

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ефанов Алексей Валерьевич

Должность: Директор Невинномысского технологического института (филиал) СКФУ

Дата подписания: 12.10.2022 09:19:08

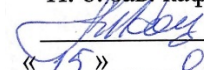
Уникальный программный ключ:

49214306dd433e7a1b0f8632f645f9d553c99e3d0

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ

И. о. зав. кафедрой ИСЭиА

 Колдаев А.И.
«15» 03 2021 г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

для проведения текущего контроля успеваемости и промежуточной аттестации
по дисциплине «Интеллектуальный анализ данных и машинное обучение»
(ЭЛЕКТРОННЫЙ ДОКУМЕНТ)

Направление подготовки/специальность **09.03.02 Информационные системы и технологии**

Направленность (профиль)/специализация **Информационные системы и технологии в бизнесе**

Квалификация выпускника **Бакалавр**

Форма обучения **заочная**

Год начала обучения **2021**

Изучается в **7, 8** семестре

Предисловие

1. Назначение: для проведения текущей и промежуточной аттестации по дисциплине «Интеллектуальный анализ данных и машинное обучение» для студентов направления подготовки 09.03.02 Информационные системы и технологии.

2. Фонд оценочных средств текущей и промежуточной аттестации на основе рабочей программы дисциплины «Интеллектуальный анализ данных и машинное обучение» в соответствии с образовательной программой по направлению подготовки 09.03.02 Информационные системы и технологии, утвержденной на заседании Учёного совета НТИ (филиал) СКФУ.

3. Разработчик: Тихонов Э.Е., доцент ИСЭиА

4. ФОС рассмотрен и утвержден на заседании кафедры ИСЭиА.

5. ФОС согласован с выпускающей кафедрой ИСЭиА.

6. Проведена экспертиза ФОС. Члены экспертной группы, проводившие внутреннюю экспертизу:

Председатель: Кузьменко В.В., и.о. директора НТИ (филиал) СКФУ, профессор кафедры гуманитарных и математических дисциплин

Члены экспертной группы:

Должикова М.В. – заместитель директора по учебно-воспитательной работе НТИ (филиал) СКФУ;

Колдаев А.И. – доцент кафедры информационных систем, электропривода и автоматике.

Эксперт, проводивший внешнюю экспертизу:

Остапенко Н.А., – кандидат технических наук, ведущий инженер-конструктор КБ модернизации ООО КИЭП «Энергомера» филиал АО «Электротехнические заводы «Энергомера»

7. Экспертное заключение: фонд оценочных средств отвечают основным требованиям федерального государственного образовательного стандарта и способствует формированию требуемых компетенций.

7. Срок действия ФОС: на срок реализации образовательной программы.

Паспорт фонда оценочных средств
для проведения текущего контроля и промежуточной аттестации

(ЭЛЕКТРОННЫЙ ДОКУМЕНТ)

Направление подготовки/специальность **09.03.02 Информационные системы и технологии**

Направленность (профиль)/специализация **Информационные системы и технологии в бизнесе**

Квалификация выпускника **Бакалавр**

Форма обучения **заочная**

Год начала обучения **2021**

Изучается в **7, 8** семестре

Код оцениваемой компетенции	Этап формирования компетенции (№ темы) (в соответствии с рабочей программой)	Средства и технологии и оценки	Вид контроля, аттестация (текущий/промежуточный)	Тип контроля (устный, письменный или с использованием технических средств)	Вид контроля	Наименование оценочного средства	Количество заданий для каждого уровня, шт.	
							Пороговый	Продвинутый
ПК-3, ПК-4	Темы 1 - 13		текущий	Устный опрос	Собеседование по результатам выполнения лабораторных работ	Вопросы для собеседования	20	15
			промежуточный					

Составитель _____
(подпись)

« ____ » _____ 2021 г.

Вопросы для собеседования

по дисциплине «Интеллектуальный анализ данных и машинное обучение»

