

Экзамен

Список экзаменационных вопросов

1. Проблема обработки данных. Матрица данных. гипотеза компактности и скрытых факторов.
2. Структура матрицы данных и задачи обработки. Матрица объект-объект и признак-признак. Расстояние и близость.
3. Измерение признаков. Отношения и их представление. Основные проблемы измерений.
4. Основные типы шкал. Проблема адекватности. Классификация данных: общая постановка задачи.
5. Решающие функции и основные подходы к их построению.
6. Классификация данных как статистическая задача.
7. Классификация в случае двух нормальных распределений с равными матрицами ковариаций.
8. Примеры построения решающих функции для нормальных распределений с равными диагональными матрицами ковариаций.
9. Ошибки классификации для случая двух нормальных распределений с равными матрицами ковариаций.
10. Апостериорная вероятность отнесения данных к классу (на основе теоремы Байеса) для случая многомерных нормальных распределений.
11. Классификация при количестве классов больше двух (нормальное распределение с равными матрицами ковариаций).
12. Классификация для случая двух нормальных распределений с разными матрицами ковариаций.
13. Линейный дискриминант Фишера.
14. Пошаговый дискриминантный анализ.
15. Кластерный анализ: общая постановка задачи, определение расстояний между объектами и кластерами, критерии кластеризации.
16. Последовательная процедура итеративной оптимизации в задачах кластер-анализа.
17. Параллельная процедура кластеризации. Алгоритм K-внутригрупповых средних.
18. Алгоритм автоматической классификации на основе алгоритма адаптивного выбора подклассов (АВП).
19. Иерархическая группировка.
20. Обучаемые классификаторы: детерминистский подход. Вероятность получения линейного разделения классов.
21. Построение линейных решающих правил персептронного типа – обучение с коррекцией ошибок.
22. Построение линейных решающих функции методом градиентной минимизации функции качества.
23. Алгоритмы оценки информативности признаков.
24. Метод главных компонент для выбора признаков.
25. Факторный анализ: общая модель.
26. Структура факторных уравнений. Неоднозначность факторного решения. Метод главных факторов.
27. Метод центроидных факторов.
28. Проблема оценки значений факторов и виды факторных моделей.
29. Оценки общностей и вращение факторов.

30. Многомерное шкалирование.

Перечень экзаменационных задач

1. Выполнить центрирование и нормирование матрицы данных.
2. Построить байесовское решающее правило для двух классов для нормального распределения.
3. Построить решающее правило для классификации двух классов на основе апостериорных вероятностей.
4. Найти уравнение линии равной плотности вероятностей $f(x) = C$, для двумерного нормального распределения.
5. Построить решающую функцию для классификации 2-х нормальных классов
6. Найти расстояние Махаланобиса для двух классов. Найти выражение для средней ошибки классификации этих классов с использованием байесовской решающей функции.
7. Построить решающее правило для классификации двух классов с разными матрицами ковариации.
8. Написать первые 3 шага персептронной процедуры обучения для классификации двух классов X_1 , X_2 , состоящих из векторов заданных построчно в матрицах X_1 , X_2 .
9. Определить расстояние между двумя кластерами C_1 , C_2 по методу ближайшего соседства.
10. Написать два шага процедуры кластеризации по методу K-средних.
11. Произвести иерархическую кластеризацию данных, заданных построчно в матрице C .
12. Найти выражения главных компонент для набора данных с заданной матрицей ковариации Σ .

From:
<http://se.moevm.info/> - **se.moevm.info**

Permanent link:
http://se.moevm.info/doku.php/courses:data_analysis_and_interpretation:exam?rev=1562844177

Last update: **2022/12/10 09:08**