Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 23.05.2024 11:29:32
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742735c18b1d6

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение высшего образования
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет информационных технологий

УТВЕРЖДЕНО
Декан факультета
Информационных технологий

_____/ Д.Г. Демидов /

«15» февраля 2024 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

АНАЛИЗ ДАННЫХ

Направление подготовки/специальность
09.03.03 Прикладная информатика

Профиль/специализация
Большие и открытые данные

Квалификация
бакалавр

Формы обучения
очная

Москва, 2024 г.

Разработчик:

К.п.н., доцент / Н.И. Царькова /

Согласовано:

Заведующий кафедрой «Прикладная информатика»,
К.э.н., профессор / С.В. Суворов /

Содержание

Оглавление

1 Цели, задачи и планируемые результаты обучения по дисциплине 4
2 Место дисциплины в структуре образовательной программы 4
3 Структура и содержание дисциплины 5
3.1 Виды учебной работы и трудоемкость 5
3.2 Тематический план изучения дисциплины 5
3.3 Содержание дисциплины 6
3.4 Тематика курсовых проектов (курсовых работ) 12
4 Учебно-методическое и информационное обеспечение 12
4.1 Нормативные документы и ГОСТы 12
4.2 Основная литература 12
4.3 Дополнительная литература 13
4.5Лицензионное и свободно распространяемое программное обеспечение 13
4.6Современные профессиональные базы данных и информационные справочные системы 13
5 Материально-техническое обеспечение 13
5.1 Требования к оборудованию и помещению для занятий 13
5.2 Требования к программному обеспечению 13
6 Методические рекомендации 13
6.1 Методические рекомендации для преподавателя по организации обучения 13
6.2 Методические указания для обучающихся по освоению дисциплины 14
7 Фонд оценочных средств 14
7.1 Методы контроля и оценивания результатов обучения 14
7.2 Шкала и критерии оценивания результатов обучения 14
7.3 Оценочные средства 17

1 Цели, задачи и планируемые результаты обучения по дисциплине

К основным целям освоения дисциплины относится:

- формирование теоретических знаний о современных принципах, методах и средствах анализа данных, практических умений и навыков по применению современных методов анализа данных в различных сферах человеческой деятельности;
- выявление, анализ и прогнозирование поведения объектов и процессов с целью управления объекта и принятия решения;
- закрепление получаемых в семестре знаний и навыков на практике;
- формирование взаимосвязей, получаемых в семестре знаний и навыков с изученными ранее и изучаемых параллельно с данной дисциплиной;
- подготовка студентов к деятельности в соответствии с квалификационной характеристикой бакалавра.

К основным задачам дисциплины относятся:

- изучение существующих технологий подготовки данных к исследованию и овладение практическими умениями и навыками реализации технологий анализа данных;
- формирование и проверка гипотез о природе и структуре данных;
- обнаружение и анализ закономерностей для построения моделей с целью прогнозирования социально-экономических явлений;
- изучение и освоение теоретического материала, как в процессе контактной, так и в ходе самостоятельной работы;
- выполнение предоставленных практических заданий различных форм, как в процессе контактной, так и в ходе самостоятельной работы;
- самостоятельная работа над тематикой дисциплины для формирования компетенций основной образовательной программы (далее, ООП).

Обучение по дисциплине «Анализ данных» направлено на формирование у обучающихся следующих компетенций:

Код и наименование компетенций	Индикаторы достижения компетенции
УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач.	**Знать**: Принципы сбора, отбора и обобщения информации. Методики системного подхода для решения профессиональных задач. **Уметь**: Анализировать и систематизировать разнородные данные. Оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности. **Владеть**: Навыками научного поиска и практической работы с информационными источниками. Методами принятия решений.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к части (части, формируемой участниками образовательных отношений) блока Б1, а именно, Б1.2.2.6 Дисциплина Модуля «Математические методы обработки данных», Б1.2.2.2 «Анализ данных».

Дисциплина взаимосвязана логически и содержательно-методически со следующими дисциплинами и практиками ООП:

- Проектная деятельность;
- Математический анализ;
- Линейная алгебра;
- Эконометрические методы исследования;
- Теория вероятностей;
- Математическая статистика;
- Методы прогнозирования;
- Интеллектуальный анализ данных.

