

Типовые вопросы к экзамену

1. Предпосылки возникновения интеллектуального анализа данных.
2. Разведочный анализ данных. Цели и основные инструменты.
3. Принцип «мягких вычислений», определение и идеология Data Mining.
4. Подходы технологии ИАД. Классификация методов.
5. Стадии интеллектуального анализа данных.
6. Основные задачи Data Mining.
7. Описательные и предсказательные задачи Data Mining.
8. Количественный анализ – суть и предназначение. Фазы количественной обработки данных.
9. Качественный анализ - суть и предназначение.
10. Обучение с учителем, обучение без учителя.
11. Практическое применение Data Mining.
12. Понятие корреляционно-регрессионного анализа данных. Этапы КРА.
13. Корреляционная зависимость. Коэффициент корреляции Пирсона. Коэффициент ранговой корреляции Спирмена. Коэффициент сопряженности Бравайса.
14. Суть простого регрессионного анализа и его этапы. Общая постановка задачи исследователя.
15. Адекватность модели регрессии, анализ остатков, гетероскедастичность, коэффициент Дарбина-Ватсона.
16. Дисперсионный анализ. Постановка задачи.
17. Сущность и техника дисперсионного анализа в зависимости от числа изучаемых независимых факторов.
18. Виды экспериментов и моделей в дисперсионном анализе.
19. Однофакторный дисперсионный анализ. Методика проведения.
20. Рассеивание по факторам, остаточное рассеивание и полная сумма квадратов отклонений отдельных наблюдений от общей средней в однофакторном дисперсионном анализе.
21. F-критерий и доли влияния учтенного и неучтенного факторов в однофакторном дисперсионном анализе.
22. Многофакторный дисперсионный анализ – суть, ограничения и основное тождество.
23. Двухфакторный дисперсионный анализ. Методика и алгоритмы.
24. Многомерный статистический анализ. Понятие и основные положения.
25. Классификация методов многомерного статистического анализа. Выбор метода для решения задач.
26. Кластерный анализ. Основные понятия.
27. Понятие расстояния между объектами в пространстве признаков. Аксиомы.
28. Метрики кластеризации.
29. Особенности кластерного анализа. Основные задачи.
30. Общая методология и этапы кластеризации.
31. Кластерный анализ. Классификация методов.

32. Кластерный анализ. Дендродрамма.
33. Кластерный анализ. Методы связи.
34. Неиерархические методы кластеризации. Алгоритм К-средних
35. Иерархические методы кластеризации – достоинства и недостатки.
36. Качество кластеризации.
37. Факторный и компонентный анализ.
38. Основные цели и особенности факторного анализа.
39. Этапы факторного анализа. Свойства гипотетических факторов.
40. Метод главных компонент.
41. Дискриминантный анализ. Определение и основные задачи.
42. Дискриминация и её цель. Каноническая дискриминантная функция.
43. Поиск наилучшей дискриминации. Коэффициенты канонической дискриминантной функции.
44. Структурные и стандартизованные коэффициенты канонической дискриминантной функции.
45. Коэффициент канонической корреляции. Остаточная дискриминация.
46. Классифицирующие функции.
47. Метод опорных векторов.
48. Решение задачи бинарной классификации при помощи метода линейных опорных векторов.
49. Проблемы классификации в SVM. Достоинства и недостатки метода.
50. Метод «ближайшего соседа».
51. Подход, основанный на прецедентах. Достоинства и недостатки.
52. Метод «ближайшего соседа» для задачи прогнозирования.
53. Суть методов кросс-табуляции.
54. Байесовская сеть. Определение и применение для анализа данных.
55. Байесовские методы. Байесовская классификация (байесовский подход и правило).
56. Байесовская классификация. Постановка задачи. Общая структура байесовского классификатора.
57. Байесовская классификация. Многомерная нормальная плотность. Алгоритм классификации.
58. Решение задачи классификации новых объектов методом k-ближайших соседей
59. Достоинства байесовских сетей как метода Data Mining
60. Суть метода наивно-байесовской классификации. Недостатки метода.
61. Оптимальный байесовский классификатор.
62. Ряды динамики. Классификация временных рядов.
63. Выявление аномальных уровней временных рядов. Метод Ирвина.
64. Компоненты (составляющие) значений уровней временных рядов. Аддитивная, мультипликативная и смешанная формы моделей уровней временных рядов.
65. Определение существования тенденции во временном ряду.
66. Основные показатели динамики и получение прогнозных значений.

- 67. Анализ дискретных временных рядов. Коэффициент автокорреляции. Автокорреляционная функция временного ряда.
- 68. Основные подходы к анализу временных рядов. Их возможности и недостатки.
- 69. Методы сглаживания и экстраполяции временных рядов.
- 70. Стационарные временные ряды.