

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Петровская Анна Викторовна  
Должность: Директор  
Дата подписания: 19.09.2024 16:22:20  
Уникальный программный ключ:  
798bda6555fbdebe827768f6f1710bd17a9070c31fdc1b6a6ac5a1f10c8c5199

Приложение 6  
к основной профессиональной образовательной программе  
по направлению подготовки 38.03.01 Экономика  
направленность (профиль) программы Финансовая безопасность

**Министерство науки и высшего образования Российской Федерации  
федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Российский экономический университет имени Г.В. Плеханова»**

**Факультет экономики, менеджмента и торговли**

**Кафедра бухгалтерского учета и анализа**

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ**

**по дисциплине Основы работы с Большими данными**

**Направление подготовки**

**38.03.01 Экономика**

**Направленность (профиль) программы**

**Финансовая безопасность**

**Уровень высшего образования**

**Бакалавриат**

**Год начала подготовки - 2023**

Краснодар – 2022 г.

Составитель: к.п.н., доцент В.В. Салий

Оценочные материалы одобрены на заседании кафедры бухгалтерского учета и анализа Протокол № 6от 10.01.2022

## ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

### По дисциплине «Основы работы с Большими данными» **ПЕРЕЧЕНЬ КОМПЕТЕНЦИЙ С УКАЗАНИЕМ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ И ЭТАПОВ ИХ ФОРМИРОВАНИЯ ПО ДИСЦИПЛИНЕ**

Формируемые компетенции (код и наименование компетенции)	Индикаторы достижения компетенций (код и наименование индикатора)	Результаты обучения (знания, умения)
<b>УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</b>	<b>УК-1.1.</b> Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи	<b>УК-1.1. З-1. Знает</b> основные методы критического анализа и основы системного подхода как общенаучного метода.
		<b>УК-1.1. У-1. Умеет</b> анализировать задачу, используя основы критического анализа и системного подхода.
		<b>УК-1.1. У-2. Умеет</b> осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации.
<b>ОПК-2. Способен осуществлять сбор, обработку и статистический анализ данных, необходимых для решения поставленных экономических задач.</b>	<b>ОПК-2.1.</b> Использует основные методы, средства получения, представления, хранения и обработки статистических данных.	<b>ОПК-2.1. З-1. Знает</b> методы поиска и систематизации информации об экономических процессах и явлениях.
		<b>ОПК-2.1. У-1. Умеет</b> работать с национальными и международными базами данных с целью поиска информации, необходимой для решения поставленных экономических задач.
		<b>ОПК-2.1.У-2. Умеет</b> рассчитывать экономические и социально-экономические показатели, характеризующие деятельность хозяйствующих субъектов на основе типовых методик и действующей нормативно-правовой базы
<b>ОПК-2.1.У-3. Умеет</b> представить наглядную визуализацию данных.		
<b>ОПК-6. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности</b>	<b>ОПК-6.1.</b> Использует соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы	<b>ОПК-6.1. З-1. Знать:</b> характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий
	<b>ОПК-6.2.</b> Понимает принципы работы современных цифровых информационных технологий, соответствующих содержанию профессиональных задач	<b>ОПК-6.1. У-1. Уметь:</b> использовать современные цифровые информационные технологии для решения задач профессиональной деятельности  <b>ОПК-6.2. З-1. Знать:</b> принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий  <b>ОПК-6.2.У-1. Умеет</b> применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных

**МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ХАРАКТЕРИЗУЮЩИЕ ЭТАПЫ  
ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ****Перечень учебных заданий на аудиторных занятиях****Темы для проведения групповых дискуссий****Тема 1. Основы Больших данных****Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

1. Что такое большие данные?
2. Методики анализа больших данных
3. Аналитический инструментарий
4. Рынки решений для управления большими данными
5. Теория и практика больших данных в отраслях
6. Как использовать аналитические данные для разработки качественных ИТ-сервисов
7. Большие данные (big data) в информационных технологиях.
8. Совокупность подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов и значительного многообразия для получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста, распределения по многочисленным узлам вычислительной сети, сформировавшихся в конце 2000-х годов, альтернативных традиционным системам управления базами данных и решениям класса Business Intelligence.
9. В данную серию включают средства массово-параллельной обработки неопределённо структурированных данных, прежде всего, решениями категории NoSQL, алгоритмами MapReduce, программными каркасами и библиотеками проекта Hadoop.
10. В качестве определяющих характеристик для больших данных отмечают три V: объём (англ. volume, в смысле величины физического объёма), скорость (англ. velocity в смыслах как скорости прироста, так и необходимости высокоскоростной обработки и получения результатов), многообразие (англ. variety, в смысле возможности одновременной обработки различных типов структурированных и полуструктурированных данных).
11. Информационно-аналитических систем.
12. Структура средств сбора и доработки данных.
13. Способы и состав средств преобразования данных.
14. Среды разработки средств сбора, доработки данных и информационных хранилищ.
15. Основные принципы построения структуры информационных хранилищ.

