

Индустриальный семинар

Артём Филатов

Общий формат собеседований

- Screening собеседование
- Behavioural собеседование
- Собеседование по коду
- Собеседование по ML/CV
- Дизайн ML/CV систем
- Общий дизайн систем

Screening собеседование

- Цель собеседования: узнать есть ли у кандидата необходимый опыт и знания для рассматриваемой позиции
- Что спрашивают: задача на программирование, вопросы про опыт
- Пример задачи: Вам дан массив целых чисел, напишите код, который проверяет, что существует подотрезок с суммой элементов равным K .

Behavioural собеседование

- Как правило встречается только в западных IT компаниях
- Цель собеседования: узнать, как кандидат справляется с нестандартными ситуациями, как он разрешает конфликты в команде, какая у него мотивация
- Пример вопроса: расскажите про ситуацию, когда у вас был конфликт в команде

Собеседование по коду

- Как правило - это просто задачи на программирование
- Цель собеседования: посмотреть, как кандидат пишет код, как он рассуждает о задаче, может ли он находить ошибки в собственном коде
- Если позиция предполагает знание определенного языка, то могут быть задачи и на знание языка
- Задача на знание языка: Реализуйте умный указатель на C++

Общий дизайн систем

- Цель собеседования: проверить понимание кандидата, как строятся современные распределенные системы
- Пример задачи: предложите дизайн Инстаграмма

Собеседование по ML/CV

- Цель собеседования: проверить знание и понимание популярных алгоритмов, метрик и задач
- Современное компьютерное зрение во многих приложениях - это Deep Learning
- Примеры вопросов:
 - Понимание классических алгоритмов (что такое логистическая регрессия?)
 - Понимание метрик (что такое точность и полнота и в чем их смысл?)
 - Понимание математики (Пусть $y = Ax + b + e$, пусть e - вектор из нормального распределения, выведите распределение над y)
 - Понимание методов оптимизации (Расскажите как устроен SGD)

Собеседование по ML/CV

- Все вопросы будут проверять не то, как вы хорошо что-то запомнили, а именно как вы понимаете, что происходит
- Например, понимаете минусы и плюсы метрики accuracy и готовы предложить альтернативу
- Среди вопросов по Deep Learning'у могут следующие
 - Что такое receptive field? Как его оценить?
 - В чем идея: attention / ResNet / сверток и т.д.
 - Вы видите, что ваша сеть выдает нули, ваши действия?

Собеседование по CV

- Обычно, если кандидат идет работать с картинками или видео, то ничего более специфичного знать не стоит
- Компании ищут кандидатов, которые быстро адаптируются и при необходимости изучат новую область
- Но на некоторые позиции, специальные знания нужны, это относится к работе с калибровками камер, 3D реконструкцией, рендерингом и т.д.
- Вопросы могут быть следующие:
 - Чем отличаются внешние и внутренние параметры калибровки камеры?
 - Что такое дисторсия камеры?
 - Что такое RANSAC?
 - Что такое z-buffer?

Дизайн ML систем

- Некоторая смесь собеседования по дизайну систем и по ML/CV
- Скорее всего попросят сделать дизайн систем в той области, куда кандидат подается
- Например: кандидат подается в команду, которая занимается построением карт в беспилотных автомобилях, в таком случае его могут попросить описать дизайн SLAM'a

