

федеральное государственное бюджетное образовательное учреждение
высшего образования "Приволжский исследовательский медицинский университет"
Министерства здравоохранения Российской Федерации



УТВЕРЖДАЮ
Проректор по учебной работе
Богомолова Е.С.

24 06 2024 г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

По дисциплине **Большие данные и их обработка**
направление подготовки **09.04.02 Информационные системы и технологии**
профиль **Информационные системы и технологии в здравоохранении**

Квалификация выпускника:

Магистр

Форма обучения:

очно-заочная

Нижний Новгород
2024

Фонд оценочных средств по дисциплине «Большие данные и их обработка» предназначен для контроля знаний по программе магистратуры по направлению подготовки 09.04.02 Информационные системы и технологии, профилю «Информационные системы и технологии в здравоохранении».

1. Паспорт фонда оценочных средств по дисциплине «Большие данные и их обработка»

Компетенция	Результаты обучения и индикаторы достижения компетенций	Виды занятий	Оценочные средства
УК-1	способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий		
	<p>Знать: ИД-3_{УК-1.3} возможности технологий анализа больших данных при проведении научных исследований.</p> <p>Уметь: ИД-6_{УК-1.6} понимать важность применения научных методов для извлечения пользы из больших массивов данных.</p> <p>Владеть: ИД-8_{УК-1.8} научными методами обработки больших данных, в том числе методами интеллектуального анализа больших данных.</p>	Лекции, практические занятия	Контрольная работа Ситуационные задачи Тестирование Собеседование
ОПК-6	способен использовать методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий		
	<p>Знать: ИД-3_{ОПК-6.3} существующие в современном мире источники и типы информации.</p> <p>Уметь: ИД-6_{ОПК-6.6} визуализировать имеющиеся данные, отбрасывать несущественную информацию, структурировать информацию в рамках поставленной задачи.</p> <p>Владеть: ИД-9_{ОПК-6.9} современными средствами визуализации, методами предварительной подготовки данных.</p>	Лекции, практические занятия	Контрольная работа Ситуационные задачи Тестирование Собеседование
ПК-6	способен разрабатывать, вводить в действие и обслуживать базы данных; дополнять, модифицировать и совершенствовать базы данных и другие хранилища информации		
	<p>Знать: ИД-2_{ПК-6.2} существующие современные технологии высоконагруженных систем хранения и обработки данных, принципы работы высоконагруженных систем.</p> <p>Уметь: ИД-6_{ПК-6.6} выбирать подходящую технологию хранения и обработки больших данных, использовать современные высоконагруженные системы хранения и обработки больших данных.</p> <p>Владеть: ИД-10_{ПК-6.10} технологией Map Reduce и ее реализацией Hadoop.</p>	Лекции, практические занятия	Контрольная работа Ситуационные задачи Тестирование Собеседование

Текущий контроль по дисциплине «Большие данные и их обработка» осуществляется в течение всего срока освоения данной дисциплины. Выбор оценочного средства для проведения текущего контроля на усмотрение преподавателя.

Промежуточная аттестация обучающихся по дисциплине «Большие данные и их обработка» проводится по итогам обучения и является обязательной.

2. Критерии и шкала оценивания

Критерии оценивания	Шкала оценивания по системе бальной оценки			
	неудовлетворительно	удовлетворительно	хорошо	отлично
Полнота знаний	Уровень знаний ниже минимальных требований. Имели место грубые ошибки	Минимально допустимый уровень знаний. Допущено много негрубых ошибки	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок
Наличие умений	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки	Продемонстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	Продемонстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами	Продемонстрированы все основные умения, решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме
Наличие навыков (владение опытом)	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами	Продемонстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов
Характеристика сформированности компетенции	Компетенция в полной мере не сформирована. Имеющихся знаний, умений, навыков недостаточно для решения профессиональных задач. Требуется повторное обучение	Сформированность компетенции соответствует минимальным требованиям. Имеющихся знаний, умений, навыков в целом достаточно для решения профессиональных задач, но требуется дополнительная	Сформированность компетенции в целом соответствует требованиям, но есть недочеты. Имеющихся знаний, умений, навыков и мотивации в целом достаточно для решения профессиональных задач, но требуется	Сформированность компетенции полностью соответствует требованиям. Имеющихся знаний, умений, навыков и мотивации в полной мере достаточно

