

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Соловьев Дмитрий Александрович

Должность: ректор ФГБОУ ВО «Саратовский университет»

Дата подписания: 03.09.2025 13:54:20

Уникальный программный ключ:

528682d78e671e6e8a6b7f91ba2172f735a12

Приложение 1



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Саратовский государственный университет
генетики, биотехнологии и инженерии имени Н.И. Вавилова»

УТВЕРЖДАЮ

И.о. заведующего кафедрой

 / Ключиков А.В./

« 12 » апреля 2024 г.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Дисциплина	АНАЛИЗ ДАННЫХ
Направление подготовки	09.03.03 Прикладная информатика
Направленность (профиль)	Проектирование информационных систем
Квалификация выпускника	Бакалавр
Нормативный срок обучения	4 года
Форма обучения	Очная
Кафедра-разработчик	Цифровое правление процессами в АПК
Ведущий преподаватель	Шибайкин В.А., доцент

Разработчик: доцент, Шибайкин В.А.



(подпись)

Саратов 2024

Содержание

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОПОП.....	3
2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания.....	5
3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы.....	9
4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.....	19

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОПОП

В результате изучения дисциплины «Анализ данных» обучающиеся, в соответствии с ФГОС ВО по направлению подготовки 09.03.03. Прикладная информатика, утвержденного приказом Министерства образования и науки РФ от 19.09.2017 № 922, формируют следующие компетенции, указанные в таблице 1.

Таблица 1

Формирование компетенций в процессе изучения дисциплины «Анализ данных»

Компетенция		Индикаторы достижения компетенций	Этапы формирования компетенции в процессе освоения ОПОП (семестр)	Виды занятий для формирования компетенции	Оценочные средства для оценки уровня сформированности компетенции
Код	Наименование				
1	2	3	4	5	6
ПК-2	Способность понимать сущность и значение информации в развитии современного информационного общества, сознавать опасности и угрозы, возникающие в этом процессе, перерабатывать большие объёмы информации, анализировать и интерпретировать геопространственные данные, проводить целенаправленный поиск в различных источниках информации по профилю деятельности	ПК-2.1. Способен очищать данные от выбросов, пропусков и дубликатов, а также преобразовывать разные форматы данных для исследования основных свойств данных и обучения моделей искусственного интеллекта	6	лекции, лабораторные занятия	собеседование, доклад, типовое задание, тестовое задание

1	2	3	4	5	6
		ПК-2.2 Осуществляет применение основных понятий, фактов, концепций, принципов математики и информатики для решения задач извлечения, преобразования и загрузки данных	6	лекции, лабораторные занятия	собеседование, доклад, типовое задание, тестовое задание
ПК-8	Способен вести базы данных и поддержку информационного обеспечения решения прикладных задач	ПК-8.2 Способен создавать, формировать и манипулировать данными из систем и сетей	6	лекции, лабораторные занятия	собеседование, доклад, типовое задание, тестовое задание

Профиль подготовки «Проектирование информационных систем»

Компетенция ПК-2 – также формируется в ходе освоения дисциплин: Информационные технологии сбора и обработки данных, Теория информации, Теория искусственного интеллекта, Технологии геопространственного анализа, Визуализация геопространственных данных, Технологическая (проектно-технологическая) практика, Преддипломная практика, Выполнение и защита выпускной квалификационной работы.

Компетенция ПК-8 – также формируется в ходе освоения дисциплин: Базы данных, Информационные системы управления производственной компанией, Информационные системы управления взаимоотношениями с клиентами, Технологическая (проектно-технологическая) практика, Технологическая (проектно-технологическая) практика, Преддипломная практика, Выполнение и защита выпускной квалификационной работы/

2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Перечень оценочных материалов*

Таблица 2

№ п/п	Наименование оценочного материала	Краткая характеристика оценочного материала	Представление оценочного средства в ОМ
1.	собеседование	средство контроля, организованное как специальная беседа педагогического работника с обучающимся на темы, связанные с изучаемой дисциплиной и рассчитанной на выяснение объема знаний обучающегося по определенному разделу	вопросы по темам дисциплины: – перечень вопросов для устного опроса – перечень вопросов для самостоятельной работы
2.	доклад	продукт самостоятельной работы студента, представляющий собой публичное выступление по представлению полученных результатов решения определенной учебно-практической, учебно-исследовательской или научной темы	темы докладов
3.	типовое задание	содержит средство проверки умений и навыков использования цифровых технологий, программного обеспечения для решения аналитических и управленческих задач.	банк типовых заданий
4.	тестирование	метод, который позволяет выявить уровень знаний, умений и навыков, способностей и других качеств личности, а также их соответствие определенным нормам путем анализа способов выполнения обучающимися ряда специальных заданий	банк тестовых заданий

