберите правильный вариант ответа и дайте обоснование:</b>            Для обеспечения воспроизводимости результатов анализа данных на другом компьютере необходимо:            А) Использовать абсолютные пути к файлам в коде            Б) Фиксировать версии всех используемых библиотек и среды выполнения            В) Отправлять коллеге полный дамп всех данных по почте</p>	Б Фиксация версий библиотек и описания среды гарантирует, что анализ будет выполнен в идентичном окружении, что исключит расхождения из-за разных версий пакетов.	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	2						
89.	<p><b>Выберите правильные варианты ответов и дайте обоснование:</b>            Какие из перечисленных инструментов помогают организовать совместную работу над кодом в рамках аналитического проекта?            А) Git и платформы типа GitHub/GitLab</p>	АВ Git позволяет контролировать версии, вести историю и ревьюить изменения.	Задание комбинированного типа с выбором нескольких ответов и обоснованием	3	2						

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	(через merge requests) Б) Общая сетевая папка с постоянно перезаписываемыми файлами В) Интерактивные notebooks на платформе типа Google Colab с историей изменений Г) Пересылка архивов кода по электронной почте	Интерактивные облачные notebooks облегчают совместный просмотр и исполнение кода.	выбор ответов								
90.	<b>Выберите правильный вариант ответа:</b> Автоматический запуск тестов и сборки проекта при каждом пуше кода в репозиторий обеспечивается практикой: А) Непрерывной интеграции (CI) Б) Непрерывного развертывания (CD) В) Ручного регрессионного тестирования	А	Задание закрытого типа с однозначным выбором варианта ответа	1	2						
91.	<b>Выберите правильный вариант ответа:</b> Какой этап жизненного цикла ML-модели отвечает за интеграцию обученной модели в существующую ИС для получения предсказаний на новых данных? А) Развертывание (Deployment) Б) Обучение (Training) В) Сбор данных (Data Collection)	А	Задание закрытого типа с однозначным выбором варианта ответа	1	3						
92.	Выберите правильные варианты ответов: Какие из перечисленных компонентов должны быть включены в пайплайн MLOps для обеспечения надежности? А) Мониторинг дрейфа данных и концепта Б) Автоматическое переобучение модели по расписанию или триггеру В) Хранение сырых обучающих данных внутри кода модели Г) Ведение реестра моделей с версионированием	АБГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	3						
93.	<b>Установите соответствие между типом тестирования ML-системы и его объектом проверки:</b> <u>Тип:</u> 1. Юнит-тестирование 2. Интеграционное тестирование 3. Тестирование на корректность данных <u>Объект проверки:</u> А) Проверка корректности работы отдельной функции предобработки Б) Проверка согласованности схемы и диапазонов входящих данных В) Проверка взаимодействия сервиса предсказаний с БД и API клиентами Запишите выбранные буквы под соответствующими цифрами.	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>А</td> <td>В</td> <td>Б</td> </tr> </table>	1	2	3	А	В	Б	Задание закрытого типа на установление соответствия	2	3
1	2	3									
А	В	Б									
94.	<b>Укажите правильную последовательность этапов при развертывании ML-модели как микросервиса (REST API):</b> А) Обертывание модели в REST-интерфейс (например, с помощью FastAPI) Б) Контейнеризация сервиса (создание Docker-образа) В) Обучение и сериализация модели (например, в pickle-файл) Г) Развертывание контейнера на production-сервере или в оркестраторе	ВАБГ	Задание закрытого типа на установление последовательности	2	3						

