



Coding + system design interview (Почта)

Вторая встреча, собеседуют 2 человека (могут быть как программисты, так и ML-инженеры), длительность: примерно 1 час.

Будет несколько вопросов и задач на темы алгоритмов, структур данных и system design.

Алгоритмы и структуры данных

Вам будут даны несколько программистских задач. Код может писаться на доске/бумаге/IDE. Обычно задачи достаточно короткие, чтобы рассказать о решении за пару минут и закодить за 10-20. Здесь мы хотим понять, как вы подходите к проблеме, структуру и качество вашего кода. Вы можете выбрать решить задачу любым способом, далее почти наверняка вас попросят оптимизировать код под ограничения времени, памяти и т.д.

Моменты, на которых стоит заострить внимание:

- Прежде, чем писать код - обсудите свою идею с интервьюерами.
- Избегайте решений с множеством if/else решений. Выбор между рекурсией и итеративностью - важный шаг в решении задачи.
- Подумайте о разных подходах к решению задачи (сортировка, "разделяй и властвуй", динамическое программирование, рекурсия, ...).
- Думайте о выборе подходящих структур данных при решении задач (массив, список, очереди, хеш-таблица, heap, дерево, bloom filter, ...).
- Вас спросят про вычислительную сложность.
- Подумайте, как вы будете тестировать свой код. Проверьте краевые случаи.

Как подготовиться?

- Как правило ваша ежедневная деятельность может отличаться вопросов на этом интервью, поэтому к этому интервью важно освежить свои знания.
- Пройдитесь по основным парадигмам решений задач, структурам данных, графам (BFS, DFS).
- Рекомендуем отличную серию курсов с [Coursera](https://www.coursera.org/), чтобы освежить/подтянуть теоретическую подготовку.



System design

Это интервью нацелено на выявление вашей способности проектировать ML продакшен-системы на объемах данных Mail.ru. Уверенные ответы на вопросы указывают на то, что вы успешно будете решать большинство прикладных задач у нас в Mail.ru. В этой части не будет кодинга, а только лишь рисование на бумаге/доске и обсуждение решения.

Примеры вопросов:

- Спроектируйте систему авто-ответа на письма в Почте
- Спроектируйте поиск по фото документов в Почте

Основная идея - взять фичу в продукте и понять, как ее можно сделать с помощью машинного обучения. Мне требуем от вас запоминания всех возможных алгоритмов, а умения с помощью имеющихся разбить проблему на компоненты и решить задачу.

Важные моменты при проектировании:

- Обязательно уточняйте ограничения системы, т.к. изначальная задача часто достаточно общей постановки.
- Можете ли вы заметить недостатке в решении с ML и предложить улучшения?
- Как оценивать и деплоить данную систему?
- Понимание ограничений по ресурсам (CPU, GPU, memory, disk, network)

Как подготовиться?

- Пройдитесь по ключевым алгоритмам и методикам машинного обучения, основным проблемам, таким как переобучения, регуляризация и т.д.
- Прочитайте блоги/статьи известных компаний, посвященных дизайну систем с [применением ML](#) и [в целом](#).
- Возьмите любую фичу любого известного приложения, которую можно улучшить с помощью машинного обучения. Подумайте над тем, как эту задачу можно решить в лоб простыми инструментами для небольшого кол-ва юзеров/объектов. Далее, как с помощью ML отмасштабироваться на миллионы пользователей?
- Подумайте, как вы будете тренировать модель? Как оценивать? Как постоянно ее дообучать и внедрять?
- Будьте готовы оценить свой подход и алгоритмы, которые вы использовали в процессе решения. Например, почему в данной ситуации вы применили логистическую регрессию, а не деревья решений?