Программа оценивания контролируемой дисциплины

Таблица 3

№ п/п	Контролируемые разделы (темы дисциплины)	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
1	2	3	4
1.	Прикладные задачи аналитики данных и сферы применения аналитики данных.	ПК-2	собеседование
2.	Предиктивная аналитика	ПК-2, ПК-8	типовое задание
3.	Прикладные задачи аналитики данных	ПК-2, ПК-8	сообщение
4.	Подготовка данных	ПК-2, ПК-8	типовое задание
5.	Прогнозирование и работа собеседование большим объемом данных	ПК-2, ПК-8	типовое задание
6.	Основные характеристики данных	ПК-2, ПК-8	типовое задание
7.	Практика машинного обучения	ПК-2, ПК-8	тестовое задание, доклад, по
8.	Практика машинного обучения	ПК-2, ПК-8	типовое задание
9.	Работа собеседование данными. Регрессионный анализ, кластерный анализ, классификация	ПК-2, ПК-8	типовое задание
10	Работа собеседование нейронными сетями	ПК-2, ПК-8	типовое задание
11	Работа собеседование динамическими нейронными сетями	ПК-2, ПК-8	типовое задание
12	Примеры использования сверточных нейронных сетей	ПК-2, ПК-8	типовое задание
13	Примеры использования SQLite	ПК-2, ПК-8	тестовое задание
14	Начало использования языка программирования в анализе данных	ПК-2, ПК-8	типовое задание
15	Использование Python для визуализации данных	ПК-2, ПК-8	типовое задание
	Основы работы с данными в облачных сервисах.	ПК-2, ПК-8	тестовое задание
	Применение методов анализа данных в различных отраслях.	ПК-2, ПК-8	типовое задание

**Описание показателей и критериев оценивания компетенций по дисциплине
«Анализ данных» на различных этапах их формирования, описание шкал
оценивания**

Таблица 4

Код компетенции, этапы освоения компетенции	Индикаторы достижения компетенций	Показатели и критерии оценивания результатов обучения			
		ниже порогового уровня (неудовлетворительно)	пороговый уровень (удовлетворительно)	продвинутый уровень (хорошо)	высокий уровень (отлично)
1	2	3	4	5	6
ПК-2 семестр	ПК-2.1. Способен очищать данные от выбросов, пропусков и дубликатов, а также преобразовывать разные форматы данных для исследования основных свойств данных и обучения моделей искусственного интеллекта	обучающийся не знает значительной части программного материала, плохо ориентируется в материале о возможности количественного и качественного анализа данных, этапов по подготовке данных к анализу технологий, не знает практику применения материала, допускает существенные ошибки	обучающийся демонстрирует знания только основного материала, но не знает деталей, допускает неточности, допускает неточности в формулировках, нарушает логическую последовательность в изложении программного материала	обучающийся демонстрирует знание материала о возможности количественного и качественного анализа данных, этапов по подготовке данных к анализу, не допускает существенных неточностей	обучающийся демонстрирует знание материала о возможности количественного и качественного анализа данных, этапов по подготовке данных к анализу, возможности таблиц и спецпрограмм для анализа количественных данных; практики применения материала, исчерпывающе и последовательно, четко и логично излагает материал, хорошо ориентируется в материале, не затрудняется с ответом при видоизменении заданий
	ПК-2.2 Осуществляет применение основных понятий, фактов, концепций, принципов	обучающийся не демонстрирует базовых знаний курса, а также не имеет представления о более	обучающийся демонстрирует базовые знания, но еще не имеет представления о более сложных функциях и	обучающийся имеет знания для начала работы над анализом данных, однако преимущественно это были знания об одном из двух разделов	обучающийся владение всеми ключевыми аспектами математики и информатики, умеет применять их в различных задачах, в том числе связанных

	математик и и информати ки для решения задач извлечения , преобразов ания и загрузки данных	сложных функциях и алгоритмах, использованн ых в частях дисциплины	алгоритмах, использованы в частях дисциплины.	(математике или информатике)	с обработкой данных.
ПК- 8 6 семе стр	ПК-8.2 Способен создавать, формирова ть и манипулир овать данными из систем и сетей	обучающийся не может создавать базовые наборы данных с минимальной помощью, не способен собирать данные из одного источника, не выполняет базовые операции по фильтрации и сортировке данных с использовани ем стандартных инструментов	обучающийся может создавать базовые наборы данных с минимальной помощью, способен собирать данные из одного или двух источников и объединять их в простую структуру, выполняет базовые операции по фильтрации и сортировке данных с использовани ем стандартных инструментов	обучающийся самостоятельно создает сложные наборы данных, включающие несколько переменных, умеет интегрировать данные из нескольких источников, используя различные форматы и структуры, применяет продвинутое методы обработки данных, такие как агрегация и трансформация, с использованием специализирован ных инструментов и языков программирован ия.	обучающийся разрабатывает комплексные наборы данных, включающие метаданные и аннотации, с учетом всех аспектов качества данных, интегрирует данные из разнородных источников, включая большие объемы данных и данные в реальном времени, с высокой точностью и надежностью, выполняет сложные операции по анализу данных, включая машинное обучение и прогнозирование, с использованием передовых технологий и алгоритмов.

