

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Чернецкая Ирина Евгеньевна
Должность: Заведующий кафедрой
Дата подписания: 01.10.2023 11:50:01
Уникальный программный ключ:
bdf214c64d8a381b0782ea566b0dce05e3f5ea2d

МИНОБРНАУКИ РОССИИ

Юго-Западный государственный университет

УТВЕРЖДАЮ:

Заведующий кафедрой

вычислительной техники

И.Е. Чернецкая И.Е. Чернецкая

« 31 » 08 2023 г.

ОЦЕНОЧНЫЕ СРЕДСТВА

Для текущего контроля успеваемости

и промежуточной аттестации обучающихся

по дисциплине

Архитектура систем обработки, анализа и интерпретации данных

09.04.01 Информатика и вычислительная техника

Курск – 2023

1 ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ

1.1 ВОПРОСЫ ДЛЯ СОБЕСЕДОВАНИЯ

Раздел (тема) дисциплины. Классификация и основы построения и функционирования многопроцессорные системы и распараллеливание обработки данных.

1. Классификация и основы построения и функционирования многопроцессорные системы.
2. Организация распараллеливание обработки данных.
3. Надежность, отказоустойчивость производительность многопроцессорных систем и систем параллельной обработки данных.
4. Показатели и характеристики вычислительных систем параллельной обработки данных.
5. Классификация Флинна.
6. Внутрипроцессорный параллелизм.
7. Параллелизм на уровне команд.
8. Внутрипроцессорная многопоточность.
9. Однокристалльные мультипроцессоры.
10. Сопроцессоры.

Раздел (тема) дисциплины. Мультипроцессоры и мультикомпьютеры.

11. Сетевые процессоры.
12. Мультимедиа-процессоры.
13. Мультипроцессоры.
14. Многоядерные процессоры.
15. Мультипроцессоры и мультикомпьютеры.
16. UMA-мультипроцессоры.
17. симметричные мультипроцессорные архитектуры.
18. NUMA-мультипроцессоры.
19. COMA-мультипроцессоры.
20. Мультикомпьютеры.
21. Коммуникационные сети.
22. Процессоры с массовым параллелизмом.
23. Векторные вычисления.
24. Кластерные вычисления.
25. Кластеры высокой готовности.
26. Отказоустойчивые кластеры.
27. Кластеры параллельных вычислений.
28. Модели массового обслуживания систем параллельной обработки данных.

Раздел (тема) дисциплины. Надежность и отказоустойчивость систем параллельной обработки.

29. Расчет времени ожидания обработки запросов в кластере.
30. Оптимизация структуры кластера.
31. Надежность и отказоустойчивость систем параллельной обработки.
32. Методы обеспечения надежности и отказоустойчивости систем параллельной обработки. Задачи оптимального резервирования.
33. Марковские модели оценки надежности восстанавливаемых систем параллельной обработки.
34. Модели невосстанавливаемых систем.
35. Модели надежности сложных систем.
36. Методы резервирования.
37. Задачи оптимального резервирования.
38. Постановка и решение задачи векторной оптимизации систем параллельной обработки.
39. Модели надежности кластерных систем.
40. Задачи векторной оптимизации систем параллельной обработки.

Раздел (тема) дисциплины. Обработка и анализ структурированных данных

41. Группировка и визуализация данных.
42. Создание сводных таблиц.
43. Формирование выборки по определенным признакам.
44. Научная графика.
45. Многомерные массивы и матрицы.
46. Операции с многомерными массивами и матрицами.
47. Структурирование данных.
48. Статистическая обработка данных.
49. Сведение структурированных данных в базу данных.
50. Использование графики для визуализации результатов обработки данных.

Критерии оценки:

«1» балл выставляется за достаточный объем знаний в рамках дисциплины, использование установленной терминологии и изложение ответов на вопросы без грубых ошибок, воспроизведение фактического и теоретического материала без обобщений и выводов, умение ориентироваться в основных концепциях и понятиях дисциплины, приводить типовые примеры.

«2» балла – достаточно полные знания по дисциплине, содержание материала излагается последовательно, точно, правильно, осмысленно, самостоятельно, грамотное использование необходимой научной терминологии, умение делать обоснованные выводы, способность выявлять главенствующие факторы при техническом анализе вопросов. Даются ответы на любые заданные вопросы с несущественными ошибками и недочетами.

