

**ГАОУ ВО «ДАГЕСТАНСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ НАРОДНОГО ХОЗЯЙСТВА»**

*Утверждены решением
Ученого совета ДГУНХ,
протокол № 12
от 30 мая 2024 г*

**КАФЕДРА «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ»**

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

ПО ДИСЦИПЛИНЕ

«ТЕХНОЛОГИИ АНАЛИТИКИ БОЛЬШИХ ДАННЫХ»

**Направление подготовки 09.04.03 Прикладная информатика,
профиль «Информационные системы в экономике и управлении»**

**Уровень высшего образования- магистратура
Формы обучения – очная, заочная**

Махачкала – 2024

УДК 681.3 (076)
ББК 32.97 я 7 П 13

Составитель – Кобзаренко Дмитрий Николаевич, доктор технических наук, профессор кафедры «Информационные технологии и информационная безопасность» ДГУНХ.

Внутренний рецензент – Раджабов Карахан Якубович, кандидат экономических наук, доцент, доцент кафедры информационных технологий и информационной безопасности ДГУНХ.

Внешний рецензент – Абдурагимов Гусейн Эльдарханович, кандидат физико-математических наук, доцент кафедры "Математические методы в экономике" Дагестанского государственного университета.

Представитель работодателя - Сайидахмедов Сайидахмед Сергеевич, генеральный директор компании «Текама».

Оценочные материалы по дисциплине «Технологии аналитики больших данных» разработаны соответствии с требованиями федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.04.03 «Прикладная информатика», утвержденного приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г., № 916, в соответствии с приказом Министерства образования и науки Российской Федерации от 06.04.2021 г. № 245 «Об утверждении Порядка организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры»

Оценочные материалы по дисциплине «Технологии аналитики больших данных» размещены на официальном сайте www.dgunh.ru

Кобзаренко Д.Н. Оценочные материалы по дисциплине «Технологии аналитики больших данных» для направления подготовки 09.04.03 Прикладная информатика, профиль «Информационные системы в экономике и управлении». – Махачкала: ДГУНХ, 2024. – 29 с.

Рекомендованы к утверждению Учебно-методическим советом ДГУНХ 28 мая 2024 г.

Рекомендованы к утверждению руководителем основной профессиональной образовательной программы высшего образования – программы магистратуры по направлению подготовки 09.04.03 Прикладная информатика, профиль «Информационные системы в экономике и управлении», к.э.н., доцентом Раджабовым К.Я.

Одобрены на заседании кафедры «Информационные технологии и информационная безопасность» 23 мая 2024 г., протокол № 10.

Содержание

НАЗНАЧЕНИЕ ОЦЕНОЧНЫХ МАТЕРИАЛОВ	4
РАЗДЕЛ 1. Перечень компетенций с указанием видов оценочных средств в процессе освоения дисциплины.....	5
1.1 Перечень формируемых компетенций.....	5
1.2 Перечень компетенций с указанием видов оценочных средств	5
РАЗДЕЛ 2. Задания, необходимые для оценки планируемых результатов обучения по дисциплине.....	11
РАЗДЕЛ 3. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания	23
РАЗДЕЛ 4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков, характеризующие этапы формирования компетенций	26

НАЗНАЧЕНИЕ ОЦЕНОЧНЫХ МАТЕРИАЛОВ

Оценочные материалы для текущего контроля успеваемости (оценивания хода освоения дисциплин), для проведения промежуточной аттестации (оценивания промежуточных и окончательных результатов обучения по дисциплине) обучающихся по дисциплине «Технологии аналитики больших данных» на соответствие их учебных достижений поэтапным требованиям образовательной программы высшего образования 09.04.03 Прикладная информатика, профиль «Информационные системы в экономике и управлении».

Оценочные материалы по дисциплине «Технологии аналитики больших данных» включают в себя: перечень компетенций с указанием видов оценочных средств в процессе освоения дисциплины; описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания; типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения ОПОП; методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.

Оценочные материалы сформированы на основе ключевых принципов оценивания:

- валидности: объекты оценки должны соответствовать поставленным целям обучения;
- надежности: использование единообразных стандартов и критериев для оценивания достижений;
- объективности: разные обучающиеся должны иметь равные возможности для достижения успеха.

Основными параметрами и свойствами оценочных материалов являются:

- предметная направленность (соответствие предмету изучения конкретной дисциплины);
- содержание (состав и взаимосвязь структурных единиц, образующих содержание теоретической и практической составляющих дисциплины);
- объем (количественный состав оценочных материалов);
- качество оценочных материалов в целом, обеспечивающее получение объективных и достоверных результатов при проведении контроля с различными целями.