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
	Ответ запишите в виде последовательности букв слева направо.				
95.	<b>Прочитайте и дополните фразу:</b> _____ подход предполагает развертывание нескольких версий ML-модели одновременно с маршрутизацией части трафика на новую версию для сравнения ее качества со старой.	A/B-тестирование	Задание открытого типа на дополнение	1	3
96.	Дайте развернутый ответ: Опишите, что такое «дрейф данных» и почему его мониторинг критически важен для работоспособности ML-системы.	Дрейф данных – это изменение статистических свойств входных данных модели со временем, по сравнению с данными, на которых модель обучалась. Дрейф ведет к постепенному снижению качества предсказаний без явных сбоев.	Задание открытого типа с развернутым ответом	2	3
97.	<b>Практико-ориентированное задание:</b> Спроектируйте систему логирования для ML-микросервиса, предсказывающего отток клиентов. Укажите, какие данные необходимо логировать для последующего анализа качества и отладки.	Необходимо логировать: 1) Входные признаки запроса (анонимизированные), 2) Предсказание и его уверенность (score), 3) Временную метку, 4) Идентификатор запроса (request_id), 5) Фактический результат (если становится известен позже, для обратной связи).	Задание комбинированного типа: практико-ориентированные задания	3	3
98.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Для хранения обученных ML-моделей, их версий и метаданных (метрики, гиперпараметры) рекомендуется использовать: А) Файловую систему production-сервера с понятными именами файлов Б) Специализированный реестр моделей (Model Registry) В) Базу данных бизнес-показателей	Б. Model Registry обеспечивает централизованное версионирование, управление stage, аннотации и удобный процесс промоута моделей, что критично для управляемого жизненного цикла ML.	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	3
99.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных практик обеспечивают воспроизводимость экспериментов с машинным обучением? А) Фиксация seed для генераторов случайных чисел Б) Запись гиперпараметров, метрик и	АБГ Фиксация seed гарантирует одинаковую инициализацию. Логирование параметров и артефактов	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов	3	3

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	версии датасета для каждого запуска В) Использование только интерактивных notebooks без скриптов Г) Логирование артефактов (весов модели, графиков) для каждого эксперимента	позволяет точно воссоздать условия эксперимента и его результаты.	ответов								
100.	<b>Прочитайте и дополните фразу:</b> Инструмент _____ позволяет отслеживать эксперименты ML, логируя параметры, код, метрики и артефакты, что облегчает выбор лучшей модели и воспроизведение результатов.	MLflow	Задание открытого типа на дополнение	1	3						
101.	<b>Выберите правильный вариант ответа:</b> Какое утверждение о программной реализации расчета медианы для больших данных, не помещающихся в память, является верным? А) Требуется специальный алгоритм (например, приближенный), так как нужен полный отсортированный набор Б) Медиану можно точно рассчитать за один проход по данным без их хранения В) Достаточно использовать встроенную функцию библиотеки Pandas без изменений	А	Задание закрытого типа с однозначным выбором варианта ответа	1	4						
102.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных библиотек Python подходят для эффективного расчета описательных статистик на распределенных больших данных? А) Pandas (для данных в оперативной памяти) Б) PySpark SQL (функции approxQuantile, describe) В) NumPy (для данных в оперативной памяти) Г) Dask (для параллельных вычислений на кластере)	БГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	4						
103.	<b>Установите соответствие между потенциальной проблемой качества кода при расчете статистик и способом ее предотвращения:</b> <u>Проблема:</u> 1. Ошибка округления при работе с числами с плавающей точкой 2. Некорректная обработка пропущенных значений (NaN) 3. Неявное приведение типов данных <u>Способ предотвращения:</u> А) Явное указание типов (dtype) при загрузке данных и использование функций, игнорирующих NaN (например, pr.nanmean) Б) Использование decimal-типа для финансовых расчетов или округление результата В) Использование методов .isnull() и .fillna() перед агрегацией Запишите выбранные буквы под соответствующими цифрами.	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>В</td> <td>А</td> </tr> </table>	1	2	3	Б	В	А	Задание закрытого типа на установление соответствия	2	4
1	2	3									
Б	В	А									
104.	<b>Укажите правильную последовательность шагов для разработки надежной функции расчета</b>	ВБАГ	Задание закрытого типа на	2	4						

