

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ИМЕНИ В.Ф. УТКИНА»

Кафедра «Экономика, менеджмент и организация производства»

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ

Б1.О.30 «Оперативный и интеллектуальный анализ данных»

Направление подготовки
27.03.05 Инноватика

Направленность (профиль) подготовки
«Технологическое предпринимательство»

Квалификация выпускника – бакалавр

1. ОБЩИЕ ПОЛОЖЕНИЯ

Оценочные материалы – это совокупность учебно-методических материалов (контрольных заданий, описаний форм и процедур), предназначенных для оценки качества освоения обучающимися данной дисциплины как части ОПОП ВО.

Цель – оценить соответствие знаний, умений и уровня приобретенных компетенций обучающихся целям и требованиям основной образовательной программы в ходе проведения промежуточной аттестации.

Промежуточная аттестация проводится в форме зачета, который включает тестирование и выполнение практико-ориентированного задания. При необходимости, проводится теоретическая беседа с обучаемым для уточнения оценки. Выполнение заданий в течение семестра (включая все виды работ, предусмотренные рабочей программой дисциплины, в том числе лабораторные / практические / контрольные / самостоятельные) является обязательным условием для допуска к промежуточной аттестации.

2. ПАСПОРТ ОЦЕНОЧНЫХ МАТЕРИАЛОВ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Контролируемые разделы (темы) дисциплины (результаты по разделам)	Код контролируемой компетенции (или её части)	Наименование оценочного средства
Общая характеристика методов оперативного и интеллектуального анализа данных	ОПК-1.2, ОПК-4.1, ОПК 4.2, ОПК-7.1, ОПК 7.2, ОПК 10.1, ОПК 10.2	Зачет
Аналитические платформы Data Mining	ОПК-1.2, ОПК-4.1, ОПК 4.2, ОПК-7.1, ОПК 7.2, ОПК 10.1, ОПК 10.2	Зачет
Методы Data Mining	ОПК-1.2, ОПК-4.1, ОПК 4.2, ОПК-7.1, ОПК 7.2, ОПК 10.1, ОПК 10.2	Зачет
Этапы проведения интеллектуального анализа. Интерпретация результатов анализа	ОПК-1.2, ОПК-4.1, ОПК 4.2, ОПК-7.1, ОПК 7.2, ОПК 10.1, ОПК 10.2	Зачет

3. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ ОЦЕНИВАНИЯ КОМПЕТЕНЦИЙ

Сформированность каждой компетенции в рамках освоения данной дисциплины оценивается по трехуровневой шкале:

- 1) пороговый уровень является обязательным для всех обучающихся по завершении освоения дисциплины;
- 2) продвинутый уровень характеризуется превышением минимальных характеристик сформированности компетенций по завершении освоения дисциплины;
- 3) эталонный уровень характеризуется максимально возможной выраженностью компетенций и является важным качественным ориентиром для самосовершенствования.

Описание критериев и шкалы оценивания промежуточной аттестации

а) описание критериев и шкалы оценивания тестирования:

Шкала оценивания	Критерий
5 баллов (эталонный уровень)	уровень усвоения материала, предусмотренного программой: процент верных ответов на тестовые вопросы от 85 до 100%
4 балла (продвинутый уровень)	уровень усвоения материала, предусмотренного программой: процент верных ответов на тестовые вопросы от 75 до 84%
3 балла (пороговый уровень)	уровень усвоения материала, предусмотренного программой: процент верных ответов на тестовые вопросы от 65 до 74%
0 баллов	уровень усвоения материала, предусмотренного программой: процент верных ответов на тестовые вопросы от 0 до 64%

б) описание критериев и шкалы оценивания практико-ориентированных заданий:

Шкала оценивания	Критерий
5 баллов (эталонный уровень)	Задание выполнено верно
3 балла (продвинутый уровень)	Задание выполнено верно, но имеются неточности в ответе
1 балла (пороговый уровень)	Задание выполнено верно, с дополнительными наводящими вопросами преподавателя
0 баллов	Задание не выполнено

На промежуточную аттестацию выносится тест из 10 вопросов и практико-ориентированное задание.

4. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ

4.1. Промежуточная аттестация

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-1: Способен анализировать задачи профессиональной деятельности на основе положений, законов и методов в области математики, естественных и технических наук	ОПК-1.2. Анализирует проблему, процессы и явления, относящиеся к сфере профессиональной деятельности, на основе знаний положений, законов и методов в области математики, естественных и технических наук

а) типовые тестовые задания закрытого и открытого типа:

- 1) Определите, в каких вопросах традиционная математическая статистика остается на роли основного инструмента анализа данных:
 - а) +Проверка заранее сформулированных гипотез**
 - б) +Предварительный разведочный анализ, составляющий основу оперативной аналитической обработки данных (OLAP)**
 - с) Поиск шаблонов (паттернов), отражающих фрагменты многоаспектных отношений в данных
 - д) Выявление ранее неизвестных, нетривиальных, практически полезных скрытых знаний в сырых данных
- 2) Определите, в каких вопросах интеллектуальный анализ данных (Data Mining) имеет приоритет по сравнению с традиционной математической статистикой:
 - а) Проверка заранее сформулированных гипотез
 - б) Предварительный разведочный анализ, составляющий основу оперативной аналитической обработки данных (OLAP)**
 - с) +Поиск шаблонов (паттернов), отражающих фрагменты многоаспектных отношений в данных**
 - д) +Выявление ранее неизвестных, нетривиальных, практически полезных скрытых знаний в сырых данных**
- 3) Сущность индексного метода

- а) Распределение суммы изменения обобщающего показателя между факторами по правилам пропорционального деления. Применяется в различных модификациях.
 - б) Этот метод используется для тех процессов и явлений, где имеется строго функциональная зависимость. Используется для измерения влияния отдельных причин на результат их взаимодействия при изменении величины обобщающего показателя.
 - в) Суть метода: при функциональной зависимости факторов, включенных в модель, производится последовательная замена одного из факторов на его фактическую величину.
 - д) **+Данный метод основывается на относительных показателях, выражающих отклонение уровня данного явления от уровня его в прошлом времени. Область применения: статистика, бухгалтерский учет, экономика.**
- 4) Какие из перечисленных информационно-аналитических систем представляют собой совокупность нескольких десятков методов прогноза динамики цен и выбора оптимальной структуры инвестиционного портфеля, основанных на различных эмпирических моделях динамики рынка и использующих несложный статистический аппарат, учитывающий сложившуюся в исследуемой сфере специфику (профессиональный язык, системы различных индексов и пр.):
- а) **+Информационно-аналитические системы технического анализа;**
 - б) Статистические пакеты;
 - в) Нейросетевые системы;
 - д) Системы рассуждений на основе аналогичных случаев (CBR).

б) типовые практико-ориентированные задания:

1. Решите задачу в MS Excel (функция КПЕР). В долг берется 300000 руб. под годовую ставку 6%. В год выплачивается по 34000 руб. Сколько лет займут эти выплаты?
 _____ (13 лет)

2. Решите задачу в MS Excel (функция БС). Есть два варианта инвестирования средств в течение 4 лет: в начале каждого года под 26% годовых или в конце каждого года под 38% годовых. Пусть ежегодно вносится 300000 руб. Определить, сколько денег окажется на счете в конце 4-го года для каждого варианта.
 _____ (2210535 руб для первого варианта, 2070742 руб. для второго варианта)

3. Решите задачу в MS Excel (функция ЧПС). Инвестиции в проект к концу первого года его реализации составят 10000руб. В последующие три года ожидаются годовые доходы по проекту 3000 руб., 4200 руб., 6800 руб. Издержки привлечения капитала 10%. Рассчитать чистую текущую стоимость проекта. _____ (1188,44 руб)

4. Решите задачу в MS Excel (функция БЗРАСПИС). По облигации номиналом 100000 рублей, выпущенной на 6 лет, предусмотрен следующий порядок начисления процентов: в первый год – 10%, в два последующих года – 20%, в оставшиеся три года – 25%. Рассчитать будущую (наращенную) стоимость облигации по сложной процентной ставке. _____ (309375 руб)

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-4: Способен осуществлять оценку эффективности систем управления, разработанных на основе математических методов	ОПК-4.1. Формулирует критерии оценки эффективности систем управления, выбирает инструментальные средства обработки информации

а) типовые тестовые задания закрытого и открытого типа:

- 1) Определите, как называется методика анализа больших данных, в которой контрольная выборка поочередно сравнивается с другими. Тем самым удастся выявить

оптимальную комбинацию показателей для достижения, например, наилучшей ответной реакции потребителей на маркетинговое предложение. При этом большие данные позволяют провести огромное количество итераций и таким образом получить статистически достоверный результат.

