

ML Roadmap

Шмидт Ян

24 мая 2025 г.

Содержание

| | | |
|----------|---|----------|
| 1 | Python | 2 |
| 2 | Math | 2 |
| 3 | ML | 3 |
| 3.1 | Classic ML | 3 |
| 3.2 | Deep Learning (DL) | 4 |
| 3.3 | MLOps | 5 |
| 4 | CV Roadmap | 6 |
| 4.1 | Classic CV | 6 |
| 4.1.1 | Базовые операции с изображениями | 6 |
| 4.1.2 | Фильтрация и улучшение изображений | 6 |
| 4.1.3 | Детекция границ, углов, текстурных признаков | 6 |
| 4.1.4 | Дескрипторы, ключевые точки, оптический поток | 6 |
| 4.2 | DL | 7 |
| 4.2.1 | Свертки | 7 |
| 4.2.2 | Механизмы в CNN | 7 |
| 4.2.3 | Основные типы моделей | 7 |
| 4.2.4 | Интерпретация признаков | 7 |
| 4.2.5 | Предобработка и аугментация изображений | 8 |
| 4.2.6 | Популярные архитектуры | 8 |
| 4.2.7 | Задачи | 8 |
| 4.2.8 | Метрики | 8 |
| 4.2.9 | Лоссы | 9 |

Дисклеймер. **Розовым цветом** буду обозначать дополнительные темы, которые не обязательно изучать, но потенциально могут встретиться где-нибудь; **оранжевым цветом** обозначаю продвинутые темы, которые точно не стоит брать в начале. **Приведенный ниже план – не монолит, всегда можно либо пропустить, либо добавить материал**

1 Python

Пройдемся по основам питона, чтобы владеть синтаксисом, знать сильные и слабые стороны языка, уметь быстро накидать MVP-решение той или иной задачи, создать скрипт, который будет удобно запускаться из терминала, и так далее.

Примерный план:

- Типы данных `int`, `str`, `float`, `bool`, `set`, `list`, `tuple`, `dict`; Изменяемые и неизменяемые типы данных
- Операторы `if`, `elif`, `else`
- Циклы `while`, `for`, вложенные циклы
- Функции. `lambda`-функции и их отличие от обычных функций
- Локальные и глобальные переменные
- Итераторы и генераторы. Отличия
- Контекстный менеджер `with`. Запись и чтение файлов
- ООП, dunder-методы; наследование, инкапсуляция, полиморфизм
- Исключения, `try`, `except`, etc.
- `threading`, `multiprocessing`
- `numpy`, `pandas`, `argparse`, `json`, etc. Написание кастомных пакетов

2 Math

По необходимости проясним моменты из

- Матстата¹
- Теорвера
- Линейной алгебры
- Численные методы, теория оптимизации и тд.

¹Из матстата наиболее вероятно могут встретиться понятия вроде таких: статистические гипотезы, ошибки 1, 2 рода, доверительные интервалы, статистические тесты, уровень значимости α

3 ML

Основы по классическому и глубокому обучению. С дальнейшей специализацией в интересующую область (CV, NLP, etc.)

3.1 Classic ML

- Задача обучения с учителем
- Типы задач в ML и их постановка
- Функции потерь, метрики бинарной классификации: Accuracy, Precision, Recall, F_1 , ROC AUC, PR-curve, confusion matrix
- Метод Наименьших Квадратов (МНК)
- Линейная классификация и регрессия. Логистическая регрессия
- KNN
- SVM – метод опорных векторов. Разные ядра SVM
- Кросс-валидация: обучение на разных фолдах
- Сокращение размерности: Lasso (L_1). Геометрический смысл отбора признаков. PCA
- Решающие деревья. Критерии ветвления
- bias-variance decomposition
- Случайный лес. Метод случайных подпространств
- Линейные ансамбли, стохастическое ансамблирование, bagging, boosting.
- Градиентный бустинг. CatBoost, LightGBM
- Stacking, blending, MoE (Mixture of Experts)
- Байесовская теория классификации. Наивный байес. Восстановление плотности распределения: 1. парzenовские окна, 2. метод максимума правдоподобия
- Геометрический смысл предположения нормальности признаков
- Задача кластеризации. Kmeans, DBSCAN; иерархическая и спектральная кластеризация
- Обобщенные линейные модели (GLM)

3.2 Deep Learning (DL)

- Строеение многослойного перцептрона. Модель МакКаллока-Питтса
- Построение простейшей нейросети. Веса. Скрытые слои
- Переобучение и недообучение и методы борьбы с ними: регуляризация
- Численная минимизация лосс-функции: метод стохастического градиента (SGD) и др.
- Адаптивные градиенты
- Предобучение, дообучение. Fine-tuning
- Разделение на bias/variance и общее понимание trade-off между ними
- Batch Norm, dropout. Layer/Group/Instance Norm
- Сингулярное разложение матриц (SVD)²
- Мультиколлинеарность и переобучение. Гребневая (Ridge (L_2)) регрессия
- Гиперпараметры, какие они бывают, их подбор
- Глубокое обучение. Сверточные нейросети: операция свертки, max-pooling.
- Затухающие градиенты. Проблемы глубоких сетей. Остаточные связи (ResNet), skip-connections.
- Аугментирование данных.
- Рекуррентные нейросетки (LSTM, GRU). Карты Кохонена
- Автокодировщики (AE): Линейный AE, sparse AE, variational AE, etc.
- GAN
- Векторные представления данных (картинок, текста, etc.) – эмбединги. CBOW. word2vec.
- Тематическое моделирование, LDA. TF-IDF, BM25.
- Трансформеры. Модели внимания, мотивация их появления. Multihead self-attention. Архитектура трансформера. Позиционное кодирование.

²Это можно к матеше отнести. Просто это важная штука в МЛ, много где встречается. Здесь заодно можно будет поговорить про собственные числа и векторы матриц

- Дистилляция моделей. Квантование. Прунинг.
- Методы отбора признаков: LASSO, PCA, UMAP, Shapley Values, TopoAE
- Многомерное шкалирование данных: SNE, t-SNE.
- PEFT: LoRA, qLoRA
- Self-supervised learning (SimCLR, BYOL)
- Zero-shot / few-shot в (особенно в контексте VLM)

3.3 MLOps

- CI/CD, Docker
- Apache Kafka
- Логгирование экспериментов: TensorBoard, WandB
- Оптимизация моделей: TensorRT, ONNX, Triton
- ClearML, MLflow

4 CV Roadmap

CV имеет несколько ответвлений: классический CV и DL-ориентированный. Из classic CV полезно знать небольшой джентельменский набор, не нужно сильно упарываться.

Что касается DL, то он намного разнообразнее и крайне динамично развивается в настоящее время. Из некоторых ответвлений DL CV можно упомянуть задачи *сегментации, классификации, детекции, трекинга, генерации*; все популярнее (на момент 2025) становится скрещение LLM и CV в лице VLM.

Вкратце приведу примерный RoadMap по CV, актуальный на 2025.

4.1 Classic CV

4.1.1 Базовые операции с изображениями

- Число каналов, цветовое пространство RGB
- Число бит для кодирования изображения
- *Аффинные преобразования: растяжение, повороты, масштабирование изображений через матрицы*
- *Проективные преобразования*
- *Преобразование Фурье (в т.ч. косинусное преобразование)*

4.1.2 Фильтрация и улучшение изображений

- Бинарная и *адаптивная* пороговая обработка
- Гауссовы и медианные фильтры
- *Алгоритм Макдона-Оушена для улучшения контраста*
- *Дизеринг по Флойду-Стейнбергу*

4.1.3 Детекция границ, углов, текстурных признаков

- Фильтры Габора, Собеля, ... любая другая матрица
- *Детектор углов Харриса*

4.1.4 Дескрипторы, ключевые точки, оптический поток

- SIFT, SURF, *ORB*
- Matching, *stiching* изображений
- *Алгоритм Лукаса-Канаде*

4.2 DL

4.2.1 Свертки

- Классическая свертка (1D, 2D, 3D, ... nD)
- Spatial свертка
- Depthwise separable
- Flattened
- Dilated
- Transposed
- Grouped / Shuffled grouped
- Factorized

4.2.2 Механизмы в CNN

- Pooling, Stride, Padding, Способы падинга (интерполяция, добивание нулями)
- Skip connections
- Attention
- CBAM, SE-block

4.2.3 Основные типы моделей

- MLP
- Сверточные сети (CNN)
- Остаточные сети (Residual, ResNet)
- Трансформеры
- GAN
- Autoencoders (в т.ч. VAE)
- Диффузионные модели

4.2.4 Интерпретация признаков

- Карты признаков
- GradCAM и аналоги

4.2.5 Предобработка и аугментация изображений

- `albumentations`
- `torchvision`

4.2.6 Популярные архитектуры

- ResNet, ConvNext, RegNet
- MobileNet, EfficientNet
- DeepLabV3, Unet, Unet++, UNETR
- ViT, CLIP, LeViT, SAM, DAViT
- Segformer, Swin Transformer
- DETR, RCNN/FastRCNN/FasterRCNN, YOLOv7–11, DINO
- Stable Diffusion

4.2.7 Задачи

- Классификация
- Детекция
- Сегментация (семантическая, инстанс, паноптическая)
- Трекинг
- Генерация

4.2.8 Метрики

- Классификация
 - Accuracy, Precision, Recall, F_1
 - ROC AUC, FPR/TPR/TNR/FNR
- Сегментация
 - Jaccard
 - Dice
 - Tversky
- Генерация изображений
 - FID
 - IS
 - KID, SSIM, PSNR, LPIPS,

- Детекция
 - mAP @ threshold
 - IoU
 - Precision, Recall, F_1
 - MR (Miss Rate, для пешеходов/объектов), LAMR, AR (Average Recall, для оценки локализации)

4.2.9 Лоссы

- Пространственные
 - Dice, IoU,
 - Tversky, Generalised Wasserstein Dice
- Distribution-based
 - KL-дивергенция (как пререквизит)
 - Cross-Entropy (CE), Focal
- Граничные лоссы
 - Hausdorf, Boundary loss
- Составные лоссы (Dice + CE, Tversky + Focal и т.п.)