
ОГЛАВЛЕНИЕ

Предисловие	14
Что такое собеседование по проектированию систем МО (ML System Design interview).....	14
Почему это важно	14
Для кого эта книга	15
Чего нет в книге	15
Дополнительные ресурсы	15
Благодарности.....	16
От издательства	18
Глава 1. Введение и общие сведения	19
Прояснение требований.....	20
Формулировка проблемы в виде задачи МО.....	21
Определение цели МО	22
Определение входных и выходных данных системы	22
Выбор подходящей категории МО	24
Темы для обсуждения	25
Подготовка данных.....	25
Инженерия данных	26
Типы данных.....	27
Конструирование признаков.....	29
Операции конструирования признаков	30
Темы для обсуждения	34
Разработка модели	34
Выбор модели.....	34
Обучение модели	36
Темы для обсуждения	40
Оценка.....	41
Автономная (offline) оценка	41
Оперативная (online) оценка	42
Темы для обсуждения	43

6 Оглавление

Развертывание и эксплуатация.....	43
Развертывание в облаке или на устройстве	43
Сжатие модели.....	44
Тестирование при эксплуатации	44
Пайплайн предсказаний	46
Темы для обсуждения	48
Мониторинг	48
Почему в действующей системе происходят сбои	49
Какие показатели нужно отслеживать.....	50
Инфраструктура.....	50
Итоги.....	51
Ссылки	51
Глава 2. Система визуального поиска	54
Прояснение требований.....	54
Формулировка проблемы в виде задачи МО.....	56
Определение цели МО	56
Определение входных и выходных данных системы.....	56
Выбор подходящей категории МО	56
Подготовка данных.....	58
Инженерия данных	58
Конструирование признаков.....	60
Разработка модели	60
Выбор модели.....	60
Обучение модели	61
Построение датасета.....	62
Выбор функции потерь	64
Оценка.....	66
Автономные метрики	66
Оперативные метрики	72
Эксплуатация.....	72
Предсказательный пайплайн	73
Пайплайн индексации.....	74
Эффективность алгоритмов поиска ближайшего соседа (NN).....	75
Какой алгоритм использовать?.....	77
Другие темы для обсуждения.....	78
Итоги.....	79
Ссылки	80

Глава 3. Система размытия в Google Street View	81
Прояснение требований.....	81
Формулировка проблемы в виде задачи МО.....	82
Определение цели МО	82
Определение входных и выходных данных системы.....	83
Выбор подходящей категории МО	83
Двухступенчатые сети.....	83
Одноступенчатые сети	84
Сравнение одноступенчатых и двухступенчатых сетей	84
Подготовка данных.....	85
Инженерия данных.....	85
Конструирование признаков.....	86
Разработка модели.....	88
Выбор модели.....	88
Обучение модели	90
Оценка.....	91
Автономные метрики	93
Оперативные метрики	95
Эксплуатация.....	95
Перекрытие ограничительных прямоугольников.....	95
Проектирование системы МО.....	96
Пайплайн пакетных предсказаний	96
Другие темы для обсуждения.....	98
Итоги	99
Ссылки	100
Глава 4. Поиск видео на YouTube	101
Прояснение требований.....	101
Формулировка проблемы в виде задачи МО.....	102
Определение цели МО	102
Определение входных и выходных данных системы.....	102
Выбор категории МО.....	103
Подготовка данных.....	105
Инженерия данных	105
Конструирование признаков.....	106
Разработка модели.....	110
Выбор модели.....	110
Обучение модели	116

8 Оглавление

Оценка.....	117
Автономные метрики.....	117
Оперативные метрики.....	118
Эксплуатация.....	119
Предсказательный пайплайн.....	120
Пайплайн индексации видео.....	121
Пайплайн индексации текста.....	121
Другие темы для обсуждения.....	121
Итоги.....	123
Ссылки.....	124
Глава 5. Обнаружение вредоносного контента.....	125
Прояснение требований.....	126
Формулировка проблемы в виде задачи МО.....	127
Определение цели МО.....	127
Определение входных и выходных данных системы.....	127
Выбор категории МО.....	131
Подготовка данных.....	137
Инженерия данных.....	137
Конструирование признаков.....	138
Разработка модели.....	143
Выбор модели.....	143
Обучение модели.....	144
Оценка.....	146
Автономные метрики.....	146
Оперативные метрики.....	147
Эксплуатация.....	148
Сервис обнаружения вредоносного контента.....	148
Сервис обработки нарушений.....	148
Сервис понижения приоритета.....	148
Другие темы для обсуждения.....	149
Итоги.....	150
Ссылки.....	151
Глава 6. Система рекомендации видео.....	153
Прояснение требований.....	154
Формулировка проблемы в виде задачи МО.....	155
Определение цели МО.....	155
Определение входных и выходных данных системы.....	156
Выбор категории МО.....	156

Подготовка данных	161
Инженерия данных	161
Конструирование признаков	163
Разработка модели	168
Матричная факторизация	168
Двухбашенная нейронная сеть	175
Работа двухбашенной нейронной сети	177
Матричная факторизация или двухбашенная нейронная сеть?	179
Оценка	179
Автономные метрики	179
Оперативные метрики	180
Эксплуатация	181
Генерация кандидатов	182
Скоринг	184
Повторное ранжирование	184
Трудности при разработке систем рекомендации видео	185
Другие темы для обсуждения	186
Итоги	187
Ссылки	188
Глава 7. Система рекомендации событий	189
Прояснение требований	190
Формулировка проблемы в виде задачи MO	191
Определение цели MO	191
Определение входных и выходных данных системы	191
Выбор подходящей категории MO	192
Списочные методы LTR	193
Подготовка данных	194
Инженерия данных	194
Конструирование признаков	196
Разработка модели	204
Выбор модели	204
Обучение модели	212
Оценка	214
Автономные метрики	214
Оперативные метрики	214
Эксплуатация	215
Пайплайн оперативного обучения	215
Предсказательный пайплайн	216

Другие темы для обсуждения.....	218
Итоги.....	219
Ссылки.....	220
Глава 8. Предсказание кликов по рекламе на социальных платформах	221
Введение.....	221
Прояснение требований.....	222
Формулировка проблемы в виде задачи МО.....	223
Определение цели МО.....	223
Определение входных и выходных данных системы.....	223
Выбор категории МО.....	224
Подготовка данных.....	225
Инженерия данных.....	225
Конструирование признаков.....	226
Признаки пользователя.....	228
Разработка модели.....	230
Выбор модели.....	230
Обучение модели.....	241
Оценка.....	242
Автономные метрики.....	242
Оперативные метрики.....	244
Эксплуатация.....	245
Пайплайн подготовки данных.....	246
Пайплайн непрерывного обучения.....	247
Предсказательный пайплайн.....	247
Другие темы для обсуждения.....	248
Итоги.....	249
Ссылки.....	250
Глава 9. Похожие объекты на платформах краткосрочной аренды жилья	251
Прояснение требований.....	252
Формулировка проблемы в виде задачи МО.....	253
Определение цели МО.....	253
Определение входных и выходных данных системы.....	253
Выбор категории МО.....	254
Подготовка данных.....	255
Инженерия данных.....	255
Конструирование признаков.....	256

Разработка модели	257
Выбор модели	257
Обучение модели	257
Построение датасета	258
Оценка	262
Автономные метрики	262
Оперативные метрики	264
Эксплуатация	264
Обучающий пайплайн	264
Индексирующий пайплайн	265
Предсказательный пайплайн	266
Другие темы для обсуждения	267
Итоги	268
Ссылки	269
Глава 10. Персонализированная лента новостей	270
Введение	270
Прояснение требований	271
Формулировка проблемы в виде задачи MO	272
Определение цели MO	272
Определение входных и выходных данных системы	274
Выбор категории MO	274
Подготовка данных	275
Инженерия данных	275
Конструирование признаков	277
Признаки пользователей	280
Разработка модели	282
Выбор модели	282
Обучение модели	285
Оценка	287
Автономные метрики	287
Оперативные метрики	287
Эксплуатация	289
Предсказательный пайплайн	289
Другие темы для обсуждения	290
Итоги	291
Ссылки	292

Глава 11. Списки возможных знакомых	293
Введение	293
Прояснение требований.....	293
Формулировка проблемы в виде задачи MO.....	294
Определение цели MO	294
Выбор категории MO.....	295
Подготовка данных.....	299
Инженерия данных.....	299
Конструирование признаков.....	301
Разработка модели.....	303
Выбор модели.....	303
Графовые нейронные сети (GNN)	303
Обучение модели	304
Оценка.....	308
Автономные метрики	308
Оперативные метрики	308
Эксплуатация.....	309
Эффективность.....	309
Проектирование системы MO.....	311
Другие темы для обсуждения.....	314
Итоги	315
Ссылки	316
Послесловие.....	317