

**Министерство науки и высшего образования РФ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Национальный исследовательский университет «МЭИ»**

Направление подготовки/специальность: 38.03.05 Бизнес-информатика

Наименование образовательной программы: Информационное и программное обеспечение бизнес-процессов

Уровень образования: высшее образование - бакалавриат

Форма обучения: Очная

**Оценочные материалы
по дисциплине
Анализ больших данных**

**Москва
2024**

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ РАЗРАБОТАЛ:

Разработчик

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Раскатова М.В.
	Идентификатор	R6bc62db2-RaskatovaMV-ead4381

М.В.
Раскатова

СОГЛАСОВАНО:

Руководитель
образовательной
программы

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Крепков И.М.
	Идентификатор	R04da5bdb-KrepkovIM-33fe3095

И.М.
Крепков

Заведующий
выпускающей кафедрой

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Невский А.Ю.
	Идентификатор	R4bc65573-NeVskyAY-0b6e493d

А.Ю.
Невский

ОБЩАЯ ЧАСТЬ

Оценочные материалы по дисциплине предназначены для оценки: достижения обучающимися запланированных результатов обучения по дисциплине, этапа формирования запланированных компетенций и уровня освоения дисциплины.

Оценочные материалы по дисциплине включают оценочные средства для проведения мероприятий текущего контроля успеваемости и промежуточной аттестации.

Формируемые у обучающегося компетенции:

1. РПК-1 Способен проводить обследование организаций, выявлять информационные потребности пользователей, формировать требования к информационной системе

ИД-1 Знает соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации

ИД-2 Умеет собирать, систематизировать, документировать и анализировать требования к информационным системам

ИД-3 Владеет методами системного анализа и моделирования для анализа архитектуры предприятий и методами сбора информации для формализации требований пользователей заказчика

и включает:

для текущего контроля успеваемости:

Форма реализации: Компьютерное задание

1. Анализ данных в MS Excel. Контрольная работа (Контрольная работа)

2. Визуальный анализ данных (Контрольная работа)

3. Деревья решений (Контрольная работа)

4. Кластерный анализ (Контрольная работа)

5. Прогнозирование экономических явлений (Контрольная работа)

БРС дисциплины

6 семестр

Раздел дисциплины	Веса контрольных мероприятий, %					
	Индекс КМ:	КМ-1	КМ-2	КМ-3	КМ-4	КМ-5
	Срок КМ:	4	8	10	12	14
Понятие больших данных. Задачи анализа данных						
Понятие больших данных. Задачи анализа данных	+	+				
Процесс Data Mining. Сферы применения						
Процесс Data Mining. Сферы применения	+					
Классификация, кластеризация. Методы кластерного анализа						
Классификация, кластеризация. Методы кластерного анализа			+			
Прогнозирование. Методы прогнозирования						

Прогнозирование. Методы прогнозирования			+		
Деревья решений. Ассоциативные правила					
Деревья решений. Ассоциативные правила				+	
Визуальный анализ данных					
Визуальный анализ данных					+
Вес КМ:	20	20	20	20	20

\$Общая часть/Для промежуточной аттестации\$

СОДЕРЖАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ТЕКУЩЕГО КОНТРОЛЯ

I. Оценочные средства для оценки запланированных результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Индекс компетенции	Индикатор	Запланированные результаты обучения по дисциплине	Контрольная точка
РПК-1	ИД-1 _{РПК-1} Знает соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации	Знать: математический аппарат и инструментальные средства для обработки и анализа данных математический аппарат, методы и основные алгоритмы для решения задач классификации и кластеризации Уметь: применять математический аппарат и инструментальные средства для обработки и анализа данных применять математический аппарат и инструментальные средства для решения задач классификации и кластеризации	Кластерный анализ (Контрольная работа) Прогнозирование экономических явлений (Контрольная работа)
РПК-1	ИД-2 _{РПК-1} Умеет собирать, систематизировать, документировать и	Знать: методы сбора и обработки данных, систематизации,	Анализ данных в MS Excel. Контрольная работа (Контрольная работа) Кластерный анализ (Контрольная работа) Прогнозирование экономических явлений (Контрольная работа)

	анализировать требования к информационным системам	документирования методы прогнозирования, сбора и систематизации данных Уметь: осуществлять сбор и обработку данных, систематизацию, документирование решать задачи прогнозирования с использованием современных методов прогнозирования	Деревья решений (Контрольная работа)
РПК-1	ИД-ЗРПК-1 Владеет методами системного анализа и моделирования для анализа архитектуры предприятий и методами сбора информации для формализации требований пользователей заказчика	Знать: методы системного анализа и моделирования для поиска ассоциативных правил при решении профессиональных задач способы решения поставленных задач с использованием визуального анализа данных Уметь: применять визуальный анализ данных при решении поставленных задач строить деревья решений, осуществлять поиск ассоциативных правил при	Анализ данных в MS Excel. Контрольная работа (Контрольная работа) Деревья решений (Контрольная работа) Визуальный анализ данных (Контрольная работа)