**Тема 2. Обработка больших данных.****Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

1. BigData и аналитика в реальном времени
2. BigData и ценовая политика в обработке данных клиентов
3. Обзор технологий хранения больших данных.
4. Базы данных. Системы управления базами данных. Модели данных.
5. Подготовка исходных данных для анализа: первичная обработка и визуализация имеющихся данных.
6. Основные понятия математической статистики.
7. Методы анализа данных: дескриптивная статистика, параметрические, непараметрические, номинальные методы (корреляционный, регрессионный, дисперсионный анализы, кластерный, дискриминантный, факторный анализы).

8. Принципы гибкой архитектуры данных и открытых систем, которыми руководствуются при создании ИАС.
9. Информационный обмен, связанный с аналитической работой.
10. Понятие о метаданных (МД), базе метаданных - репозитории, используемых в информационном хранилище.
11. Принципы построения информационных хранилищ.
12. Требования к качеству данных и способы его обеспечения при загрузке в информационное хранилище.
13. Проблемы, разрешаемые при приведении данных к единой структуре информационного хранилища.
14. Концепции построения структур хранилищ данных.
15. Назначение, состав и выполняемые функции базы метаданных - репозитория ИХ.
16. Принципы создания репозитория ИХ.
17. Элементы моделей данных ИХ (факт-таблица, таблицы измерений, консольные таблицы).

### **Тема 3. Анализ и извлечение данных**

**Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

1. Цели факторного анализа.
2. Программные средства анализа данных: Statistica, SPSS, Excel и другие; их преимущества и недостатки.
3. Представление исходных данных в программе R-Studio.
4. SQL over BigData. Hive
5. Проблема смещения данных в обработке больших данных
6. Применение SQL в IT индустрии
7. Сравнение решений Hive и MapReduce на примере задач анализа логов
8. Beyond MapReduce. Spark
9. Машинное обучение на больших данных
10. Алгоритмы для работы с большими данными
11. Методы онлайн обучения
12. Градиентный спуск
13. Решение задач кластеризации на больших данных
14. Задача подсчета слов в датасете (WordCount)
15. Задачи и содержание оперативного (OLAP) анализа.
16. Классификация ИТ-анализа по режиму и темпу.
17. Функции, выполняемые процедурами сечение и поворот.
18. Функции, выполняемые процедурами свертка и развертка.
19. Функции, выполняемые процедурами проекция и построение трендов.

### **Тема 4. Технология Data Mining**

**Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

1. Введение в интеллектуальный анализ данных
2. Субъектные и объектные системы.
3. Первая модель управления (внешнее управление).
4. Сравнительный анализ подходов к понятию управления: как средство достижения цели, как процесс управляющих воздействий, как процесс обработки информации.
5. Управление как процесс выбора и принятия решений.
6. Вторая модель управления - автоматическое управление
7. Предварительная обработка данных
8. Задача классификации с обучением

9. Основные принципы построения архитектуры
10. Области применения методов интеллектуального анализа.
11. Системы рассуждений на основе аналогичных случаев.
12. Классификационные и регрессионные деревья решений.
13. Байесовское обучение (ассоциации).
14. Генетические алгоритмы.
15. Эволюционное программирование и алгоритмы ограниченного перебора.

#### **Критерии оценки (в баллах):**

**20 баллов** выставляется обучающемуся, если он демонстрирует высокий уровень владения материалом по всем темам дискуссий, превосходное умение формулировать свою позицию, отстаивать её в споре, задавать вопросы, обсуждать дискуссионные положения, высокий уровень этики ведения дискуссии. Уровень сформированности компетенций соответствует продвинутому уровню;

**15- баллов** выставляется обучающемуся, если он демонстрирует владение материалом по всем темам дискуссий на уровне выше среднего, умение отстаивать её в споре, задавать вопросы, обсуждать дискуссионные положения, знание этики ведения дискуссии. Уровень сформированности компетенций соответствует повышенному уровню;

**10 баллов** выставляется обучающемуся, если он демонстрирует владение материалом по всем темам дискуссий не в полном объеме, умение задавать вопросы, обсуждать дискуссионные положения, знание этики ведения дискуссии. Уровень сформированности компетенций соответствует базовому уровню.