1.2 ВОПРОСЫ И ЗАДАНИЯ В ТЕСТОВОЙ ФОРМЕ

Тема 1. Понятие технологии программирования и основные этапы ее развития

1. 4GL означает ...

2. В основу разработки графического пользовательского интерфейса GUI положены метафоры, из перечисленного:

а) «видишь, что получил»

б) «работаешь с тем, что видишь»

в) «рабочий стол»

3. Соотнесите

В стандарте UML для избежания повторов в нескольких вариантах следует применять связь:	использования
В стандарте UML при описании изменений в нормальном поведении системы следует применять связь:	расширения

4. Расположите в правильном порядке этапе жизненного цикла программного обеспечения

Тема 2. Проблемы разработки сложных программных систем. Блочный-иерархический подход к проектированию ПО

1. Факторы сложности определения требований к программным системам ...

2. Процесс разбиения сложного объекта на сравнительно независимые части получил название

а) декомпозиции;

б) детализации;

в) актуализации;

г) иерархизации

3. Соотнесите

Непротиворечивость	контроль согласованности элементов между собой
Полнота	контроль на присутствие лишних элементов
Формализация	строгость методического подхода
Повторяемость	необходимость выделения одинаковых блоков для удешевления и ускорения разработки

4. Определите порядок:

Постановка задачи, построение модели, разработка алгоритма

Тема 3. Классификация и основы построения и функционирования многопроцессорные системы (МВС) и распараллеливание обработки данных

1. Назовите аппаратно-программные особенности компьютера, влияющие в реальных условиях на выполнение конкретной программы:

а) версия операционной системы;

б) наличие средств визуализации;
в) особенности структуры процессора, эффективность работы компиляторов.

2. Производительность МВС – это: ...

3. Соотнесите

области применения МВС в науке:	
области применения МВС в бизнесе	

- 1) расчеты, требующие значительных вычислительных ресурсов
- 2) поддержка работы электронной почты
- 3) поиск в Internet
- 4) бухгалтерский учет
- 5) обработка транзакций, управление базами данных

Тема 4. Мультипроцессоры и мультикомпьютеры

1. Принцип, предопределяющий формирование вычислительной системы из унифицированных элементов ...– это:

- а) модульность
- б) близкодействие
- в) асинхронность функционирования ВС
- г) множество элементов сети, называемых переходами

2. Векторный процессор – это ...

3. Соотнесите

множественный поток команд и одиначный поток данных	

- 1) SIMD
- 2) MIMD
- 3) MISD
- 4) SISD
4. На какие подгруппы распались Машины SIMD:

Шкала оценивания: 16 балльная.

Критерии оценивания:

Каждый вопрос (задание) в тестовой форме оценивается по дихотомической шкале: выполнено – 1 балл, не выполнено – 0 баллов.

Применяется следующая шкала перевода баллов в оценку по 5-балльной шкале:

- 14-16 баллов соответствует оценке «отлично»;
- 11-13 баллов соответствует оценке «хорошо»;
- 8-10 баллов соответствует оценке «удовлетворительно»;
- 7 баллов и менее – оценке «неудовлетворительно».

2 ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ

2.1 БАНК ВОПРОСОВ И ЗАДАНИЙ В ТЕСТОВОЙ ФОРМЕ

1 Вопросы в закрытой форме

1.1 «Общее описание ПО информационной системы» должно содержать

- а) описание операционной системы
- б) основные сведения о всех видах обеспечения
- в) структуру ПО
- г) функции частей ПО

1.2 Описание функционирования ПС» в техническом проекте должно содержать

- а) исходные данные
- б) функции частей ПС
- в) характеристика функциональности
- г) характеристика переменных

1.3 Из перечисленного ODMG поддерживает наборы значений

- а) куча
- б) массив
- в) список
- г) матрица

1.4 Из перечисленного жизненный цикл ПО в соответствии с подходом RAD включает стадии

- а) анализ и планирование требований
- б) внедрение
- в) проектирование
- г) реализация

1.5 Вычислить Из перечисленного литералы объектно-ориентированной базы данных могут быть следующего вида:

- а) наборы
- б) простые
- в) специальная величина null
- г) структуры

1.6 Из перечисленного методы объекта могут быть описаны как:

- а) виртуальные
- б) динамические
- в) методы-сообщения
- г) статические

1.7 Из перечисленного на этапе проектирования программы по каждому модулю разрабатываются спецификации:

- а) вход/выход
- б) имя/цель

- в) обзор действий
- г) ссылки между модулями

1.8 Из перечисленного на этапе реализации алгоритма происходит:

- а) интеграция
- б) кодирование
- в) тестирование
- г) объединения

1.9 Из перечисленного объект обладает:

- а) индивидуальностью
- б) поведением
- в) состоянием
- г) атрибутами

1.10 Из перечисленного объектно-ориентированный подход к проектированию программных продуктов основан на

- а) выделении классов объектов
- б) создании иерархии классов
- в) установлении характерных свойств объектов и методов их обработки
- г) определении объектов

1.11 Из перечисленного основными компонентами диаграмм потоков данных являются:

- а) внешние сущности
- б) накопители данных
- в) процессы
- г) актеры

1.12 Из перечисленного основными элементами объектной модели являются

- а) абстрагирование
- б) иерархия
- в) инкапсуляция
- г) модульность

1.13 Из перечисленного составляющими информационного моделирования предметной области являются

- а) детальное конструирование
- б) информационный анализ бизнес-областей
- в) процедура обработки данных
- г) системное проектирование функций обработки данных

1.14 Из перечисленного спецификация в объектно-ориентированной базе данных определяет:

- а) исключительные ситуации
- б) операции
- в) состояния данного типа
- г) начальное значение переменной

1.15 Из перечисленного стандарт UML версии 1.1 предлагает диаграммы для моделирования:

- а) вариантов использования

- б) классов
- в) поведения системы
- г) состояний

1.16 Из перечисленного технология построения модели при индуктивном способе на эмпирическом этапе включает:

- а) гипотезу
- б) интуицию
- в) предположение
- г) умозаключение

1.17 Из перечисленного типами связей между функциями при методе SADT являются:

- а) временная
- б) логическая
- в) процедурная
- г) случайная

1.18 Из перечисленного цели оценивания и обеспечения корректности ПС достигаются посредством:

- а) анализ
- б) выполнения тестовых процедур
- в) просмотров
- г) разработки тестовых процедур 2

1.19 Из перечисленного, видами точек останова являются:

- а) по адресу
- б) по загрузке модуля
- в) по обращению к данным
- г) условная

1.20 Из перечисленного, документ «Описание алгоритма» должен содержать:

- а) используемая информация и ее характеристики
- б) математическое описание алгоритма
- в) назначение и характеристики алгоритма
- г) результаты решения, полученные в результате реализации алгоритма

1.21 Из перечисленного, конструктивными характеристиками качества ПС являются:

- а) защищенность
- б) корректность
- в) надежность
- г) практичность

1.22 Пиковая производительность системы определяется:

- а) временем выполнения тестовых задач
- б) временем выполнения реальных задач

в) произведением пиковой производительности одного процессора на число процессоров в системе

- г) число импульсов, генерируемых в единицу времени

1.23 Назовите принцип формирования иерархической памяти:

а) при повышении уровня памяти скорость обработки данных должна увеличиваться, а объем уровня памяти – уменьшаться

б) при повышении уровня памяти скорость обработки данных должна уменьшаться, а объем уровня памяти – увеличиваться

в) при повышении уровня памяти скорость обработки данных и объем уровня памяти должны увеличиваться

г) высокая производительность, масштабируемость, минимально допустимое время простоя

1.24 Под архитектурой высокопроизводительной системы в частности понимают:

а) устройство микросхем

б) внешний вид высокопроизводительной системы

в) топологию связи между процессорами

г) систему различных видов памяти

1.25 Характеристика класса SIMD:

а) множественный поток команд и множественный поток данных

б) одиночный поток команд и множественный поток данных

в) множественный поток команд и одиночный поток данных

г) одиночный поток команд и одиночный поток данных

2 Вопросы в открытой форме

2.1 Производительность МВС – это:...

2.2 Единица производительности МВС ...

2.3 Дайте определение понятию потока.

2.4 Какая цель преследуется при разработке новой архитектуры высокопроизводительной системы?

2.5 Почему существует несколько классификаций архитектур вычислительных систем?

2.6 Назовите главную особенность систем с архитектурой SMP:

2.7 Назовите основной недостаток систем с архитектурой SMP:

2.8 Кэши являются когерентными, если ...

2.9 Назовите главную особенность систем с архитектурой MPP:

2.10 Назовите основной недостаток систем с архитектурой MPP:

2.11 МВС-1000 является системой:

2.12 Назовите главную особенность архитектуры NUMA:

2.13 Почему архитектура NUMA называется гибридной?