- a) **+А/Б-тестирование**
 - b) Распознавание образов
 - c) Краудсорсинг
- 2) Определите, как называется набор методик, который позволяет анализировать комментарии пользователей социальных сетей и сопоставлять с результатами продаж в режиме реального времени
- a) **+Интеграция данных (data fusion)**
 - b) Распознавание образов
 - c) Краудсорсинг
- 3) Парадигма для предоставления возможности сетевого доступа к масштабируемому и эластичному пулу общих физических или виртуальных ресурсов с предоставлением самообслуживания и администрированием по требованию – это...
- a) **+Облачные вычисления**
 - b) Большие данные
 - c) Наука о данных
- 4) Отметьте ключевые характеристики больших данных согласно ГОСТ Р ИСО/МЭК 20546-2021:
- a) **+Объем**
 - b) **+Скорость обработки**
 - c) **+Разнообразие**
 - d) **+Вариативность**
 - e) Стоимость
 - f) Масштабируемость
- 5) Укажите способы обеспечения безопасности больших данных согласно ГОСТ Р ИСО/МЭК 20546-2021:
- a) **+Шифрование**
 - b) **+Маскирование**
 - c) **+Управление доступом на основе ролей**
 - d) Защита персональных данных
 - e) Вертикальное и горизонтальное масштабирование
 - f) Создание метаданных

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-4: Способен осуществлять оценку эффективности систем управления, разработанных на основе математических методов	ОПК-4.2. Выполняет оценку эффективности систем управления на основе математических методов

а) типовые тестовые задания закрытого и открытого типа:

- 1) Какой из перечисленных методов интеллектуального анализа данных в большей мере применим для решения бизнес-задач в розничной торговле для ответа на вопросы типа: «Если сегодня покупатель приобрёл автомобиль, то через какое время он вероятнее всего купит новый аккумулятор?»:
- a) **+Исследование временных шаблонов**
 - b) Анализ покупательской корзины
 - c) Сегментация клиентов
- 2) Какой из перечисленных методов интеллектуального анализа данных в большей мере применим для решения бизнес-задач в маркетинге для выбора наиболее удачного местоположения товаров и/или при планировании перекрестных продаж?

- a) **+ Метод ассоциативных правил**
 - b) Кластерный анализ
 - c) Метод деревьев решений
- 3) Какой из перечисленных методов интеллектуального анализа данных в большей мере применим для решения бизнес-задач в банковском деле для выявления мошенничества с кредитными картами:
- a) **+Поиск ассоциативных правил и стереотипов поведения**
 - b) Анализ покупательской корзины (анализ сходства)
 - c) Сегментация клиентов и выявление признаков лояльности
- 4) Применение методов Data Mining на практике предполагает прохождение нескольких стадий. Определите их правильную последовательность:
- a) **+свободный поиск и выявление закономерностей – прогностическое моделирование – анализ исключений, выявление и объяснение аномалий**
 - b) свободный поиск и выявление закономерностей – анализ исключений, выявление и объяснение аномалий – прогностическое моделирование
 - c) прогностическое моделирование – свободный поиск и выявление закономерностей – анализ исключений, выявление и объяснение аномалий
- 5) Прочитайте описание ситуации и определите, какой из методов интеллектуального анализа целесообразно применить: «Квартиры в различных частях города будут стоить по-разному. Тип дома, район, количество и площадь комнат и другие признаки являются факторами, которые участвуют в формировании цены. Эксперт не применяет каких-либо формул, а делает оценку на основе личного опыта и знания цен на аналогичные дома. При этом его знание не статично: он может распознавать и оценивать тенденции на рынке недвижимости, приводя оценку в соответствие с ситуацией. Итак, объекты недвижимости описываются определенным набором стандартных признаков, рассматриваемых экспертом и формирующих цену. Поэтому оценка недвижимости хорошо формализуется для решения методами _____».
- a) **+Нейросетевой регрессии**
 - b) Генетического алгоритма
 - c) Поиска ассоциативных правил
- 6) Как правило, выделяют пять стандартных типов закономерностей, которые позволяют относить используемые методы к методам интеллектуального анализа данных: ассоциация, последовательность, классы, кластеры, временные ряды. Определите, какой из перечисленных типов закономерностей имеет место, если исследование показывает, что 75% покупателей, приобретавших кукурузные чипсы, приобретают и «колу».
_____ (ассоциация).
- 7) Как правило, выделяют пять стандартных типов закономерностей, которые позволяют относить используемые методы к методам интеллектуального анализа данных: ассоциация, последовательность, классы, кластеры, временные ряды. Определите, какой из перечисленных типов закономерностей имеет место, если исследование показывает, что после покупки дома в 45% случаев в течение месяца приобретается и новая кухонная плита, а в пределах двух недель 60% новоселов обзаводятся холодильником _____ (последовательность).
- 8) Характерной особенностью интеллектуального анализа данных является активное использование методов классификации, кластеризации, прогнозирования, используемых для выявления неявных закономерностей и свойств, присутствующих в данных. Определите, какой из методов следует применить, если перед маркетологом стоит задача разделения всех клиентов некоторого бизнеса на неопределенное количество групп по характеристикам условного сходства – социальному и географическому положению, основным мотивам покупки, базовым товарам персональной потребительской корзины, размеру чека и т.п.
_____ (кластеризация).