Пороговый уровень

1. Задача анализа данных. Интеллектуальный анализ данных и машинное обучение. Отличия от других видов анализа данных.
2. Постановка и порядок решения задачи интеллектуального анализа данных на примере информационной модели торговой компании.
3. Закономерности в данных: неочевидные, объективные, полезные. Области применения интеллектуального анализа данных.
4. Подготовка данных для анализа. Обработка пропущенных и недостающих данных. Анализ исключений.
5. Основные концепции баз данных. Модель данных. СУБД. Использование баз данных в интеллектуальном анализе данных.
6. Постановка и порядок решения задачи интеллектуального анализа данных на примере хранилища слабоструктурированных текстовых документов.
7. Введение в многомерные базы данных. Особенности базы данных для оперативной аналитической обработки данных. Порядок и особенности построения хранилищ данных.
8. Виды и особенности шкал измерений данных.
9. Введение в OLAP. Типовая модель данных для OLAP. Особенности приложений для оперативной аналитической обработки данных.
10. Постановка и порядок решения задачи интеллектуального анализа данных на примере документной базы данных.
11. Особенности использования данных при интеллектуальном анализе данных. Данные, информация и знания.
12. Основные задачи интеллектуального анализа данных.
13. Стадии интеллектуального анализа данных. Начальные этапы. Вычисление и анализ показателей описательной статистики.
14. Характеристики временных рядов, используемые в ходе интеллектуального анализа данных.
15. Построение и использование моделей для интеллектуального анализа данных.
16. Решение задач классификации в ходе интеллектуального анализа данных: цель и основные методы.
17. Преимущественные особенности методов анализа данных: статистических, машинного обучения, оперативного анализа, интеллектуального анализа.
18. Решение задач кластеризация в ходе интеллектуального анализа данных: цель и основные методы.
19. Сущность и методы решения задачи прогнозирования при

интеллектуальном анализе данных.

20. Комплексный подход к внедрению Data Mining, OLAP и хранилищ данных в СППР

Продвинутый уровень

1. Data Mining
2. Data mining и базы данных
3. Data mining и искусственный интеллект
4. Задачи Data mining
5. Интеллектуальный анализ данных и машинное обучение
6. Архитектура Модели интеллектуального анализа данных
7. Определение Модели интеллектуального анализа данных
8. Свойства Модели интеллектуального анализа данных
9. Классификация данных
10. Кластеризация данных
11. Ассоциативные правила
12. Прогнозирование
13. Деревья решений
14. Области применения Data mining
15. Средства интеллектуального анализа данных

Описание шкалы оценивания

В рамках рейтинговой системы успеваемость студентов по каждой дисциплине оценивается в ходе текущего контроля и промежуточной аттестации.

Текущий контроль

Рейтинговая оценка знаний студента

№ п/п	Вид деятельности студентов	Сроки выполнения	Количество баллов
4 семестр			
1	Практическое занятие 8	15	30
2	Лабораторная работа 8	15	25
	Итого за 4 семестр:		55
	Итого:		55

Максимально возможный балл за весь текущий контроль устанавливается равным **55**. Текущее контрольное мероприятие считается сданным, если студент получил за него не менее 60% от установленного для этого контроля максимального балла. Рейтинговый балл, выставляемый студенту за текущее контрольное мероприятие, сданное студентом в установленные графиком контрольных мероприятий сроки, определяется следующим образом:

<i>Уровень выполнения контрольного задания</i>	<i>Рейтинговый балл (в % от максимального балла за контрольное задание)</i>
<i>Отличный</i>	<i>100</i>
<i>Хороший</i>	<i>80</i>
<i>Удовлетворительный</i>	<i>60</i>
<i>Неудовлетворительный</i>	<i>0</i>

Промежуточная аттестация

Промежуточная аттестация в форме экзамена предусматривает проведение

обязательной экзаменационной процедуры и оценивается 40 баллами из 100. В случае если рейтинговый балл студента по дисциплине по итогам семестра равен 60, то программой автоматически добавляется 32 премиальных балла и выставляется оценка «отлично». Положительный ответ студента на экзамене оценивается рейтинговыми баллами в диапазоне от 20 до 40 ($20 \leq S_{\text{экз}} \leq 40$), оценка меньше 20 баллов считается неудовлетворительной.

Шкала соответствия рейтингового балла экзамена 5-балльной системе

Рейтинговый балл по дисциплине	Оценка по 5-балльной системе
35 – 40	Отлично
28 – 34	Хорошо
20 – 27	Удовлетворительно

Итоговая оценка по дисциплине, изучаемой в одном семестре, определяется по сумме баллов, набранных за работу в течение семестра, и баллов, полученных при сдаче экзамена:

*Шкала пересчета рейтингового балла по дисциплине
в оценку по 5-балльной системе*

Рейтинговый балл по дисциплине	Оценка по 5-балльной системе
88-100	Отлично
72-87	Хорошо
53-71	Удовлетворительно
<53	Неудовлетворительно

Критерии оценки ответов при собеседовании:

Оценка «зачтено» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения. Допускаются некоторые неточности, недостаточно правильные формулировки в изложении программного материала, затруднения при выполнении практических работ.

Оценка «не зачтено» выставляется студенту, если он не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы.