2.14 SGI Origin3000 является системой:

2.15 Назовите главную особенность систем с архитектурой PVP

2.16 Почему системы, имеющие PVP-архитектуру, можно назвать системами общего назначения?

2.17 Критическим параметром, влияющим на величину производительности кластерной системы, является:

2.18 Что представляет собой вычислительный кластер?

2.19 Основной недостаток систем, имеющих PVP-архитектуру:

2.20 Что в большей мере определяет производительность кластерной системы?

2.21 Что представляет собой узел вычислительного кластера?

2.22 Какая схема соединения процессоров в кластере является наиболее эффективной?

2.23 Ассоциативные системы относятся к классу:

2.24 Конвейерная технология предполагает:

2.25 Что позволяет осуществлять многомодальная логика?

3 Вопросы на установление соответствия

3.1 Соотнесите

главная особенность систем с архитектурой SMP	неоднородный доступ к памяти
главная особенность систем с архитектурой MPP	наличие общей физической памяти, разделяемой всеми процессорами
главная особенность архитектуры NUMA	память физически разделена
	сверхвысокая производительность
	наличие векторно-конвейерных процессоров

3.2 Соотнесите

основной недостаток систем с архитектурой SMP	плохая масштабируемость
основной недостаток систем с архитектурой MPP	высокая цена
	сложность для программирования
	низкая скорость межпроцессорного обмена

3.3 Соотнесите

MBC-1000	NUMA
SGI Origin3000	MPP
	SMP
	cc-NUMA

3.4 Соотнесите

главная особенность систем с архитектурой PVP	наличие общей физической памяти
основной недостаток систем, имеющих PVP-архитектур	наличие специальных векторно-конвейерных процессоров
	неоднородный доступ к памяти
	плохая масштабируемость
	дороговизна

	низкая скорость межпроцессорного обмена
--	---

3.5 Соотнесите

PVP-архитектура	SGI Origin3000
	Fujitsu-VPP5000
	MBC-1000

3.6 Соотнесите

схема соединения процессоров в кластере	гиперкуб
	плоская решетка
	куб

3.7 Соотнесите

Узлы соединяются с помощью колец в коммуникационной технологии	Myrinet
	Raceway
	SCI

3.8 Соотнесите

Основные преимущества технологии Myrinet	высокая скорость передачи, малое время задержки, прямая коммутация, умеренная стоимость
Основные преимущества интерфейса SCI	эта технология представляет собой комбинацию шины и локальной сети
	обеспечивает реализацию когерентности кэш-памяти
	простота концепции
	дешевизна

3.9 Соотнесите

Коммутируемая сеть	Ethernet
	FDDI
	Myrinet

3.10 Соотнесите

Ассоциативный способ обработки данных предполагает	обработку только тех данных, которые удовлетворяют определенным критериям отбора
Конвейерная технология предполагает:	обработку нескольких команд одновременно
	доступ к данным по указателям
	последовательную обработку команд
	обработку команд, удовлетворяющих

	определенным критериям
	обработку всех данных

3.11 Соотнесите

Ассоциативные системы относятся к классу	SISD
Матричные системы относятся к классу	MISD
Конвейерные системы относятся к классу	SIMD
Потоковые процессоры принадлежат к архитектуре:	MIMD

3.12 Соотнесите

Что является командой в ДНК–процессоре	световой сигнал
Что является структурой ДНК–процессор	электрический импульс
	биохимическая реакция
	структура гена
	структура клеточной мембраны
	структура молекулы ДНК

3.13 Соотнесите

преимущества биокomпьютеров	более простая, по сравнению с производством полупроводников, технология изготовления
недостатки биокomпьютеров	легкость в обработке результатов вычислений
	высокая точность вычислений
	сложная технология изготовления
	высокое энергопотребление
	низкая точность вычислений

3.14 Соотнесите

клеточные компьютеры	компьютеры, собранные из отдельных блоков, называемых «клетками»
коммуникационные процессоры	компьютер, схема которого имитирует работу живой клетки
	самоорганизующиеся колонии микроорганизмов, активизирующиеся в присутствии определенного вещества
	микрочипы, представляющие собой нечто среднее между жесткими специализированными интегральными микросхемами и гибкими процессорами общего назначения

3.15 Соотнесите

особенностью нейронных процессоров является	способность нейрона передавать электрический импульс
основные принципы	изменение состояния нейронов в процессе