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
	<p><b>квартилей в распределенной среде:</b></p> <p>А) Протестировать функцию на небольших контрольных наборах данных</p> <p>Б) Реализовать распределенный алгоритм (например, с использованием приближенных методов)</p> <p>В) Определить требования: точность, скорость, объем данных</p> <p>Г) Оптимизировать функцию по памяти и производительности на реалистичных данных</p> <p>Ответ запишите в виде последовательности букв слева направо.</p>		установление последовательности		
105.	<p><b>Прочитайте и дополните фразу:</b></p> <p>При разработке библиотеки для анализа данных важно обеспечить _____ результатов, чтобы одинаковый код, запущенный на одинаковых данных, давал идентичный результат независимо от платформы или времени запуска.</p>	воспроизводимос ть	Задание открытого типа на дополнение	1	4
106.	<p><b>Дайте развернутый ответ:</b></p> <p>Опишите, почему при разработке функции для расчета стандартного отклонения в Spark необходимо учитывать этап агрегации и этап свертки.</p>	Стандартное отклонение нельзя рассчитать за один проход простым суммированием по партициям. Необходимо использовать алгоритмы, которые вычисляют промежуточные агрегаты на каждой партиции, а затем финально сворачивают эти агрегаты для получения общего результата, что учитывает распределенную природу данных.	Задание открытого типа с развернутым ответом	3	4
107.	<p><b>Практико-ориентированное задание:</b></p> <p>Напишите юнит-тест для функции, рассчитывающей среднее арифметическое, который проверяет корректность обработки пустого списка, списка с отрицательными числами и списка, содержащего только нули.</p>	<pre>def test_calculate_mean(): assert calculate_mean([]) is None # или выброс исключения assert calculate_mean([- 5, -1, -10]) == - 5.333... assert calculate_mean([0 , 0, 0]) == 0.0</pre>	Задание комбинированного типа: практико-ориентированные задания	3	4
108.	<p><b>Выберите правильный вариант ответа и дайте обоснование:</b></p> <p>Для отладки функции расчета статистик, которая выдает странные результаты на больших данных, наиболее эффективно:</p> <p>А) Временно запустить функцию на</p>	А Запуск на малой выборке позволяет быстро локализовать	Задание комбинированного типа с выбором одного ответа и	2	4

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	небольшой, но репрезентативной выборке данных Б) Добавить больше print-операторов в production-код В) Увеличить объем логирования для всех записей в продакшене	ошибку, использовать отладчик и интерактивный анализ, не затрачивая время и ресурсы на полный прогон.	обоснованием выбора ответа								
109.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных принципов важно соблюдать при проектировании API библиотеки для статистических вычислений? А) Единообразии имен функций и порядка аргументов Б) Возврат понятных сообщений об ошибках при некорректных входных данных В) Использование глобального состояния для кэширования результатов Г) Предоставление как скалярных результатов, так и промежуточных агрегатов при необходимости	АБГ Единообразие и понятные ошибки улучшают UX разработчика. Возможность получить агрегаты расширяет сценарии использования.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	4						
110.	<b>Выберите правильный вариант ответа:</b> Какой метод позволяет эффективно оценить приближенное значение медианы для огромного набора данных в Spark без его полной сортировки? А) Метод approxQuantile Б) Метод collect() с последующей сортировкой в драйвере В) Метод orderBy с лимитом	А	Задание закрытого типа с однозначным выбором варианта ответа	1	4						
111.	<b>Выберите правильный вариант ответа:</b> При реализации алгоритма выборки (sampling) из распределения Пуассона для симуляции данных важно обеспечить: А) Детерминированность результата при любом seed Б) Возможность потоковой (online) генерации без хранения всех данных В) Использование только встроенного генератора случайных чисел ОС	А	Задание закрытого типа с однозначным выбором варианта ответа	1	5						
112.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных аспектов необходимо протестировать в библиотеке для генерации случайных чисел? А) Соответствие ожидаемому распределению (гистограммный тест) Б) Независимость последовательных значений В) Скорость генерации на больших объемах Г) Красоту визуализации сгенерированных данных	АБВ	Задание закрытого типа с многозначным выбором вариантов ответа	2	5						
113.	<b>Установите соответствие между задачей, связанной с вероятностью, и подходящим инструментом/библиотекой для ее реализации:</b> <u>Задача:</u> 1. Моделирование цепей Маркова 2. Вычисление плотности вероятности для нормального распределения 3. Проверка гипотезы о соответствии	<table border="1"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>Б</td> <td>А</td> <td>В</td> </tr> </table>	1	2	3	Б	А	В	Задание закрытого типа на установление соответствия	2	5
1	2	3									
Б	А	В									