Критерии оценивания	Шкала оценивания по системе бальной оценки			
	неудовлетворительно	удовлетворительно	хорошо	отлично
		практика по большинству практических задач	дополнительная практика по некоторым профессиональным задачам	для решения сложных профессиональных задач
Уровень сформированности компетенций	Низкий	Ниже среднего	Средний	Высокий

3. Оценочные средства

3.1. Текущий контроль

3.1.1. Контролируемый раздел дисциплины «Фундамент анализа данных Big Data»

Перечень вопросов

1. Каковы функции и задачи направления *Dig Data*?
2. В чем смысловое значение характеристики *Volume*?
3. В чем смысловое значение характеристики *Velocity*?
4. В чем смысловое значение характеристики *Variety*?
5. Методы класса *Data Mining* в анализе Больших данных.
6. В чем проявляется разнообразие больших данных *Big Data*?
7. В чем потенциал объединенной технологий Блокчейн и *Big Data*?
8. Перечислите основные особенности, входящие в определение термина Большие данные.
9. Каковы отличительные характеристики распределенной файловой системы *Hadoop* (хадуп)?

3.3.2. Контролируемый раздел дисциплины «Сбор, хранение и обработка данных»

Перечень вопросов

1. Роль машинного обучения в анализе Больших данных.
2. Какие статистические методы используются в анализе Больших данных.
3. В чем суть принципа горизонтальной масштабируемости в качестве базового принципа обработки Больших данных.
4. В чем специфика статистических методов обработки в применении к Большим данным?
5. Какая файловая система предназначена для хранения файлов больших размеров с возможностью потокового доступа к информации, поблочно распределённой по узлам вычислительного кластера?
6. Из каких основных четырех модулей состоит Архитектурная концепция *Hadoop*?
7. Опишите назначение и алгоритм работы Загрузчик данных (*Data Loader*).
8. Опишите алгоритм работы второго этапа обработки больших объемов данных под названием *Reduce*.
9. Опишите методы обработки больших данных. В чем особенность методов интеллектуального анализа больших данных.
10. В чем состоит основной принцип организации работы *HadoopDB* ?

3.1.2. Контролируемый раздел дисциплины «Математика для Data Science»

Перечень вопросов

1. Что из себя представляет модель распределенной обработки данных *MapReduce*?
2. Парадигма *MapReduce* как принцип работы с большими данными.
3. Какие языки программирования используются при работе с Большими данными.
4. Какие стадии включает в себя модель *MapReduce*?
5. Что из себя представляет пакетно-ориентированная система обработки данных *Apache Hadoop*?
6. Опишите принцип работы технологии *Map Reduce*.
7. Опишите алгоритм работы первого этапа обработки больших объемов данных под названием *Map*.
8. Что представляет собой технология *Hadoop*, связанная с обработкой больших данных?
9. В чем отличие технологии *Framework* от библиотеки программных модулей?
10. Что из себя представляет операция *Свертки списка*.

3.2. Промежуточный контроль

Билеты для зачета

Билет №1
<ol style="list-style-type: none"> 1. Методы класса <i>Data Mining</i> в анализе Больших данных. 2. Опишите алгоритм работы второго этапа обработки больших объемов данных под названием <i>Reduce</i>. 3. В чем специфика статистических методов обработки в применении к Большим данным?
Билет №2
<ol style="list-style-type: none"> 1. В чем смысловое значение характеристики <i>Volume</i>? Каковы отличительные характеристики распределенной файловой системы <i>Hadoop</i> (хадуп)? 2. Что из себя представляет модель распределенной обработки данных <i>MapReduce</i>? 3. Что представляет собой технология <i>Hadoop</i>, связанная с обработкой больших данных?
Билет №3
<ol style="list-style-type: none"> 1. Каковы основные функции и общие принципы построения электронной истории болезни (ЭИБ)? Какие задачи решает ЭИБ? 2. Роль машинного обучения в анализе Больших данных. 3. Парадигма <i>MapReduce</i> как принцип работы с большими данными.
Билет №4
<ol style="list-style-type: none"> 1. В чем смысловое значение характеристики <i>Velocity</i>? 2. В чем суть принципа горизонтальной масштабируемости в качестве базового принципа обработки Больших данных. 3. Опишите алгоритм работы первого этапа обработки больших объемов данных под названием <i>Map</i>.
Билет №5

<ol style="list-style-type: none"> 1. Методы класса <i>Data Mining</i> в анализе Больших данных. 2. Опишите назначение и алгоритм работы Загрузчик данных (<i>Data Loader</i>). 3. В чем отличие технологии <i>Framework</i> от библиотеки программных модулей?
Билет №6
<ol style="list-style-type: none"> 1. В чем состоит высокотехнологичный потенциал объединенной технологий Блокчейн и <i>Big Data</i>? 2. В чем состоит основной принцип организации работы <i>HadoopDB</i> ? 3. Что из себя представляет пакетно-ориентированная система обработки данных <i>Apache Hadoop</i>?
Билет №7
<ol style="list-style-type: none"> 1. В чем смысловое значение характеристики <i>Variety</i>? 2. Опишите методы обработки больших данных. В чем особенность методов интеллектуального анализа больших данных. 3. Что из себя представляет операция <i>Свертки списка</i>.
Билет №8
<ol style="list-style-type: none"> 1. В чем проявляется разнообразие больших данных <i>Big Data</i>? 2. Опишите файловую систему, которая предназначена для хранения файлов больших размеров с возможностью потокового доступа к информации, поблочно распределённой по узлам вычислительного кластера ? 3. Какие стадии включает в себя модель <i>MapReduce</i>?
Билет №9
<ol style="list-style-type: none"> 1. Опишите принцип работы технологии <i>Map Reduce</i>. 2. Основные области применения <i>Hadoop</i>. 3. Из каких основных четырех модулей состоит Архитектурная концепция <i>Hadoop</i>?
Билет №10
<ol style="list-style-type: none"> 1. Перечислите основные особенности, входящие в определение термина Большие данные. 2. Какие статистические методы используются в анализе Больших данных. 3. Какие языки программирования используются при работе с Большими данными.

Тестовые вопросы

Тестовые вопросы и варианты ответов	Компетенция, формируемая тестовым вопросом
1. КОНКРЕТНАЯ ФУНКЦИЯ, КОТОРАЯ ИСПОЛЬЗУЕТСЯ ДЛЯ ОБРАБОТКИ БОЛЬШИХ НАБОРОВ РАСПРЕДЕЛЕННЫХ ЗАДАЧ С ИСПОЛЬЗОВАНИЕМ БОЛЬШОГО КОЛИЧЕСТВА КОМПЬЮТЕРОВ, ОБРАЗУЮЩИХ КЛАСТЕР	УК-1