б) типовые практико-ориентированные задания:

- 9) Перечислите типовые вопросы, на которые можно найти ответы с применением методов интеллектуального анализа данных в розничной торговле.

_____ (Кто ваш покупатель? Как сегментировать клиентов? На какую целевую аудиторию сделать акцент? Какие факторы влияют на решение о покупке? Какова значимость каждого из этих факторов? Какие товары предлагать на совместных акциях? Какие существуют закономерности в поведении клиентов? На какой объем спроса будем ориентироваться?)

- 10) Перечислите типовые вопросы, на которые можно найти ответы с применением методов интеллектуального анализа данных для помощи специалистам кадровых служб

_____ (Каким требованиям должен удовлетворять соискатель? Кто и как часто совершает ошибки? Сколько компания теряет из-за ошибок сотрудников? Кто из сотрудников, вероятнее всего, является мошенником? Как оптимизировать штат и нагрузку?)

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-7: Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-7.1. Понимает принципы сбора, хранения и обработки больших объёмов данных, применяет методы и технологии обработки и визуализации данных

а) типовые тестовые задания закрытого и открытого типа:

- Как называется алгоритм кластеризации, позволяющий визуализировать ее результаты в виде двумерных карт, где расстояния между объектами соответствуют расстояниям между их векторами в многомерном пространстве, а сами значения признаков отображаются различными оттенками?
 - +Самоорганизующиеся карты признаков SOM**
 - Деревья решений
 - Генетический алгоритм
 - Алгоритмы ограниченного перебора
- Как называется древовидная диаграмма, содержащая n уровней, каждый из которых соответствует одному из шагов процесса последовательного укрупнения кластеров?
 - +Дендрограмма**
 - Скаттерплот
 - Каскадная диаграмма
 - Паукообразная диаграмма
- Как называется ориентированный граф (вершины, связанные стрелками или дугами), вершины которого – понятия, а дуги – отношения между ними)?
 - +Семантическая сеть**
 - Диаграмма Венна
 - Дерево вывода
 - Блок-схема (структурная схема)
- По описанию определите, о каком инструменте визуализации аналитической информации идет речь: «Это подробное графическое представление структуры рассуждений, в котором упор сделан на логические взаимосвязи и осуществляемые при рассуждении элементарные операции, а не на используемые в ней информационные структуры. Состоит из множества элементов различной формы, соединенных совокупностью направленных связей. Связь показывает передачу управления, а форма элемента характеризует особенности выполняемых действий и принимаемых решений. Для описания действий и логических операций внутри элементов применяется произвольная форма записи, типичными вариантами являются псевдокод и естественный язык».
 - +Блок-схема (структурная схема)**

- b) Диаграмма Венна
 - c) Дерево вывода
 - d) Семантическая сеть
- 5) Аналитическая интерактивная панель, которая включает чарты и другие элементы для представления и анализа данных, сгруппированные для лучшего восприятия и понимания информации на одном экране, - это...
- a) +Дашборд
 - b) Маркетплейс
 - c) Датасет
- 6) Каждый вид диаграммы выполняет определённую задачу, поэтому нельзя произвольно выбирать тот или иной тип диаграммы без видимой причины. В различных программах и сервисах для проведения бизнес-аналитики становятся доступными всё большее число различных видов диаграмм. Если сгруппировать диаграммы по целям использования, получатся следующие группы: сравнение категорий, соотношение категорий, отражение частей целого, распределение величины, изменение данных во времени, изменение данных в пространстве, детализация данных. Укажите виды диаграмм, которые следует применять для цели сравнение категорий.