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы						
	<p>распределению Инструмент/библиотека: А) Функции из scipy.stats (например, norm.pdf, kstest) Б) Специализированные библиотеки (например, romegrante, hmmlearn) В) Написание кастомного кода с использованием базовых генераторов Запишите выбранные буквы под соответствующими цифрами.</p> <table border="1" data-bbox="352 510 836 568"> <tr> <td data-bbox="352 510 517 539">1</td> <td data-bbox="517 510 681 539">2</td> <td data-bbox="681 510 836 539">3</td> </tr> <tr> <td data-bbox="352 539 517 568"></td> <td data-bbox="517 539 681 568"></td> <td data-bbox="681 539 836 568"></td> </tr> </table>	1	2	3							
1	2	3									
114.	<p><b>Укажите правильную последовательность шагов при внедрении вероятностной модели в систему для прогноза событий:</b> А) Создание эффективного API для прогноза Б) Валидация качества модели на hold-out выборке В) Реализация или дообучение модели на актуальных данных Г) Интеграция API в бизнес-процесс и настройка мониторинга Ответ запишите в виде последовательности букв слева направо.</p>	ВБАГ	Задание закрытого типа на установление последовательности	2	5						
115.	<p><b>Прочитайте и дополните фразу:</b> _____ тестирование используется для проверки того, что реализованный алгоритм генерации случайных чисел действительно производит последовательность с заявленными статистическими свойствами.</p>	Статистическое	Задание открытого типа на дополнение	1	5						
116.	<p><b>Дайте развернутый ответ:</b> Объясните, почему при распараллеливании симуляции методом Монте-Карло на кластере важно правильно управлять seed для генераторов случайных чисел на каждом узле.</p>	Чтобы избежать корреляции между потоками данных и обеспечить воспроизводимость результатов. Каждому параллельному процессу или потоку должен быть назначен уникальный, детерминированно вычисляемый seed.	Задание открытого типа с развернутым ответом	3	5						
117.	<p><b>Практико-ориентированное задание:</b> Разработайте схему кэширования для сервиса, который рассчитывает вероятности на основе часто запрашиваемых, но редко меняющихся параметров распределений. Укажите, что кэшировать, когда инвалидировать кэш и как избежать проблем с памятью.</p>	Кэшировать пары (параметры распределения, аргумент функции) – рассчитанное значение вероятности. Инвалидировать кэш при изменении справочника параметров или по TTL (времени жизни). Использовать LRU-кэш с	Задание комбинированного типа: практико-ориентированные задания	3	5						

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
		ограничением по количеству записей или памяти для предотвращения утечек.			
118.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Для оценки вероятности редкого события (например, сбоя кластера) на основе исторических данных лучше использовать: А) Наивную частоту (количество событий / общее время) Б) Байесовский подход с априорным распределением В) Игнорирование события из-за малого количества данных	Б При малом количестве наблюдений (редкие события) байесовский подход, который объединяет наблюдаемые данные с априорными ожиданиями (экспертной оценкой), дает более стабильную и обоснованную оценку, чем простая частота, которая может быть очень неустойчивой.	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	5
119.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных практик улучшают сопровождаемость кода, реализующего сложные вероятностные вычисления? А) Использование самодокументируемых имен переменных (mean, variance, lambda_) Б) Подробные комментарии с ссылками на формулы или статьи В) Инкапсуляция сложной логики в отдельные функции/классы с четкими интерфейсами Г) Хранение всех промежуточных вычислений в глобальных переменных	АБВ Четкие имена, комментарии и хорошая структура кода делают логику понятной для других разработчиков.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	5
120.	<b>Прочитайте и дополните фразу:</b> При разработке системы, использующей вероятностные модели, важно обеспечить _____ входящих данных, чтобы неверные значения (например, отрицательная дисперсия) не вызывали сбоя или бессмысленные предсказания.	валидацию	Задание открытого типа на дополнение	1	5
121.	<b>Выберите правильный вариант ответа:</b> При разработке библиотеки для A/B-тестирования важно, чтобы расчет p-value был: А) Детерминированным при одинаковых входных данных Б) Разным при каждом запуске для учета случайности В) Всегда равным 0.05 для удобства	А	Задание закрытого типа с однозначным выбором варианта ответа	1	6
122.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных метрик необходимо собирать и логировать в процессе проведения онлайн-эксперимента (A/B-теста)? А) Размеры групп (количество	АБВ	Задание закрытого типа с многозначным выбором вариантов ответа	2	6