РАЗДЕЛ 1. Перечень компетенций с указанием видов оценочных средств в процессе освоения дисциплины

1.1 Перечень формируемых компетенций

Код компетенции	Формулировка компетенции
ОПК	ОБЩЕПРОФЕССИОНАЛЬНЫЕ КОМПЕТЕНЦИИ
ОПК-2	Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач

1.2 Перечень компетенций с указанием видов оценочных средств

Формируемые компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Уровни освоения компетенций	Критерии оценивания сформированности компетенций	Виды оценочных средств
ОПК-2. Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	ОПК-2.1 Имеет представление о науке работы с данными и понимает базовые алгоритмы машинного обучения	<u>Знать:</u> основы науки о данных, в т.ч. основные понятия больших данных, базовые методы машинного обучения	Пороговый уровень	Обучающийся слабо (частично) знает основы науки о данных, в т.ч. основные понятия больших данных, базовые методы машинного обучения	Блок А – задания репродуктивного уровня - тестовые задания; - вопросы для обсуждения.
			Базовый уровень	Обучающийся с незначительными ошибками и отдельными пробелами знает основы науки о данных, в т.ч. основные понятия больших данных, базовые методы машинного обучения	
			Продвинутый уровень	Обучающийся с требуемой степенью полноты и точности знает основы науки о данных, в т.ч. основные понятия больших дан-	

Формируемые компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Уровни освоения компетенций	Критерии оценивания сформированности компетенций	Виды оценочных средств
				ных, базовые методы машинного обучения	
		<u>Уметь:</u> получать новые знания на основе методов анализа данных	Пороговый уровень	Обучающийся слабо (частично) умеет получать новые знания на основе методов анализа данных	Блок В – задания для реконструктивного уровня - рефераты.
	Базовый уровень		Обучающийся с незначительными затруднениями умеет получать новые знания на основе методов анализа данных		
	Продвинутый уровень		Обучающийся умеет получать новые знания на основе методов анализа данных		
		<u>Владеть:</u> навыками обоснованно предлагать базовые алгоритмы машинного обучения для решения практически задач в области больших данных в экономике и управлении.	Пороговый уровень	Обучающийся слабо (частично) владеет навыками обоснованно предлагать базовые алгоритмы машинного обучения для решения практически задач в области больших данных в экономике и управлении.	Блок С – задания для практико-ориентированного уровня - лабораторные задания.
	Базовый уровень		Обучающийся с незначительными затруднениями владеет навыками обоснованно предлагать базовые алгоритмы машинного обучения для решения практически задач в области больших данных в		

Формируемые компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Уровни освоения компетенций	Критерии оценивания сформированности компетенций	Виды оценочных средств
				экономике и управлении.	
			Продвинутый уровень	Обучающийся свободно владеет навыками обоснованно предлагать базовые алгоритмы машинного обучения для решения практически задач в области больших данных в экономике и управлении.	
	ОПК-2.2 Использует современные интеллектуальные технологии для разработки, обучения и развертывания моделей машинного обучения для решения профессиональных задач	<u>Знать:</u> современные интеллектуальные технологии для решения профессиональных задач.	Пороговый уровень	Обучающийся слабо (частично) знает современные интеллектуальные технологии для решения профессиональных задач.	Блок А – задания репродуктивного уровня - тестовые задания; - вопросы для обсуждения.
Базовый уровень			Обучающийся с незначительными ошибками и отдельными пробелами знает современные интеллектуальные технологии для решения профессиональных задач.		
Продвинутый уровень			Обучающийся с требуемой степенью полноты и точности знает современные интеллектуальные технологии для решения профессиональных задач.		
		<u>Уметь:</u> обосновывать выбор современных интеллектуальных технологий и программной среды при	Пороговый уровень	Обучающийся слабо (частично) умеет обосновывать выбор современных интеллектуальных технологий и программной среды	Блок В – задания для репродуктивного уровня - тесты

Формируемые компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Уровни освоения компетенций	Критерии оценивания сформированности компетенций	Виды оценочных средств
		разработке оригинальных программных средств для решения профессиональных задач		при разработке оригинальных программных средств для решения профессиональных задач	-задачи.
	Базовый уровень		Обучающийся с незначительными затруднениями умеет обосновывать выбор современных интеллектуальных технологий и программной среды при разработке оригинальных программных средств для решения профессиональных задач		
	Продвинутый уровень		Обучающийся умеет обосновывать выбор современных интеллектуальных технологий и программной среды при разработке оригинальных программных средств для решения профессиональных задач		
		<u>Владеть:</u> навыками разработки оригинальных программных средств, в том числе с использованием современных технологий, для решения	Пороговый уровень	Обучающийся слабо (частично) владеет навыками разработки оригинальных программных средств, в том числе с использованием современных технологий, для решения профессиональных задач.	Блок С – задания для практико-ориентированного уровня - лабораторные задания.

Формируемые компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Уровни освоения компетенций	Критерии оценивания сформированности компетенций	Виды оценочных средств
		профессиональных задач.	Базовый уровень	Обучающийся с незначительными затруднениями владеет навыками разработки оригинальных программных средств, в том числе с использованием современных технологий, для решения профессиональных задач.	
			Продвинутый уровень	Обучающийся свободно владеет навыками разработки оригинальных программных средств, в том числе с использованием современных технологий, для решения профессиональных задач.	
	ОПК-2.3 Разрабатывает модели бизнес-процессов, анализирует и оптимизирует бизнес-процессы с применением инструментальных средств их визуального представления	<u>Знать:</u> современные инструментальные средства разработки бизнес-процессов.	Пороговый уровень	Обучающийся слабо (частично) знает современные инструментальные средства разработки бизнес-процессов.	Блок А – задания репродуктивного уровня - тестовые задания; - вопросы для обсуждения.
Базовый уровень			Обучающийся с незначительными ошибками и отдельными пробелами знает современные инструментальные средства разработки бизнес-процессов.		
Продвинутый уровень			Обучающийся с требуемой степенью полноты и точности знает современные инструментальные средства разработки бизнес-процессов.		

Формируемые компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Уровни освоения компетенций	Критерии оценивания сформированности компетенций	Виды оценочных средств	
		<u>Уметь:</u> обосновывать выбор современных инструментальных средств для решения профессиональных задач	Пороговый уровень	Обучающийся слабо (частично) умеет обосновывать выбор современных инструментальных средств для решения профессиональных задач	Блок В – задания для реконструктивного уровня - тесты -задачи.	
			Базовый уровень	Обучающийся с незначительными затруднениями умеет обосновывать выбор современных инструментальных средств для решения профессиональных задач		
			Продвинутый уровень	Обучающийся умеет обосновывать выбор современных инструментальных средств для решения профессиональных задач		
		<u>Владеть:</u> навыками разработки, анализа и оптимизации бизнес-процессов	Пороговый уровень	Обучающийся слабо (частично) владеет навыками разработки, анализа и оптимизации бизнес-процессов		Блок С – задания для практико-ориентированного уровня - лабораторные задания.
			Базовый уровень	Обучающийся с незначительными затруднениями владеет навыками разработки, анализа и оптимизации бизнес-процессов		
			Продвинутый уровень	Обучающийся свободно владеет навыками разработки, анализа и оптимизации бизнес-процессов		

РАЗДЕЛ 2. Задания, необходимые для оценки планируемых результатов обучения по дисциплине

Для проверки сформированности компетенции ОПК-2. Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач, **индикатор достижения компетенции ОПК-2.1** Имеет представление о науке работы с данными и понимает базовые алгоритмы машинного обучения

Блок А. Задания репродуктивного уровня («знать»)

А.1 Тестовые задания

1. Суммарная оперативная память IBM Watson составляет порядка:

- А) 100 гигабайт
- Б) 5000 терабайт
- В) 10 зетабайт
- Г) 15 терабайт

2. Кто ввел термин Большие данные?

- А) Клиффорд Линч
- Б) Алан Тьюринг
- В) Бьерн Страуструп
- Г) Дональд Кнут

3. Какие данные занимают больше мировой памяти относительно остальных?

- А) Structured Data
- Б) Unstructured Data
- В) Semi-Structured Data
- Г) Quasi-Structured Data

4. BigData – это ...

- А) Представление фактов, понятий или инструкций в форме, приемлемой для интерпретации, или обработки.
- Б) Комплексный набор методов обработки структурированных и неструктурированных данных колоссальных объемов.
- В) Колоссальный объем данных, собранных человечеством.
- Г) Класс в Java, предназначенный для хранения данных от 100 Гб

5. Какая компания создала технологию MapReduce?

- А) Google
- Б) Yahoo
- В) EMC
- Г) Oracle

6. Данные текстовых файлов с определенными паттернами для их обработки (например, XML) являются:

- А) Структурированными
- Б) Полуструктурированными
- В) Квазиструктурированными
- Г) Неструктурированными

7. Что означает термин «Big Data» в информационных технологиях?

- А) Комплексный набор методов для создания файлов большого объема
- Б) Комплексный набор методов обработки структурированных и неструктурированных данных колоссальных объемов.
- В) Файлы с большим количеством данных.
- Г) Представление времени, дня, месяца и года в качестве значения количества миллисекунд, прошедших с начала нашей эры.

8. Данные имеющие определенный тип, формат и структуру (например, транзакционные данные) являются:

- А) Структурированными
- Б) Полуструктурированными
- В) Квазиструктурированными
- Г) Неструктурированными

9. Чему примерно равен объем всей существующей на земле информации (в байтах)?

- А) 10^{11}
- Б) 10^{21}
- В) $10^{1010101}$
- Г) 10^{171}

10. В каком году впервые был введен термин Большие данные?

- А) 2002
- Б) 2004
- В) 2006
- Г) 2008

11. Что является средством анализа в BI?

- А) Карты показателей;
- Б) Совместная работа и управление рабочими процессами;
- В) Информационные панели;
- Г) BI инфраструктура.

12. Основное умение исследователя данных?

- А) Умение находить наиболее важные элементы в хранимой информации
- Б) Уметь прогнозировать исход работы системы
- В) Находить скрытые логические связи в системе собранной информации

Г) Отличать неструктурированные данные от структурированных

13. Какой язык программирования из перечисленных является наиболее важным для аналитика?

- А) C++
- Б) PHP
- В) F#
- Г) R

14. Что означает термин «Business Intelligence» в информационных технологиях?

- А) Комплексный набор методов для создания бизнес планов.
- Б) Методы и инструменты для перевода необработанной информации в осмысленную, удобную для восприятия форму.
- В) Файлы, содержащие информацию о бизнес плане.
- Г) Технологии, направленные на развитие бизнеса.

15. Языком, на котором был разработан RabbitMQ, является:

- А) Java
- Б) Python
- В) C++
- Г) Erlang

16. Что является главным результатом процесса Business Intelligence?

- А) Возможность принятия решений для бизнеса
- Б) Результаты интеллектуального анализа данных
- В) Возможность использования искусственного интеллекта
- Г) Получение структуризации данных после выполнения всех шагов процесса

17. Что из перечисленного не является средством анализа?

- А) Продвинутая визуализация
- Б) Reporting
- В) Predictive Modelling
- Г) Data Mining

18. Что относится к средствам предоставления информации в «Business Intelligence»?