функционирования искусственной нейронной сети	функционирования
	наличие когерентного кэша
	высокая тактовая частота
	конвейерная технология обработки данных
	принцип обработки информации

3.16 Соотнесите

Процессоры баз данных предназначены для	создания пользовательских интерфейсов
	управления базами данных
	осуществления транзакций

3.17 Соотнесите

достоинства коммутаторов с пространственным разделением	высокая надежность
недостатки коммутаторов с пространственным разделением	возможность одновременного контакта со всеми устройствами
	низкая сложность
	большая задержка
	высокая сложность
	невозможность одновременного контакта со всеми устройствами

3.18 Соотнесите

«очередь FIFO»	первым обрабатывается первый запрос в очереди
«очередь LIFO»	первым обрабатывается последний запрос в очереди
	запросы обрабатываются в случайном порядке

3.19 Соотнесите

Топология связей узлов в структурном модуле системы МВС-1000	кольцо (одномерный тор)
а структура объединения блоков системы SPP1000	двумерный тор
	куб
	гиперкуб

3.20 Соотнесите

Что в первую очередь мешает повышению производительности вычислительных систем	недостаточная производительность процессор
Что может ограничивать возможность масштабирования вычислительных систем?	низкая пропускная способность междузловых соединений

	ограниченный объем оперативной памяти
	соотношение стоимость/производительность
	архитектура вычислительной системы
	отношение скорости процессора к скорости связи

3.21 Соотнесите

способы и методы повышения отказоустойчивости системы	кластеризация системы
способы и методы повышения надежности системы	использование RAID-систем хранения
	дублирование аппаратных компонентов
	развитие новых принципов и подходов в создании электронных компонентов системы
	обеспечение оптимальных условий эксплуатации систем
	повышение квалификации системных программистов

3.22 Соотнесите

алгоритмам кластеризации	Метод ближайшего соседа
	Минимальное покрывающее дерево
	k-Means алгоритм
	Генетические алгоритмы

3.23 Соотнесите

верные утверждения	Данные в традиционной базе данных структурированы.
	Данные в базе больших данных структурированы
	Данные в традиционной базе данных полуструктурированы или неструктурированы
	Данные в базе больших данных полуструктурированы или неструктурированы

3.24 Соотнесите

верные утверждения	Модель хранения и обработки данных в традиционной базе данных - вертикальная модель
	Модель хранения и обработки данных в базе больших данных - вертикальная модель
	Модель хранения и обработки данных в

	традиционной базе данных - горизонтальная модель
	Модель хранения и обработки данных в базе больших данных - горизонтальная модель

3.25 Соотнесите

Автоматическое разбиение элементов некоторого множества (объекты, данные, вектора характеристик) на группы по принципу схожести	кластеризация
	классификация
	интерполяция
	модуляция

4 Вопросы на установление последовательности

4.1 Расположите в правильном порядке уровни современной информационно-аналитической системы

- а) извлечение, преобразование и загрузка данных
- б) сбор и первичная обработка данных
- в) складирование данных
- г) анализ данных
- д) представление данных в витринах данных
- е) Web-портал

4.2 Расположите Этапы обучения в правильном порядке

- а) Сбор данных;
- б) Подготовка данных (фильтрация, дополнение, кодирование);
- в) Постановка задачи анализа;
- г) Подбор параметров модели и алгоритма обучения;
- д) Выбор модели (алгоритма анализа данных);
- е) Обучение модели (автоматический поиск остальных параметров модели);

4.3 Определяющими характеристиками Big data являются скорость, многообразие, достоверность, структурированность, квазиструктурированность, неструктурированность, уязвимость

4.4 Основными принципами работы с Big Data являются: распределенность, горизонтальная масштабируемость, отказоустойчивость, локальность данных, интерпретация данных в процессе их обработки, достоверность, значимость, скорость, надежность

4.5 Распределенные файловые системы должны удовлетворять следующим требованиям:

сетевая прозрачность, прозрачность размещения, высокая доступность, мобильность пользователя, отказоустойчивость, масштабируемость, достоверность, актуальность,

4.6 Целью кластеризации является: упрощение дальнейшей обработки данных, сокращение объема хранимых данных, выделение нетипичных объектов, построение иерархии множества объектов, построение критических диаграмм, выделение основных

параметров выборки, проверка подчинения предполагаемому закону распределения,

4.7 Представьте правильные последовательности:

Структурированные данные, Полуструктурированные данные, Квазиструктурированные данные.