Критерии оценивания лабораторной или практической работы

Оценка «отлично» ставится, если обучающийся выполняет работу в полном объеме с соблюдением необходимой последовательности проведения опытов и измерений; самостоятельно и рационально монтирует необходимое оборудование; все опыты проводит в условиях и режимах, обеспечивающих получение правильных результатов и выводов; соблюдает требования правил техники безопасности; правильно и аккуратно выполняет все записи, таблицы, рисунки, чертежи, графики, вычисления.

Оценка «хорошо» ставится, если выполнены требования к оценке 5, но было допущено два-три недочета.

Оценка «удовлетворительно» ставится, если работа выполнена не полностью, но объем выполненной её части позволяет получить правильный результат и вывод; или если в ходе проведения опыта и измерения были допущены ошибки.

Оценка «неудовлетворительно» ставится, если работа выполнена не полностью или объем выполненной части работ не позволяет сделать правильных выводов; или если опыты, измерения, вычисления, наблюдения производились неправильно.

Составитель _____
(подпись)

« ____ » _____ 2021 г.

Вопросы на экзамен

по дисциплине «Интеллектуальный анализ данных и машинное обучение»

Вопросы к экзамену (7 семестр)

Вопросы (задача, задание) для проверки уровня обученности

Знать

1. Данные и модели их представления.
2. Системы поддержки принятия решений (СППР).
3. Роль и место интеллектуального анализа данных в СППР.
4. Задачи ИАД.
5. Алгебра матриц.
6. Функции многих переменных.
7. Необходимые и достаточные условия существования экстремумов применительно к квадратичным формам.
8. Типы шкал.
9. Допустимые преобразования в шкалах.
10. Проверка истинности утверждений.
11. Статистическая выборка.
12. Числовые характеристики распределений.
13. Комплексные числа и их применение при визуализации многомерных данных.
14. Методы и алгоритмы оцифровки графиков.
15. Методы и алгоритмы обработки изображений.
16. Простые и сложные признаки и способы оценки их информативности.
17. Алгоритмы поиска систем информативных признаков.
18. Матрица объект-признак и её статистические характеристики.
19. Проблема сжатия данных.
20. Разнотипные данные и методы их обработки.
21. Задача поиска логических закономерностей.
22. Методы классификации и прогнозирования.
23. Задачи кластерного анализа.
24. Иерархические и итеративные методы кластеризации.
25. Особенности кластеризации в качественных и количественных шкалах.
26. Кластеризация данных по матрице объект-признак.
27. Кластеризация данных по матрице матрице связи.
28. Назначение компонентного и факторного анализа.
29. Сходство и различие компонентного и факторного анализа.

Уметь,
владеть

30. Применение компонентного и факторного анализа к задачам ИАД.
31. Методы распознавания образов с учителем и без учителя.
32. Задачи принятия решений.
33. Метод анализа иерархий.
34. Модификации метода анализа иерархий в интересах реализации интеллектуальных подсказок пользователям.

35. Основные понятия когнитивного моделирования.
36. Инструментальные средства ИАД применительно к задачам СППР.
37. Направления развития ИАД.
38. Направления развития современных информационных технологий в СППР.
39. Краткая история нейрокомпьютинга.
40. Задачи ИАД на основе искусственных нейронных сетей.
41. Место нейронных сетей среди других методов решения задач
42. Информационный подход к моделированию нейрона.
43. Биологический подход к моделированию нейрона.
44. Структура искусственной нейронные сети.
45. Структура двухкровневого персептрона, многоуровневого персептрона (МСП).
46. Особенности структуры нейронных сетей и ее влияние на свойства сети.
47. Алгоритм решения задач с помощью МСП.
48. Классификация задач решаемых с помощью МСП.
49. Постановка задач распознавания, аппроксимации, прогнозирования. Примеры задач.
50. Топологии нейронных сетей.

Вопросы к экзамену (8 семестр)