		решении профессиональных задач	
--	--	-----------------------------------	--

II. Содержание оценочных средств. Шкала и критерии оценивания

КМ-1. Анализ данных в MS Excel. Контрольная работа

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Контрольная работа

Вес контрольного мероприятия в БРС: 20

Процедура проведения контрольного мероприятия: Выполняется задание согласно варианту. По результатам выполнения задания составляется отчет. При защите работы задаются вопросы, ответы на которые оцениваются

Краткое содержание задания:

Выполнить задание по подбору параметров и разработки сценария, используя инструменты Анализа данных среды Microsoft Excel

Контрольные вопросы/задания:

Знать: методы системного анализа и моделирования для поиска ассоциативных правил при решении профессиональных задач	1.Что такое сценарий в Excel 2.Назовите инструменты анализа данных в Excel
Уметь: осуществлять сбор и обработку данных, систематизацию, документирование	1.Найдите максимальную прибыль, используя алгоритм поиска решений 2.Используя различные алгоритмы поиска решений, выполните поставленную задачу

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 90

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно

Оценка: 4

Нижний порог выполнения задания в процентах: 80

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если большинство вопросов раскрыто. выбрано верное направление для решения задач

Оценка: 3

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено

Оценка: 2

Описание характеристики выполнения знания: Оценка "неудовлетворительно" выставляется если задание выполнено неверно или преимущественно не выполнено

КМ-2. Кластерный анализ

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Контрольная работа

Вес контрольного мероприятия в БРС: 20

Процедура проведения контрольного мероприятия: Выполняется задание согласно варианту. По результатам выполнения задания составляется отчет. При защите работы задаются вопросы, ответы на которые оцениваются

Краткое содержание задания:

Провести классификацию объектов, каждый из которых характеризуется двумя признаками

Контрольные вопросы/задания:

Знать: математический аппарат, методы и основные алгоритмы для решения задач классификации и кластеризации	1.Опишите метод “ближайшего соседа” разбиения объектов на кластеры 2.Опишите метод “дальнего соседа” разбиения объектов на кластеры
Знать: методы прогнозирования, сбора и систематизации данных	1.Что такое кластерный анализ 2.Назовите методы кластерного анализа данных
Уметь: применять математический аппарат и инструментальные средства для решения задач классификации и кластеризации	1.Вычислите расстояние между кластерами, используя Евклидово расстояние 2.Разбейте объекты на кластеры, используя метод “ближайшего соседа” 3.Разбейте объекты на кластеры, используя метод “дальнего соседа”

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 90

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно

Оценка: 4

Нижний порог выполнения задания в процентах: 80

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если большинство вопросов раскрыто. выбрано верное направление для решения задач

Оценка: 3

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено

Оценка: 2

Описание характеристики выполнения знания: Оценка "неудовлетворительно" выставляется если задание выполнено неверно или преимущественно не выполнено

КМ-3. Прогнозирование экономических явлений

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Контрольная работа

Вес контрольного мероприятия в БРС: 20

Процедура проведения контрольного мероприятия: Выполняется задание согласно варианту. По результатам выполнения задания составляется отчет. При защите работы задаются вопросы, ответы на которые оцениваются

Краткое содержание задания:

Решить задачу прогнозирования по линейной регрессионной модели

Контрольные вопросы/задания:

Знать: математический аппарат и инструментальные средства для обработки и анализа данных	<ol style="list-style-type: none"> 1. Дайте определение прогнозу, назовите его виды 2. Назовите основные составляющие временного ряда 3. Назовите виды прогнозов 4. С какой задачей сходна задача прогнозирования 5. Чем отличаются краткосрочный и долгосрочный прогнозы
Уметь: применять математический аппарат и инструментальные средства для обработки и анализа данных	<ol style="list-style-type: none"> 1. Постройте график уравнения линейной регрессии 2. Выполните прогнозирование различными функциями Excel
Уметь: решать задачи прогнозирования с использованием современных методов прогнозирования	<ol style="list-style-type: none"> 1. Найти коэффициенты уравнения регрессии, используя функции Excel 2. Проведите анализ полученных результатов