данные определяющие конкретную предметную область, упорядоченные специальным образом и организованные так, чтобы над такими данными можно было выполнить анализ, не соответствуют чёткой структуре таблиц и отношений в реляционных базах данных, однако такие данные содержат специальные теги и иные маркёры, позволяющие отделить семантические элементы, текстовые данные с неустойчивым форматом, которые для обработки инструментами требуют больших временных затрат на преобразование

4.8 Представьте правильные последовательности

Метод обучение ассоциативным правилам, Метод классификации, Метод кластерного анализа

статистический метод классификации объектов, который приводит к разделению разнообразных групп на более мелкие группы подобных (сходных) объектов, для которых критерий подобия заранее не известен, методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным, метод, используемый для обучения машин способам обнаружения зависимостей между данными в больших базах данных,

4.9 Представьте правильные последовательности

Хранилище данных, Распределенная система, Hadoop, MapReduce

технология работы с Big data, которая представляет собой предметно-ориентированную информационную базу данных, специально разработанную и предназначенную для подготовки отчётов и анализа данных с целью поддержки принятия решений в организации; технология работы с Big data, которая представляет собой множество компьютеров, взаимодействующих по сети и объединенных для решения общей вычислительной задачи; технология работы с Big data, которая представляет собой свободно распространяемый набор утилит, библиотек и фреймворк для разработки и выполнения распределённых программ, работающих на кластерах из сотен и тысяч; технология работы с Big data, которая представляет собой модель распределённых вычислений, используемая для параллельных вычислений над очень большими, вплоть до нескольких петабайт, наборами данных в компьютерных кластерах

4.10 Представьте правильные последовательности

Сетевой анализ, Распознавание образов, Кластерграмма, Машинное обучение набор методов, используемых для описания и анализа отношений между дискретными узлами в графе или сети; набор методов машинного обучения, развивающих основы и методы классификации и идентификации предметов, явлений, процессов, сигналов, ситуаций и т. п. объектов, которые характеризуются конечным набором некоторых свойств и признаков; метод

визуализации, использующийся при кластерном анализе и показывающий, как отдельные элементы множества данных соотносятся с кластерами по мере изменения их количества; класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач

4.11 Представьте правильные последовательности

Сетевая прозрачность распределенной файловой системы; Прозрачность размещения распределенной файловой системы; Высокая доступность распределенной файловой системы; Мобильность пользователя распределенной файловой системы

Способы доступа и обработки должны быть одинаковыми для локальных и удаленных файлов; имя файла не должно определять его местоположения в сети и не должно меняться при изменении его физического размещения; выполнение запросов клиентов за время, приемлемое для решаемых задач; пользователи должны иметь возможность обращаться к разделяемым файлам из любого узла сети

4.12 Представьте правильные последовательности

Имитационное моделирование; Пространственный анализ; Анализ временных рядов; Прогнозная аналитика

метод исследования, при котором изучаемая система заменяется моделью, с достаточной точностью описывающей реальную систему, с которой проводятся эксперименты, с целью получения информации об этой системе; набор методов, которые анализируют топологические, геометрические или географические свойства, представленные в наборе; совокупность математико-статистических методов анализа, предназначенных для выявления структуры временных рядов и для их прогнозирования; класс методов анализа данных, концентрирующийся на прогнозировании будущего поведения объектов и субъектов с целью принятия оптимальных решений

4.13 Представьте правильные последовательности

Value (ценность накопленной информации); Validity (пригодность, обоснованность); Visualization (визуализация); Veracity (достоверность данных)

значимость данных с точки зрения прикладных задач; характеризует, в какой мере располагаемые данные являются точными и правильными с точки зрения их предполагаемого использования; представление данных в более удобном для восприятия виде; свойство, которое характеризует надежность данных.