- А) Генератор нерегламентированных запросов
- Б) Совместная работа и управление рабочими процессами
- В) Предиктивное моделирование и Data Mining
- Г) Карты показателей

19. Процессом создания и выбора модели для предсказания вероятности наступления некоторого события является:

- А) OLAP

- Б) Data Mining
- В) Predictive Modelling
- Г) Data Science

20. Что не является целью процесса Business Intelligence?

- А) Интерпретация большого количества данных;
- Б) Моделирование исходов различных вариантов действий;
- В) Модификация существующего программного обеспечения;
- Г) Отслеживание результатов решений.

А.2 Вопросы для обсуждения

1. Что означает термин «Big Data» в информационных технологиях?
2. Что является основной целью обработки Big Data?
3. Кто и в каком году впервые ввел термин «Big Data»?
4. Какие главные характеристики Big Data?
5. Какие данные занимают больше мировой памяти относительно остальных?
6. Какие понятия содержит в себе принцип трех "V"?
7. С какого года Большие данные изучаются как академический предмет в вузовских программах по науке о данных?
8. Что является примером квази-структурированных данных?
9. Как назывался первый суперкомпьютер, оснащенный вопросно-ответной системой искусственного интеллекта?
10. Чем характеризуются "Большие данные"?
11. Что является главным результатом процесса Business Intelligence?
12. Что означает термин «Business Intelligence» в информационных технологиях?
13. Расшифруйте аббревиатуру OLAP.
14. Что относится к средствам предоставления информации в Business Intelligence?
15. Что относится к средствам интеграции в «Business Intelligence»?
16. Какие цели ставит перед собой Data Science?
17. Что такое жизненный цикл аналитики данных?
18. Дайте определение термину «предиктивное моделирование»?
19. Что такое ETL?
20. Какова роль BI-аналитика в проекте?

Блок В. Задания реконструктивного уровня («уметь»)

В.1 Рефераты (темы)

1. Современные тенденции развития Big Data в мировой компьютерной индустрии.
2. Современные тенденции развития Big Data в Российской Федерации.
3. Предпосылки развития Big Data в Республике Дагестан.
4. Современное состояние операционной системы Hadoop.

5. Отрасли народного хозяйства Республики Дагестан, в которых может быть задействована операционная система Hadoop.
6. Перечень народно-хозяйственных задач в Республике Дагестан, которые целесообразно решать с помощью технологии Map Reduce.
7. Применение технологий NoSQL в современных бизнес проектах.
8. Перспективы применения среды R для анализа статистических данных, необходимых для принятия решений в управлении городским округом Махачкала.
9. Отрасли народного хозяйства Республики Дагестан, в которых могут быть задействованы системы класса Data Mining.
10. Перспективы применения кластерного анализа и его результатов для управления экономикой Республики Дагестан.

Блок С. Задания практико-ориентированного уровня для диагностирования сформированности компетенций («владеть»)

С.1 Лабораторные работы

Лабораторная работа №1-2

Задание 1:

Познакомиться со средой R-Studio (установка и настройка)

Задание 2:

Научиться работать с данными и готовить их в текстовом редакторе. Освоить команды для чтения данных из текстового файла.

Задание 3:

Изучить графические команды и научиться строить графики. Построить графики динамики курса иностранной валюты за последние две недели.

Блок Д. Задания для использования в рамках промежуточной аттестации

Д.1 Перечень вопросов на экзамен

1. Термин Big Data, цель и главные характеристики больших данных.
2. Принцип трех "V".
3. Структурированные, неструктурированные и квази-структурированные данные.
4. Business Intelligence.
5. Цели и задачи Data Science.
6. Что такое жизненный цикл аналитики данных.

Для проверки сформированности компетенции ОПК-2. Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач, **индикатор достижения компетенции ОПК-2.2** Использует современные интеллектуальные технологии для разработки, обучения и развертывания моделей машинного обучения для решения профессиональных задач

Блок А. Задания репродуктивного уровня («знать»)

А.1 Тестовые задания

1. Что из этого не является реализацией Hadoop?

- А) Google MapReduce
- Б) Phoenix
- В) GreenMint
- Г) Qizmt

2. Какие из перечисленных пунктов являются достоинствами MapReduce?

- А) Оптимальная производительность
- Б) Эффективное применение в маленьких кластерах с небольшим объемом данных
- В) Масштабируемость
- Г) Отказоустойчивость

3. Какие компоненты являются частями MapReduce?

- А) Task Tracker
- Б) Name Node и Data Node
- В) Job Tracker и Task Tracker
- Г) Job Tracker, Task Tracker, Name Node и Data Node

4. Что такое Spark?

- А) Инструмент для кластерных вычислений
- Б) Графический движок
- В) Библиотека для работы с графами
- Г) Технология распределенных вычислений

5. Дайте определение Map Reduce...

- А) Модель распределенных вычислений, предназначенная для параллельных вычислений над очень большими (до нескольких петабайт) объемами данных
- Б) Набор компонентов и интерфейсов для распределенных файловых систем и общего ввода-вывода
- В) Распределенная файловая система, работающая на больших кластерах типовых машин
- Г) Распределенный сервис для коллекционирования, сбора, и перемещения больших массивов данных

6. Что из этого является недостатком MapReduce?

- А) Фиксированный алгоритм обработки данных
- Б) Масштабируемость
- В) Отказоустойчивость
- Г) Возможность автоматического распараллеливания

7. Какое API было добавлено в Hadoop v2.0?

- А) YAWN
- Б) YARN
- В) SARN
- Г) DARN

8. Какая цель у NameNode в HDFS?