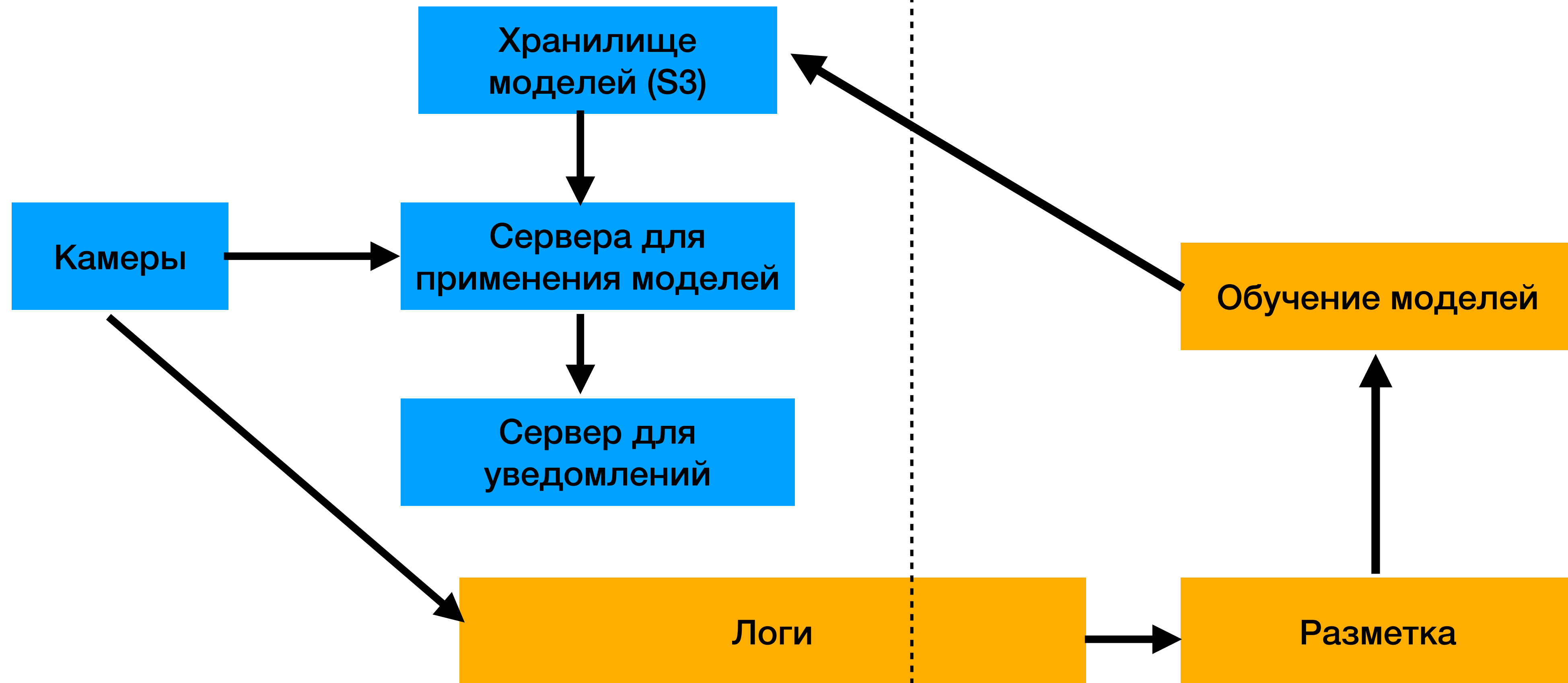
Дизайн ML систем

- Кандидат должен продемонстрировать понимание различных аспектов работы с CV алгоритмами:
 - Применение в продакшене
 - Метрики
 - Обучение
 - Тестирование

Дизайн ML систем

- Вопрос может быть такой: Мы строим систему детекции пожаров по городским камерам в городе N, расскажите как бы вы построили такую систему?
- Кандидат должен рассказать:
 - Как он будет собирать данные (например отправит исторические данные на ручную разметку, чтобы отметить на каких видео есть пожар)
 - Как модель будет работать в продакшене (M серверов анализирующих каждый 10 кадр с камер)
 - Как будет обучаться модель
 - Какая будет метрика (например F-метрика для задачи классификации)

Дизайн ML систем



Общение с командами

- После основных этапов интервью происходит общение с командами
- Цель этого собеседования, понять, что кандидат подходит команде и ему интересны задачи этой команды
- Как правило кандидату предложат порассуждать над актуальной задачей этой команды
- Также кандидату могут рассказать несколько примеров актуальных задач и “продавать” ему эти задачи

Состояние индустрии

Что ожидать на работе?

- В CV как правило основная цель - это разрабатывать/улучшать алгоритмы компьютерного зрения
- Из чего состоит работа:
 - Эксперименты (написание/запуск/анализ)
 - Подготовка данных
 - Выкатка в production
 - Чтение статей
- Комбинация зависит от команды/позиции/направления

Что ожидать на работе?

- Эксперименты часто быстро пишутся, но долго работают
- Также у вас есть ограничения по вычислительным мощностям
- Значит сотрудник должен генерировать гипотезы имеющие наибольшую вероятность оказаться успешными
- Чтобы генерировать хорошие гипотезы, сотрудник должен иметь хорошее понимание и интуицию за теми алгоритмами, которые он разрабатывает

Что ожидать на работе?

- Очень часто возникает ситуация: ваш алгоритм ошибается в определенных местах, что делать?
 - Нужно провести анализ метрик и ошибок, выдвинуть гипотезы: почему алгоритм ошибается?
 - Нужно проверить эти гипотезы
 - Сформулировать модификации алгоритма
 - Поставить эксперименты

Что ожидает работодатель?

- Гибкий ум: сотрудник должен адаптироваться к новой информации
- Неудачи в экспериментах не должны демотивировать - это часть работы
- Сотрудник не должен бояться выходить за область своей ответственности:
 - Например, сотрудник увидел, что код для применения моделей можно ускорить, но он занимается обычно экспериментами
- Знание дополнительных языков (C++) и технологий открывает больше возможностей к тому, чтобы принести пользу

На что обратить внимание при подготовке студентов

- Проектная работа позволит разобраться, как строить CV системы с нуля
- Давайте студентам читать статьи
- Объясняйте интуицию за алгоритмами
- Большие задания на программирование позволят набить студентам руку
- В домашних заданиях студенты должны пробовать и экспериментировать
- Соревнования по машинному обучению здорово прокачивают навыки построения экспериментов

Состояние отрасли

- Работа с нейронными сетями
 - Классификация
 - Сегментация
 - Детекция
 - Генеративные модели

Состояние отрасли

- Работа с 3D реконструкцией
 - Фотограмметрия
 - SLAM
- Работа с камерами

Перспективы отрасли

- Генеративные модели становятся лучше, применение в бизнесе тоже будет расти
- Методы 3D реконструкции получили новый виток развития
- Применение методов детекции и сегментации до сих пор себя не исчерпало