



Тест по темам: смеси, ЕМ, обнаружение аномалий

В тестовых заданиях первая галочка — правильный ответ, вторая галочка — выбранный ответ. Цвет обозначает, правильно ли в данном пункте поставлена галочка. Если все пункты верные (галочки совпадают / все пункты зеленые), то за задание ставится полный балл, в противном случае ставится 0 баллов.

1. В методе изолирующего леса (isolation forest) точки, находящиеся в листьях, более близких к корню дерева, считаются

☐ ☐ более типичными

☒ ☒ более аномальными

Балл: 2.0

Комментарий к правильному ответу:

2. Увеличение параметра n_i в одноклассовом методе опорных векторов приводит к

☐ ☐ уменьшению числа обнаруженных выбросов

☒ ☒ увеличению числа обнаруженных выбросов

Балл: 2.0

Комментарий к правильному ответу:

3. ЕМ алгоритм в общем виде

☒ ☒ приводит к неубыванию значений правдоподобия

☐ ☐ правдоподобие может как увеличиваться, так и уменьшаться в процессе алгоритма

Балл: 2.0

Комментарий к правильному ответу:

4. Может ли метод кластеризации DB-SCAN детектировать объекты-выбросы?

☒ ☒ да

☐ ☐ нет

Балл: 2.0

Комментарий к правильному ответу:

5. Выберите верные ответы про ЕМ-алгоритм в оценке смеси Гауссиан?

☐ ☐ кластеры могут иметь произвольную, в том числе невыпуклую, форму



- ☒ ☒ кластеры могут иметь вытянутую эллипсоидальную форму
- ☒ ☒ делает нечёткую кластеризацию (объект может относиться к нескольким кластерам с разными вероятностями)
- ☐ ☐ в методе автоматически определяется число компонент (кластеров)

Балл: 2.0

Комментарий к правильному ответу:

6. Пусть мы оцениваем фиксированную выборку несовпадающих наблюдений смесью К Гауссиан, $K \geq 2$. Является ли максимизируемый функционал правдоподобия вогнутым (в зависимости от параметров смеси)?

☒ ☐ нет

☐ ☒ да

Балл: 0

Комментарий к правильному ответу:

перенумерация компонент смеси приводит к решению с таким же значением функционала, но их усреднение ухудшает качество

7. Выберите статистики, оценивающие центр распределения устойчиво к выбросам:

☒ ☒ среднее из 25% и 75% персентили

☒ ☒ медиана

☐ ☐ среднее

Балл: 2.0

Комментарий к правильному ответу:

8. На М шаге EM алгоритма распределение над латентными переменными

☐ ☐ зависит от параметров, по которым производится максимизация М-шага

☒ ☒ фиксировано и не зависит от параметров, по которым производится максимизация М-шага

Балл: 2.0

Комментарий к правильному ответу: