Развернутый перечень вопросов по части «Машинное обучение»

1. Основные понятия машинного обучения: модели, выборки, целевые значения, прогнозы. Задачи классификации, регрессии и кластеризации, обучение с учителем и без учителя. Простые примеры метрик качества в этих задачах (точность, среднеквадратичная ошибка, среднее внутрикластерное и межкластерное расстояние) и алгоритмов (метод k ближайших соседей в задачах классификации и регрессии, наивный байесовский классификатор, метод k средних).
2. Метрики качества в задачах классификации и регрессии (accuracy, precision, recall, F-мера, ROC-AUC, log loss, MSE, MAE, квантильные потери, MAPE, SMAPE), их особенности и отличия друг от друга, как выбирать подходящую функцию потерь. Работа с признаками: извлечение признаков из текстов, изображений и аудио, кодирование категориальных признаков (one-hot-encoding, hashing trick, mean encoding).
3. Линейные методы классификации и регрессии. Функции потерь и регуляризаторы. Метод стохастического градиентного спуска. Разреживание в l1 регуляризаторе – без строгого доказательства, но прокомментировать стандартные «нестрогие» объяснения этого факта, и почему они не убедительны. Оптимизационная задача в логистической регрессии, оценка вероятности принадлежности к классу и связь логистической регрессии с log loss.
4. Линейные методы классификации и регрессии. Функции потерь и регуляризаторы. Метод стохастического градиентного спуска. Оптимизационная задача в методе опорных векторов и ее связь с максимизацией ширины разделяющей полосы. Двойственная задача и ее решение. Опорные векторы.
5. Решающие деревья в задаче классификации и задаче регрессии. Критерии при построении разбиений в решающем дереве (gini, энтропийный критерий, MSE). Ансамбли решающих деревьев: случайный лес и градиентный бустинг над деревьями. Как решается задача регрессии и как решается задача классификации с помощью градиентного бустинга.
6. Решающие деревья в задаче классификации и задаче регрессии. Bias-variance trade-off (без вывода). Анализ бустинга и бэггинга с помощью bias-variance trade-off.
7. Нейронные сети, обучение (backprop), слои для сверточных сетей (dense, conv, pooling, batchnorm), нелинейности (relu vs sigmoid, softmax), функции потерь (logloss, l2, hinge)
8. Рекуррентные нейросети, обучение (backprop tt), отличие от сверточных, разновидности рекуррентных слоев (RNN, LSTM, GRU), примеры использования: аннотация изображений, перевод
9. Задача кластеризации. Примеры некорректности задачи кластеризации. Примеры кластерных структур. Иерархическая кластеризация, формула Ланса-Уилльямса. Алгоритм Ланса-Уилльямса, свойство редуктивности. Статистические методы кластеризации. Кластеризация с помощью EM-алгоритма (без вывода M-шага). Алгоритм k-Means, недостатки алгоритма k-Means.
10. Задача снижения размерности пространства признаков. Идея метода главных компонент (PCA). Связь PCA и сингулярного разложения матрицы признаков (SVD). Идея методов SNE, tSNE, принципиальные отличия от PCA.