Вопросы (задача, задание) для проверки уровня обученности

Знать

1. 1. Data Mining и аналитические информационные технологии.
2. Цели, задачи, математический и программный инструментария ИАД.
3. ИАД в задачах математического моделирования сложных систем.
4. Применение ИАД в задачах моделирования сложных динамических систем, включая технологические и экономические процессы. Проблема неопределенности.
5. Вероятностный дескрипт.
6. Проблема неопределенности, как главная задача анализа данных в интересах задач управления.
7. Вероятностные методы описания неопределенности.
8. Принципы статистического анализа данных средствами ИАД.
9. Статистические методы и алгоритмы ИАД. Особенности применения. Ограничения. Восстановление зависимостей.
10. Линейная регрессия в задачах анализа данных.
11. Восстановление зависимостей и связей одна из основных задач ИАД.
12. Методы регрессионного анализа данных. Статистические решения: Методы проверки гипотез. Статистические решения как терминальная задача ИАД.
13. Решения на основе метода проверки гипотез.
14. Задачи классификации.
15. Дискриминантный анализ данных.
16. Классификационные задачи ИАД. Классификация «с учителем». Алгоритмы дискриминантного анализа данных.
17. Современные технологии прогнозирования.
18. Прогнозирование развития сложных многомерных процессов методами ИАД.
19. Введение в искусственные нейронные сети (ИНС). Модели нейронов. Персептроны.
20. Математическое описание ИНС.
21. Элементы и простейшие модели ИНС.

22. Эволюционные алгоритмы прогноза и распознавания.
23. Прогнозирование и распознавание состояния сложных систем на основе метода эволюционного моделирования.
24. Работа с массивами данных. Вектора, матрицы, операции, логика, множества.
25. Базовые категории и понятия для работы в программной среде Матлаб (R) Двухоконный режим.
26. Графические отображения в среде Матлаб (R). Работа в программном окне.
27. Способы графического отображения в среде Матлаб (R).
28. Основы программирования в среде Матлаб (R).
29. Технология программирования в среде Матлаб (R). Циклы, условные операторы. Экспертная система.
30. Deskриптивный анализ данных. Полный статистический дескрипт вероятностной структуры и параметров данных.
31. Полиномиальные и стохастические модели. Оценка параметров движения.
32. Моделирование динамических стохастических процессов в среде Матлаб (R).
33. Регрессионная оценка параметров движения.
34. Метод статистических испытаний.
35. Имитационное моделирование случайных событий и процессов. Метод Монте-Карло.
36. Многомерные задачи классификации и распознавания.
37. Основные технологии статистической классификации многомерных случайных объектов средствами ИАД.
38. Экстраполяционный прогноз.
39. Прогнозирование на основе фильтра Калмана.
40. Прогнозирование состояния сложных динамических систем статистическими и другими средствами ИАД.
41. ИНС с обратным распространением ошибки.
42. Моделирование двухслойной ИНС с обратным распространением ошибки и применение в задачах распознавания.
43. Модели и прогнозирование хаотических процессов.
44. Моделирование хаотических процессов средствами ИАД.
45. Анализ возможности построения прогноза в хаотических средах.

Уметь,
владеть

1. 1. Основные понятия машинного обучения. Основные постановки задач. Примеры прикладных задач.
2. Линейные пространства. Векторы и матрицы. Линейная независимость. Обратная матрица.
3. Производная и градиент функции. Градиентный спуск. Выпуклые функции.
4. Случайные величины. Дискретные и непрерывные распределения. Примеры.
5. Оценивание параметров распределений, метод максимального правдоподобия. Бутстрэппинг.
6. Линейные методы классификации и регрессии: функционалы качества, методы настройки, особенности применения.
7. Метрики качества алгоритм регрессии и классификации.
8. Оценивание качества алгоритмов. Отложенная выборка, ее недостатки. Оценка полного скользящего контроля. Кросс-валидация. Leave-one-out.
9. Деревья решений. Методы построения деревьев. Их регуляризация.

10. Композиции алгоритмов. Разложение ошибки на смещение и разброс.
11. Случайный лес, его особенности.
12. Методы поиска выбросов в данных. Методы восстановления пропусков в данных. Работа с несбалансированными выборками.
13. Задача анализа потребительской корзины. Поддержка и достоверность. Частые, замкнутые и максимальные частые множества. Алгоритм Априори.
14. Задача кластеризации. Алгоритм K-Means. Оценки качества кластеризации.

Материалы для текущего контроля в виде

тестирования

Комплект тестов 1.

1. Какие науки включены в Data Mining?
 - a. **статистика, базы данных, искусственный интеллект;**
 - b. информатика, базы данных, статистика;
 - c. искусственный интеллект, базы данных, базы знаний;
 - d. информатика, базы данных, хранилища данных.

2. Каким образом технология Data Mining используется в интернет?
 - a. для создания сайтов;
 - b. для организации поисковых систем;**
 - c. для отображения web-страниц.

3. Какие задачи решаются в СППР?
 - a. ввод данных, преобразование данных, вывод данных;
 - b. ввод данных, модификация данных, передача данных;
 - c. ввод данных, хранение данных, анализ данных.**

4. Какой класс задач анализа отвечает за выполнение запросов?
 - a. **информационно-поисковый;**
 - b. оперативно-аналитический;
 - c. интеллектуальный.