(столбчатые диаграммы с горизонтальным или вертикальным расположением полос)

7) Каждый вид диаграммы выполняет определённую задачу, поэтому нельзя произвольно выбирать тот или иной тип диаграммы без видимой причины. В различных программах и сервисах для проведения бизнес-аналитики становятся доступными всё большее число различных видов диаграмм. Если сгруппировать диаграммы по целям использования, получатся следующие группы: сравнение категорий, соотношение категорий, отражение частей целого, распределение величины, изменение данных во времени, изменение данных в пространстве, детализация данных. Укажите виды диаграмм, которые следует применять для цели отражения частей целого.

(круговые диаграммы, нормированные столбчатые, нормированные линейчатые)

8) Каждый вид диаграммы выполняет определённую задачу, поэтому нельзя произвольно выбирать тот или иной тип диаграммы без видимой причины. В различных программах и сервисах для проведения бизнес-аналитики становятся доступными всё большее число различных видов диаграмм. Если сгруппировать диаграммы по целям использования, получатся следующие группы: сравнение категорий, соотношение категорий, отражение частей целого, распределение величины, изменение данных во времени, изменение данных в пространстве, детализация данных. Укажите виды диаграмм, которые следует применять для цели отражения изменений во времени.

(линейный график, нормированная диаграмма с областями, диаграмма с областями)

9) Проектирование и создание хорошего дашборда носит итерационный характер. Нужно продумать контекст и то, как будут выглядеть все базовые элементы дашборда. Правильный заголовок вовлекает пользователя в диалог и ориентирует его уже на самых ранних этапах анализа данных. В какой части дашборда следует располагать заголовок?

(в верхнем левом углу).

10) Перечислите, на основе каких виджетов формируется дашборд в сервисе Yandex Datalens.

(чарты, селекторы, заголовок, блоки с дополнительным текстом)

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-7: Способен понимать принципы работы	ОПК-7.2. Применяет современные

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
современных информационных технологий и использовать их для решения задач профессиональной деятельности	информационные технологии для решения задач профессиональной деятельности

а) типовые тестовые задания закрытого и открытого типа:

- 1) Разновидность АИС, предназначенная для генерации новых знаний, не содержащихся в исходных данных в явном виде называется:
 - a) автоматизированная информационно-поисковая система
 - b) автоматизированная система обработки данных
 - c) автоматизированная система управления
 - d) **автоматизированная интеллектуальная информационная система.**

- 2) Хранящаяся на поисковом сервере база данных, по которой осуществляется поиск запрошенной пользователем информации, и которая, как правило, содержит ссылки на проиндексированные ресурсы и краткую информацию о них, называется:
 - a) поисковая система
 - b) **индекс поисковой системы**
 - c) индекс сервера
 - d) информационный портал

- 3) Специализированный сервер Интернета, рекламирующий веб-ресурс пользователя в обмен на размещение им рекламы других участников сети, – это:
 - a) **баннерообменная система**
 - b) спам
 - c) баннерная реклама
 - d) Интернет-маркетинг

- 4) Чтобы начать набор текста функции или подпрограммы на языке VBA необходимо:
 - a) выполнить команду Debug/CompileVBAProject
 - b) выполнить команду Insert/Procedure
 - c) **выполнить команду Insert/Module**

- 5) Чтобы создать новую форму в VBA необходимо:
 - a) выполнить команду Insert/Module;
 - b) выполнить команду Insert/Procedure;
 - c) **выполнить команду Insert/UserForm**