- А) Хранить индекс того, какая часть данных находится в каком узле
- Б) Хранить имя файла, хранящегося в конкретном узле
- В) Хранить индекс узла, в котором хранится имя файла
- Г) Хранить имена узлов

9. Вертикальное масштабирование...

- А) Требует изменений в прикладных программах, работающих на таких системах
- Б) Не требует никаких изменений в прикладных программах, работающих на таких системах
- В) Уменьшает производительность каждого компонента БД
- Г) Увеличивает скорость загрузки данных

10. Что, согласно теореме CAP (теореме Брюера), возможно обеспечить в любой реализации распределённых вычислений?

- А) Только согласованность данных
- Б) Только доступность данных
- В) Согласованность данных, доступность данных, устойчивость к разделению
- Г) Не более двух свойств из трёх вышеприведённых

11. Что означает термин NoSQL?

- А) Не SQL
- Б) Не только SQL
- В) Без SQL
- Г) SQL – плохо

12. Разбиение системы на более мелкие структурные компоненты и разнесение их по отдельным физическим машинам (или их группам), и (или) увеличение количества серверов, параллельно выполняющих одну и ту же функцию, это:

- А) Горизонтальное масштабирование
- Б) Вертикальное масштабирование
- В) Master- slave репликация
- Г) Peer-to-peer репликация

13. Выделение таблицы или группы таблиц на отдельный сервер это...

- А) Горизонтальное масштабирование
- Б) Вертикальное масштабирование
- В) Горизонтальный шардинг
- Г) Вертикальный шардинг

А.2 Вопросы для обсуждения

1. В чем преимущества решений на базе Hadoop?
2. Что такое MapReduce, какими достоинствами и недостатками обладает MapReduce?
3. Какому основному принципу следует HDFS и какой размер блока по умолчанию в HDFS?
4. Какие функции выполняет NameNode в HDFS?
5. Какой узел отвечает за репликацию данных в Hadoop?
6. Что является отличительной особенностью NoSQL?
7. Что, согласно теореме CAP, возможно обеспечить в любой реализации распределённых вычислений?
8. Что относится к плюсам репликации?
9. Что относится к преимуществам нереляционных БД?
10. Что такое шардинг?
11. Какие достоинства у Amazon S3?
12. Какие отличия и основные возможности у языка R?
13. В чем особенности хранения в Amazon S3?

Блок В. Задания реконструктивного уровня («уметь»)

В.1 Тестовые задания

Задача №1. Что отобразится на экране консоли R после ввода следующих строк и нажатия Enter?

```
> v1 <- c(seq(4,9,2))
> v2 <- c("D", "B")
> mylist <- list(ch=v2, n=v1)
> v1 <- sort(mylist[[1]], decreasing=FALSE)
> v1
```

Ответ: "B" "D"

Задача №2. Что отобразится на экране консоли R после ввода следующих строк и нажатия Enter?

```
> v1 <- c("Мага", "Патя", "Зурик", "Асхаб", "Зухра")
> v2 <- c("М", "Ж", "М", "М", "Ж")
> v3 <- c(22, 18, 21, 30, 32)
> MyTable <- data.frame(Name=v1, Sex=v2, Age=v3)
> sum(MyTable$Age[(MyTable$Sex == "М") & (MyTable$Age < 29)])
```

Ответ: 43

Задача №3. Что отобразится на экране консоли R после ввода следующих строк и нажатия Enter?

```
> f <- function(n) {
+ n <- n * n + 1
```

```
+ return(n)
+ }
> a <- factor(c(2,1,3,1,1,2), ordered=TRUE)
> levels(a) <- c("Махачкала", "Каспийск", "Кизляр")
> b <- length(a[a == "Махачкала"])
> f(b)
```

Ответ: 10

Блок С. Задания практико-ориентированного уровня для диагностирования сформированности компетенций («владеть»)

С.1 Лабораторные работы

Лабораторная работа №3

Задание 1:

Вычисление общих характеристик выборки данных. Научиться оценивать общую тенденцию, находить ошибки в данных и пропуски, выполнять одномерные статистические тесты.

Лабораторная работа №4

Задание 1:

Анализ временного ряда на предмет прогнозирования будущего с помощью R-Studio. Тренд и период колебаний, построение временного ряда, прогнозирование. Выполнить на примере колебания курса иностранной валюты.

Лабораторная работа №5

Задание 1:

Анализ связей двух выборок данных с помощью R-Studio. Научиться выполнять анализ двух таблиц, строить таблицы сопряженности, выполнять анализ корреляций, выполнять регрессионный анализ. На примере анализа данных цены на нефть и курса доллара США.

Для проверки сформированности компетенции ОПК-2. Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач, **индикатора достижения компетенции ОПК-2.3** Разрабатывает модели бизнес-процессов, анализирует и оптимизирует бизнес-процессы с применением инструментальных средств их визуального представления

Блок А. Задания репродуктивного уровня («знать»)

А.1 Тестовые задания

1. Какая из БД на 100% совместима с интерфейсом языка R?

А) MySQL R

Б) Oracle R

- В) PostgreSQL R
- Г) NoSQL R

2. Функция округления до единиц вверх в языке «R»:

- А) Ceiling(x)
- Б) Floor(x)
- В) Trunc(x)
- Г) Round(x,2)

3. Какой тип лицензии у языка R?