4.14 Представьте правильные последовательности

Смешение и интеграция данных; Обучение ассоциативным правилам; Краудсорсинг; Искусственные нейронные сети

набор методов, позволяющих интегрировать и анализировать разнородные данные из разнообразных источников для глубинного анализа более точно и эффективно, чем из единственного источника данных; совокупность методов для анализа необходимых взаимосвязей, то есть «ассоциативных правил», среди переменных в больших базах данных; метод сбора, категоризация и

обогащение данных силами широкого круга лиц, привлечённых на основании публичной оферты, без вступления в трудовые отношения, обычно посредством использования сетевых медиа; математическая модель, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма

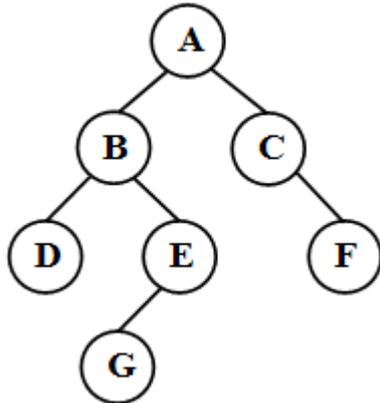
4.15 Представьте правильные последовательности

Метрический классификатор; Метод ближайших соседей; Метод k ближайших соседей; Метод взвешенных ближайших соседей; Линейный классификатор

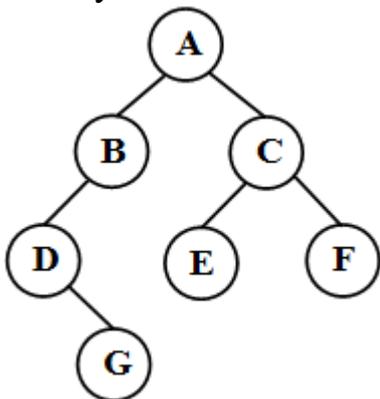
классифицируемый объект, относят к классу, состоящему из схожих с ним объектов, классифицируемый объект x принадлежит тому классу, которому принадлежит его ближайший объект выборки; берется k число ближайших соседей, и объект относится к тому классу, к которому относится большинство его соседей; убывающей с ростом ранга соседям приписывается вес, для того, чтобы не возникало неоднозначности при количестве классов более двух; алгоритм классификации, который основан на построении разделяющей линейной поверхности; некий алгоритм, который вычисляет оценку сходства предоставленных объектов

4.16 К динамическим структурам данных относятся:

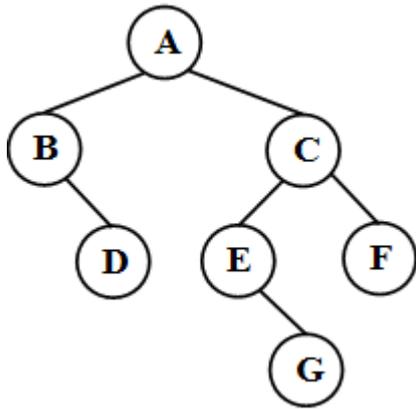
4.17 Укажите представление дерева во входном потоке, если каждой вводимой пустой связи соответствует символ звездочка '*':



4.18 Укажите представление дерева во входном потоке, если каждой вводимой пустой связи соответствует символ звездочка '*':



4.19 Укажите представление дерева во входном потоке, если каждой вводимой пустой связи соответствует символ звездочка '*':



4.20 Представьте правильные последовательности

Найденные в процессе использования технологии Data Mining закономерности должны обладать такими свойствами

4.21 Представьте правильные последовательности

Атрибут также называют

4.22 Представьте правильные последовательности

Действия, в рамках стадии свободного поиска, выполняются при помощи

4.23 Представьте правильные последовательности

Технология Web mining применяет технологию Data Mining для анализа

4.24 расположите в порядке перечисления: задача интеграции, задача преобразования, задача уменьшения размера и задача уменьшения значения исходных данных

Частый анализ; классификация и прогноз; предварительная обработка данных; анализ потока данных

4.25 Выберите масштабируемые алгоритмы кластеризации

CURE, DENCURE, CLIQUE, OPOSSUM

Шкала оценивания результатов тестирования: в соответствии с действующей в университете балльно-рейтинговой системой оценивание результатов промежуточной аттестации обучающихся осуществляется в рамках 100-балльной шкалы, при этом максимальный балл по промежуточной аттестации обучающихся по очной форме обучения составляет 36 баллов, по очно-заочной и заочной формам обучения – 60 баллов (установлено положением П 02.016).

Максимальный балл за тестирование представляет собой разность двух чисел: максимального балла по промежуточной аттестации для данной формы обучения (36 или 60) и максимального балла за решение компетентностно-ориентированной задачи (6).

Балл, полученный обучающимся за тестирование, суммируется с баллом, выставленным ему за решение компетентностно-ориентированной задачи.