- 6) Технология работы со специальными отчетами, включая программное обеспечение, для получения и анализа многомерных структурированных данных. Одним из популярных программных продуктов, реализующих эту технологию, является SQL Server Analysis Server - _____ (**OLAP-технологии**).
- 7) _____ - первоначально специализированное устройство, позже компьютерная программа, используемая для набора, сохранения, редактирования и печати текста. Современные _____ имеют также функции компоновки макета текста и предварительного просмотра документов в том виде, в котором они будут напечатаны (**текстовые процессоры**).
- 8) Комплекс программных средств для математической, статистической и графической обработки текстовых и числовых данных в табличном виде – это _____ (**табличный процессор**).
- 9) Средства _____ технологий – это математические методы и модели решения задач, алгоритмы обработки данных, инструментальные средства моделирования бизнес-процессов, данных проектирования информационных систем, разработки программ, собственно программные продукты, разнообразные информационные ресурсы, технические средства обработки данных (**информационных**).

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-10: Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения	ОПК-10.1. Понимает сущность алгоритмов и принципы их построения

а) типовые тестовые задания закрытого и открытого типа:

- 1) Для обнаружения фактов мошенничества со счетами кредитных карт, подозрительных сделок в страховании применяется метод обнаружения аномалий, по сути противоположный кластеризации и реализуемый при наличии больших данных о типичном профиле исследуемого объекта. Какие из перечисленных методов интеллектуального анализа могут применяться для выявления аномалий?
 - а) Калмановская фильтрация
 - б) Метод опорных векторов SVM
 - в) Нейросети
 - д) +Все ответы верные**
- 2) Верно ли утверждение: Qlik может научить интеллектуальному анализу данных за счет ассоциативного движка.
 - а) +Неверно, Qlik нужен для анализа данных и создания дашбордов. Освоение принципов и методов интеллектуального анализ данных – отдельная задача.**
 - б) Верно, у Qlik есть такая функция.
 - в) Неверно, такая функция есть только в Qlik View.
- 3) Может ли CRM-система разделить потоки данных между отделами?
 - а) Нет, CRM-система нужна только для приема звонков и ведения карточек клиентов
 - б) +Конечно, в этом и есть особенность хороших CRM**
- 4) При использовании информационных технологий для интеллектуального анализа данных возникают проблемы, связанные с качеством исходных данных. При получении данных из разных источников и из единственного источника возникающие проблемы имеют специфику. Какие из перечисленных проблем возникают в ситуации, когда данные получены из единственного источника?
 - а) +Отсутствие ограничений целостности, плохая архитектура базы данных**
 - б) +Ошибки ввода данных**
 - в) Гетерогенные модели данных и архитектура баз данных
 - д) Перекрытие и противоречивость данных
- 5) Какая из перечисленных информационных систем соответствует описанию: «распределенная система управления базами данных, относящаяся к классу NoSQL-систем и рассчитанная на создание высоко масштабируемых и надежных хранилищ огромных массивов данных, представленных в виде хэша. Известно, что данная СУБД используется такими компаниями, как Cisco, IBM, Cloudkick, Reddit, Digg, Rackspace, Apple и Twitter. Написана на языке Java, реализует распределенную hash-систему, что обеспечивает практически линейную масштабируемость при увеличении объема данных. Использует модель хранения данных на базе семейства столбцов, чем отличается от систем, которые хранят данные только в связке «ключ – значение», имеет возможность организовать хранение хэшей с несколькими уровнями вложенности. Относится к категории отказоустойчивых»:
 - а) +Apache Cassandra**
 - б) ApacheYARN
 - в) MapReduce
- 6) К информационным технологиям интеллектуального анализа данных сегодня относят оперативный анализ данных, исследование данных и извлечение знаний. Какому из этих типов соответствует описываемая технология: «Технология представляет новое направление в области ИАД, где процесс поиска закономерностей в данных рассматривается как процесс

машинного обучения. Технология объединяет в себе вопросы моделирования закономерностей и зависимостей в базах данных и определяет математические методы построения систем "открытия" новых данных на основе методов классификации, кластеризации, построения деревьев решений и др.» _____ (**извлечение знаний**)

б) типовые практико-ориентированные задания:

- 7) Приведите примеры современных СУБД, которые включают поддержку функциональности интеллектуального анализа данных.

_____ (**Microsoft SQL Server Analysis Services(Microsoft Corp.);Oracle Business Intelligence(Oracle Corp.);IBM DB2 Intelligent Miner(IBM)**)

- 8) Изучите пример применения одного из методов интеллектуального анализа данных – генетического алгоритма. Закончите описание формулировкой результатов применения метода. «Если нам надо выбрать совокупность фиксированного числа параметров рынка, наиболее выраженно влияющих на его динамику, это будет набор имен этих параметров. Об этом наборе можно говорить как о совокупности хромосом, определяющих качества индивида - данного решения поставленной задачи. Значения параметров, определяющих решение, будут тогда называться генами. Поиск оптимального решения при этом похож на эволюцию популяции индивидов, представленных их наборами хромосом. В этой эволюции действуют три механизма: во-первых, отбор сильнейших - наборов хромосом, которым соответствуют наиболее оптимальные решения; во-вторых, скрещивание - производство новых индивидов при помощи смешивания хромосомных наборов отобранных индивидов; и, в-третьих, мутации - случайные изменения генов у некоторых индивидов популяции. В результате...?»

(В результате смены поколений в конце концов вырабатывается такое решение поставленной задачи, которое уже не может быть далее улучшено)

- 9) Известно, что, если данных мало, а описывающая их модель сложна и включает много степеней свободы, то всегда можно подогнать эту модель под данные, даже если это целиком случайные числа. Какой в связи с этим возникает риск применения информационных технологий для интеллектуального анализа данных в условиях, когда срок существования исследуемой компании не превышает 2-3 года?

(Опасность тут состоит в получении статистически незначимых моделей и принятии на их основе неверных решений. Тот факт, что методы ИАД отлично работают, когда нужно объяснить то, что было в прошлом, но непригодны для принятия решений «на будущее» в условии ограниченной накопленной истории, может стать барьером применения систем data mining для решения реальных задач из сферы бизнеса и финансов)

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-10: Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения	ОПК-10.2. Формулирует практические задачи, требующие алгоритмизации, выявляет возможности применения компьютерных технологий для их решения

а) типовые тестовые задания закрытого и открытого типа:

- 1) Сопоставьте уровни знаний, получаемых в результате обработки и анализа информации, и соответствующие аналитические инструменты:
 - а) Поверхностный: Язык простых запросов, Неглубокий: Оперативная аналитическая обработка, Скрытый: Интеллектуальный анализ данных
 - б) Поверхностный: Интеллектуальный анализ данных, Неглубокий: Язык простых запросов, Скрытый: Оперативная аналитическая обработка
 - в) Поверхностный: Оперативная аналитическая обработка, Неглубокий: Язык простых

- запросов, Скрытый: Интеллектуальный анализ данных.
- 2) _____ - это пользовательский интерфейс для отображения многомерных данных. С помощью данного интерфейса можно группировать, сортировать, фильтровать и менять расположение данных с целью получения различных аналитических выборок (**сводный отчет, сводная таблица**).
- 3) По описанию определите, о каком методе анализа идет речь: «Данный метод пригоден только для решения задач классификации, и поэтому весьма ограниченно применяется в области финансов и бизнеса, где чаще встречаются задачи численного прогноза. В результате применения этого метода к обучающей выборке данных создается иерархическая структура классифицирующих правил типа "ЕСЛИ... ТО..."» _____
(**деревья решений**)
- 4) Укажите функции управления предприятием, которые поддерживают современные информационные системы _____
(**планирование, учет, анализ, регулирование**)
- 5) Какой информационной системе соответствует следующее определение: программно-аппаратный комплекс, способный объединять в одно целое предприятия с различной функциональной направленностью (производственные, торговые, кредитные и др. организации).
_____ (**корпоративная информационная система**)
- 6) Как называется интегрированная система, обеспечивающая планирование и управление всеми ресурсами предприятия, его снабжением, сбытом, кадрами и заработной платой, производством, научно-исследовательскими и конструкторскими работами?
_____ (**ERP-система**)
- 7) К какому типу прямой или обратной задач относится формирование информации для управленческих решений? _____ (**обратная**)
- 8) Какие инструменты используются для формирования управленческого решения в условиях риска? _____ (**дерево решений, платежная матрица**)

Оператор ЭДО ООО "Компания "Тензор"

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ

СОГЛАСОВАНО

ФГБОУ ВО "РГРТУ", РГРТУ, Евдокимова Елена Николаевна, Заведующий
кафедрой ЭМОП

Простая подпись