



## Тест по темам: смеси, EM, 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. Возможно ли использовать ансамблирование (ensemble learning) при детектировании выбросов:

☐ ☐ да, но только если алгоритмы ансамбля принадлежат разным семействам

☐ ☐ нет

☒ ☒ да

**Балл: 2.0**

**Комментарий к правильному ответу:**

2. Local Outlier (LOF) метод решает задачу:

☒ ☒ обнаружения выбросов

☐ ☐ реализации EM-алгоритма

☐ ☐ кластеризации

☐ ☐ классификации

**Балл: 2.0**

**Комментарий к правильному ответу:**

3. В чём отличие задач outlier detection от novelty detection?

☒ ☒ в наличии выбросов в обучающей выборке

☐ ☐ в способности выбросов быть детектированными

☐ ☐ в степени аномальности выбросов, которые необходимо найти

**Балл: 2.0**

**Комментарий к правильному ответу:**

4. Может для некоторого распределения выборки объект-выброс совпадать со средним значением по всем объектам обучающей выборки?

☒ ☒ да

☐ ☐ нет

**Балл: 2.0**

**Комментарий к правильному ответу:**



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

☐ ☒ да

☒ ☐ нет

**Балл:** 0

**Комментарий к правильному ответу:**

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

6. Выберите верные ответы про EM-алгоритм в оценке смеси Гауссиан?

☒ ☒ кластеры могут иметь вытянутую эллипсоидальную форму

☐ ☐ кластеры могут иметь произвольную, в том числе невыпуклую, 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:**

7. Пусть мы оцениваем фиксированную выборку несовпадающих наблюдений смесью  $K$  Гауссиан,  $K \geq 2$ . Можно ли сделать функцию правдоподобия сколь угодно большой специальным выбором параметров смеси?

☐ ☐ да, но только при частных видах обучающей выборки

☐ ☐ нет

☒ ☒ да

**Балл:** 2.0

**Комментарий к правильному ответу:**

$\mu(k) = x(k)$ , а  $\sigma(k)$  стремим к нулю

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

☐ ☐ нет

☒ ☒ да



**Балл: 2.0**

**Комментарий к правильному ответу:**