3 Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет _5_ зачетных(е) единиц(ы) (_180_ часов).

3.1 Виды учебной работы и трудоемкость (по формам обучения)

3.1.1 Очная форма обучения

№ п/п	Вид учебной работы	Количество часов	Семестры	
			3	
1	**Аудиторные занятия**	**72**		
	В том числе:			
1.1	Лекции		36	
1.2	Семинарские/практические занятия			
1.3	Лабораторные занятия		36	
2	**Самостоятельная работа**	**108**	108	
3	**Промежуточная аттестация**			
	Зачет/диф.зачет/экзамен		зачет	
	Итого:	**180**		

3.2 Тематический план изучения дисциплины (по формам обучения)

3.2.1 Очная форма обучения

№ п/п	Разделы/темы дисциплины	Трудоемкость, час					
		Всего	Аудиторная работа				Самостоятельная работа
			Лекции	Семинарские/ практические занятия	Лабораторные занятия	Практическая подготовка	
1	Раздел 1.	180	36		36		108
1.1	Тема 1. *Введение в эконометрические методы исследования*	6	2				8
1.2	Тема 2. *Распределение вероятностей.*	6	2				8
1.3	Тема 3. Статистика выводов	12	4				12

1.4	Тема 4. Анализ табличных данных	36	6		12		20
1.5	Тема 5. Основы регрессионного анализа	24	4		8		16
1.6	Тема 6. Корреляция	6	2				4
1.7	Тема 7. Аппарат множественной регрессии	24	6		6		16
1.8	Тема 8. Дисперсионный анализ	28	4		10		16
1.9	Тема 9.Кластерный анализ	12	6				8
Итого		**180**	**36**		**36**		**108**

3.3 Содержание дисциплины

Раздел 1.

Тема 1. *Введение в анализ данных*

- Роль методов анализа данных в научно-исследовательской и практической деятельности
- Основные отличия классических и современных методов анализа данных.
- Основные особенности методов классической математической статистики.
- Основные типы статистических задач.
- Типы данных
- Матрицу данных.
- Пример пространственной выборки.
- Различные классификации показателей.
- Основные свойства шкалы измерения.
- Основные шкалы измерения.

Тема 2. Распределение вероятностей.

- Совокупность данных
- Генеральная совокупность
- Выборочная совокупность
- Основные виды статистических оценок.
- Закон распределения случайной величины.
- Закон распределения дискретной случайной величины
- Основные свойства точечных оценок.

Тема 3. Статистика выводов

- Основные числовые характеристики дискретных случайных величин:
- Математическое ожидание, дисперсия, стандартное отклонение, ковариация и
- Свойства математического ожидания, дисперсии,
- Ковариации и коэффициента корреляции.
- Математическое ожидание функции от дискретной случайной величины..

Тема 4. Анализ табличных данных

- Данные в экономике. Объекты, признаки, таблицы.
- Гистограмма, как способ визуализации данных.
- Условное форматирование, его типы и правило
- . Графики и диаграммы рассеяния.
- Инструменты описательной статистики
- Сводные таблицы и сводные диаграммы для визуализации качественных признаков.
- Предварительная обработка данных.

Тема 5. Основы регрессионного анализа

- Линейная регрессия.

- Квадратичная функция потерь и предположение о нормальном распределении шума.
- Метод наименьших квадратов: аналитическое решение и оптимизационный подход.
- Линейная классификация. Аппроксимация дискретной функции потерь.
- Логистическая регрессия:

Тема 6. Корреляция

- Методы описательной статистики?
- Корреляция и ее свойства.
- Коэффициент корреляции и его свойства.
- Что такое регрессия?
- Суть метода наименьших квадратов.
- Основные характеристики качества регрессионной модели

Тема 6. Аппарат множественной регрессии

- Случайное отклонение в регрессионной модели.
- Спецификация модели. Выборочный коэффициент корреляции между переменными уравнения регрессии.
- Проверка качества уравнения регрессии.
- Коэффициент детерминации R2.
- Условия оптимальности метода наименьших квадратов (теорема Гаусса-Маркова).
- Интервальные оценки коэффициентов линейного уравнения регрессии.
- Доверительные интервалы для зависимой переменной.
- Гетероскедастичность. Автокорреляция. Мультиколлинеарность.
- Множественная линейная регрессия. Стандартизированная форма множественной регрессии.
- Проверка предпосылок метода наименьших квадратов.
- Статистика Дарбина-Уотсона.

Тема 7. Дисперсионный анализ

- Модель однофакторного дисперсионного анализа
- Факторы.
- Результативный признак.
- Общая вариация.
- Межгрупповая вариация.
- Детерминированная модель двухфакторного дисперсионного анализа с повторениями.

Тема 8. Кластерный анализ

- Методы кластерного анализа.
- Иерархические методы
- Иерархические агломеративные методы
- Итеративные методы
- Проверка результатов кластеризации

3.4.1 Семинарские/практические занятия

Указываются темы занятий.

3.4.2 Лабораторные занятия

Лабораторная работа 1 по теме № _4__

- Визуализация данных с помощью диаграмм
- Построение диаграмм математических функций
- Постройте окружность $x^2 + y^2 = 4$.

- Постройте поверхность z=Sin (x) *Cos (y), x изменяется в диапазоне от -3 до 0 с шагом 0,15 и y изменяется в диапазоне от 2 до 5 с шагом 0,15.
- Визуализация данных в MS Excel

Лабораторная работа 2 по теме №__4_
Сводные таблицы и среды для описания данных

- . Компания Contoso, Ltd. производит микрочипы. При производстве было выявлено пять типов дефектов (обозначенных 1—5). Чипы производятся двумя операторами (A и B) на четырех станках (1—4). Имеется выборка данных о дефектных чипах, включающая тип дефекта, имя оператора, номер станка и день недели, в который дефект был обнаружен. На основе этих данных определите действия, которые требуется совершить для повышения качества продукции. Для расслоения дефектов с учетом типа дефекта, дня недели, применяемого станка и работавшего оператора следует использовать сводные таблицы. Возможно, потребуется разбивка данных по станкам, операторам и т. д. Предполагается, что каждый оператор и каждый станок произвели одинаковое количество продукции.
- 2. Вы, владелец ресторана быстрого питания, произвели маркетинговое исследование в целях изучения клиентов. В случайной выборке для клиентов указаны доход, пол и сколько дней в неделю они посещают рестораны быстрого питания. На основе этой информации определите, как пол и доход влияют на частоту посещения ресторанов быстрого питания.
- 3. Студенты Школы изящных искусств подали заявления на изучение или английского языка или методов научного исследования. Необходимо определить, проводит ли Школа изящных искусств политику дискриминации в отношении женщин при приеме студентов на обучение с учетом их выбора. О студентах Школы изящных искусств имеются следующие данные:• женщина (Female) или мужчина (Male);• предпочитаемая дисциплина: английский (Eng) или методы научного исследования (Sci);
- • принят? (Yes или No).При условии, что женщины обладают равной с мужчинами квалификациейв каждой из дисциплин, свидетельствуют ли эти данные о дискриминации женщин? Используйте всю доступную информацию.
- Вам поручено оценить качество медицинской помощи, оказанной пациентам с сердечными приступами в больницах Новосибирска и Владивостока. За последние месяцы получены следующие данные о пациентах:• больница (Новос или Влад);• категория риска (Высокая или Низкая). Пациенты с высокой категорией риска выживают с меньшей вероятностью, чем пациенты с низкой категорией риска;• исход для пациента (Live или Die).На основе этих данных определите, в какой из больниц качество оказания медицинской помощи пациентам с сердечными приступами выше. Подсказка: используйте все доступные данные.
- Дан ежемесячный уровень индекса Доу-Джонса для 1985—2020 гг. Отражают ли эти данные какие-либо необычные сезонные закономерности доходности акций? Подсказка: месяц (январь, февраль и т. д.) можно извлечь с помощью формулы =ТЕКСТ(A4,"MMM"), скопированной в любой столбец.
- Имеется информацию о продажах косметики. Для каждой сделки имеются следующие данные • имя продавца;• дата продажи;• продукт;• количество;• выручка;• регион.Создайте сводную таблицу, объединяющую следующую информацию:• число сделок для каждого продавца;• общую выручку по продуктам для каждого продавца;• на основе предыдущего создайте функцию, которая всегда возвращает данные о продаже помады);• общую выручку для каждого продавца с разбивкой по регионам;• общую выручку по каждому продавцу и году. (Подсказка: необходимо сгруппировать данные по годам.)