5. Какой класс задач анализа отвечает за обобщение данных?
 - a. информационно-поисковый;
 - b. оперативно-аналитический;**
 - c. интеллектуальный.

6. Какой класс задач анализа отвечает за построение моделей?
 - a. информационно-поисковый;
 - b. оперативно-аналитический;
 - c. интеллектуальный.**

7. Какой класс задач анализа отвечает за поиск закономерностей в данных?
 - a. **информационно-поисковый;**

- b. оперативно-аналитический;
 - c. интеллектуальный.
8. Какая подсистема СППР отвечает за ввод данных?
- a. **OLTP;**
 - b. хранилище данных;
 - c. SQL;
 - d. OLAP;
 - e. Data Mining.
9. Какая подсистема СППР отвечает за хранение данных?
- a. OLTP;
 - b. **хранилище данных;**
 - c. SQL;
 - d. OLAP;
 - e. Data Mining.
10. Какая подсистема СППР отвечает за информационно-поисковый анализ данных?
- a. OLTP;
 - b. хранилище данных;
 - c. **SQL;**
 - d. OLAP;
 - e. Data Mining.
11. Какая подсистема СППР отвечает за оперативный анализ данных?
- a. OLTP;
 - b. хранилище данных;
 - c. SQL;
 - d. **OLAP;**
 - e. Data Mining.
12. Какая подсистема СППР отвечает за Интеллектуальный анализ данных и машинное обучение?
- a. OLTP;
 - b. хранилище данных;
 - c. SQL;
 - d. OLAP;
 - e. **Data Mining.**
13. Как реализуется подсистема ввода данных?
- a. с помощью технологии Data Mining;
 - b. с помощью базы данных;
 - c. **с помощью СУБД;**
 - d. с помощью хранилища данных;
 - e. с помощью витрины данных.
14. Какие данные могут храниться в системе анализа?

- a. детализированные;
 - b. обобщенные;**
 - c. детализированные и обобщенные.
15. Могут ли в системе анализа храниться данные в разных форматах?
- a. могут;
 - b. не могут;**
 - c. могут только данные в текстовых форматах.
16. Какие данные могут храниться в системе анализа?
- a. нормализованные;
 - b. избыточные (денормализованные);**
 - c. частично нормализованные.

Комплект тестов 2.

1. Как называются структуры данных, предназначенные для ввода, модификации и поиска?
- a. оперативные источники данных;**
 - b. хранилища данных;
 - c. базы данных;
 - d. файлы данных.
2. Как называются структуры данных, предназначенные для анализа?
- a. оперативные источники данных;
 - b. хранилища данных;**
 - c. базы данных;
 - d. витрины данных.
3. Предметно-ориентированный, интегрированный, неизменяемый, поддерживающий хронологию набор данных, организованный для целей поддержки принятия решений – это
- a. оперативный источник данных;**
 - b. хранилище данных;
 - c. база данных;
 - d. файл данных.
4. Что является главным недостатком виртуального хранилища данных?
- a. большое время обработки запросов;
 - b. значительные ресурсы компьютера;
 - c. разные форматы и кодировки данных в разных ОИД;
 - d. практическая невозможность получения данных за длительный период времени.**
5. Как называются данные, непосредственно переносимые из ОИД?

- a. метаданные;
- b. агрегированные данные;

c. **детальные.**

6. Какие данные отражают сущность события?

- a. измерения;
- b. метаданные;

c. **факты.**

7. Какие данные описывают события?

- a. **измерения;**
- b. метаданные;
- c. факты.

8. Как называются числовые фактические данные, которые могут быть просуммированы по всем измерениям?

- a. **аддитивные;**
- b. полуаддитивные;
- c. неаддитивные.

9. Как называются числовые фактические данные, которые могут быть просуммированы только по определенным измерениям?

- a. аддитивные;
- b. **полуаддитивные;**
- c. неаддитивные.

10. Как называются фактические данные, которые не могут быть просуммированы ни по одному измерению?

- a. аддитивные;
- b. **полуаддитивные;**

c. **неаддитивные.**

11. На какие вопросы должны отвечать метаданные?

- a. **что, кто, где, как, когда, почему;**
- b. что, кто, зачем, как, когда, почему;
- c. что, кто, где, как, по какой причине, почему;
- d. что, кто, где, как, когда, сколько.

12. Репозиторий – это

- a. словарь терминов;
- b. **хранилище метаданных;**
- c. каталог с файлами.

