

Интеллектуальный анализ данных

Вопросы к экзамену весна 2017 г.

Теоретически вопросы

1. Науки, на которых основан интеллектуальный анализ данных.
2. Понятие Data Mining.
3. Данные, используемые в интеллектуальном анализе данных.
4. Типы наборов данных.
5. Базы данных. Основные положения.
6. Система управления базой данных.
7. Метаданные.
8. Понятие временного ряда.
9. Операторы, используемые в анализе временных рядов.
10. Модель линейного фильтра.
11. Модель авторегрессии.
12. Модель скользящего среднего.
13. Смешанные модели авторегрессии – скользящего среднего.
14. Нестационарные модели временных рядов.
15. Модели многомерных временных рядов.
16. Авторегрессионная модель многомерного временного ряда.
17. Фильтр Калмана.
18. Деревья решений.
19. Преимущества деревьев решений.
20. Процесс конструирования дерева решений.
21. Алгоритмы, реализующие деревья решений.
22. Понятие нейронных сетей.
23. Задачи Data Mining, решаемых с помощью нейронных сетей.
24. Элементы нейронных сетей.
25. Активационные функции нейронных сетей.
26. Архитектура нейронных сетей.
27. Обучение нейронных сетей.
28. Модели нейронных сетей.
29. Программное обеспечение для работы с нейронными сетями.
30. Понятие кластерного анализа.
31. Задачи кластерного анализа.
32. Математические характеристики кластера.
33. Методы кластерного анализа.
34. Иерархические методы кластерного анализа.
35. Меры сходства.
36. Итеративные методы кластерного анализа.
37. Алгоритм k –средних.
38. Алгоритм PAM.
39. Факторный анализ кластеров.
40. Алгоритм BIRCH.
41. Алгоритм WaveCluster.
42. Алгоритм CLARA.

Вопросы по лабораторным работам

1. Как рассчитывается частный коэффициент корреляции?

2. Как рассчитывается множественный коэффициент корреляции?
3. Для чего используется коэффициент детерминации?
4. Как проверяется значимость уравнения регрессии?
5. Как проверяется значимость коэффициентов регрессии?
6. Как строятся интервальные оценки коэффициентов уравнения регрессии?
7. Как разделить общую выборку для целей дискриминантного анализа и проверки?
8. Как определяют статистическую значимость дискриминантного анализа?
9. Какой стохастический процесс называется «белый шумом»?
10. Какие методы применяются для распознавания стационарности временных рядов?
11. В каких случаях процессы AR являются стационарными?
12. Как используется автокорреляционная функция для идентификации модели стационарного стохастического процесса?
13. Как используется частная автокорреляционная функция для идентификации модели стационарного стохастического процесса?
14. Как осуществляется прогнозирование ARMA-процессов?
15. Что может служить признаком нестационарности временного ряда?
16. Для чего применяются Тесты Дики-Фуллера?
17. Охарактеризуйте процессы ARIMA.
18. Как осуществляется прогнозирование ARMA-процессов?
19. Основные задачи, которые решают кластерный и дискриминантный анализ?
20. Что такое тестовая выборка в кластерном анализе?
21. Какие существуют вероятности принадлежности отдельных значений выбранной переменной к той или иной группе в кластерном анализе?
22. Что такое функция классификации в кластерном анализе?
23. Что такое кросс-проверка в кластерном анализе?
24. Какие существуют меры расстояний между объектами при древовидной кластеризации?
25. Правила объединения или связи для двух кластеров?