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы												
	<p>пользователей в контроле и тесте)  Б) Значения целевой метрики для каждого пользователя/события  В) Временные метки всех событий  Г) Личные данные пользователей в открытом виде</p>																
123.	<p><b>Установите соответствие между этапом обработки данных в А/В-тесте и потенциальной программной ошибкой, которую нужно предотвратить:</b>  <b>Этап:</b>  Распределение пользователей по группам  Агрегация результатов  Статистический расчет  <b>Потенциальная ошибка:</b>  А) Несбалансированность групп из-за неслучайного seed или коллизий хэша  Б) Путаница в типах данных при суммировании, приводящая к переполнению  В) Использование неверной формулы для выбранного статистического критерия  Запишите выбранные буквы под соответствующими цифрами.</p> <table border="1" data-bbox="352 898 839 958"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td></td> <td></td> <td></td> </tr> </table>	1	2	3				<table border="1" data-bbox="868 405 1066 461"> <tr> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>А</td> <td>Б</td> <td>В</td> </tr> </table>	1	2	3	А	Б	В	Задание закрытого типа на установление соответствия	2	6
1	2	3															
1	2	3															
А	Б	В															
124.	<p><b>Укажите правильную последовательность действий при построении отчета по результатам А/В-теста в автоматизированной системе:</b>  А) Визуализация результатов (графики, доверительные интервалы)  Б) Загрузка и очистка сырых данных по эксперименту  В) Расчет статистической значимости и мощности  Г) Формирование итоговой таблицы с основными метриками и выводом  <b>Ответ запишите в виде последовательности букв слева направо.</b></p>	БВГА	Задание закрытого типа на установление последовательности	2	6												
125.	<p><b>Прочитайте и дополните фразу:</b>  _____ система позволяет автоматически запускать А/В-тесты, собирать данные, рассчитывать статистику и предоставлять дашборды без ручного вмешательства аналитика.</p>	Экспериментальная	Задание открытого типа на дополнение	1	6												
126.	<p><b>Дайте развернутый ответ:</b>  Опишите, почему при разработке системы А/В-тестирования важно реализовать механизм проверки правильности сплитования (А/А-тесты).</p>	А/А-тест (когда обе группы получают одинаковый опыт) позволяет проверить корректность работы платформы: случайность распределения, отсутствие систематических ошибок в сборе данных и корректность статистических расчетов. Если в А/А-тесте обнаруживается	Задание открытого типа с развернутым ответом	2	6												

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
		«статистически значимая» разница, это сигнал о проблемах в системе, а не о реальном эффекте.			
127.	<b>Практико-ориентированное задание:</b> Спроектируйте схему хранения сырых событий для A/B-теста в БД. Укажите ключевые поля таблицы и типы данных, учитывая необходимость последующей эффективной агрегации.	Таблица experiment_event s: event_id (UUID), user_id (INT/BIGINT), experiment_name (VARCHAR), variant (VARCHAR — 'control'/'test'), metric_value (DECIMAL/DOUBLE, nullable), event_timestamp (TIMESTAMP), additional_context (JSON). Индексы на (experiment_name, variant, event_timestamp).	Задание комбинированного типа: практико-ориентированные задания	3	6
128.	<b>Выберите правильный вариант ответа и дайте обоснование:</b> Для минимизации влияния временных эффектов (день недели, время суток) при анализе результатов A/B-теста следует: А) Проводить тест ровно одни сутки Б) Включать в анализ ковариаты (covariates) или использовать стратификацию В) Игнорировать эти эффекты, так как они случайны	Б Использование ковариат (например, день недели как признак в модели) или стратификация при сплитовании позволяет статистически учесть и скорректировать влияние известных временных паттернов, повышая точность оценки эффекта от изменения.	Задание комбинированного типа с выбором одного ответа и обоснованием выбора ответа	2	6
129.	<b>Содержание задания:</b> Выберите правильные варианты ответов и дайте обоснование: Какие из перечисленных функций должны быть покрыты автотестами в библиотеке для проверки статистических гипотез? А) Корректность расчета p-value для краевых случаев (пустые данные, нулевая дисперсия) Б) Скорость работы на выборках из миллиарда записей В) Соответствие результатов эталонным вычислениям (например, из SciPy) Г) Красота графиков, генерируемых для отчетов	АВ Тестирование краевых случаев предотвращает падения. Тесты производительно сти гарантируют, что библиотека справляется с большими данными. Сравнение с SciPy – это интеграционный тест, а не юнит-тест. Тестирование	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	6

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
		визуализаций – не основная задача библиотеки вычислений.			
130.	<b>Выберите правильный вариант ответа:</b> При последовательном мониторинге A/B-теста (peeking) без корректировки p-value повышается риск: А) Ложноположительного результата (ошибки первого рода) Б) Ложноотрицательного результата (ошибки второго рода) В) Увеличения времени проведения теста	А	Задание закрытого типа с однозначным выбором варианта ответа	1	6
131.	<b>Выберите правильный вариант ответа:</b> При разработке интерактивного дашборда на веб-технологиях (например, Dash, Plotly) ответственность за безопасность (санкционированный доступ) лежит на: А) Frontend-библиотеке визуализации Б) Backend-сервисе, обслуживающем API дашборда В) Пользователе, который открывает дашборд	Б	Задание закрытого типа с однозначным выбором варианта ответа	1	7
132.	<b>Выберите правильные варианты ответов:</b> Какие из перечисленных аспектов критичны для производительности дашборда, работающего с большими данными? А) Использование агрегатов и предварительно рассчитанных материализованных представлений Б) Загрузка всех детализированных данных в память браузера пользователя В) Настройка индексов в базе данных, используемой как источник Г) Применение пагинации или бесконечного скролла для таблиц	АВГ	Задание закрытого типа с многозначным выбором вариантов ответа	2	7
133.	<b>Укажите правильную последовательность действий при построении отчета по результатам A/B-теста в автоматизированной системе:</b> А) Визуализация результатов (графики, доверительные интервалы) Б) Загрузка и очистка сырых данных по эксперименту В) Расчет статистической значимости и мощности Г) Формирование итоговой таблицы с основными метриками и выводом Ответ запишите в виде последовательности букв слева направо.	БВГА	Задание закрытого типа на установление последовательности	2	7
134.	<b>Укажите правильную последовательность этапов при разработке новой виджета (компонента) для дашборда:</b> А) Согласование макета и функционала с бизнес-потребителем Б) Написание кода компонента на выбранном фреймворке В) Тестирование компонента на разных данных и в разных браузерах Г) Интеграция компонента в общий каркас дашборда и настройка взаимодействия Ответ запишите в виде последовательности букв слева направо.	АБВГ	Задание закрытого типа на установление последовательности	2	7

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
135.	<b>Прочитайте и дополните фразу:</b> Низкоуровневые _____, такие как D3.js, предоставляют низкоуровневый контроль над визуализацией, но требуют больше усилий для разработки по сравнению с высокоуровневыми, такими как Plotly или Chart.js.	библиотеки	Задание открытого типа на дополнение	1	7
136.	<b>Дайте развернутый ответ:</b> Опишите, как подход «mobile first» влияет на процесс разработки дашбордов и визуализаций.	Подход «mobile first» требует начинать проектирование интерфейса с мобильных устройств, что накладывает ограничения на размер экрана и способ взаимодействия (тач). Это приводит к необходимости: 1) Приоритизации информации – отображения только самого важного. 2) Адаптивного дизайна – перестроения layout при изменении размера экрана. 3) Упрощения интерактивных элементов для удобного нажатия пальцем.	Задание открытого типа с развернутым ответом	1	7
137.	<b>Практико-ориентированное задание:</b> Составьте план тестирования веб-интерфейса нового дашборда для отдела аналитики. Включите в план не менее трех проверок, которые должен выполнить тестировщик перед передачей дашборда в эксплуатацию.	1) Проверить корректность отображения всех графиков и таблиц при разных разрешениях экрана. 2) Убедиться, что фильтры и элементы управления реагируют на действия пользователя и обновляют данные. 3) Проверить, что время загрузки дашборда и отклик интерфейса соответствуют техническому заданию.	Задание комбинированного типа: практико-ориентированные задания	2	7
138.	<b>Выберите правильный вариант ответа и дайте обоснование:</b>	Б Alt-текст	Задание комбинированн	2	7

