ТЕСТ ПО ТЕМЕ: ОБУЧЕНИЕ БЕЗ УЧИТЕЛЯ

Вопрос 1.

Что является основной особенностью обучения без учителя (Unsupervised Learning)?

A. Использование размеченных данных для обучения модели

**B. Отсутствие целевых переменных в обучающих данных**

C. Обязательное разделение данных на train/validation/test

D. Использование функции потерь для оптимизации параметров

Вопрос 2.

Какой алгоритм кластеризации требует предварительного указания количества кластеров?

A. DBSCAN

B. Mean Shift

**C. K-Means**

D. Hierarchical Clustering (с автоматическим определением)

Вопрос 3.

Что измеряет метрика Silhouette Score?

A. Расстояние между центроидами кластеров

B. Количество точек в каждом кластере

**C. Насколько точки похожи на свой кластер по сравнению с другими кластерами**

D. Время выполнения алгоритма кластеризации

Вопрос 4.

Какое значение Silhouette Score указывает на идеальную кластеризацию?

A. 0

B. 0.5

**C. 1**

D. -1

Вопрос 5.

Что делает алгоритм DBSCAN с точками, которые не принадлежат ни к одному кластеру?

A. Удаляет их из dataset'а

B. Создает для них отдельные кластеры

**C. Помечает их как шум (noise)**

D. Присваивает их к ближайшему кластеру

Вопрос 6.

Какие два основных параметра требуются для работы алгоритма DBSCAN?

A. k (количество кластеров) и max\_iter (максимальное количество итераций)

**B. eps (радиус окрестности) и min\_samples (минимальное количество точек)**

C. learning\_rate (скорость обучения) и batch\_size (размер батча)

D. n\_components (количество компонент) и random\_state (случайное состояние)

Вопрос 7.

Что означает метрика Adjusted Rand Index (ARI) равная 0?

A. Идеальное соответствие истинным меткам

B. Полное несоответствие истинным меткам

**C. Случайная кластеризация**

D. Невозможно вычислить метрику

Вопрос 8.

Для какой задачи лучше всего подходит иерархическая кластеризация?

A. Когда заранее известно точное количество кластеров

**B. Когда нужно получить дендрограмму и изучить структуру данных на разных уровнях**

C. Когда данные имеют сферическую форму кластеров

D. Когда требуется максимальная скорость работы алгоритма

Вопрос 9.

Что показывает метод локтя (Elbow Method) в контексте K-Means?

A. Оптимальное значение параметра eps для DBSCAN

B. Количество выбросов в данных

**C. Оптимальное количество кластеров на основе внутрикластерной суммы квадратов**

D. Время выполнения алгоритма для разных параметров

Вопрос 10.

В каком случае Supervised Learning имеет преимущество перед Unsupervised Learning?

A. Когда нужно исследовать структуру неизвестных данных

**B. Когда есть размеченные данные и нужно делать предсказания на новых данных**

C. Когда требуется найти скрытые паттерны в данных

D. Когда нужно снизить размерность данных

## ОСНОВНЫЕ ТЕМЫ, ОХВАЧЕННЫЕ В ТЕСТЕ:

1. Основы обучения без учителя

2. Алгоритмы кластеризации (K-Means, DBSCAN, Hierarchical)

3. Метрики оценки качества кластеризации

4. Параметры алгоритмов и их настройка

5. Сравнение Supervised vs Unsupervised подходов

6. Практическое применение алгоритмов