<ol style="list-style-type: none"> 1) <i>Map Reduce</i>; 2) <i>Data Mining</i>; 3) <i>NET Framework</i>; 4) <i>Machine Learning</i>. 	
<p>2. БОЛЬШОЙ СТЕК ТЕХНОЛОГИЙ, ТАК ИЛИ ИНАЧЕ СВЯЗАННЫХ С ОБРАБОТКОЙ БОЛЬШИХ ДАННЫХ НЕ ТОЛЬКО ПРИ ПОМОЩИ <i>MAPREDUCE</i></p> <ol style="list-style-type: none"> 1) <i>Deep Learning</i>; 2) <i>Big Data</i>; 3) <i>Hadoop</i>; 	УК-1
<p>3. ФАЙЛОВАЯ СИСТЕМА ДЛЯ ХРАНЕНИЯ ФАЙЛОВ БОЛЬШИХ РАЗМЕРОВ С ВОЗМОЖНОСТЬЮ ПОТОКОВОГО ДОСТУПА К ИНФОРМАЦИИ, ПОБЛОЧНО РАСПРЕДЕЛЁННОЙ ПО УЗЛАМ ВЫЧИСЛИТЕЛЬНОГО КЛАСТЕРА</p> <ol style="list-style-type: none"> 1) <i>FAT</i>; 2) <i>NTFS</i>; 3) <i>HDFS</i> 	УК-1
<p>4. ОСНОВНЫЕ ПРИЗНАКИ VVV, КОТОРЫЕ ХАРАКТЕРИЗУЮТ СВОЙСТВА БОЛЬШИХ ДАННЫХ</p> <ol style="list-style-type: none"> 1) <i>Vocabulary, Velocity, Variety</i>, 2) <i>Volume, Velocity, Variety</i>, 3) <i>Visuality, Velocity, Variety</i>, 4) <i>Virtualiry, Visuality, Vocabulary</i> 	УК-1
<p>5. ЯЗЫКИ ПРОГРАММИРОВАНИЯ, КОТОРЫЕ ИСПОЛЬЗУЮТСЯ ДЛЯ РАБОТЫ С БОЛЬШИМИ ДАННЫМИ?</p> <ol style="list-style-type: none"> 1) <i>ALGOL</i> 2) <i>Visual BASIC for applications</i> 3) <i>FORTRAN</i> 4) <i>Phyton</i> 5) <i>C⁺⁺</i> 	УК-1
<p>6. ТЕХНОЛОГИЯ, ПРЕДСТАВЛЯЮЩАЯ СОБОЙ СВОБОДНО РАСПРОСТРАНЯЕМЫЙ НАБОР УТИЛИТ, БИБЛИОТЕК И ФРЕЙМВОРК ДЛЯ РАЗРАБОТКИ И ВЫПОЛНЕНИЯ РАСПРЕДЕЛЁННЫХ ПРОГРАММ, РАБОТАЮЩИХ НА КЛАСТЕРАХ ИЗ СОТЕН И ТЫСЯЧ УЗЛОВ</p>	УК-1