№ задания	Содержание задания	Ответ на задание	Тип задания	Уровень сложности (балл)	№ Темы
	Для обеспечения доступности (accessibility) дашборда для пользователей с нарушениями зрения необходимо: А) Использовать только черно-белую палитру Б) Добавлять текстовые описания (alt-текст) к графикам и обеспечивать навигацию с клавиатуры В) Уменьшать количество информации на экране	позволяет скринридерам озвучить содержание графика. Навигация с клавиатуры критична для пользователей, которые не могут использовать мышь. Это основные требования стандартов веб-доступности (WCAG).	ого типа с выбором одного ответа и обоснованием выбора ответа		
139.	<b>Выберите правильные варианты ответов и дайте обоснование:</b> Какие из перечисленных практик DevOps применимы к жизненному циклу дашбордов? А) Версионирование конфигураций дашбордов и скриптов их развертывания Б) Автоматическое тестирование корректности отображения данных при изменении кода В) Использование feature flags для постепенного раскрытия новых виджетов пользователям Г) Ручное копирование файлов на production-сервер по FTP	АБВ Версионирование обеспечивает отслеживаемость изменений. Автотесты (например, скриншотные) предотвращают регрессии. Feature flags позволяют управлять рисками при внедрении. Ручное развертывание не является DevOps-практикой.	Задание комбинированного типа с выбором нескольких ответов и обоснованием выборов ответов	3	7
140.	<b>Прочитайте и дополните фразу:</b> _____ инструменты, такие как Apache Superset или Redash, позволяют бизнес-пользователям создавать визуализации через веб-интерфейс без написания кода, но требуют настройки и поддержки со стороны IT-команды.	BI (Business Intelligence)	Задание открытого типа на дополнение	1	7

**Методические материалы, определяющие процедуру оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующие процессы формирования компетенций**  
**Характеристика процедуры текущей и промежуточной аттестации по дисциплине**

Оценивание знаний, умений, навыков и опыта деятельности проводятся на основе сведений, приводимых в матрице соответствия оценочных средств запланированным результатам обучения.

Цель текущего контроля успеваемости по учебным дисциплинам в семестре – проверка приобретаемых обучающимися знаний, умений, навыков в контексте формирования установленных образовательной программой компетенций в течение семестра. Текущий контроль осуществляется через систему оценки преподавателем всех видов работ обучающихся, предусмотренных рабочей программой дисциплины и учебным планом.

Критерии и шкала оценивания результатов изучения дисциплины во время занятий (текущий контроль успеваемости):

Оценка	Критерии оценки тестовых заданий	Количество верных ответов, %
«Отлично»	глубокое знание учебного материала, умение свободно выполнять задания, усвоивший взаимосвязь основных понятий дисциплины; способный самостоятельно приобретать новые знания и умения; способный самостоятельно использовать углубленные знания	86 – 100

«Хорошо»	полное знание учебного материала, успешно выполняющий предусмотренные программой задания, показывающий систематический характер знаний по дисциплине и способный к их самостоятельному пополнению и обновлению в ходе дальнейшего обучения в вузе и в будущей профессиональной деятельности	71 – 85
«Удовлетворительно»	обнаруживший знание основного учебного материала в объеме, необходимом для дальнейшего обучения, выполняющего задания, предусмотренные программой, допустившим неточности в ответе, но обладающим необходимыми знаниями для их устранения	50 – 70
«Неудовлетворительно»	имеющему пробелы в знаниях основного учебно-программного материала, допустившему принципиальные ошибки в выполнении предусмотренных программой заданий	0-50

**Критерии и шкала оценивания результатов изучения дисциплины на промежуточной аттестации**

Промежуточная аттестация результатов изучения дисциплины проводится в виде зачета и экзамена.