13. Если поток образуется данными, копируемыми из ОИД, то он называется

- a. **входным потоком;**
- b. потоком обобщения;

- c. архивным потоком;
- d. потоком метаданных;
- e. обратным потоком.

14. Если поток образуется агрегированными данными, то он называется

- a. входным потоком;

b. потоком обобщения;

- c. архивным потоком;
- d. потоком метаданных;
- e. обратным потоком.

15. Если поток образуется детальными данными, количество обращений к которым снизилось, то он называется

- a. входным потоком;
- b. потоком обобщения;

c. архивным потоком;

- d. потоком метаданных;
- e. обратным потоком.

16. Если поток образуется данными, переносимыми в репозиторий, то он называется

- a. входным потоком;
- b. потоком обобщения;
- c. архивным потоком;

d. потоком метаданных;

- e. обратным потоком.

17. Если поток образуется очищенными данными, записываемыми в ОИД, то он называется

- a. входным потоком;
- b. потоком обобщения;
- c. архивным потоком;
- d. потоком метаданных;

e. обратным потоком.

Комплект тестов 3.

1. Многомерный анализ – это:

- a. **одновременный анализ по нескольким измерениям;**
- b. одновременный анализ по нескольким параметрам;
- c. одновременный анализ по нескольким данным.

2. Формирование подмножества многомерного массива данных, соответствующего единственному значению одного или нескольких элементов измерений, не входящих в это подмножество, называется

- a. гиперкубом;

b. срезом гиперкуба;

- c. базой данных;
- d. витриной данных.

3. Технология оперативной аналитической обработки данных, использующая методы и средства для сбора, хранения и анализа многомерных данных в целях поддержки процессов принятия решений, называется

- a. OLTP;
- b. OLAP;**
- c. Data Mining.

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

- a. MOLAP;**
- b. ROLAP;
- c. HOLAP;
- d. DOLAP;
- e. JOLAP.

5. Если для реализации многомерной модели используют реляционные базы данных, то способ реализации гиперкуба называется

- a. MOLAP;
- b. ROLAP;**
- c. HOLAP;
- d. DOLAP;
- e. JOLAP.

6. Если для реализации многомерной модели используют и многомерные, и реляционные базы данных, то способ реализации гиперкуба называется

- a. MOLAP;
- b. ROLAP;
- c. HOLAP;**
- d. DOLAP;
- e. JOLAP.

7. Настольная OLAP, предназначенная для локального анализа и представления данных, называется

- a. MOLAP;
- b. ROLAP;
- c. HOLAP;
- d. DOLAP;**
- e. JOLAP.

8. OLAP, предназначенная для создания и управления данными и метаданными, называется

- a. MOLAP;
- b. ROLAP;
- c. HOLAP;
- d. DOLAP;**

e. JOLAP.

9. В каком отношении находятся таблица фактов и таблица измерений?

- a. «один-к-одному»;
- b. «один-ко-многим»;**
- c. «многие-ко-многим».

10. Исследование и обнаружение машиной (алгоритмами, средствами искусственного ин- теллекта) в сырых данных скрытых знаний, которые: ранее не были известны, нетриви- альны, практически полезны, доступны для интерпретации человеком, называется

- a. OLTP;
- b. хранилищем данных;
- c. OLAP;

d. Data Mining.

11. Какие операции над данными включены в ETL-процесс?

- a. ввод, модификация, вывод;
- b. чтение, изменение, запись;
- c. извлечение, преобразование, загрузка;**
- d. получение, хранение, анализ.

12. Гиперкуб – это ...

- a. объект, все измерения которого имеют одинаковую размерность;
- b. поликуб;
- c. объект, все измерения которого имеют разную размерность;**
- d. многомерный куб;
- e. многомерная база данных.

13. Многомерный просмотр данных основан на ...

- a. многомерной базе данных;**
- b. технологии мультимедиа;
- c. многослойной базе;
- d. сетевой технологии;
- e. гипертекстовой технологии.

14. Какую технологию используют большинство хранилищ данных?

- a. концептуальную БД;
- b. реляционную БД;**
- c. иерархическую БД;
- d. физическую БД.

15. Какое конструирование у хранилищ данных?

- a. физическое;
- b. логическое;**
- c. логическое и физическое;

- d. иерархическое.
16. Концептуальную модель хранилища данных можно представить в виде...
- a. таблицы;
 - b. графического рисунка;**
 - c. геометрических фигур;
 - d. схемы.

Комплект тестов 4.