## **ЗАДАНИЯ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ**

### **Контрольная работа**

#### **По теме 4. Технология Data Mining**

**Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

1. Приведите формальную постановку основных задач Data Mining.
2. Перечислите и охарактеризуйте основные задачи Data Mining.
3. К основным задачам Data Mining относятся:
4. Постановка задачи классификации.
5. Постановка задачи регрессии.
6. Постановка задачи кластеризации.
7. Постановка задачи поиска ассоциативных правил.
8. Постановка задачи поиска последовательности.
9. Постановка задачи прогнозирования.
10. Чем отличается классификация от кластеризации?
11. Приведите варианты, сходство и отличия задач группировки и их графическую иллюстрацию.
12. Приведите отличия и сходство задачи прогнозирования от задачи поиска ассоциативных правил с привлечением общей системной модели решения задач и на основе их формальной постановки.
13. Опишите сущность формального концептуального анализа и его применение.
14. Сформулируйте постановку и приведите примеры задачи лингвистического резюмирования.
15. Охарактеризуйте задачу поиска аномалий и сопоставьте ее с кластеризацией данных на разных уровнях представления.

16. Подберите адекватные формальные постановки задач Data Mining для решения следующих проблем:
- Разделить проекты на проекты, имеющие высокую степень успешности, среднюю и неуспешные, и определить их метрики.
  - Найти аналогичные проекты по набору требований в виде гетерогенных параметров.
  - Определить риски проектирования и реализации программного продукта по диаграмме выгорания.
  - Спрогнозировать время выполнения проекта исходя из набора требований, КРІ персонала и объема финансирования.
  - Определить, какую квалификацию имеет «Большинство разработчиков».
17. В чем отличие технологии Data Mining от других технологий анализа больших объемов данных?
18. Назовите наиболее распространенные методы анализа структурированных данных с использованием технологии Data Mining. Какие типы скрытых зависимостей они выявляют?
19. Почему при создании человеко-машинных информационных систем большое внимание уделяется зависимостям, полученным в виде правил «если, то»?
20. Кратко опишите суть метода построения дерева решений. Какие ограничения имеются у этого метода?
21. Кратко опишите суть метода поиска ассоциативных правил. Какие ограничения имеются у этого метода?
21. В чем суть алгоритма Apriori, каковы его основные шаги? Каков смысл свойства анти-монотонности?
22. Что такое обобщенное ассоциативное правило? В чем отличие алгоритма вычисления обобщенных ассоциативных правил от алгоритма Apriori?
23. Какие улучшения требуются алгоритму поиска обобщенных ассоциативных правил для обеспечения приемлемых технических характеристик?
24. Какие прогрессивные методы Data Mining предполагается развивать в ближайшее время для получения скрытых зависимостей?
25. Какие программные продукты, реализующие технологию Data Mining, Вам известны? Чем вызвано такое их разнообразие?
26. Какие программные продукты позволяют получать скрытые зависимости в виде правил «если, то»?
27. Приведите примеры инцидентов и проблем, возникающих при управлении IT службами организаций или компаний.
28. Почему после выяснения и устранения проблемы необходимо еще некоторое время отслеживать ранее возникавшие инциденты?
29. Как объяснить разное число правил, найденных для одного массива данных методами построения деревьев и поиска ассоциативных зависимостей?
30. Почему вид правил, найденных различными методами при обработке одного и того же набора данных, различается и «условными» и «следственными» частями?
31. Определите главные отличия результатов, получаемых методами построения деревьев и поиска ассоциативных зависимостей?

**Критерии оценки (в баллах):**

**10 баллов** выставляется обучающемуся, если он правильно ответил на 80% вопросов (компетенция сформирована на продвинутом уровне);

**9 баллов** выставляется обучающемуся, если он правильно ответил на 70% вопросов (компетенция сформирована на продвинутом уровне);

**8 баллов** выставляется обучающемуся, если он правильно ответил на 60% вопросов теста (компетенция сформирована на повышенном уровне);

**7 баллов** выставляется обучающемуся, если он правильно ответил на 55% вопросов теста (компетенция сформирована на повышенном уровне);

**6 баллов** выставляется обучающемуся, если он правильно ответил на 50% вопросов теста (компетенция сформирована на базовом уровне)

### **Комплект тестовых заданий по дисциплине**

**Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

**1. Какие науки включены в Data Mining?**

- а. статистика, базы данных, искусственный интеллект;
- б. информатика, базы данных, статистика;
- в. искусственный интеллект, базы данных, базы знаний;
- г. информатика, базы данных, хранилища данных.

**2. Какая подсистема СППР отвечает за интеллектуальный анализ данных?**

- а. OLTP;
- б. хранилище данных;
- в. SQL;
- г. OLAP;
- д. Data Mining.

**3. Репозиторий – это...**

- а. словарь терминов;
- б. хранилище метаданных;
- в. каталог с файлами.