- Для 1985—1992 гг. даны ежемесячные ставки процента по облигациям, по которым деньги выплачиваются через год после покупки. Часто предполагают, что высокие ставки процента являются более подвижными. Соответствуют ли ваши данные этому утверждению? (Подсказка: в сводных таблицах можно отобразить стандартные отклонения.)
- На основе данных в примере с продовольственными товарами подготовьте диаграмму, отображающую тенденцию продаж по годам для каждого магазина.
- Для примера с продовольственными товарами создайте вычисляемое поле для расчета средней цены единицы продукции для каждого продукта.
- Для примера с продовольственными товарами создайте сводную диаграмму, суммирующие продажи каждого продукта в каждом магазине за 2018 и 2019гг.
- На основе данных создайте вычисляемые объекты, суммирующие продажи десертов (пирожные + пудинги) и фруктов (яблоки + виноград).
- Для примера с производством микрочипов создайте сводную таблицу, суммирующую ежемесячные продажи первого и третьего чипов во Франции и США.
- В примере со сводной таблицей для клиентов покажите 15 самых крупных клиентов в одной таблице и 5 самых мелких клиентов в другой таблице.
- Данные содержат сведения о продажах различных запчастей. Код для каждой запчасти начинается или с Part (для комплектующей) или с Comp (для компьютера). Создайте сводную таблицу, отображающую продажи только для комплектующих. (Подсказка: используйте фильтр по подписи.)
- На основе данных к заданию 14 просуммируйте общие продажи для комплектующих и компьютеров.
- Данные содержат сведения выборки для американцев: возраст, информацию о том, курят ли они и что именно, сигареты или сигары, а также умерли ли они в текущем году. Какие выводы можно сделать на основании этих данных?
- 17. В данных находится следующая информация о студентах, подавших заявления в магистратуру Московского политехнического университета: пол, желательная специализация, было ли заявление принято или отклонено. Созданная должным образом сводная таблица показывает, что женщин было принято меньше, чем мужчин. Существует ли в этом университете дискриминация женщин?
- 18. Данные содержат ответы по шкале от 1 до 7 на различные вопросы анкеты о качестве обучения. Создайте сводную диаграмму, отображающую долю каждой оценки (1—7) в ответе на каждый вопрос. Отфильтруйте диаграмму для любого набора из трех вопросов на выбор.
- 19. Для данных пункта 10 создайте сводную таблицу, в которой показано приведенное отличие каждой ежемесячной продажи от продажи за предыдущий месяц для 2019 г.

Лабораторная работа 3 по теме №_5__

Парная регрессия и корреляция

Цель выполнения лабораторной работы: Ознакомление с парной регрессией и корреляцией и получения уравнения парной регрессии методом наименьших квадратов

Результат: Уравнение парной регрессии

Порядок выполнения лабораторной работы:

- Подготовка к выполнению к работе, в том числе:
  - изучение парной регрессии и корреляции
- Выбор вида математической функции для парной регрессии.
- Оценка параметров парной регрессии методом наименьших квадратов.
- Проверка уравнения с помощью функции ЛИНЕЙН.
- Защита лабораторной работы.

Контрольные вопросы:

1.Дайте определение модели парной регрессии.

2. Какие виды моделей парной регрессии вы знаете?

3. Какими способами в парной регрессии может быть осуществлен выбор вида математической функции?

4. Какими методами можно найти параметры линейной парной регрессии?

5. Поясните смысл коэффициента регрессии.

6. С помощью какого коэффициента можно оценить тесноту связи между результатом и фактором?

7. Какова концепция F – критерия Фиш

ера?

8. Каким образом оценивается значимость параметров уравнения регрессии?