- А) Adware
- Б) Commercial CC
- В) Open source
- Г) Shareware

4. Преподнесение какой-либо полезной информации в форме интересного рассказа – это...

- А) Сторителлинг
- Б) Инфографика
- В) Бизнес аналитика
- Г) Картограмма

5. Основная цель статистического анализа:

- А) Поиск генеральной совокупности
- Б) Выяснение свойств генеральной совокупности
- В) Сравнение генеральных совокупностей
- Г) Выявление последовательности входного набора

6. Определённое предположение о распределении вероятностей, лежащем в основе наблюдаемой выборки данных, - это...

- А) Статистический критерий
- Б) Статистическая выборка
- В) Статистическая гипотеза
- Г) Задача кластеризации

7. С помощью какого алгоритма можно найти ассоциативное правило?

- А) Алгоритм apriori
- Б) Алгоритм k-means
- В) Алгоритм c-means
- Г) Иерархический алгоритм

А.2 Вопросы для обсуждения

1. Что такое дедупликация данных?
2. Какие типы визуализации можно выделить?
3. Какие основные типы Data Mining?

4. Какие категории Web Mining можно выделить?
5. Что изучает статистика?
6. К каким алгоритмам классификации относится метод ближайших соседей?
7. Что является целью кластеризации?

Блок В. Задания реконструктивного уровня («уметь»)

В.1 Тестовые задания

Задача №1. Что отобразится на экране консоли R после ввода следующих строк и нажатия Enter?

```
> k <- seq (from = 1, to = 6, by = 2)
> m <- c(-1, 6)
> v <- c(m, k)
> v <- v+1
> sum (v[v > 3])
```

Ответ: 17

Задача №2. Что отобразится на экране консоли R после ввода следующих строк и нажатия Enter?

```
> f1 <- c(seq(2,4))
> f2 <- c(rep(-1,3))
> f3 <- c(0,0,0)
> m <- rbind(f3,f1,f2)
> m [,3] <- m[,2]+2
> sum(m[3,])
```

Ответ: -1

Блок С. Задания практико-ориентированного уровня для диагностирования сформированности компетенций («владеть»)

С.1 Лабораторные работы

Лабораторная работа №6

Задание 1:

Научиться работать с матрицами списками и таблицами данных в R-Studio.

Лабораторная работа №7

Задание 1:

Вычисление общих характеристик выборки данных. Научиться оценивать общую тенденцию, находить ошибки в данных и пропуски, выполнять одномерные статистические тесты.

Блок Д. Задания для использования в рамках промежуточной аттестации

Д.1 Перечень вопросов на экзамен

1. Технология Apache Hadoop.
2. Технология MapReduce, достоинства и недостатки.
3. Файловая система HDFS.
4. Хранилища NoSQL главные отличия от обычных БД.
5. Теорема CAP и ее следствия.
6. Amazon S3.
7. Язык R.
8. Основные задачи, требования и типы визуализации.
9. Data Mining и основные типы.
10. Классификация и кластеризация.

РАЗДЕЛ 3. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Балльно-рейтинговая система является базовой системой оценивания сформированности компетенций обучающихся очной формы обучения.

Итоговая оценка сформированности компетенций обучающихся в рамках балльно-рейтинговой системы осуществляется в ходе текущего контроля успеваемости, промежуточной аттестации и определяется как сумма баллов, полученных обучающимися в результате прохождения всех форм контроля.

Оценка сформированности компетенций по дисциплине складывается из двух составляющих:

✓ первая составляющая – оценка преподавателем сформированности компетенций в течение семестра в ходе текущего контроля успеваемости (максимум 100 баллов). Структура первой составляющей определяется технологической картой дисциплины, которая в начале семестра доводится до сведения обучающихся;

✓ вторая составляющая – оценка сформированности компетенций обучающихся на экзамене (максимум – 30 баллов)

Для студентов очно-заочной и заочной форм обучения применяются 4-балльная и бинарная шкалы оценивания результатов текущего контроля успеваемости и промежуточной аттестации обучающихся.

уровни освоения компетенций	продвинутый уровень	базовый уровень	пороговый уровень	допороговый уровень
100 – балльная шкала	85 и \geq	70 – 84	51 – 69	0 – 50
4 – балльная шкала	«отлично»	«хорошо»	«удовлетворительно»	«неудовлетворительно»

Шкала оценок при текущем контроле успеваемости по различным показателям

<i>Показатели оценивания сформированности компетенций</i>	<i>Баллы</i>	<i>Оценка</i>
выполнение тестов	0-30	«неудовлетворительно» «удовлетворительно» «хорошо» «отлично»
ответы на устные вопросы	0-15	«неудовлетворительно» «удовлетворительно» «хорошо» «отлично»
выполнение лабораторных заданий	0-15	«неудовлетворительно» «удовлетворительно»

		«хорошо» «отлично»
выполнение рефератов	0-10	«неудовлетворительно» «удовлетворительно» «хорошо» «отлично»

Соответствие критериев оценивания уровню освоения компетенций по текущему контролю успеваемости