**4. Если для реализации многомерной модели используют многомерные базы данных, то способ реализации гиперкуба называется**

- а. MOLAP;
- б. ROLAP;
- в. HOLAP;
- г. DOLAP;
- д. JOLAP.

**5 Задача классификации сводится к ...**

- а. нахождению частых зависимостей между объектами или событиями;
- б. определению класса объекта по его характеристикам;
- в. определению по известным характеристикам объекта значения некоторого его параметра;
- г. поиску независимых групп и их характеристик во всем множестве анализируемых данных.

**6. Целью поиска ассоциативных правил является ...**

- а. нахождение частых зависимостей между объектами или событиями;
- б. определение класса объекта по его характеристикам;
- в. определение по известным характеристикам объекта значения некоторого его параметра;
- г. поиск независимых групп и их характеристик во всем множестве анализируемых данных.

**7. Назовите основной этап работы с хранилищами данных?**

- а. этап очистки данных;
- б. этап обновления;
- в. этап нормализации.

**8. Что называют кубом OLAP?**

- а. структуру, в которой хранятся совокупности данных, полученные путем всех возможных сочетаний измерений в таблице измерений;



б. структуру, в которой хранятся совокупности данных, полученные путем всех возможных сочетаний измерений в таблице фактов;

в. таблицу размерностей.

**9. Информационные хранилища созданы для удобства ...**

а. руководителей всех уровней для принятия решений;

б. стратегического планирования;

в. реорганизации бизнеса;

г. предметных приложений;

д. редактирования данных.

**10. Информационные хранилища размещаются на ...**

а. библиотеках-автоматах;

б. сетевых серверах;

в. мейнфреймах;

г. серверах и кластерах серверов;

д. файл-серверах.

**11. Ассоциация — ...**

а. установление зависимости непрерывной выходной переменной от входных переменных;

б. группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;

в. выявление закономерностей между связанными событиями;

г. установление зависимости дискретной выходной переменной от входных переменных.

**12. Сумма расстояний между двумя точками, принадлежащими к одному кластеру, которые строго меньше, чем расстояние между двумя точками, не принадлежащими к одному кластеру определяет**

а. Качество кластеризации

б. Робастность классификации

в. Производительность кластеризации

**13. Точность классификации объектов определяется точностью по кроссвалидации**

**в виде**

а. Процента правильно классифицированных объектов

б. Отношения количества правильно классифицированных объектов к количеству неправильно классифицированных объектов

в. Дисперсией значений признаков

**14. Методы CART, C4.5, Randomforest относят к методам**

а. Кластеризации

б. Классификации

в. Ассоциации

**15. Задача поиска редких и нетипичных объектов относится к задаче**

а. Поиска ассоциативных правил

б. Поиска аномалий

в. Кластеризации

**16. Основные сквозные цифровые технологии Программы «Цифровая экономика Российской Федерации»**

а. робототехника; автономные системы; облачные вычисления

б. искусственный интеллект; робототехника; анализ Больших данных; интернет вещей и индустриальный интернет

в. робототехника; умные устройства; интернет вещей и индустриальный интернет

**17. Большие данные характеризуются наиболее точно**

а. Большим объемом данных в хранилищах

б. Большим объемом разнообразных данных, имеющих высокую скорость изменчивости

в. Большим объемом слабоструктурированных данных

18. Для выявления групп схожих паттернов в данных по множеству признаков применяют

а. Методы ассоциации

б. Методы прогнозирования

в. Методы классификации

г. Методы кластеризации

19. Дано множество данных обучающих примеров, для каждого из которых известны классы принадлежности, требуется для новых данных, для которых неизвестны их классы, определить эти классы. Какие модели для решения указанной задачи применимы

а. Модели обучения с учителем

б. Модели обучения без учителя

в. Модели смешанного обучения

г. Модели кластерного обучения

20. Для решения задач предиктивной аналитики в экономике применяют

а. Модели иерархической кластеризации

б. Модель на основе дерева решений

в. Модель CRISP-DM

г. Модель Hadoop

21. Прогнозирование оттока клиентов в бизнесе относится к задачам

а. Предиктивной аналитики

б. Диагностической аналитики

в. Дескриптивной аналитики

г. Когнитивной аналитики.

#### **Критерии оценки (в баллах):**

**10 баллов** выставляется обучающемуся, если он правильно ответил на 10 вопросов теста (компетенция сформирована на продвинутом уровне);

**9 баллов** выставляется обучающемуся, если он правильно ответил на 9 вопросов теста (компетенция сформирована на продвинутом уровне);

**8 баллов** выставляется обучающемуся, если он правильно ответил на 8 вопросов теста (компетенция сформирована на повышенном уровне);

**7 баллов** выставляется обучающемуся, если он правильно ответил на 7 вопросов теста (компетенция сформирована на повышенном уровне);

**6 баллов** выставляется обучающемуся, если он правильно ответил на 6 вопросов теста (компетенция сформирована на базовом уровне)

**5 баллов** выставляется обучающемуся, если он правильно ответил на 5 вопросов теста (компетенция сформирована на базовом уровне).