3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

3.1. Входной контроль

Примерный перечень вопросов

1. Сформулируйте определение информатики и информационных технологий
2. Tensorflow – это язык программирования?
3. Группировки множества объектов на подмножества - это ...
4. Python – это ...
5. Выберите «три кита» Объектно-Оrientированного Программирования
6. Семантическая сеть - это ...
7. Каков результат следующего выражения: $(3^2 - 2^3 + 2^4 - 4^2) \wedge 4 = ?$
8. Напишите чему равно $4! - 3!$
9. Гамбургер и кола вместе стоят 1 доллар и 10 центов, известно, что гамбургер стоит на доллар дороже. Сколько стоит гамбургер?
10. Массив (матрица) – это ...
11. Что такое сортировка пузырьком?
12. В чём смысл двоичного (бинарного) поиска?

3.2. Доклады

Рекомендуемая тематика докладов по дисциплине приведена в таблице 5.

Таблица 5

Темы докладов, рекомендуемые к написанию при изучении дисциплины «Анализ данных»

№ п/п	Темы докладов
1	2
1	Принцип Анализа текстовой информации — Text Mining. Инструменты.
2	Добыча данных — Data Mining
3	Общие сведения о платформе Loginom. Интерфейс платформы, основные понятия. Компоненты и узлы
4	Методы многомерного анализа данных
5	Понятие, цели и задачи визуализации. Этапы визуализации

3.3. Кейс-задания

При проверке кейса использовалась сто бальная шкала. Каждый ответ весит определенное количество баллов, а именно максимальное количество баллов за каждый ответ: Вопрос 1 - 20 баллов; Вопрос 2 - 50 баллов; Вопрос 3 - 30 баллов;. Критерии оценки ответов: Полнота ответа с использованием всей информации из

описания ситуации, Обоснованность, Умение оперировать терминами и понятиями в сфере управления персоналом, Использование теоретических моделей и концепций, Представленность нескольких точек зрения на проблему, Отсутствие фактических ошибок.

Пример Кейс-задания Тема Основные характеристики данных

Выбрать критерий позволяющий с определенной (определяется обучающимся) достоверностью определить различия в среднем распределении вектора скорости для 3 разных конфигураций. Предварительно посчитать длину вектора скорости акселерометра по формуле $|\vec{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}$, оценить и оценить соответствующие характеристики выборки (определяется обучающимися).

Пример данных

Конфигурация вентилятора	x	y	z
1	1,004	-0,043	-0,344
1	0,949	0,203	-0,289
1	0,988	-0,035	-0,102
1	0,965	0,16	0,129
1	1,051	-0,555	-0,145
1	0,988	-0,047	-0,063
1	1,031	-0,125	-0,258

Вопросы к кейсу:

1. Сформулируйте план Ваших действий по проекту.
2. Сформулируйте методику проверки поставщиков с использованием цифровых технологий: основные характеристики выборки для конфигураций.
3. Сформулируйте минимум 2 критерия для определения различий в конфигурациях. Определите являются выборки зависимыми или независимыми. Как вы определите величину достоверности?

3.4. Типовое задание

Тематика типовых заданий устанавливается в соответствии с разделами дисциплины. Типовое задание подразделяется на два этапа. На первом этапе проводится теоретически обзор по теме занятия. На втором этапе решается задача по данной теме. Для каждой темы предусмотрен вариант задания.

Пример типового задания. Тема: Примеры использования SQLite

В таблице загруженной в базу данных есть список атрибутов: ProductID – Код продукта, Name – Наименование продукта, Producer – Производитель продукта, Price – Цена продукта, Material – Материал продукта, Color – Цвет продукта, ProductCategory – Категория продукта. Запрос: Получить название, производителя и цвет всех товаров

	ProductID	Name	Producer	Price	Material	Color	ProductCategory
0	1	Блузка	Lorelei	25	Хлопок	Бежевый	Особый
1	2	Рубашка	SuperB	19	Полиэстер	Розовый	Обычный
2	3	Блузка	SommerZeit	18	Полиэстер	Розовый	Обычный
3	4	Блузка	Lorelei	28	Хлопок	Белый	Обычный
4	5	Рубашка	Lorelei	21	Хлопок	Голубой	Обычный
5	6	Юбка	Hallo-Hallo	37	Полиэстер	Коричневый	Особый
6	7	Кофта	SuperB	17	Трикотаж	Бежевый	Обычный
7	8	Кофта	Hallo-Hallo	32	Шерсть	Жёлтый	Обычный
8	9	Рубашка	Geschenk	45	Хлопок	Белый	Особый

Составить запрос: получить название, производителя и цвет всех товаров

Проанализировать результаты и сделать выводы

3.6. Тестовые задания

По дисциплине «Анализ данных » предусмотрено проведение следующих видов тестирования: письменное, компьютерное и т.п.