<ul style="list-style-type: none"> 1) <i>Hadoop</i> 2) <i>Software</i> 3) <i>Big Data.</i> 	
<p>7. ТЕХНОЛОГИЯ, ПРЕДСТАВЛЯЮЩАЯ СОБОЙ СПОСОБНОСТЬ СИСТЕМЫ, СЕТИ ИЛИ ПРОЦЕССА СПРАВЛЯТЬСЯ С УВЕЛИЧЕНИЕМ РАБОЧЕЙ НАГРУЗКИ (УВЕЛИЧИВАТЬ СВОЮ ПРОИЗВОДИТЕЛЬ-НОСТЬ) ПРИ ДОБАВЛЕНИИ РЕСУРСОВ (ОБЫЧНО АППАРАТНЫХ), ВОЗМОЖНОСТЬ НАРАЩИВАНИЯ ДОПОЛНИТЕЛЬНЫХ РЕСУРСОВ БЕЗ СТРУКТУРНЫХ ИЗМЕНЕНИЙ ЦЕНТРАЛЬНОГО УЗЛА СИСТЕМЫ</p> <ul style="list-style-type: none"> 1) <i>Ускоряемость</i> 2) <i>Масштабируемость</i> 3) <i>Мультиплексность ;</i> 4) <i>Овертомность.</i> 	ОПК-6
<p>8. ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ, ФРЕЙМВОРК, СОСТОЯЩИЙ ИЗ НАБОРА УТИЛИТ ДЛЯ РАЗРАБОТКИ И ВЫПОЛНЕНИЯ ПРОГРАММ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ</p> <ul style="list-style-type: none"> 1) <i>Hadoop</i> 2) <i>Crowdsourcing</i> 3) <i>Data Mining</i> 4) <i>Data Reducing</i> 	ОПК-6
<p>9. УКАЖИТЕ 3 ЭТАПА, НА КОТОРЫЕ ДЕЛИТСЯ РАБОТА С БОЛЬШИМИ ДАННЫМИ</p> <ul style="list-style-type: none"> 1) <i>сбор, хранение, обработка</i> 2) <i>анализ, преобразование, использование</i> 3) <i>интеграция, управление, анализ</i> 	ОПК-6
<p>10. ЭТАП РАБОТЫ С БОЛЬШИМИ ДАННЫМИ, НА КОТОРОМ КОМПАНИЯ ИНТЕГРИРУЕТ В СВОЮ РАБОТУ ТЕХНОЛОГИИ И СИСТЕМЫ, ПОЗВОЛЯЮЩИЕ СОБИРАТЬ БОЛЬШИЕ ОБЪЕМЫ ИНФОРМАЦИИ ИЗ РАЗНЫХ ИСТОЧНИКОВ; ВНЕДРЯЮТСЯ МЕХАНИЗМЫ ОБРАБОТКИ И ФОРМАТИРОВАНИЯ ДАННЫХ ДЛЯ УПРОЩЕНИЯ РАБОТЫ АНАЛИТИКОВ С «BIG DATA»</p> <ul style="list-style-type: none"> 1) <i>консолидация</i> 2) <i>интеграция</i> 	ОПК-6

<p>3) <i>систематизация</i> 4) <i>автоматизация</i></p>	
<p>11. ЭТАП РАБОТЫ С БОЛЬШИМИ ДАННЫМИ, НА КОТОРОМ ДО НАЧАЛА РАБОТЫ С ДАННЫМИ ПРИНИМАЕТСЯ РЕШЕНИЕ О МЕСТЕ ИХ ХРАНЕНИЯ НА ОСНОВЕ МНОЖЕСТВА КРИТЕРИЕВ, ГЛАВНЫМИ ИЗ КОТОРЫХ СЧИТАЮТСЯ ПРЕДПОЧТЕНИЯ ПО ФОРМАТУ И ТЕХНОЛОГИИ ОБРАБОТКИ</p> <p>1) <i>обоснование</i> 2) <i>управление</i> 3) <i>оформление</i> 4) <i>согласование</i></p>	ОПК-6
<p>12. ЗАКЛЮЧИТЕЛЬНЫЙ 3-й ЭТАП РАБОТЫ С БОЛЬШИМИ ДАННЫМИ, НА КОТОРОМ ЭТИ ДАННЫЕ НАЧИНАЮТ ПРИНОСИТЬ ПОЛЬЗУ В РЕЗУЛЬТАТЕ ПРИМЕНЕНИЯ ВЫЧИСЛИТЕЛЬНЫХ ТЕХНОЛОГИЙ ДЛЯ ПОЛУЧЕНИЯ НАИБОЛЕЕ ЦЕННОГО РЕЗУЛЬТАТА ДЛЯ ПРЕДМЕТНОЙ ОБЛАСТИ ИХ ПРИМЕНЕНИЯ</p> <p>1) <i>анализ</i> 2) <i>внедрение</i> 3) <i>приспособление</i> 4) <i>таргетирование</i></p>	ОПК-6
<p>13. УКАЖИТЕ ПРАВИЛЬНЫЙ ТЕРМИН, ХАРАКТЕРИЗУЮЩИЙ ДОСТОВЕРНОСТЬ КАК САМОГО НАБОРА ДАННЫХ, ТАК И РЕЗУЛЬТАТОВ ЕГО АНАЛИЗА</p> <p>1) <i>Veracity</i> 2) <i>Verbality</i> 3) <i>Verify</i> 4) <i>Vtrification</i></p>	ОПК-6
<p>14. УКАЖИТЕ ПРАВИЛЬНЫЙ ТЕРМИН, ХАРАКТЕРИЗУЮЩИЙ ИЗМЕНЧИВОСТЬ ПОТОКОВ ДАННЫХ, У КОТОРЫХ БЫВАЮТ СВОИ ПИКИ И СПАДЫ ПОД ВЛИЯНИЕМ СЕЗОНОВ ИЛИ СОЦИАЛЬНЫХ ЯВЛЕНИЙ, ЧТО В ЗНАЧИТЕЛЬНОЙ СТЕПЕНИ ОБУСЛАВЛИВАЕТ. СЛОЖНОСТЬ ИХ АНАЛИЗА</p> <p>1) <i>Variability</i></p>	ОПК-6