Баллы	Оценка	Уровень освоения компетенций	Критерии оценивания
0-50	«неудовлетворительно»	Допороговый уровень	Обучающийся не приобрел знания, умения и не владеет компетенциями в объеме, закрепленном рабочей программой дисциплины
51-69	«удовлетворительно»	Пороговый уровень	Не менее 50% заданий, подлежащих текущему контролю успеваемости, выполнены без существенных ошибок
70-84	«хорошо»	Базовый уровень	Обучающимся выполнено не менее 75% заданий, подлежащих текущему контролю успеваемости, или при выполнении всех заданий допущены незначительные ошибки; обучающийся показал владение навыками систематизации материала и применения его при решении практических заданий; задания выполнены без ошибок
85-100	«отлично»	Продвинутый уровень	100% заданий, подлежащих текущему контролю успеваемости, выполнены самостоятельно и в требуемом объеме; обучающийся проявляет умение обобщать, систематизировать материал и применять его при решении практических заданий; задания выполнены с подробными пояснениями и аргументированными выводами

Шкала оценок по промежуточной аттестации

Наименование формы промежуточной аттестации	Баллы	Оценка
Экзамен	0-30	«неудовлетворительно» «удовлетворительно» «хорошо» «отлично»

Соответствие критериев оценивания уровню освоения компетенций по промежуточной аттестации обучающихся

<i>Баллы</i>	<i>Оценка</i>	<i>Уровень освоения компетенций</i>	<i>Критерии оценивания</i>
0-9	«неудовлетворительно»	Допороговый уровень	Обучающийся не приобрел знания, умения и не владеет компетенциями в объеме, закрепленном рабочей программой дисциплины; обучающийся не смог ответить на вопросы
10-14	«удовлетворительно»	Пороговый уровень	Обучающийся дал неполные ответы на вопросы, с недостаточной аргументацией, практические задания выполнены не полностью, компетенции, осваиваемые в процессе изучения дисциплины сформированы не в полном объеме.
15-17	«хорошо»	Базовый уровень	Обучающийся в целом приобрел знания и умения в рамках осваиваемых в процессе обучения по дисциплине компетенций; обучающийся ответил на все вопросы, точно дал определения и понятия, но затрудняется подтвердить теоретические положения практическими примерами; обучающийся показал хорошие знания по предмету, владение навыками систематизации материала и полностью выполнил практические задания
18-20	«отлично»	Продвинутый уровень	Обучающийся приобрел знания, умения и навыки в полном объеме, закрепленном рабочей программой дисциплины; терминологический аппарат использован правильно; ответы полные, обстоятельные, аргументированные, подтверждены конкретными примерами; обучающийся проявляет умение обобщать, систематизировать материал и выполняет практические задания с подробными пояснениями и аргументированными выводами

РАЗДЕЛ 4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков, характеризующие этапы формирования компетенций

В вопросы на экзамен включены два теоретических вопроса и практическое задание, соответствующие содержанию формируемых компетенций. Экзамен проводится в письменной форме. На ответ и решение задачи студенту отводится 60 минут. За ответ на теоретические вопросы студент может получить максимально – 20 баллов, за решение задачи максимально – 10 баллов.

Тестирование имеет статус контрольной работы, выполняется в письменном виде на тестовых бланках. На выполнение тестирования дается 90 минут.

Методика оценивания выполнения тестов

Баллы	Оценка	Показатели	Критерии
25-30	«отлично»	1. Полнота выполнения тестовых заданий; 2. Своевременность выполнения;	Выполнено 85 % заданий предложенного теста, в заданиях открытого типа дан полный, развернутый ответ на поставленный вопрос.
19-24	«хорошо»	3. Правильность ответов на вопросы; 4. Самостоятельность тестирования.	Выполнено 70 % заданий предложенного теста, в заданиях открытого типа дан полный, развернутый ответ на поставленный вопрос; однако были допущены неточности в определении понятий, терминов и др.
12-18	«удовлетворительно»		Выполнено 50 % заданий предложенного теста, в заданиях открытого типа дан неполный ответ на поставленный вопрос, в ответе не присутствуют доказательные примеры, текст со стилистическими и орфографическими ошибками.
0-11	«неудовлетворительно»		Выполнено до 50% заданий предложенного теста, на поставленные вопросы ответ отсутствует или неполный, допущены существенные ошибки в теоретическом материале (терминах, понятиях).

Устный опрос проводится на практических занятиях. Студенту дается вопрос на основе пройденного материала из перечня вопросов на экзамен. Студент должен выйти к доске и дать развернутый ответ на поставленный вопрос.

Методика оценивания ответов на устные вопросы

Баллы	Оценка	Показатели	Критерии
14-15	«отлично»	1. Полнота данных ответов; 2. Аргументированность данных ответов; 3. Правильность	Полно и аргументировано даны ответы по содержанию задания. Обнаружено понимание материала, может обосновать свои суждения, применить знания на практике, привести необходимые примеры не только по учебнику, но и самостоятельно составлен-

		ответов на вопросы.	ные. Изложение материала последовательно и правильно.
10-13	«хорошо»		Студент дает ответ, удовлетворяющий тем же требованиям, что и для оценки «отлично», но допускает 1-2 ошибки, которые сам же исправляет.
6-9	«удовлетворительно»		Студент обнаруживает знание и понимание основных положений данного задания, но: 1) излагает материал неполно и допускает неточности в определении понятий или формулировке правил; 2) не умеет достаточно глубоко и доказательно обосновать свои суждения и привести свои примеры; 3) излагает материал непоследовательно и допускает ошибки.
0-5	«неудовлетворительно»		Студент обнаруживает незнание ответа на соответствующее задание, допускает ошибки в формулировке определений и правил, искажающие их смысл, беспорядочно и неуверенно излагает материал; отмечаются такие недостатки в подготовке студента, которые являются серьезным препятствием к успешному овладению последующим материалом.

