

Теоретические вопросы к коллоквиуму по курсу Machine Learning

Правила коллоквиума

- Билет состоит из трех вопросов: двух по теории, которые берутся из списка ниже, и одного по практике, который заранее не известен.
- В течении всего экзамена должны быть включены камера, микрофон, показ экрана.
- В случае обнаружения факта попытки списать коллоквиум заканчивается оценкой неуд.

Теоретические вопросы

1 Задачи машинного обучения

Обучение с учителем. Обучение без учителя. Регрессия. Классификация. Кластеризация.

2 Приближение функции многочленами

Задача, модель, функция потерь. Проблема переобучения на примере интерполяции. Зачем нужно разбиение выборки на тренировочную и тестовую часть. Кросс-валидация.

3 Линейная регрессия

Задача и модель линейной регрессии. Регуляризация.

4 Логистическая регрессия

Задача и модель бинарной логистической регрессии.

5 Метрики классификации

Метрики классификации: accuracy, precision, recall, F1, ROC-AUC. Multi-label классификация.

6 Кластеризация

Задача кластеризации. K-means. DBSCAN. Агломеративная кластеризация.

7 Уменьшение размерности

Задача уменьшения размерности. SVD. PCA. T-SNE.

8 Решающие деревья

Алгоритм построения решающего дерева. Кросс-энтропия.

9 Ансамблирование

Ансамбль алгоритмов. Бэггинг. Случайный лес. Градиентный бустинг.

10 Обработка текстов

Токенизация, Лемматизация, CountVectorizer, OneHotVectorizer, Стемминг.
Word2vec (идея, какие полезные свойства эмбедингов дает).

11 Релевантность документа

Задача проверки на релевантность документа на запрос. TF-IDF.

12 Рекомендательные системы

Постановка задачи. Бейзлайны. Фильтрация на основе содержания (content-based approach). Коллаборативная фильтрация.

13 Теория вероятности

Вероятность. Условная вероятность. Формула Байеса. Формула полной вероятности. Распределения вероятностей. ММП. Наивный байесовский классификатор.

14 Линейная алгебра, матрицы

Определитель, обратная матрица, алгоритмы обращения матриц, собственные значения и векторы матрицы, SVD, матричное дифференцирование, аналитическое решение задачи линейной регрессии