Письменное тестирование.

Письменное тестирование рассматривается как рубежный контроль успеваемости и проводится после изучения раздела дисциплины **Консолидация данных** и раздела **Повышение вовлеченности сотрудников**.

Результаты тестирования учитываются при проведении промежуточной аттестации

Пример тестового задания занятие Практика машинного обучения.

Группа _____ ФИО тестируемого

Тест 1 Консолидация данных

- В какой шкале с зарегистрированными данными можно выполнять только операцию проверки их совпадения**
 - или несовпадения?
 - Абсолютная шкала
 - Шкала отношений
 - Номинальная шкала
 - Шкала порядка
- Верно ли утверждение, что если коэффициент корреляции равен 0.6, то такая взаимосвязь всегда будет**
 - статистически достоверна (уровень значимости обязательно будет меньше 0,05)?
 - Да, вероятность получить значимую взаимосвязь случайно практически равна 0.
 - Нет, $r=0,6$ не обязательно означает статистически значимую взаимосвязь
- Какой метод проверки статистических гипотез необходимо выбрать, если:**
 - значения признака могут быть представлены в любой шкале, начиная от шкалы наименований;

- b. **распределение признака может быть любым и совпадение его с каким-либо теоретическим законом**
 - c. **распределения необязательно и не нуждается в проверке.**
 - d. Непараметрические методы
 - e. Параметрические методы;
4. **Если наблюдаемая величина подчиняется нормальному закону распределения, гистограмма числового ряда (несколько вар) будет иметь форму:**
- a. Унимодальную
 - b. Острове́ршинную
 - c. Симметричную
 - d. Все вышесказанное
 - e. Ассиметричную
5. **Проанализируйте 2 гистограммы и выберите верное высказывание:**
- a. Размах примерно одинаковый в обоих случаях
 - b. В первом случае размах больше, чем во втором
 - c. Во втором случае размах больше, чем в первом
6. **Укажите верные высказывания:**
- a. Нулевая гипотеза - это гипотеза об отсутствии различий.
 - b. Альтернативная гипотеза - это гипотеза о значимости различий.
 - c. Противоположная гипотеза - это гипотеза о независимости выборок.
 - d. Нулевая гипотеза - это гипотеза о влиянии друг на друга исследуемых выборок
7. **Укажите верные высказывания (тема регрессионный анализ):**
- a. Мы рассчитываем коэффициент регрессионной прямой так, чтобы максимизировать сумму квадратов остатков
 - b. Коэффициент b всегда равен только коэффициенту корреляции
 - c. Мы рассчитываем коэффициент регрессионной прямой так, чтобы минимизировать сумму квадратов остатков
8. **Как вы считаете, в отрицательно асимметричных распределениях:**
- a. Среднее значение больше медианы
 - b. Среднее значение меньше моды
 - c. Медиана больше моды
9. **Проведено исследование на 100 пациентах с целью выявления степени воздействия медицинского препарата**
- a. на течение ОРВИ. На какую совокупность можно распространить наши выводы?
 - b. Мужчины и женщины во возрасте от 25 до 50 лет
 - c. Пациенты с ОРВИ
 - d. Пациенты, у которых выявлено то либо иное заболевание
10. **Представьте, что в выборку из 40 наблюдений (значения признака в диапазоне от 1 до 45) добавили еще 1**
- a. число, значение которого равно 3500.
 - b. Как вы считаете, какая из мер центральной тенденции изменится самым значительным образом?
 - c. Мода
 - d. Среднее значение $\square$
 - e. Медиана

3.7. Рубежный контроль

Вопросы рубежного контроля № 1

Вопросы, рассматриваемые на аудиторных занятиях

1. Какие основные прикладные задачи решает аналитика данных в бизнесе?
2. Приведите примеры сфер применения аналитики данных.
3. Что такое предиктивная аналитика и как она отличается от описательной аналитики?
4. Какие методы используются в предиктивной аналитике
5. Какие задачи аналитики данных актуальны для телекоммуникационных компаний?
6. Приведите примеры использования аналитики данных в телекоммуникациях.
7. Почему важен этап подготовки данных перед анализом?
8. Какие методы используются для сбора данных?
9. Какие этапы включает в себя процесс прогнозирования?
10. Какие модели машинного обучения чаще всего используются для прогнозирования?
11. Какие основные характеристики данных необходимо учитывать при анализе?
12. Что такое выбросы в данных и как с ними работать?
13. Какие библиотеки Python используются для машинного обучения?
14. Приведите пример задачи, которую можно решить с помощью машинного обучения в Python
15. Какие этапы включает в себя процесс построения модели машинного обучения?
16. Как оценивается качество модели машинного обучения?