Основанием для определения оценки на зачете служит уровень освоения обучающимся материала и формирования компетенция, предусмотренных учебным планом.

Успеваемость на зачете определяется оценками: зачтено; не зачтено.

Оценка	Критерии оценивания	Балльно-рейтинговая оценка, %
«Зачтено»	Выставляется, если сформированность заявленных образовательных результатов компетенций оценивается не ниже «удовлетворительно» при условии отсутствия критерия «неудовлетворительно». Выставляется, когда обучающийся показывает хорошие знания изученного учебного материала; самостоятельно, логично и последовательно излагает и интерпретирует материалы учебного курса; полностью раскрывает смысл предлагаемого вопроса; владеет основными терминами и понятиями изученного курса; показывает умение переложить теоретические знания на предполагаемый практический опыт	51-100
«Не зачтено»	Выставляется, если при ответе обучающегося выявились существенные пробелы в знаниях основных положений фактического материала, неумение с помощью преподавателя получить правильное решение конкретной практической задачи из числа предусмотренных рабочей программой учебной дисциплины.	0- 50

Ответы и решения обучающихся оцениваются по следующим общим критериям: распознавание проблем; определение значимой информации; анализ проблем; аргументированность; использование стратегий; творческий подход; выводы; общая грамотность.

Обучающиеся обязаны сдавать все задания в сроки, установленные преподавателем. Оценка «Удовлетворительно» по дисциплине, может выставляться и при неполной сформированности компетенций в ходе освоения отдельной учебной дисциплины, если их формирование предполагается продолжить на более поздних этапах обучения, в ходе изучения других учебных дисциплин.

**Успеваемость на экзамене определяется оценками: отлично, хорошо, удовлетворительно, неудовлетворительно.**

Оценка	Критерии оценивания	Балльно-рейтинговая оценка, %
«Отлично»	выставляется, если сформированность заявленных образовательных результатов компетенций оценивается критериями «хорошо» и «отлично», при условии отсутствия оценки «неудовлетворительно»: студент показал прочные знания основных положений фактического материала, умение самостоятельно решать конкретные практические задачи повышенной сложности, свободно использовать справочную литературу, делать обоснованные выводы из результатов анализа конкретных ситуаций;	80-100

<b>«Хорошо»</b>	выставляется, если сформированность заявленных образовательных результатов компетенций оценивается критериями «хорошо» и «отлично», при условии отсутствия оценки «неудовлетворительно», допускается оценка «удовлетворительно»: обучающийся показал прочные знания основных положений фактического материала, умение самостоятельно решать конкретные практические задачи, предусмотренные рабочей программой, ориентироваться в рекомендованной справочной литературе, умеет правильно оценить полученные результаты анализа конкретных ситуаций;	60-79
<b>«Удовлетворительно»</b>	выставляется, если сформированность заявленных образовательных результатов компетенций оценивается критериями «удовлетворительно», «хорошо» и «отлично»: обучающийся показал знание основных положений фактического материала, умение получить с помощью преподавателя правильное решение конкретной практической задачи из числа предусмотренных рабочей программой, знакомство с рекомендованной справочной литературой;	50-59
<b>«Неудовлетворительно»</b>	выставляется, если при ответе обучающегося выявились существенные пробелы в знаниях основных положений фактического материала, неумение с помощью преподавателя получить правильное решение конкретной практической задачи из числа предусмотренных рабочей программой учебной дисциплины.	0-50

Ответы и решения обучающихся оцениваются по следующим общим критериям: распознавание проблем; определение значимой информации; анализ проблем; аргументированность; использование стратегий; творческий подход; выводы; общая грамотность.

Обучающиеся обязаны сдавать все задания в сроки, установленные преподавателем. Оценка «Удовлетворительно» по дисциплине, может выставляться и при неполной сформированности компетенций в ходе освоения отдельной учебной дисциплины, если их формирование предполагается продолжить на более поздних этапах обучения, в ходе изучения других учебных дисциплин.