#### **Кейсы**

#### **Тема 2. Обработка больших данных**

**Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

##### *Задание 1.*

Поиск источников информации в сети Интернет: открытые и закрытые источники данных. Портал открытых данных РФ. Сохранение данных в программе Excel. Преобразование и первичная обработка данных.

##### *Задание 2.*

Представление исходных данных в программе R-Studio (векторы, массивы, матрицы, списки, таблицы). Статистическая обработка данных в программах Excel и R-Studio: подсчет описательных статистик, графическое представление данных. Группировка данных, обнаружение значимых корреляций, зависимостей и тенденций в результате анализа имеющейся информации, выявления отношений между данными различного типа. Применение различных методов выделения, извлечения и группировки данных, которые позволяют выявить систематизированные структуры данных и вывести из них правила для принятия решений и прогнозирования их последствий (регрессионный, дисперсионный, кластерный, дискриминантный, факторный анализы).

### **Тема 3. Анализ и извлечение данных**

**Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

#### *Задание 1.*

Возможности графического представления информации в программе R-Studio: графические функции отображения одномерных и многомерных данных, графический вывод с использованием графических параметров.

#### *Задание 2*

Исходная база Больших данных содержит «сырые» многомерные числовые данные.

Требуется создать формальную постановку задачи, извлечь и описать данные, применить метод кластеризации для выборочных данных, сформировать группы похожих объектов и вычислить признаки типичного объекта для каждой группы. Подготовить отчет.

#### **Критерии оценки (в баллах):**

**5 баллов** выставляется обучающемуся, если он в полном объеме и правильно выполнил задание (компетенция сформирована на продвинутом уровне);

**4 балла** выставляется обучающемуся, если он в полном объеме и с незначительными замечаниями выполнил задание (компетенция сформирована на повышенном уровне);

**3 балла** выставляется обучающемуся, если он на базовом уровне, ошибками выполнил задание (компетенция сформирована на базовом уровне).

### **Задания для творческого рейтинга**

#### **Тематика докладов**

**Индикаторы достижения: УК-1.1., ОПК-2.1, ОПК-6.1., ОПК-6.2.**

1. Критерии оценки программных средств Data Mining.
2. Классификация программных средств Data Mining.
3. Программное обеспечение Data Mining для поиска ассоциативных правил.
4. Программное обеспечение для решения задач кластеризации и сегментации.
5. Программное обеспечение для решения задач классификации.
6. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования.
7. Обзор программного средства RapidMiner.
8. Обзор Statistica.
9. Обзор Deductor.
10. Обзор KNIME.
11. Обзор Weka.
12. Обзор SAS Enterprise Miner.
13. Обзор IBM SPSS Modeler.
14. Обзор Oracle Data Mining ODM.

15. Поддержка принятия решений на основе анализа Больших данных
16. Большие данные и искусственный интеллект
17. Экосистема Hadoop
18. Архитектуры систем для обработки Больших данных
19. Методы TextMining
20. Применение технологий Большие данные для решения задач в микроэкономике
21. Применение технологий Большие данные для решения задач в макроэкономике
22. Ограничения технологии Большие данные
23. Большие данные и хранилища данных.
24. Применение Больших данных для извлечения новых знаний о клиентах
25. Этапы обработки Больших данных
26. Озера данных и Большие данные
27. Облачные решения Больших данных
28. Интеграция Больших данных с моделями машинного обучения
29. Большие данные и DataScience
30. Большие данные и интернет-вещей
31. Большие данные как дальнейшее развитие информационной экономики
32. Цифровая экономика и цифровая трансформация, взаимодействие с большими данными
33. Распределенные вычисления и хранилище данных (облачное хранение)
34. Проблема создания и размещения дата-центров
35. Большие данные и принятие решений.
36. Искусственный интеллект
37. Понятие bigdata. Новые подходы к накоплению и обработке данных в экономике и финансах на микро- и макроуровнях.
38. Открытые данные компьютерных поисковых систем и социальных сетей. GoogleTrends. YandexWorstat.
39. Цифровая трансформация модели потребления.
40. Тенденции в области цифрового контента.
41. Применение методов машинного обучения к большим данным
42. Применение методов Data Mining к большим данным
43. Применение нейронных сетей к анализу больших данных.
44. Поиск ассоциативных правил в больших данных.
45. Программные средства анализа «больших данных»

#### **Критерии оценки (в баллах):**

**20 баллов** выставляется обучающемуся, если он без ошибок подготовил доклад. Содержание и оформление доклада соответствует требованиям в полном объеме. Уровень сформированности компетенций соответствует продвинутому уровню;

**19-14 балла** выставляется обучающемуся, если он с незначительными замечаниями по содержанию или оформлению подготовил отчет по реферату. Уровень сформированности компетенций соответствует повышенному уровню;

**13-10 балла** выставляется обучающемуся, если он с ошибками подготовил доклад. Содержание и оформление доклада соответствует требованиям не менее, чем на 50 процентов. Уровень сформированности компетенций соответствует базовому уровню.