Вопросы для самостоятельного изучения

1. Как аналитика данных помогает в принятии управленческих решений?
2. Какие метрики используются для оценки эффективности аналитики данных?
3. Какие ограничения существуют у предиктивной аналитики?
4. Как интерпретировать результаты предиктивной аналитики?
5. Какие данные собирают телекоммуникационные компании для аналитики?
6. Как аналитика данных помогает улучшить качество обслуживания в телекоммуникациях?
7. Какие инструменты используются для автоматизации сбора данных?
8. Как обрабатывать пропущенные значения в данных?
9. Какие методы используются для оценки точности прогноза?
10. Как выбрать оптимальную модель для прогнозирования?
11. Как определить распределение данных?
12. Что такое корреляция и как её измерять?
13. Как работать с категориальными данными в Python?
14. Как настроить гиперпараметры модели машинного обучения?
15. Какие методы используются для борьбы с переобучением модели?

16. Как интерпретировать результаты машинного обучения?

Вопросы рубежного контроля № 2

Вопросы, рассматриваемые на аудиторных занятиях

1. *Что такое нейронная сеть и как она работает?*
2. *Какие основные архитектуры нейронных сетей существуют?*
3. *Как происходит процесс обучения нейронной сети?*
4. *Что такое динамические нейронные сети и как они отличаются от статических?*
5. *Какие задачи решают динамические нейронные сети?*
6. *Приведите пример алгоритма динамической нейронной сети.*
7. *Что такое сверточные нейронные сети (CNN)?*
8. *Как работает сверточный слой в CNN?*
9. *В каких задачах чаще всего применяются сверточные нейронные сети?*
10. *Какие методы используются для анализа текстовых данных?*
11. *Что такое токенизация и зачем она нужна?*
12. *Как работают модели на основе рекуррентных нейронных сетей (RNN) для анализа текста?*
13. *Что такое SQLite и в каких случаях его используют?*
14. *Какие основные команды SQL используются для работы с базой данных?*
15. *Как выполнить запрос на выборку данных в SQLite?*
16. *Почему Python популярен для анализа данных?*
17. *Какие библиотеки Python используются для анализа данных?*
18. *Как установить необходимые библиотеки для анализа данных в Python?*
19. *Какие библиотеки Python используются для визуализации данных?*
20. *Как построить график с использованием Matplotlib?*
21. *Что такое Seaborn и как он отличается от Matplotlib?*
22. *Как читать данные из файла в Python?*
23. *Как записать данные в файл с использованием Python?*
24. *Какие форматы файлов чаще всего используются для хранения данных?*
25. *Какие основные облачные сервисы используются для работы с данными?*
26. *Что такое облачное хранилище данных?*
27. *Какие преимущества и недостатки у работы с данными в облаке?*
28. *В каких отраслях чаще всего применяется анализ данных?*
29. *Какие методы анализа данных используются в маркетинге?*
30. *Как анализ данных помогает в здравоохранении?*

Вопросы для самостоятельного изучения

1. Как выбрать архитектуру нейронной сети для конкретной задачи?
2. Что такое переобучение и как с ним бороться?
3. Какие методы используются для оценки качества нейронной сети?
4. Как обучать динамические нейронные сети?
5. Какие проблемы могут возникнуть при работе с динамическими нейронными сетями?
6. Как интерпретировать результаты работы динамической нейронной сети?
7. Как настроить гиперпараметры сверточной нейронной сети?
8. Какие методы используются для улучшения производительности CNN?
9. Как работает пулинг в сверточных нейронных сетях?
10. Как обрабатывать многозначность в текстовых данных?
11. Какие метрики используются для оценки качества моделей анализа текста?
12. Что такое векторизация текста и как она выполняется?
13. Как создать базу данных в SQLite?
14. Как выполнить сложный запрос с использованием нескольких таблиц?
15. Как оптимизировать запросы в SQLite?
16. Как работать с большими объемами данных в Python?
17. Какие инструменты используются для параллельной обработки данных в Python?
18. Как управлять памятью при работе с данными в Python?
19. Как создать интерактивные визуализации в Python?
20. Какие библиотеки используются для создания дашбордов в Python?
21. Как выбрать подходящий тип графика для визуализации данных?
22. Как работать с большими файлами в Python?
23. Как обрабатывать ошибки при чтении и записи файлов?
24. Как работать с различными кодировками файлов?
25. Как перемещать данные между локальными и облачными хранилищами?
26. Какие меры безопасности необходимо учитывать при работе с данными в облаке?
27. Как оптимизировать затраты на использование облачных сервисов?
28. Как анализ данных используется в финансовом секторе?
29. Какие методы анализа данных применяются в производстве?
30. Как анализ данных помогает в оптимизации цепочек поставок?