Общий балл по промежуточной аттестации суммируется с баллами, полученными обучающимся по результатам текущего контроля успеваемости в течение семестра; сумма баллов переводится в оценку по дихотомической шкале следующим образом:

Сумма баллов по 100-балльной шкале	Оценка по дихотомической шкале
100 – 50	зачтено
49 и менее	не зачтено

Критерии оценивания результатов тестирования: Каждый вопрос (задание) в тестовой форме оценивается по дихотомической шкале: выполнено – 2 балла, не выполнено – 0 баллов.

2.2 КОМПЕТЕНТНОСТНО-ОРИЕНТИРОВАННЫЕ ЗАДАЧИ

Компетентностно-ориентированная задача №1

Рассмотрим набор значений $\{1,2,3,4,5,90\}$, усеченное среднее значение которых ($p = 20\%$)

Компетентностно-ориентированная задача №2

Энтропия - это количество информации, необходимое для устранения неопределенности. Энтропия бросания равномерной правильной шестигранной кости – это

Компетентностно-ориентированная задача №3

Предположим, что максимальные и минимальные значения дохода атрибута составляют 12 000 юаней и 98 000 юаней соответственно. Используйте метод максимальной и минимальной нормализации, чтобы отобразить значение атрибута в диапазоне от 0 до 1. 73 600 юаней за атрибуты дохода будут конвертированы в

Компетентностно-ориентированная задача №4

Предположим, что данные, используемые для анализа, содержат атрибут возраста. Значения возраста в кортеже данных следующие (в порядке возрастания): 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 30, 33, 33, 35, 35, 36, 40, 45, 46, 52, 70, проблема: используйте метод сглаживания среднего значения, чтобы сгладить вышеупомянутые данные, глубина блока составляет 3. Значение второго поля:

Компетентностно-ориентированная задача №5

Рассмотрим набор значений $\{12243324556826\}$, квартильный диапазон которого:

Компетентностно-ориентированная задача №6

Пусть $X = \{1, 2, 3\}$ - частые наборы предметов, тогда X может генерировать _____ правила ассоциации

Шкала оценивания решения компетентностно-ориентированной задачи: в соответствии с действующей в университете балльно-рейтинговой системой оценивание результатов промежуточной аттестации обучающихся осуществляется в рамках 100-балльной шкалы, при этом максимальный балл

по промежуточной аттестации обучающихся по очной форме обучения составляет 36 баллов, по очно-заочной и заочной формам обучения – 60 (установлено положением П 02.016).

Максимальное количество баллов за решение компетентностно-ориентированной задачи – 6 баллов.

Балл, полученный обучающимся за решение компетентностно-ориентированной задачи, суммируется с баллом, выставленным ему по результатам тестирования.

Общий балл промежуточной аттестации суммируется с баллами, полученными обучающимся по результатам текущего контроля успеваемости в течение семестра; сумма баллов переводится в оценку по дихотомической шкале следующим образом:

Сумма баллов по 100-балльной шкале	Оценка по дихотомической шкале
100 – 50	зачтено
49 и менее	не зачтено

Критерии оценивания решения компетентностно-ориентированной задачи:

6-5 баллов выставляется обучающемуся, если решение задачи демонстрирует глубокое понимание обучающимся предложенной проблемы и разностороннее ее рассмотрение; свободно конструируемая работа представляет собой логичное, ясное и при этом краткое, точное описание хода решения задачи (последовательности (или выполнения) необходимых трудовых действий) и формулировку доказанного, правильного вывода (ответа); при этом обучающимся предложено несколько вариантов решения или оригинальное, нестандартное решение (или наиболее эффективное, или наиболее рациональное, или оптимальное, или единственно правильное решение); задача решена в установленное преподавателем время или с опережением времени.

4-3 балла выставляется обучающемуся, если решение задачи демонстрирует понимание обучающимся предложенной проблемы; задача решена типовым способом в установленное преподавателем время; имеют место общие фразы и (или) несущественные недочеты в описании хода решения и (или) вывода (ответа).

2-1 балла выставляется обучающемуся, если решение задачи демонстрирует поверхностное понимание обучающимся предложенной проблемы; осуществлена попытка шаблонного решения задачи, но при ее решении допущены ошибки и (или) превышено установленное преподавателем время.

0 баллов выставляется обучающемуся, если решение задачи демонстрирует непонимание обучающимся предложенной проблемы, и (или) значительное место занимают общие фразы и голословные рассуждения, и (или) задача не решена.