

Вопросы к экзамену

1. Постановка задачи машинного обучения с учителем. Схема работы наивного байесовского классификатора.
2. Линейная регрессия, логистическая регрессия: когда есть единственный глобальный минимум функции потерь? Определение и свойства функции softmax.
3. Граф вычислений полносвязной двухслойной нейросети для классификации. Вычисление градиента в нем. Простой стохастический градиентный спуск (vanilla SGD). Сверточный слой: forward pass. Pooling, stride, варианты padding.
4. Vanishing/exploding градиенты в простой рекуррентной сети, как LSTM (частично) решает проблему.
5. Методы оптимизации SGD+momentum, RMSprop. Dropout и батч-нормализация.
6. GAN, дискриминатор/генератор, функции потерь дискриминатора/генератора. Алгоритм обучения, ошибки при обучении (коллапс моды, затухающие градиенты), оценка модели, метрика FID, приемы стабилизации обучения. Какие бывают GAN'ы.
7. Извлечение признаков CNN: ResNet (схема работы), ResNeXt (grouped convolutions).
8. Извлечение признаков CNN: DenseNet, MobileNetV1 (depthwise separable convolutions).
9. Постановки задач object detection, semantic segmentation, instance segmentation. Оценка качества. Модель Fast R-CNN, RoI pooling.
10. Модели Faster R-CNN и Mask R-CNN, RoI Align.
11. Модели YOLOv1 и SSD. Post-processing (non-maximum suppression).
12. Какие недостатки shallow graph embedding преодолевают графовые нейронные сети. Модели GCN и GAT.
13. Векторные представления слов Word2Vec Skip-gram и FastText. Модель DeepWalk для shallow graph embedding.
14. Языковая модель ELMo. Архитектура Transformer: энкодер и декодер.
15. Языковые модели BERT, GPT, GPT-2.
16. Дистилляция знаний. Подход из статьи Хинтона (softmax with temperature). Промежуточные учителя (teacher assistants). Алгоритм FitNet.
17. Абстрактное синтаксическое дерево программы. Модель Code2seq для генерации описания программного кода на естественном языке (три подзадачи из статьи).