3.8 Промежуточная аттестация

В соответствии с учебным планом по направлению подготовки 09.03.03

Прикладная информатика вид промежуточной аттестации – экзамен.

Целью проведения экзамена по дисциплине является определение фактического уровня теоретических знаний и навыков обучающихся. В билетах для экзамена расчетные задания отсутствуют.

Вопросы, выносимые на экзамен

1. *Какие основные прикладные задачи решает аналитика данных в бизнесе?*
2. *Приведите примеры сфер применения аналитики данных.*
3. *Что такое предиктивная аналитика и как она отличается от описательной аналитики?*
4. *Какие методы используются в предиктивной аналитике*
5. *Какие задачи аналитики данных актуальны для телекоммуникационных компаний?*
6. *Приведите примеры использования аналитики данных в телекоммуникациях.*
7. *Почему важен этап подготовки данных перед анализом?*
8. *Какие методы используются для сбора данных?*
9. *Какие этапы включает в себя процесс прогнозирования?*
10. *Какие модели машинного обучения чаще всего используются для прогнозирования?*
11. *Какие основные характеристики данных необходимо учитывать при анализе?*
12. *Что такое выбросы в данных и как с ними работать?*
13. *Какие библиотеки Python используются для машинного обучения?*
14. *Приведите пример задачи, которую можно решить с помощью машинного обучения в Python*
15. *Какие этапы включает в себя процесс построения модели машинного обучения?*
16. *Как оценивается качество модели машинного обучения?*
17. *Как аналитика данных помогает в принятии управленческих решений?*
18. *Какие метрики используются для оценки эффективности аналитики данных?*
19. *Какие ограничения существуют у предиктивной аналитики?*
20. *Как интерпретировать результаты предиктивной аналитики?*
21. *Какие данные собирают телекоммуникационные компании для аналитики?*
22. *Как аналитика данных помогает улучшить качество обслуживания в телекоммуникациях?*
23. *Какие инструменты используются для автоматизации сбора данных?*
24. *Как обрабатывать пропущенные значения в данных?*
25. *Какие методы используются для оценки точности прогноза?*

26. Как выбрать оптимальную модель для прогнозирования?
27. Как определить распределение данных?
28. Что такое корреляция и как её измерять?
29. Как работать с категориальными данными в Python?
30. Как настроить гиперпараметры модели машинного обучения?
31. Какие методы используются для борьбы с переобучением модели?
32. Как интерпретировать результаты машинного обучения?
33. Что такое нейронная сеть и как она работает?
34. Какие основные архитектуры нейронных сетей существуют?
35. Как происходит процесс обучения нейронной сети?
36. Что такое динамические нейронные сети и как они отличаются от статических?
37. Какие задачи решают динамические нейронные сети?
38. Приведите пример алгоритма динамической нейронной сети.
39. Что такое сверточные нейронные сети (CNN)?
40. Как работает сверточный слой в CNN?
41. В каких задачах чаще всего применяются сверточные нейронные сети?
42. Какие методы используются для анализа текстовых данных?
43. Что такое токенизация и зачем она нужна?
44. Как работают модели на основе рекуррентных нейронных сетей (RNN) для анализа текста?
45. Что такое SQLite и в каких случаях его используют?
46. Какие основные команды SQL используются для работы с базой данных?
47. Как выполнить запрос на выборку данных в SQLite?
48. Почему Python популярен для анализа данных?
49. Какие библиотеки Python используются для анализа данных?
50. Как установить необходимые библиотеки для анализа данных в Python?
51. Какие библиотеки Python используются для визуализации данных?
52. Как построить график с использованием Matplotlib?
53. Что такое Seaborn и как он отличается от Matplotlib?
54. Как читать данные из файла в Python?
55. Как записать данные в файл с использованием Python?
56. Какие форматы файлов чаще всего используются для хранения данных?
57. Какие основные облачные сервисы используются для работы с данными?
58. Что такое облачное хранилище данных?
59. Какие преимущества и недостатки у работы с данными в облаке?
60. В каких отраслях чаще всего применяется анализ данных?
61. Какие методы анализа данных используются в маркетинге?
62. Как анализ данных помогает в здравоохранении?