9. Приведите пример моделей, нелинейных относительно включаемых переменных

Лабораторная работа 4 по теме №_5__

Анализ данных _РЕГРЕССИЯ_в EXCEL

- Вычислить остатки; найти остаточную сумму квадратов; факторную сумму квадратов и общую сумму квадратов, оценить дисперсию на одну степень свободы; построить графики остатков. (Все конечные вычисления оформить в виде таблицы).
- Осуществить проверку значимости параметров уравнения регрессии с помощью t-критерия Стьюдента (а = 0,1).
- Проверить значимость уравнения регрессии с помощью F - критерия Фишера (а = 0,1). Сделать вывод о качестве модели.
- Выполнить лабораторную работу, используя опцию «анализ данных»
- Осуществить прогнозирование среднего значения показателя У при уровне значимости а = 0,05 и а = 0,1, если прогнозное значения фактора Х составит 80% от его максимального значения.
- 6.Проверить выполнение предпосылок МНК.

Контрольные вопросы для лабораторной работы № 2

1.Сформулировать основную идею дисперсионного анализа и записать в виде формулы.

2.Изобразить схему дисперсионного анализа.

3.Какие характеристики 'используются для оценки точности предсказаний по уравнению регрессии? Опишите каждую из них.

3. Сформулируйте пять предпосылок МНК.

4. Что такое дисперсия?

5. Что такое математическое ожидание?

6. Что такое ковариация?

7.Сформулировать основные свойства оценок.

8.Что такое линейный коэффициент корреляции? Его экономическая интерпретация.

9.Что такое линейный коэффициент детерминации? Его экономическая интерпретация.

10 Для чего нужны расчеты доверительных интервалов?

11. В «анализе данных» в опции «регрессия» укажите значения всех показателей.

Лабораторная работа 5 по теме №_7__

Множественная регрессия и корреляция

Цель выполнения лабораторной работы: Изучение множественной регрессии и корреляции и построения уравнения множественной регрессии.

Результат: Уравнение множественной регрессии.

Порядок выполнения лабораторной работы:

- Подготовка к выполнению к работе, в том числе:
  - изучение Множественной регрессии и корреляции.
  - изучение мультиколлинеарности факторов;
  - изучение предпосылок МНК.
- Получение варианта задания.
- Оценка параметров множественной регрессии, согласно общей схеме..
- Проверка значимости уравнения регрессии и значимости всех параметров.
- Защита лабораторной работы.

Контрольные вопросы:

1. В чем состоит спецификация модели множественной регрессии?

2. Какие требования предъявляют к факторам для включения их в модель множественной регрессии?

3. Чем вызывается явление мультиколлинеарности в многофакторных эконометрических моделях?

4. Каким образом можно устранить мультиколлинеарность факторов?

5. Раскройте экономическую интерпретацию коэффициентов чистой регрессии.

6. Чем являются коэффициенты в степенной модели множественной регрессии ? 7. Каким образом можно оценить параметры уравнения множественной регрессии?

8. Какие коэффициенты используются для оценки сравнительной силы воздействия факторов на результат?

9. Какие переменные называют фиктивными?

10. При каких условиях строится уравнение множественной регрессии с фиктивными переменными?

11. Каким образом можно проверить наличие гомо- или гетероскедастичность остатков?

12. В чем суть обобщенного метода наименьших квадратов?

Лабораторная работа 6 по теме №_8__

Однофакторный дисперсионный анализ

Задание для лабораторной работы №4

- . Для пациентов четырех кардиологов известно количество дней, в течение которых пациенты оставались в больнице после операции на отключенном сердце.
- • Существуют ли доказательства того, что врачи проводят различную политику выписки из лечебного учреждения?
- В течение скольких дней пациент первого кардиолога будет оставаться в больнице с вероятностью 95%?
- Лекарственный препарат может быть изготовлен в термостате с рабочей температурой 400, 300 или 200 градусов. Дано количество полученного препарата в фунтах для различных партий, изготовленных при разных температурах.
- • Влияет ли температура на результативность процесса?
- • Укажите интервал для количества продукта в фунтах, в котором с вероятностью 95% окажется выход продукта из термостата с рабочей температурой 200 градусов?
- • Если считать, что давление внутри корпуса также влияет на результативность процесса, остается ли анализ действительным?