### **МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ХАРАКТЕРИЗУЮЩИЕ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ ВО ВРЕМЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ**

#### **Структура зачетного задания**

<i>Наименование оценочного материала</i>	<i>Максимальное количество баллов</i>
<i>Вопрос 1</i>	<i>15</i>
<i>Вопрос 2</i>	<i>15</i>
<i>Практическое задание 1</i>	<i>10</i>

### **Задания, включаемые в зачетное задание**

#### ***Перечень вопросов к зачету:***

1. Понятие и назначение Big Data.
2. Технологии обработки больших объемов данных.
3. Особенности высокоскоростной обработки данных.
4. Типы слабоструктурированных данных и особенности их обработки.
5. Основные принципы работы Hadoop.
6. Средства сбора и доработки данных.
7. Визуализация больших данных.
8. Задачи администрирования в системах обработки больших данных.
9. Принципы построения систем хранения и обработки больших данных.
10. Методы обеспечения качества данных.
11. Применение BigData в различных сферах деятельности.
12. Применение BigData в решении бизнес-задач.
13. Источники больших данных.
14. Платформы больших данных.
15. Процесс изучения данных.
16. Экосистемы больших данных.
17. Какие данные называют «Большими»?
18. Почему Большие данные нецелесообразно хранить в обычной реляционной базе данных?
19. Какой момент считается точкой отсчета науки о Больших данных?
20. Методы обработки Больших данных.
21. Приведите пример неочевидной полезной закономерности, полученной с помощью методов анализа Больших данных.
22. Модели машинного обучения для обработки Больших данных.
23. Как анализ Больших данных может быть использован для решения задач биржевой торговли?
24. Как анализ Больших данных используется в современной медицине?
25. Этапы обработки Больших данных
26. Что представляет из себя интернет вещей и как он связан с методами анализа Больших данных?
27. Создание «прозрачной» информации посредством анализа экономических данных Большого объема.
28. Методы анализа Больших данных в задачах принятия математически обоснованных управленческих решений.
29. Узкое сегментирование клиентов с учетом персональных пожеланий на основе технологий обработки Больших данных.
30. Увеличение скорости в принятии решений за счет сложной аналитики Больших данных в режиме реального времени.
31. Аналитика Больших данных в задачах развития товаров и услуг следующего поколения.
32. Оцифровка текста на естественном языке.

33. Какие программные продукты могут использоваться для семантического анализа текстов на естественных языках?
34. Основные принципы работы современных систем машинного перевода текстов на естественных языках.
35. Наиболее эффективные программные продукты для распознавания речи.
36. Основные базы знаний WolframDataPlatform.
37. Пример геоинформационных систем, работа которых существенно зависит от эффективности алгоритмов анализа Больших данных?
38. Архитектуры систем для обработки Больших данных
39. Статистические модели для решения задач обработки Больших данных
40. Этапы обработки Больших данных
41. Подходы к представлению результатов обработки Больших данных
42. Виды аналитики Больших данных
43. Методы KnowledgeDiscovery для извлечения дескриптивных знаний.
44. Методы KnowledgeDiscovery для извлечения предиктивных знаний.
45. Понятие и основные определения Data Mining
46. Методы и стадии Data Mining.
47. Задачи Data Mining.
48. Сферы применения Data Mining.
49. Зависимость структуры хранилища данных от предметной области.
50. Сравнительная характеристика OLTP и OLAP.

### ***Практические задания к зачету***

1. Используя методологию DataMining, решите задачу кластеризации данных в объекте исследования «Финансовая компания».
2. Используя методологию DataMining, решите задачу классификации данных в объекте исследования «Финансовая компания».
3. Используя методологию DataMining, решите задачу прогнозирования данных в объекте исследования «Финансовая компания».
4. Используя методологию DataMining, решите задачу поиска аномалий в данных в объекте исследования «Финансовая компания».
5. Используя методологию DataMining, решите задачу кластеризации данных в объекте исследования «Интернет-магазин».
6. Используя методологию DataMining, решите задачу классификации данных в объекте исследования «Интернет-магазин».
7. Используя методологию DataMining, решите задачу прогнозирования данных в объекте исследования «Интернет-магазин».
8. Используя методологию DataMining, решите задачу поиска аномалий в данных в объекте исследования «Интернет-магазин».
9. Используя методологию DataMining, решите задачу кластеризации данных в объекте исследования «Рекламная компания».
10. Используя методологию DataMining, решите задачу классификации данных в объекте исследования «Рекламная компания».
11. Используя методологию DataMining, решите задачу прогнозирования данных в объекте исследования «Рекламная компания».
12. Используя методологию DataMining, решите задачу поиска аномалий в данных в объекте исследования «Рекламная компания».
13. Используя методологию DataMining, решите задачу кластеризации социально-экономических данных субъектов РФ.