1. К классу описательных задач относятся:
- a. кластеризация и классификация;
 - b. кластеризация и поиск ассоциативных правил;**
 - c. классификация и регрессия;
 - d. классификация и поиск ассоциативных правил.
2. К классу предсказательных задач относятся:
- a. кластеризация и классификация;
 - b. кластеризация и поиск ассоциативных правил;
 - c. классификация и регрессия;**
 - d. классификация и поиск ассоциативных правил.
3. К классу задач *supervised learning* (обучение с учителем) относятся:
- a. кластеризация и классификация;
 - b. кластеризация и поиск ассоциативных правил;
 - c. классификация и регрессия;**
 - d. классификация и поиск ассоциативных правил.
4. К классу задач *unsupervised learning* (обучение без учителя) относятся:
- a. кластеризация и классификация;
 - b. кластеризация и поиск ассоциативных правил;**
 - c. классификация и регрессия;
 - d. классификация и поиск ассоциативных правил.
5. Задача классификации сводится к ...
- a. нахождению частых зависимостей между объектами или событиями;
 - b. определению класса объекта по его характеристикам;**
 - c. определению по известным характеристикам объекта значения некоторого его параметра;
 - d. поиску независимых групп и их характеристик во всем множестве анализируемых данных.
6. Задача регрессии сводится к ...
- a. нахождению частых зависимостей между объектами или событиями;
 - b. определению класса объекта по его характеристикам;

с. определению по известным характеристикам объекта значения некоторого его параметра;

d. поиску независимых групп и их характеристик во всем множестве анализируемых данных.

7. Задача кластеризации заключается в ...

a. нахождении частых зависимостей между объектами или событиями;

b. определении класса объекта по его характеристикам;

c. определении по известным характеристикам объекта значения некоторого его параметра;

d. поиске независимых групп и их характеристик во всем множестве анализируемых данных.

8. Целью поиска ассоциативных правил является ...

a. **нахождение частых зависимостей между объектами или событиями;**

b. определение класса объекта по его характеристикам;

c. определение по известным характеристикам объекта значения некоторого его параметра;

d. поиск независимых групп и их характеристик во всем множестве анализируемых данных.

9. К описательным моделям относятся следующие модели данных:

a. **модели классификации и последовательностей;**

b. регрессивные, кластеризации, исключений, итоговые и ассоциации;

c. модели классификации, кластеризации, исключений, итоговые и ассоциации;

d. модели классификации, последовательностей и исключений.

10. Модели классификации описывают ...

a. **правила или набор правил, в соответствии с которыми можно отнести опи- сание любого нового объекта к одному из классов;**

b. функции, которые позволяют прогнозировать изменения непрерывных числовых параметров;

c. функциональные зависимости между зависимыми и независимыми показателями и переменными в понятной человеку форме;

d. группы, на которые можно разделить объекты, данные о которых подвергаются анализу.

11. Модели последовательностей описывают ...

a. правила или набор правил, в соответствии с которыми можно отнести описание любого нового объекта к одному из классов;

b. функции, которые позволяют прогнозировать изменения непрерывных числовых параметров;

c. функциональные зависимости между зависимыми и независимыми показателями и переменными в понятной человеку форме;

d. группы, на которые можно разделить объекты, данные о которых подвергаются анализа.

12. Регрессивные модели описывают ...

- a. правила или набор правил в соответствии с которыми можно отнести описание любого нового объекта к одному из классов;
- b. функции, которые позволяют прогнозировать изменения непрерывных числовых параметров;
- c. функциональные зависимости между зависимыми и независимыми показателями и переменными в понятной человеку форме;**
- d. группы, на которые можно разделить объекты, данные о которых подвергаются анализу.

Комплект тестов 5.

1. Какова основная цель создания и использования хранилищ данных:
 - a. анализ данных для принятия управленческих решений;**
 - b. надежное хранение, накопленных данных;
 - c. резервное копирование данных.

2. OLAP - это:
 - a. технология онлайн-быстрой аналитической обработки сложных запросов к базе данных;**
 - b. технология онлайн-обработки небольших по размерам, но идущих большим потоком транзакций, требующих максимально быстрого ответа.

3. Какие из вариантов ответов являются характерными требованиями к хранению данных для принятия решений в хранилищах данных?
 - a. данные ориентированы на приложения;
 - b. данные управляются транзакциями;**
 - c. данные обобщены либо очищены.**

4. Назовите основной этап работы с хранилищами данных?
 - a. этап очистки данных;**
 - b. этап обновления;
 - c. этап нормализации.