Лабораторная работа 7 по теме №_8__

Двухфакторный дисперсионный анализ

- 1. Считается, что давление (высокое, среднее или низкое) и температура (высокая, средняя или низкая) влияют на продуктивность процесса. Приняв это за основу, решите следующие задачи.
- • Исходя из данных на листе 1, определите, температура и/или давление влияет на продуктивность процесса.
- • В каких пределах с вероятностью 95% находится выход продукции при высоком давлении и низкой температуре?
- . Требуется определить, как отдельный торговый агент и число коммерческих звонков врачу (один, три и пять) влияют на сумму (в тысячах у.е), на которую каждый врач выписывает лекарственный препарат. Исходя из данных на листе 2, найдите ответы на следующие вопросы:
- • Как агент и число коммерческих звонков влияют на объем продаж?
- • Если третий агент сделал пять звонков врачу, то в пределах какой суммы в долларах с вероятностью 95% врач выпишет рецепты?
- Ответьте на вопросы к заданию 2, исходя из данных на листе 2 .
- .Файл содержит информацию о продажах арахисового масла в течение нескольких недель, когда покупателям выдавался (или не выдавался) купон и когда реклама была (или не была) напечатана в воскресной газете. Опишите, как купон и реклама влияют на продажи арахисового масла (смотрите теорию выше). Постройте графики зависимости продаж.

.

3.4 Тематика курсовых проектов (курсовых работ)

Не предусмотрено

4 Учебно-методическое и информационное обеспечение

4.1 Нормативные документы и ГОСТы

1. Федеральный государственный образовательный стандарт высшего образования (уровень бакалавриата) по направлению подготовки 09.03.03 «Прикладная информатика», утвержденный приказом Министерства образования и науки Российской Федерации от 19.09.2017 №922 «Об утверждении федерального государственного стандарта высшего образования – бакалавриат по направлению подготовки 09.03.03 Прикладная информатика. https://fgos.ru/fgos/fgos-09-03-03-prikladnaya-informatika-922/

4.2 Основная литература

1. Котиков, П. Е. Анализ данных : учебно-методическое пособие / П. Е. Котиков. — Санкт-Петербург : СПбГПМУ, 2019. — 48 с. — ISBN 978-5-907184-46-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/174498 (дата обращения: 16.10.2024).

2. Кадочникова, Е. И. Статистический анализ пространственных данных : учебное пособие / Е. И. Кадочникова, Ю. А. Варламова. — Казань : КФУ, 2024. — 140 с. — ISBN 978-5-00130-700-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/332354 (дата обращения: 16.10.2024).

…

4.3 Дополнительная литература

1. Кузьмин, В. И. Методы анализа данных : учебное пособие / В. И. Кузьмин, А. Ф. Гадзаов. — 2-е изд., перераб. и доп. — Москва : РТУ МИРЭА, 2020. — 155 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/171433 (дата обращения: 16.10.2024).

2. Крутиков, В. И. Анализ данных : учебное пособие / В. И. Крутиков, В. В. Мешечкин. — Кемерово : КемГУ, 2014. — 138 с. — ISBN 978-5-8353-1770-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/61396 (дата обращения: 16.10.2024).

4.4 Электронные образовательные ресурсы

1.ЭОР «АД» https://online.mospolytech.ru/course/view.php?id=4928

4.5Лицензионное и свободно распространяемое программное обеспечение

1. Операционная система, Windows 11 (или ниже) - Microsoft Open License
2. Офисные приложения, Microsoft Office 2013(или ниже) - Microsoft Open License

4.6Современные профессиональные базы данных и информационные справочные системы

1.не предусмотрено

5 Материально-техническое обеспечение

5.1 Требования к оборудованию и помещению для занятий

Лабораторные работы и самостоятельная работа студентов должны проводиться в специализированной аудитории, оснащенной современной оргтехникой и персональными компьютерами с программным обеспечением в соответствии с тематикой изучаемого материала. Число рабочих мест в аудитории должно быть достаточным для обеспечения индивидуальной работы студентов. Рабочее место преподавателя должно быть оснащены современным компьютером с подключенным к нему проектором на настенный экран, или иным аналогичным по функциональному назначению оборудованием.

5.2 Требования к программному обеспечению

Для выполнения лабораторных работ и самостоятельной работы необходимо следующее программное обеспечение:

Microsoft Windows.

Веб-браузер, Chrome.