14. Используя методологию DataMining, решите задачу классификации социально-экономических данных субъектов РФ.
15. Используя методологию DataMining, решите задачу прогнозирования данных социально-экономических данных региона.
16. 16.Используя методологию DataMining, решите задачу поиска аномалий в данных социально-экономических данных региона.
17. Осуществить поиск информации по заданной теме в сети Интернет. Сохранить данные. Загрузить данные в СУБД Access и распределить по таблицам, выстроив связи между объектами. Сделать отчет на основе полученной информации
18. Используя методологию Data Mining, решите задачу кластеризации данных в объекте исследования «Страховая компания».
19. Используя методологию Data Mining, решите задачу классификации данных в объекте исследования «Страховая компания».
20. Используя методологию Data Mining, решите задачу прогнозирования данных в объекте исследования «Страховая компания».
21. Используя методологию Data Mining, решите задачу поиска аномалий в данных в объекте исследования «Страховая компания».
22. Используя методологию Data Mining, решите задачу кластеризации данных в объекте исследования «Банк».
23. Используя методологию Data Mining, решите задачу классификации данных в объекте исследования «Банк».
24. Используя методологию Data Mining, решите задачу прогнозирования данных в объекте исследования «Банк».
25. Используя методологию Data Mining, решите задачу поиска аномалий в данных в объекте исследования «Банк».

**Показатели и критерии оценивания планируемых результатов освоения компетенций и результатов обучения,  
шкала оценивания**

Шкала оценивания		Формируемые компетенции	Индикатор достижения компетенции	Критерии оценивания	Уровень освоения компетенций
85 – 100 баллов	«отлично»/ «зачтено»	УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1. Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи	Знает верно и в полном объеме основные методы критического анализа и основы системного подхода как общенаучного метода.  Умеет верно и в полном объеме анализировать задачу, используя основы критического анализа и системного подхода; осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации; использовать современный инструментарий и интеллектуальные информационно-аналитические системы	Продвинутый
		ОПК-2. Способен осуществлять сбор, обработку и статистический	ОПК-2.1. Использует основные методы, средства	Знает методы поиска и систематизации информации об экономических процессах и явлениях.	



		анализ данных, необходимых для решения поставленных экономических задач.	получения, представления, хранения и обработки статистических данных.	<b>Умеет верно и в полном объеме</b> работать с национальными и международными базами данных с целью поиска информации, необходимой для решения поставленных экономических задач; рассчитывать экономические и социально-экономические показатели, характеризующие деятельность хозяйствующих субъектов на основе типовых методик и действующей нормативно-правовой базы; представить наглядную визуализацию данных.	
		ОПК-6. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-6.1. Использует соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы	<b>Знает верно и в полном объеме::</b> характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий	
<b>Умеет верно и в полном объеме:</b> использовать современные цифровые информационные технологии для решения задач профессиональной деятельности					
<b>Знает верно и в полном объеме:</b> принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий					
			ОПК-6.2. Понимает принципы работы современных цифровых информационных технологий	<b>Умеет верно и в полном</b>	

			технологий, соответствующих содержанию профессиональных задач	<b>объеме:</b> применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий	
<b>70 – 84 баллов</b>	<b>«хорошо»/ «зачтено»</b>	<b>УК-1.</b> Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	<b>УК-1.1.</b> Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи	<b>Знает с незначительными замечаниями</b> основные методы критического анализа и основы системного подхода как общенаучного метода.	<b>Повышенный</b>
				<b>Умеет с незначительными замечаниями</b> анализировать задачу, используя основы критического анализа и системного подхода; осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации; использовать современный инструментарий и интеллектуальные информационно-аналитические системы	
		<b>ОПК-2.</b> Способен осуществлять сбор, обработку и анализ данных, необходимых для решения поставленных управленческих задач, с	<b>ОПК-2.1.</b> Определяет источники информации и осуществляет их поиск на основе поставленных целей для решения профессиональных	<b>Знает с незначительными замечаниями:</b> методы сбора информации, способы и вид ее представления, применяя современное программное обеспечение	
				<b>Умеет с незначительными замечаниями :</b> использовать современный инструментарий и	

		использованием современного инструментария и интеллектуальных информационно-аналитических систем	х задач	интеллектуальные информационно-аналитические системы	
		ОПК-6. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-6.1. Использует соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы	Знает с незначительными замечаниями: характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий	
	Умеет с незначительными замечаниями: использовать современные цифровые информационные технологии для решения задач профессиональной деятельности				
	ОПК-6.2. Понимает принципы работы современных цифровых информационных технологий, соответствующих содержанию профессиональных задач			Знает с незначительными замечаниями: принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий	Умеет с незначительными замечаниями: применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий
<b>50 – 69 баллов</b>	<b>«удовлетворительно»/ «зачтено»</b>	<b>УК-1. Способен осуществлять</b>	<b>УК-1.1. Осуществляет</b>	<b>Знает на базовом уровне с ошибками основные методы</b>	<b>Базовый</b>

		поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	поиск необходимой информации, опираясь на результаты анализа поставленной задачи	критического анализа и основы системного подхода как общенаучного метода. <b>Умеет с незначительными замечаниями</b> анализировать задачу, используя основы критического анализа и системного подхода; осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации; использовать современный инструментарий и интеллектуальные информационно-аналитические системы	
		<b>ОПК-2.</b> Способен осуществлять сбор, обработку и анализ данных, необходимых для решения поставленных управленческих задач, с использованием современного инструментария и интеллектуальных информационно-аналитических систем	<b>ОПК-2.1.</b> Определяет источники информации и осуществляет их поиск на основе поставленных целей для решения профессиональных задач	<b>Знает на базовом уровне, с ошибками:</b> методы сбора информации, способы и вид ее представления, применяя современное программное обеспечение <b>Умеет на базовом уровне, с ошибками:</b> использовать современный инструментарий и интеллектуальные информационно-аналитические системы	
		<b>ОПК-6.</b> Способен понимать	<b>ОПК-6.1.</b> Использует	<b>Знает на базовом уровне, с ошибками:</b> характеристики	

		<p>принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности</p>	<p>соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы</p> <p><b>ОПК-6.2.</b> Понимает принципы работы современных цифровых информационных технологий, соответствующих содержанию профессиональных задач</p>	<p>соответствующих содержанию профессиональных задач современных цифровых информационных технологий</p> <p><b>Умеет на базовом уровне, с ошибками:</b> использовать современные цифровые информационные технологии для решения задач профессиональной деятельности</p> <p><b>Знает на базовом уровне, с ошибками:</b> принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий</p> <p><b>Умеет на базовом уровне, с ошибками:</b> применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий</p>	
<p><b>менее 50 баллов</b></p>	<p><b>«неудовлетворительно»</b> / <b>«не зачтено»</b></p>	<p><b>УК-1.</b> Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</p>	<p><b>УК-1.1.</b> Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи</p>	<p><b>Не знает на базовом уровне</b> основные методы критического анализа и основы системного подхода как общенаучного метода.</p> <p><b>Не умеет на базовом уровне</b> анализировать задачу, используя основы критического анализа и системного подхода; осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая</p>	<p><b>Компетенции не сформированы</b></p>

				надежность различных источников информации; использовать современный инструментарий и интеллектуальные информационно-аналитические системы	
		<b>ОПК-5.</b> Способен использовать при решении профессиональных задач современные информационные технологии и программные средства, включая управление крупными массивами данных и их интеллектуальный анализ	<b>ОПК-5.2.</b> Применяет современные информационные технологии и системы для постановки и решения задач управления, включая управление крупными массивами данных и их интеллектуальный анализ	<b>Не знает на базовом уровне:</b> особенности использования современных информационных технологий и систем для постановки и решения задач управления, включая управление крупными массивами данных и их интеллектуальный анализ	
				<b>Не умеет на базовом уровне:</b> решать задачи управления на основе использования современных информационных технологий и систем	
		<b>ОПК-6.</b> Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	<b>ОПК-6.1.</b> Использует соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы	<b>Не знает на базовом уровне::</b> характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий	
				<b>Не умеет на базовом уровне:</b> использовать современные цифровые информационные технологии для решения задач профессиональной деятельности; принципы работы соответствующих содержанию профессиональных	

			<p>ОПК-6.2.  Понимает  принципы работы  современных  цифровых  информационных  технологий,  соответствующих  содержанию  профессиональн  ых задач</p>	<p>задач современных цифровых  информационных технологий;  применять принципы работы  соответствующих содержанию  профессиональных задач  современных цифровых  информационных технологий</p>	
--	--	--	---	---	--