5. Что называют кубом OLAP?
 - a. структуру, в которой хранятся совокупности данных, полученные путем всех возможных сочетаний измерений в таблице измерений;**
 - b. структуру, в которой хранятся совокупности данных, полученные путем всех возможных сочетаний измерений в таблице фактов;
 - c. таблицу размерностей.

6. Информационные хранилища созданы для удобства ...
 - a. руководителей всех уровней для принятия решений;**
 - b. стратегического планирования;**
 - c. реорганизации бизнеса;
 - d. предметных приложений;
 - e. редактирования данных.

7. Информационные хранилища размещаются на ...
- a. библиотеках-автоматах;
 - b. сетевых серверах;**
 - c. мейнфреймах;
 - d. серверах и кластерах серверов;
 - e. файл-серверах.
8. При слиянии данных в информационное хранилище из внутренних и внешних источников обеспечивается ...
- a. предметная ориентация данных;**
 - b. выбор требуемых сведений из предметных приложений по наименованиям;
 - c. гипертекстовый просмотр данных;
 - d. согласование данных по наименованию;
 - e. хранение данных по предметным областям.
9. Интеллектуальный выбор данных из информационного хранилища – это ...
- a. реализация методов искусственного интеллекта;
 - b. выбор по заданному алгоритму;**
 - c. реализация самообучающихся систем;**
 - d. реализация экономико-статистических методов.
10. В процессе погружения в информационное хранилище данные ...
- a. очищаются от ненужной для анализа информации;
 - b. агрегируются;**
 - c. преобразуются из разных типов данных предметных приложений в единую структуру хранения;**
 - d. индексируются;
 - e. синхронизируются.

Комплект тестов 6.

1. Классификация — ...
- a. некоторый набор операций над базой данных, который рассматривается как единственное завершённое, с точки зрения пользователя, действие над некоторой информацией, обычно связано с обращением к базе данных;
 - b. разновидность систем хранения, ориентированная на поддержку процесса анализа данных, обеспечивающая непротиворечивость и хронологию данных, а также высокую скорость выполнения аналитических запросов;
 - c. высокоуровневые средства отражения информационной модели и описания структуры данных;
 - d. это установление зависимости дискретной выходной переменной от входных переменных.**
2. Регрессия — ...
- a. это установление зависимости непрерывной выходной переменной от входных переменных;**
 - b. это группировка объектов (наблюдений, событий) на основе данных,

описывающих свойства объектов;

- c. выявление закономерностей между связанными событиями;
- d. это установление зависимости дискретной выходной переменной от входных переменных.

3. Кластеризация — ...

- a. это установление зависимости непрерывной выходной переменной от входных переменных;

- b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;**

- c. выявление закономерностей между связанными событиями;
- d. это установление зависимости дискретной выходной переменной от входных переменных.

4. Ассоциация — ...

- a. это установление зависимости непрерывной выходной переменной от входных переменных;

- b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;**

- c. выявление закономерностей между связанными событиями;**

- d. это установление зависимости дискретной выходной переменной от входных переменных.

5. Машинное обучение — ...

- a. специализированное программное решение (или набор решений), который включает в себя все инструменты для извлечения закономерностей из сырых данных;

- b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;**

- c. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат;

- d. подразделение искусственного интеллекта, изучающее методы построения алгоритмов, способных обучаться на данных.**

6. Аналитическая платформа — ...

- a. специализированное программное решение (или набор решений), который включает в себя все инструменты для извлечения закономерностей из сырых данных;**

- b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;**

- c. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат;

- d. подразделение искусственного интеллекта, изучающее методы построения алгоритмов, способных обучаться на данных.

7. Обучающая выборка — ...

- a. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;

б. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр и соответствующий ему правильный выходной результат;

с. выявление в сырых данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

8. Ошибка обучения — ...

а. **это ошибка, допущенная моделью на учебном множестве;**

б. это ошибка, полученная на тестовых примерах, то есть вычисляется по тем же формулам, но для тестового множества;

с. имена, типы, метки и назначения полей исходной выборки данных;

д. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат.

9. Ошибка обобщения — ...

а. это ошибка, допущенная моделью на учебном множестве;

б. это ошибка, полученная на тестовых примерах, то есть вычисляется по тем же формулам, но для тестового множества;

с. имена, типы, метки и назначения полей исходной выборки данных;

д. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат.

Критерии оценки ответов на экзамене:

Оценка «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, использует в ответе материал монографической литературы, правильно обосновывает принятое решение, владеет разносторонними навыками и приемами выполнения практических задач.

Оценка «хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.

Оценка «удовлетворительно» выставляется студенту, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала, испытывает затруднения при выполнении практических работ.

Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы.