

# «Теоретический минимум»

## январь 2025

1. Матричное дифференцирование. Определение  $Df(x)[\Delta x]$ . Связь  $Df(x)[\Delta x]$  с градиентом  $(\frac{\partial f}{\partial x})$  в случаях:
  - 1)  $f : \mathbb{R}^{n \times 1} \rightarrow \mathbb{R}$
  - 2)  $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$
2. Градиентный спуск (GD), стохастический градиентный спуск (SGD), метод Ньютона.
3. Настройка гиперпараметров с помощью кросс-валидации. K-fold cross-validation, leave-one-out, кросс-валидация в работе с временными рядами.
4. Формула классификации и регрессии в модели KNN. Примеры весов. Примеры метрик.
5. Определение линейного классификатора. Понятие отступа (Margin). Примеры верхних оценок на долю ошибок. Как обучаются веса?
6. Подходы one-vs-one и one-vs-rest для линейных классификаторов.
7. Переобучение. Способы борьбы с переобучением.
8. Матрица ошибок (confusion matrix). Определение accuracy, precision, recall, f1-measure.
9. Рок кривая, AUC-ROC.
10. Определение решающего дерева. Критерии расщепления в случае задачи классификации и регрессии.
11. Постановка задачи PCA. Как выбирать оптимальную размерность маломерного пространства?
12. Постановка задачи кластеризации. Алгоритм K-means.
13.  $l_1$ ,  $l_2$  регуляризация в задаче линейной регрессии и классификации. Какая из них позволяет отбирать признаки и почему?
14. Постановка задачи обучения логистической регрессии.

15. Постановка задачи SVM (Hard и Soft margin).
16. Алгоритм Word2Vec. Какой функционал и как оптимизируется в варианте Skip-Gram Negative Sampling.
17. Байесовский подход в машинном обучении. Априорное/Апостериорное распределение на параметры. Связь априорного распределения с регуляризацией. Виды точечных оценок (MAP-оценка/байесовская оценка)
18. EM-алгоритм. На какие компоненты раскладывается неполное правдоподобие. Как выглядят шаги EM-алгоритма.
19. LFM модели в рекомендациях на основе коллаборативной фильтрации (SVD, ALS). Формулы обновления параметров в ALS (для эмбедингов пользователя и объекта).
20. Bias-variance-decomposition формула разложения. Интерпретация компонент.
21. Бэггинг и случайный лес.
22. Алгоритм градиентного бустинга. Особенность градиентного бустинга над деревьями.
23. Оценка качества кластеризации.
24. Многослойный персептрон. Основные функции активации.
25. Функции активации на выходном слое в случае задачи классификации (бинарной/многоклассовой) и регрессии. Типичные функции потерь для задачи классификации и регрессии.
26. Chain rule для производной сложной функции. Алгоритм Backpropagation.
27. Конволюционная нейронная сеть. Устройство свёртки (какие есть параметры). Стандартный свёрточный слой (свёртка, нелинейность, пулинг). Идея transfer learning'a.
28. Концептуальное устройство RNN. Почему в базовом варианте затухают/взрываются градиенты?
29. Устройство self-attention. Чем отличается self-attention в transformer-encoder и transformer-decoder частях?