Темы рефератов предлагаются на выбор в начале семестра. Студент выполняет реферат и высылает его на проверку преподавателю. После проверки соответствия реферата и содержимого теме назначается день защиты. Защита реферата происходит в рамках регламента: 10-15 минут выступление, далее ответы на вопросы по теме реферата.

Методика оценивания выполнения рефератов

Баллы	Оценка	Показатели	Критерии
9-10	«отлично»	1. Полнота выполнения рефератов; 2. Своевременность выполнения; 3. Правильность ответов на вопросы.	Выполнены все требования к написанию и защите реферата: обозначена проблема и обоснована её актуальность, сделан краткий анализ различных точек зрения на рассматриваемую проблему и логично изложена собственная позиция, сформулированы выводы, тема раскрыта полностью, выдержан объём, соблюдены требования к внешнему оформлению, даны правильные ответы на дополнительные вопросы.
6-8	«хорошо»		Основные требования к реферату и его защите выполнены, но при этом допущены недочеты. В частности, имеются неточности в изложении материала; отсутствует логическая последовательность в суждениях; не выдержан объём реферата; имеются упущения в оформлении; на дополнительные вопросы при защите даны неполные ответы.

3-5	«удовлетворительно»		Имеются существенные отступления от требований к реферированию. В частности: тема освещена лишь частично; допущены фактические ошибки в содержании реферата или при ответе на дополнительные вопросы.
0-2	«неудовлетворительно»		Тема реферата не раскрыта, обнаруживается существенное непонимание проблемы.

Лабораторные задания выполняются на лабораторных занятиях. Задания выполняются на основе методических указаний к лабораторным работам. Студент должен выполнить лабораторное задание во временном интервале 60 минут. Остальные 30 минут даются на защиту лабораторного задания в рамках которой проверяется правильность и самостоятельность его выполнения.

Методика оценивания выполнения лабораторных заданий

Баллы	Оценка	Показатели	Критерии
14-15	«отлично»	1. Полнота выполнения лабораторного задания; 2. Своевременность выполнения задания; 3. Последовательность и рациональность выполнения задания; 4. Самостоятельность решения.	Задание выполнено самостоятельно. При этом составлен правильный алгоритм решения, в логических рассуждениях и в решении нет ошибок, получен верный ответ, задание решено рациональным способом.
10-13	«хорошо»		Задание выполнено с помощью преподавателя. При этом составлен правильный алгоритм решения, в логических рассуждениях и решении нет существенных ошибок; есть объяснение решения, но задание решено нерациональным способом или допущено не более двух несущественных ошибок, получен верный ответ.
6-9	«удовлетворительно»		Задание выполнено с подсказками преподавателя. При этом задание понято правильно, в логическом рассуждении нет существенных ошибок, но допущены существенные ошибки в алгоритмах и/или расчетах; задание решено не полностью или в общем виде.
0-5	«неудовлетворительно»		Задание не выполнено или выполнено неправильно.

Процедура оценивания – порядок действий при подготовке и проведении аттестационных испытаний и формировании оценки.

Процедура промежуточной аттестации проходит в соответствии с Положением о промежуточной аттестации знаний студентов и учащихся ДГУНХ.

- Аттестационные испытания проводятся преподавателем (или комиссией преподавателей – в случае модульной дисциплины), ведущим лекционные занятия по данной дисциплине, или преподавателями, ведущими практические и лабораторные занятия (кроме устного экзамена). Присутствие посторонних лиц в ходе проведения аттестационных испытаний без разрешения ректора

или проректора по учебной работе не допускается (за исключением работников университета, выполняющих контролирующие функции в соответствии со своими должностными обязанностями). В случае отсутствия ведущего преподавателя аттестационные испытания проводятся преподавателем, назначенным письменным распоряжением по кафедре (структурному подразделению).

- Инвалиды и лица с ограниченными возможностями здоровья, имеющие нарушения опорно-двигательного аппарата, допускаются на аттестационные испытания в сопровождении ассистентов-сопровождающих.
- Во время аттестационных испытаний обучающиеся могут пользоваться программой дисциплины, а также с разрешения преподавателя справочной и нормативной литературой, непрограммируемыми калькуляторами.
- Время подготовки ответа при сдаче зачета/экзамена в устной форме должно составлять не менее 40 минут (по желанию обучающегося ответ может быть досрочным). Время ответа – не более 15 минут.
- При подготовке к устному экзамену экзаменуемый, как правило, ведет записи в листе устного ответа, который затем (по окончании экзамена) сдается экзаменатору.
- При проведении устного экзамена экзаменационный билет выбирает сам экзаменуемый в случайном порядке.
- Экзаменатору предоставляется право задавать обучающимся дополнительные вопросы в рамках программы дисциплины текущего семестра, а также, помимо теоретических вопросов, давать задачи, которые изучались на практических занятиях.
- Оценка результатов устного аттестационного испытания объявляется обучающимся в день его проведения. При проведении письменных аттестационных испытаний или компьютерного тестирования – в день их проведения или не позднее следующего рабочего дня после их проведения.
- Результаты выполнения аттестационных испытаний, проводимых в письменной форме, форме итоговой контрольной работы или компьютерного тестирования, должны быть объявлены обучающимся и выставлены в зачётные книжки не позднее следующего рабочего дня после их проведения.