<p>2) <i>Virtuality</i> 3) <i>Verify</i> 4) <i>Voluntary</i></p>	
<p>15. МОДЕЛЬ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ, КОТОРАЯ ВКЛЮЧАЕТ ДВА ЭТАПА: 1) СНАЧАЛА ДАННЫЕ ФИЛЬТРУЮТСЯ ПО УСЛОВИЯМ, КОТОРЫЕ ЗАДАЕТ ИССЛЕДОВАТЕЛЬ, СОРТИРУЮТСЯ И РАСПРЕДЕЛЯЮТСЯ МЕЖДУ ОТДЕЛЬНЫМИ КОМПЬЮТЕРАМИ (УЗЛАМИ КЛАСТЕРА СЕТИ); 2) ЗАТЕМ ЭТИ УЗЛЫ ПАРАЛЛЕЛЬНО РАССЧИТЫВАЮТ СВОИ БЛОКИ ДАННЫХ И ПЕРЕДАЮТ РЕЗУЛЬТАТ ВЫЧИСЛЕНИЙ НА СЛЕДУЮЩУЮ ИТЕРАЦИЮ 1) <i>MapReduce</i> 2) <i>Machine Learning</i> 3) <i>Data Manufacturing</i> 4) <i>Data Management</i></p>	ОПК-6
<p>16. ИНТЕРНЕТ-МОДЕЛЬ ВЫЧИСЛЕНИЙ, КОТОРАЯ В ЗНАЧИТЕЛЬНОЙ СТЕПЕНИ ОБЕСПЕЧИВАЕТ ДОСТУП К ВЫЧИСЛИТЕЛЬНЫМ РЕСУРСАМ, КОТОРЫЕ ВКЛЮЧАЮТ В СЕБЯ МНОЖЕСТВО ВЕЩЕЙ, ВРОДЕ ПРИКЛАДНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЕ, ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ, СЕРВЕРОВ, ЦЕНТРОВ ОБРАБОТКИ ДАННЫХ И Т. Д. 1) <i>Облачные Вычисления</i> 2) <i>Всепогодные Вычисления</i> 3) <i>Суперскоростные Вычисления</i></p>	ПК-6
<p>17. ТЕХНОЛОГИЯ, КОТОРАЯ УЧИТСЯ НА ОПЫТЕ (ДАННЫХ) ПРЕДСКАЗЫВАТЬ БУДУЩЕЕ ПОВЕДЕНИЕ ИНДИВИДОВ, ОБЪЕКТОВ С ПОМОЩЬЮ ПРОГНОСТИЧЕСКИХ МОДЕЛЕЙ, КОТОРЫЕ ВКЛЮЧАЮТ В СЕБЯ ИХ ХАРАКТЕРИСТИКИ В КАЧЕСТВЕ ВХОДНЫХ ДАННЫХ И ПРОИЗВОДИТ ОЦЕНКУ В КАЧЕСТВЕ ВЫХОДНЫХ.. 1) <i>прогнозная аналитика</i> 2) <i>практическая аналитика</i> 3) <i>предсказательная аналитика</i></p>	ПК-6
<p>18. ТЕХНОЛОГИЯ, КОТОРАЯ ОБОБЩАЕТ ДАННЫЕ, УДЕЛЯЯ МЕНЬШЕ ВНИМАНИЯ ТОЧНЫМ ДЕТАЛЯМ КАЖДОЙ ИХ ЧАСТИ, А ВМЕСТО</p>	ПК-6