ПО, предоставленное преподавателем.

6 Методические рекомендации

6.1 Методические рекомендации для преподавателя по организации обучения

1. При подготовке к занятиям следует предварительно проработать материал занятия, предусмотрев его подачу точно в отведенное для этого время занятия. Следует подготовить необходимые материалы – теоретические сведения, задачи и др. При проведении занятия

следует контролировать подачу материала и решение заданий с учетом учебного времени, отведенного для занятия.

2. При проверке работ и отчетов следует учитывать не только правильность выполнения заданий, но и оптимальность выбранных методов решения, правильность выполнения всех его шагов.

6.2 Методические указания для обучающихся по освоению дисциплины

Изучение дисциплины осуществляется в строгом соответствии с целевой установкой в тесной взаимосвязи учебным планом. Основой теоретической подготовки студентов являются аудиторные занятия, семинары и практики.

В процессе самостоятельной работы студенты закрепляют и углубляют знания, полученные во время аудиторных занятий, дорабатывают конспекты и записи, готовятся к промежуточной аттестации, а также самостоятельно изучают отдельные темы учебной программы.

На занятиях студентов, в том числе предполагающих практическую деятельность, осуществляется закрепление полученных, в том числе и в процессе самостоятельной работы, знаний. Особое внимание обращается на развитие умений и навыков установления связи положений теории с профессиональной деятельностью будущего специалиста.

Самостоятельная работа осуществляется индивидуально. Контроль самостоятельной работы организуется в двух формах:

самоконтроль и самооценка студента;

контроль со стороны преподавателей (текущий и промежуточный).

Текущий контроль осуществляется на аудиторных занятиях, промежуточный контроль осуществляется на зачете в письменной (устной) форме.

Критериями оценки результатов самостоятельной работы студента являются:

уровень освоения студентом учебного материала;

умения студента использовать теоретические знания при выполнении практических задач;

сформированность компетенций;

оформление материала в соответствии с требованиями..

7 Фонд оценочных средств

7.1 Методы контроля и оценивания результатов обучения

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций:

Лабораторные работы, экзамен.

7.2 Шкала и критерии оценивания результатов обучения

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине.

Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**
УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач				

УК-1.1. Знать: Принципы сбора, отбора и обобщения информации. Методики системного подхода для решения профессиональных задач. **УК-1.2. Уметь:** Анализировать и систематизировать разнородные данные. Оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности. **УК-1.3. Владеть:** Навыками научного поиска и практической работы с информационными источниками. Методами принятия решений.	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3).	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3).	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.
ОПК-2. Способен использовать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности.				
ОПК-2.1. Знать: Современные информационные технологии и программные средства, в том числе отечественного производства. **ОПК-2.2. Уметь:** Использовать современные информационные технологии и программные	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п.	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Допускаются значительные ошибки, проявляется	Обучающийся демонстрирует частичное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Но допускаются незначительные ошибки,	Обучающийся демонстрирует полное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Свободно оперирует приобретенными знаниями.

средства. **ОПК-2.3. Владеть:** Навыками использования современных информационных технологий и программных средств для решения профессиональных задач.	3).	недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	неточности, затруднения при аналитических операциях.	

Шкала оценивания результатов промежуточной аттестации и её описание:

Форма промежуточной аттестации: экзамен.

Промежуточная аттестация обучающихся в форме экзамена проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине, при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине проводится преподавателем, ведущим занятия по дисциплине методом экспертной оценки. По итогам промежуточной аттестации по дисциплине выставляется оценка «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно».

Шкала оценивания	**Описание**
Отлично	Выполнены все **обязательные условия подготовки студента к промежуточной аттестации**, предусмотренные программой дисциплины. Студент демонстрирует соответствие знаний, умений, навыков показателям, приведенным в таблицах, оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности. При этом могут быть допущены незначительные ошибки, неточности, затруднения при переносе знаний и умений на новые, нестандартные задачи.
Хорошо	Выполнены все **обязательные условия подготовки студента к промежуточной аттестации**, предусмотренные программой дисциплины. Студент демонстрирует соответствие знаний, умений, навыков показателям, приведенным в таблицах, оперирует приобретенными знаниями, умениями, навыками. В то же время при ответе допускает несущественные погрешности, задачи решает с недочетами, не влияющими на общий ход решения.
Удовлетворительно	Выполнены все **обязательные условия подготовки**