63. Как выбрать архитектуру нейронной сети для конкретной задачи?
64. Что такое переобучение и как с ним бороться?
65. Какие методы используются для оценки качества нейронной сети?
66. Как обучать динамические нейронные сети?
67. Какие проблемы могут возникнуть при работе с динамическими нейронными сетями?
68. Как интерпретировать результаты работы динамической нейронной сети?
69. Как настроить гиперпараметры сверточной нейронной сети?
70. Какие методы используются для улучшения производительности CNN?
71. Как работает пулинг в сверточных нейронных сетях?
72. Как обрабатывать многозначность в текстовых данных?
73. Какие метрики используются для оценки качества моделей анализа текста?
74. Что такое векторизация текста и как она выполняется?
75. Как создать базу данных в SQLite?
76. Как выполнить сложный запрос с использованием нескольких таблиц?
77. Как оптимизировать запросы в SQLite?
78. Как работать с большими объемами данных в Python?
79. Какие инструменты используются для параллельной обработки данных в Python?
80. Как управлять памятью при работе с данными в Python?
81. Как создать интерактивные визуализации в Python?
82. Какие библиотеки используются для создания дашбордов в Python?
83. Как выбрать подходящий тип графика для визуализации данных?
84. Как работать с большими файлами в Python?
85. Как обрабатывать ошибки при чтении и записи файлов?
86. Как работать с различными кодировками файлов?
87. Как перемещать данные между локальными и облачными хранилищами?
88. Какие меры безопасности необходимо учитывать при работе с данными в облаке?
89. Как оптимизировать затраты на использование облачных сервисов?
90. Как анализ данных используется в финансовом секторе?
91. Какие методы анализа данных применяются в производстве?
92. Как анализ данных помогает в оптимизации цепочек поставок?

Пример экзаменационного билета

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«ФГБОУ ВО Саратовский государственный университет
генетики, биотехнологии и инженерии имени Н.И. Вавилова»**

Кафедра «Цифровое управление процессами в АПК

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 1

по дисциплине «Анализ данных»

1. Вопрос 1
2. Вопрос 2.

Зав. кафедрой

дата.

ФИО

4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

4.1 Процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Контроль результатов обучения студентов, этапов и уровня формирования компетенций по дисциплине «Анализ данных » осуществляется через проведение входного, текущего, рубежных, выходного контролей и контроля самостоятельной работы

Формы текущего, промежуточного и итогового контроля и контрольные задания для текущего контроля разрабатываются кафедрой исходя из специфики дисциплины, и утверждаются на заседании кафедры.

4.2 Критерии оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Описание шкалы оценивания достижения компетенций по дисциплине приведено в таблице 6.

Таблица 6

Уровень освоения компетенции	Отметка по пятибалльной системе (Экзамен)	Описание
высокий	«отлично»	Обучающийся обнаружил всестороннее, систематическое и глубокое знание учебного материала, умеет свободно выполнять задания, предусмотренные программой, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой. Как правило, обучающийся проявляет творческие способности в понимании, изложении и использовании материала
базовый	«хорошо»	Обучающийся обнаружил полное знание учебного материала, успешно выполняет предусмотренные в программе задания, усвоил основную литературу, рекомендованную в программе
пороговый	«удовлетворительно»	Обучающийся обнаружил знания основного учебного материала в объеме, необходимом для дальнейшей учебы и предстоящей работы по профессии, справляется с выполнением практических заданий, предусмотренных программой, знаком с основной литературой, рекомендованной программой, допустил погрешности в ответе на экзамене и при выполнении экзаменационных заданий, но обладает необходимыми знаниями для их устранения под руководством преподавателя
—	«неудовлетворительно»	Обучающийся обнаружил пробелы в знаниях основного учебного материала, допустил принципиальные ошибки в выполнении предусмотренных программой практических заданий, не может продолжить обучение или приступить к профессиональной деятельности по окончании образовательной организации без дополнительных занятий

4.2.1. Критерии оценки устного ответа при промежуточной аттестации

При ответе на вопрос обучающийся демонстрирует:

знания: теоретические основы поставленных вопросов.

умения: формулировать методы применения анализа данных в задачах профессиональной деятельности.

владение навыками: навыками работы в интегрированной среде разработки программ с использованием механизмов ООП, для решения профессиональных задач

Критерии оценки*

отлично	обучающийся демонстрирует:
----------------	----------------------------

	– знание материала х, практики применения материала, исчерпывающе и последовательно, четко и логично излагает материал, хорошо ориентируется в материале, не затрудняется с ответом при видоизменении заданий; – умение <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> – успешное и системное владение навыками чтения и оценки задач
хорошо	обучающийся демонстрирует: – знание материала, не допускает существенных неточностей; – в целом успешное, но содержащие отдельные пробелы, умение <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> – в целом успешное, но содержащее отдельные пробелы или сопровождающееся отдельными ошибками владение навыками чтения и оценки задач.
удовлетворительно	обучающийся демонстрирует: – знания только основного материала, но не знает деталей, допускает неточности, допускает неточности в формулировках, нарушает логическую последовательность в изложении программного материала; – в целом успешное, но не системное умение формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;; – в целом успешное, но не системное владение навыками чтения и оценки задач..
неудовлетворительно	обучающийся: – не знает значительной части программного материала, плохо ориентируется в материале по современным технологиям анализа данных; – не умеет <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> – обучающийся не владеет навыками чтения и оценки <i>Задач</i>

4.2.2. Критерии оценки доклада

При написании доклада обучающийся демонстрирует:

знания: по вопросам доклада

умения: анализировать и содержательно интерпретировать полученные результаты исследований.

владение навыками: поиска информации в традиционных библиотеках и информационных ресурсах.

Критерии оценки доклада

отлично	обучающийся демонстрирует: - знание исследуемой темы (доклад структурирован; использованы различные точки зрения по рассматриваемому вопросу, сделаны и аргументированы основные выводы, прослушивается самостоятельность суждений, основные понятия вопроса изложены подробно) - логичность и структурированность изложения материала; - расширенную электронную презентацию к докладу на 5 слайдов
хорошо	обучающийся демонстрирует:

	- знание темы доклада (доклад структурирован; использованы различные точки зрения по рассматриваемому вопросу, сделаны и аргументированы основные выводы) - расширенную электронную презентацию к докладу менее 5 слайдов
удовлетворительно	обучающийся демонстрирует: - неполное знание материала (в материале представлена одна точка зрения, отсутствует самостоятельность суждений) - не представлена электронная презентация
неудовлетворительно	обучающийся: не выполнил доклад

4.2.3. Критерии оценки выполнения типового задания

При выполнении контрольных (самостоятельных) работ обучающийся демонстрирует:

знания: современных методов и программных средств выявления тенденций изменения признаков, современных языков запросов

умения: использовать интегрированную среду разработки, применять язык SQL для работы, по построению запросов, сбора и подготовки данных и визуализации данных;

владение навыками: навыками использования технологии анализа данных.

отлично	обучающийся демонстрирует: - знания современных методов и программных средств выявления тенденций изменения признаков - умения анализировать и правильно интерпретировать, применять инструментарий программы, использовать интегрированную среду разработки, применять язык SQL для работы. - владеет навыками использования технологии анализа данных.
хорошо	обучающийся демонстрирует: - современных методов и программных средств выявления тенденций изменения признаков, - умения анализировать и правильно интерпретировать, применять инструментарий программы, использовать интегрированную среду разработки, применять язык SQL для работы, - владеет навыками использования технологии анализа данных
удовлетворительно	обучающийся демонстрирует: - знания последовательности решения задания, не всех используемых современных методов и программных средств выявления тенденций изменения признаков. - умения применять инструментарий программы, проводить расчеты, не приводящие к правильному числовому ответу. - владеет навыками современных методов и программных средств выявления тенденций изменения признаков, не может самостоятельно составить выводов по результатам решения задачи.
неудовлетворительно	обучающийся: - не знает современных методов и программных средств выявления тенденций изменения признаков - не умеет применять инструментарий программы, проводить расчеты,

4.2.4. Критерии оценки выполнения тестовых заданий

При выполнении тестовых заданий обучающийся демонстрирует:

знания: основ функционирования информационно аналитической платформы программ, инструментарий осуществления аналитической деятельности на основе Статистических и эконометрических методов.

Критерии оценки выполнения тестовых заданий

отлично	обучающийся демонстрирует: – 85 % правильных ответов
хорошо	обучающийся демонстрирует: – 60 % правильных ответов
удовлетворительно	обучающийся демонстрирует: – 50 % правильных ответов
неудовлетворительно	обучающийся: – Дал менее 45 % правильных ответов

4.2.5. Критерии оценки сообщения

При устном сообщении обучающийся демонстрирует:

знания: о статистическом анализе, о машинном обучении, о типах данных и базовых алгоритмических конструкциях, языках для работы с массивами и строками, базовых понятиях ООП.

умения: проводить сравнительный анализ методов и инструментальных средств анализа больших данных; планировать аналитические работы, использовать современные информационные технологии аналитики для информационно-аналитического сопровождения деятельности.

владение навыками: анализом решений с точки зрения достижения целевых показателей решений.

Критерии оценки сообщения

отлично	обучающийся демонстрирует: – высокий уровень знаний в области информационных технологий в части машинного обучения и анализа данных, тема при выполнении сообщения раскрыта полностью; – умение проводить сравнительный анализ методов и инструментальных средств анализа данных. – владеет навыками поиска современных средств обработки информации.
Хорошо	обучающийся демонстрирует: – обучающийся показывает хороший уровень в области информационных технологий в части машинного обучения и анализа данных, тема при выполнении сообщения раскрыта полностью, но содержит неточности; – умение представлять информационные технологии и программные средства для анализа данных; – владеет навыками использования специальной терминологии
Удовлетворительно	обучающийся демонстрирует: – обучающийся показывает средний уровень знаний по теме сообщения, тема раскрыта на 50 % – умение представлять информационных технологий и программных средств для анализа данных

	– владеет навыками малой части использования специальных терминов.
Неудовлетворительно	обучающийся: – обучающийся показывает низкий уровень знаний по теме сообщения, тема при выполнении сообщения не раскрыта, содержит недостоверную информацию, отсутствует специальная терминология

Разработчик: доцент, Шибайкин В.А.



(подпись)