<p>ЭТОГО СОСРЕДОТАЧИВАЯСЬ НА ОБЩЕМ ПОВЕСТВОВАНИИ</p> <ol style="list-style-type: none"> 1) <i>Описательная аналитика</i> 2) <i>Повествовательная аналитика</i> 3) <i>Познавательная аналитика</i> 4) <i>Предсказательная аналитика</i> 	
<p>19. ИНФОРМАЦИОННАЯ СТРУКТУРА ХРАНЕНИЯ ДАННЫХ, КОТОРАЯ ОПРЕДЕЛЯЕТСЯ КАК АРХИТЕКТУРА, КОТОРАЯ ПОЗВОЛЯЕТ ЛПР И РУКОВОДИТЕЛЯМ ОРГАНИЗАЦИЙ СИСТЕМАТИЧЕСКИ ОРГАНИЗОВЫВАТЬ, ПОНИМАТЬ И ИСПОЛЬЗОВАТЬ СВОИ ДАННЫЕ ДЛЯ ПРИНЯТИЯ СТРАТЕГИЧЕСКИХ РЕШЕНИЙ</p> <ol style="list-style-type: none"> 1) <i>Хранилище данных</i> 2) <i>Хранилище информации</i> 3) <i>Хранилище BigData</i> 4) <i>Вместилище экспертных решений</i> 	ПК-6
<p>20. ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ, ОСНОВАННАЯ НА ИСПОЛЬЗОВАНИИ СПЕЦИФИЧЕСКИХ АЛГОРИТМОВ ДЛЯ ИЗВЛЕЧЕНИЯ ПАТТЕРНОВ, СОДЕРЖАЩИХСЯ В ВЫЯВЛЯЕМЫХ СВЯЗЯХ АНАЛИЗИРУЕМЫХ НАБОРОВ ДАННЫХ</p> <ol style="list-style-type: none"> 1) <i>интеллектуальный анализ данных</i> 2) <i>потенциальный анализ паттернов</i> 3) <i>машинно-обучаемый анализ паттернов</i> 4) <i>объектно-ориентированный анализ паттернов</i> 	ПК-6
<p>21. ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ, КОТОРАЯ ЯВЛЯЕТСЯ МЕЖДИСЦИПЛИНАРНЫМ ПО СВОЕЙ ПРИРОДЕ ПОДХОДОМ, ИСПОЛЬЗУЕТ МЕТОДЫ ИЗ ОБЛАСТИ КОМПЬЮТЕРНЫХ НАУК, СТАТИСТИКИ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА; КОТОРАЯ ИСПОЛЬЗУЕТСЯ ДЛЯ ПРОГНОСТИЧЕСКОГО АНАЛИЗА И РАСПОЗНАВАНИЯ ОБРАЗОВ В БОЛЬШИХ ДАННЫХ.</p> <ol style="list-style-type: none"> 1) <i>машинное обучение</i> 2) <i>интеллектуальное обучение</i> 3) <i>прогностический анализ</i> 4) <i>автоматизированный анализ данных</i> 	ПК-6
<p>22. РАСПРЕДЕЛЕННАЯ ФАЙЛОВАЯ СИСТЕМА</p>	ПК-6