	студента к промежуточной аттестации, предусмотренные программой дисциплины. Студент демонстрирует соответствие знаний, умений, навыков показателям, приведенным в таблицах, оперирует приобретенными знаниями, умениями, навыками. Но показывает неглубокие знания, при ответе не допускает грубых ошибок или противоречий, однако в формулировании ответа отсутствует должная связь между анализом, аргументацией и выводами, в решении задач могут содержаться грубые ошибки. Для получения правильного ответа требуются уточняющие вопросы.
Неудовлетворительно	Не выполнены обязательные **условия подготовки студента к промежуточной аттестации**, предусмотренные программой дисциплины, ИЛИ студент демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах, допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент испытывает значительные затруднения при оперировании знаниями и умениями.

7.3 Оценочные средства

7.3.1 Текущий контроль

Защита Лабораторных работ

7.3.2 Промежуточная аттестация

Зачет

Вопросы для подготовки к зачету
по дисциплине «АНАЛИЗ ДАННЫХ»

1. *Дайте определение понятию «данные».*
2. *Назовите типы шкал с примерами?*
3. *Чем дискретные данные отличаются от непрерывных?*
4. *Что такое анализ данных?*
5. *Назовите этапы анализа данных.*
6. *Чем качественный этап анализа данных отличается от количественного?*
7. *Опишите основе задачи при качественном анализе данных.*
8. *. Дайте определение понятию «вероятность».*
9. *Чем относительная частота отличается от вероятности?*
10. *Назовите основные свойства распределения вероятности?*
11. *Чему равна общая площадь под кривой функции плотности вероятности?*
12. *Что такое случайная переменная?*
13. *Дайте определение понятию «нормальная метка»?*
14. *Какие действие позволяет произвести формула Байеса.*
15. *Основные компоненты статистки выводов.*
16. *Что такое доверительный интервал?*
17. *Из чего состоит алгоритм проверки гипотез?*
18. *Что такое нулевая гипотеза?*

19. *Перечислите типы ошибок?*
20. *Виды непараметрического теста?*
21. *Перечислите наиболее общие правила при использовании статистики выводов?*
22. *Что такое сводная таблица?*
23. *Что такое степень свободы?*
24. *Как избавиться от редких данных?*
25. *Что мера ассоциации?*
26. *Перечислите типы ошибок?*
27. *Дайте определение понятию «регрессия».*
28. *Назовите причины использования регрессионного анализа.*
29. *На основе какого метода вычисляются коэффициенты уравнения регрессии?*
30. *Что такое величина достоверности аппроксимации?*
31. *Из чего состоит проверка достоверности модели регрессии?*
32. *Дайте определение понятию «корреляция».*
33. *Что означает корреляция, равная нулю?*
34. *Приведите пример положительной и отрицательной корреляции.*
35. *Что такое матрица корреляции?*
36. *Как читается матрица точечных диаграмм корреляции?*
37. *В чем особенность множественной регрессии?*
38. *Почему аппарат множественной регрессии может быть реализован в экспертной системе?*
39. *Как проверить результаты множественной линейной регрессии?*
40. *Что такое пошаговая регрессия?*
41. *Какие правила остановки в пошаговой регрессии?*
42. *Что такое логит модель?*
43. *Что такое дисперсионный анализ?*
44. *Назовите назначение переменных - индикаторов?*
45. *Какая связь между регрессионным и дисперсионным анализом?*
46. *Основные компоненты дисперсионного анализа?*
47. *Охарактеризуйте результат дисперсионного анализа?*
48. *Кластерный анализ*
49. *Терминология кластерного анализа.*
50. *Расстояние между объектами (метрика).*
51. *Плотность и локальность кластеров.*
52. *Расстояние между кластерами.*
53. *Методы группировки.*
54. *Особенности иерархических агломеративных методов.*
55. *Особенности итерационных методов кластеризации.*
56. *Эвристические графовые алгоритмы.*
57. *Статистические алгоритмы.*
58. *Кластеризация признаков.*
59. *Устойчивость и качество кластеризации.*