Описание шкалы оценивания:*Оценка: 5**Нижний порог выполнения задания в процентах: 90**Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно**Оценка: 4**Нижний порог выполнения задания в процентах: 80**Описание характеристики выполнения знания: Оценка "хорошо" выставляется если большинство вопросов раскрыто. выбрано верное направление для решения задач**Оценка: 3**Нижний порог выполнения задания в процентах: 65**Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено**Оценка: 2**Описание характеристики выполнения знания: Оценка "неудовлетворительно" выставляется если задание выполнено неверно или преимущественно не выполнено***КМ-4. Деревья решений****Формы реализации:** Компьютерное задание**Тип контрольного мероприятия:** Контрольная работа**Вес контрольного мероприятия в БРС:** 20**Процедура проведения контрольного мероприятия:** Выполняется задание согласно варианту. По результатам выполнения задания составляется отчет. При защите работы задаются вопросы, ответы на которые оцениваются**Краткое содержание задания:**

Решить задачу классификации, используя деревья решений. Описать предметную область и построить дерево решений

Контрольные вопросы/задания:

Знать: методы сбора и обработки данных, систематизации, документирования	<ol style="list-style-type: none"> 1. Опишите процесс конструирования дерева решений 2. Как называется конструкция вида «если : то : » при использовании деревьев решений
Уметь: строить деревья решений,	<ol style="list-style-type: none"> 1. Построить бинарное дерево решений

осуществлять поиск ассоциативных правил при решении профессиональных задач	2. Напишите правила для заданного дерева решений 3. Укажите преимущества использования деревьев решений 4. Объясните основные алгоритмы, реализующие деревья решений
--	--

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 90

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно

Оценка: 4

Нижний порог выполнения задания в процентах: 80

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если большинство вопросов раскрыто. выбрано верное направление для решения задач

Оценка: 3

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено

Оценка: 2

Описание характеристики выполнения знания: Оценка "неудовлетворительно" выставляется если задание выполнено неверно или преимущественно не выполнено

КМ-5. Визуальный анализ данных

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Контрольная работа

Вес контрольного мероприятия в БРС: 20

Процедура проведения контрольного мероприятия: Выполняется задание согласно варианту. По результатам выполнения задания составляется отчет. При защите работы задаются вопросы, ответы на которые оцениваются

Краткое содержание задания:

Построить заданные графики и поверхности в пакете MathCad

Контрольные вопросы/задания:

Знать: способы решения поставленных задач с использованием визуального анализа данных	1. Что такое лица Чернова, для чего они применяются 2. С какой задачей анализа данных сходна задача визуализации? 3. Назовите методы визуального представления информации
Уметь: применять визуальный анализ данных при решении поставленных задач	1. Проведите форматирование графика 2. Постройте график функции в трехмерном пространстве 3. Постройте график поверхности, заданной параметрически

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 90

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно

Оценка: 4

Нижний порог выполнения задания в процентах: 80

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если большинство вопросов раскрыто. выбрано верное направление для решения задач

Оценка: 3

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено

Оценка: 2

Описание характеристики выполнения знания: Оценка "неудовлетворительно" выставляется если задание выполнено неверно или преимущественно не выполнено

СОДЕРЖАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

6 семестр

Форма промежуточной аттестации: Зачет с оценкой

Пример билета

1. Основные понятия анализа данных.
2. Практическое задание: построить график функции в трехмерном пространстве $g(x,y)=x^2-y^2$

Процедура проведения

Зачет проводится по совокупности результатов всех контрольных мероприятий

1. Перечень компетенций/индикаторов и контрольных вопросов проверки результатов освоения дисциплины

1. Компетенция/Индикатор: ИД-1_{РПК-1} Знает соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации

Вопросы, задания

- 1.1. Основные этапы процесса Data Mining
2. Практическое задание на тему: интерполирование функции
- 2.1. Очистка данных, этапы и инструменты очистки данных
2. Практическое задание на тему: прогнозирование
- 3.1. Этапы процесса классификации: конструирование модели и ее использование
2. Практическое задание на тему: построение графика функции в трехмерном пространстве
- 4.1. Кластеризация. Постановка задачи. Процесс кластеризации
2. Практическое задание на тему: анализ данных в Excel

Материалы для проверки остаточных знаний

1. Алгоритм кластеризации k-средних строит k кластеров, расположенных:

Ответы:

1. на возможно меньших расстояниях друг от друга
2. на возможно больших расстояниях друг от друга
3. на равных расстояниях друг от друга
4. на произвольных расстояниях друг от друга
5. на пропорциональных расстояниях друг от друга

Верный ответ: 2

2. Методы кластеризации, в которых происходит последовательное объединение исходных элементов и уменьшение числа кластеров - это

Ответы:

1. иерархические дивизимные методы
2. иерархические делимые методы
3. иерархические агломеративные методы
4. неиерархические методы
5. метод k - средних

Верный ответ: 3

3. Какие данные не являются «грязными»:

Ответы:

1. выбросы
2. дубликаты
3. пропущенные значения
4. шумы
5. нулевые значения

Верный ответ: 5

4. Что не входит в этапы очистки данных процесса Data Mining:

Ответы:

1. анализ данных
2. подтверждение данных
3. преобразования данных
4. определение необходимого количества данных
5. противоток данных

Верный ответ: 4

2. Компетенция/Индикатор: ИД-2_{РПК-1} Умеет собирать, систематизировать, документировать и анализировать требования к информационным системам

Вопросы, задания

- 1.1. Понятие данных, классификация видов данных, метаданные. Измерения, шкалы, типы шкал
2. Практическое задание на тему: построение дерева решений
 - 2.1. Задачи анализа данных
 2. Практическое задание на тему: анализ данных в Excel
- 3.1. Задача прогнозирования. Сравнение задач прогнозирования и классификации
 2. Практическое задание на тему: кластерный анализ
- 4.1. Методы прогнозирования. Решение задачи прогнозирования
 2. Практическое задание на тему: дерево решений

Материалы для проверки остаточных знаний

1. Для какой шкалы применимы только такие операции как равно и не равно

Ответы:

1. интервальная шкала
2. порядковая шкала
3. номинальная шкала
4. относительная шкала
5. дихотомическая шкала

Верный ответ: 3

2. Data Mining - это процесс обнаружения в сырых данных знаний, необходимых для

Ответы:

1. замены аналитика в процессе принятия решений
2. принятия решений в различных сферах человеческой деятельности
3. увеличения стоимости анализа данных
4. уменьшения стоимости анализа данных
5. замены эксперта в процессе принятия решений

Верный ответ: 2

3. Задача прогнозирования сходна с задачей:

Ответы:

1. кластеризации
2. классификации
3. ассоциации
4. визуализации
5. оценивания

Верный ответ: 2

4. Основная единица времени, на которую делается прогноз, называется:

Ответы:

1. период прогнозирования
2. интервал прогнозирования
3. горизонт прогнозирования
4. время прогнозирования
5. дата прогнозирования

Верный ответ: 1

3. Компетенция/Индикатор: ИД-ЗрПК-1 Владеет методами системного анализа и моделирования для анализа архитектуры предприятий и методами сбора информации для формализации требований пользователей заказчика

Вопросы, задания

- 1.1. Метод деревьев решений. Преимущества деревьев решений
2. Практическое задание на тему: аппроксимация функции
- 2.1. Алгоритмы, реализующие деревья решений: CART, ID3, C4.5, CHAID
2. Практическое задание на тему: построение графика поверхности, заданной параметрически
- 3.1. Визуализация данных. Постановка задачи. Основные этапы визуального анализа данных
2. Практическое задание на тему: кластерный анализ
- 4.1. Методы и средства визуального представления информации, характеристики средств визуализации
2. Практическое задание на тему: прогнозирование экономических явлений

Материалы для проверки остаточных знаний

1. График отклонения прогнозируемых при помощи модели значений от реальных - это:

Ответы:

1. таблица сопряженности
2. график изменения величины ошибки
3. линия тренда
4. диаграмма рассеивания
5. дендрограммы

Верный ответ: 4

2. Для отображения соотношения частей и целого применяют:

Ответы:

1. линейный график
2. гистограмму
3. круговую диаграмму
4. векторную диаграмму
5. секторную диаграмму

Верный ответ: 3

3. Целью поиска ассоциативных правил является:

Ответы:

1. нахождение закономерностей между несвязанными событиями в БД
2. нахождение закономерностей между объектами в кластерах
3. нахождение закономерностей между связанными событиями в БД
4. нахождение закономерностей между объектами в классах
5. объективных закономерностей в предметной области

Верный ответ: 3

4. Полезнее ли правило случайного угадывания показывает величина:

Ответы:

1. понятность
2. достоверность
3. полезность
4. улучшение
5. поддержка

Верный ответ: 4

II. Описание шкалы оценивания

Оценка: 5

Нижний порог выполнения задания в процентах: 70

Описание характеристики выполнения задания: Работа выполнена в рамках "продвинутого" уровня. Ответы даны верно, четко сформулированные особенности практических решений

Оценка: 4

Нижний порог выполнения задания в процентах: 60

Описание характеристики выполнения задания: Работа выполнена в рамках "базового" уровня. Большинство ответов даны верно. В части материала есть незначительные недостатки

Оценка: 3

Нижний порог выполнения задания в процентах: 50

Описание характеристики выполнения задания: Работа выполнена в рамках "порогового" уровня. Основная часть задания выполнена верно. на вопросы углубленного уровня

Оценка: 2

Описание характеристики выполнения задания: Работа не выполнена или выполнена преимущественно неправильно

III. Правила выставления итоговой оценки по курсу

Оценка определяется в соответствии с Положением о балльно-рейтинговой системе для студентов НИУ «МЭИ» на основании семестровой и аттестационной составляющих