<p>ХРАНЕНИЯ ДАННЫХ, ПОЗВОЛЯЮЩАЯ ОБЕСПЕЧИВАТЬ ВЫСОКО-ПРОИЗВОДИТЕЛЬНЫЙ ДОСТУП К ДАННЫМ, И ПОДХОДИТ ДЛЯ РАБОТЫ С БОЛЬШИМИ ОБЪЕМАМИ ДАННЫХ</p> <ol style="list-style-type: none"> 1) <i>HDFS</i> 2) <i>NTFS</i> 3) <i>BigdataFS</i> 4) <i>DataNodeFS</i> 	
<p>23. ВЫБЕРИТЕ ТЕХНОЛОГИЮ ПОТОКОВОЙ ОБРАБОТКИ СОБЫТИЙ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ</p> <ol style="list-style-type: none"> 5) <i>Apache Hadoop</i> 6) <i>Spark Streaming</i> 7) <i>Apache Kafka</i> 8) <i>MapReduce</i> 	ПК-6
<p>24. БОЛЬШИЕ ДАННЫЕ – ЭТО:</p> <ol style="list-style-type: none"> 1) <i>Данные объемом более 1Тб</i> 2) <i>Данные объемом более 10Тб</i> 3) <i>Данные объемом более 100Тб</i> 4) <i>Нет ограничений на минимальный объем</i> 	ПК-6
<p>25. HADOOP – ЭТО:</p> <ol style="list-style-type: none"> 1) <i>набор утилит, и программный каркас для выполнения распределённых программ, работающих на кластерах</i> 2) <i>распределённая СУБД, позволяющая обрабатывать большие данные</i> 3) <i>язык выполнения заданий в парадигме MapReduce</i> 4) <i>распределённая файловая система, предназначенная для хранения файлов большого объёма</i> 	ПК-6
<p>26. ДЛЯ МАШИННОГО ОБУЧЕНИЯ ПОДХОДЯТ ДАННЫЕ</p> <ol style="list-style-type: none"> 1) <i>Предварительно подготовленные, очищенные от ошибок, пропусков и выбросов, а также нормализованные и представленные в виде числовых векторов</i> 2) <i>Бинарные</i> 3) <i>Любых форматов в цифровом виде</i> 4) <i>Числовые типа int</i> 	ПК-6
<p>27. ФОРМАТ PARQUET СЧИТАЕТСЯ</p>	ПК-6

<ul style="list-style-type: none"> 1) <i>полуструктурированным</i> 2) <i>колоночным (столбцовым)</i> 3) <i>строковым</i> 4) <i>неструктурированным</i> 	
<p>28. ОСНОВНЫЕ ПОНЯТИЯ ИЕРАРХИЧЕСКОЙ БД:</p> <ul style="list-style-type: none"> 1) <i>таблица, столбец, строка</i> 2) <i>уровень, узел, связь</i> 3) <i>отношение, атрибут, кортеж</i> 	ПК-6
<p>29. ДАННЫЕ, ИМЕЮЩИЕ НАИБОЛЬШИЙ ОБЪЕМ НА СЕГОДНЯШНИЙ ДЕНЬ</p> <ul style="list-style-type: none"> 1) <i>структурированные</i> 2) <i>квазиструктурированные</i> 3) <i>полуструктурированные</i> 4) <i>неструктурированные</i> 	ПК-6
<p>30. К BIG DATA НЕ ПРИМЕНЯТСЯ ОДИН ИЗ НИЖЕ ПЕРЕЧИ-СЛЕННЫХ ПРИНЦИПОВ РАБОТЫ</p> <ul style="list-style-type: none"> 1) <i>отказоустойчивость</i> 2) <i>вертикальная масштабируемость</i> 3) <i>локальность данных</i> 4) <i>горизонтальная масштабируемость</i> 	ПК-6

Эталоны ответов

<i>Номер тестового задания</i>	<i>Номер эталона ответа</i>
1	1)
2	3)
3	3)
4	2)
5	4)
6	1)
7	2)
8	1)
9	3)
10	2)
11	2)

12	1)
13	1)
14	1)
15	1)
16	1)
17	1)
18	1)
19	1)
20	1)
21	1)
22	1)
23	1)
24	4)
25	1)
26	1)
27	2)
28	2)
29	4)
30	2)