

Questions

Natural Language Processing

1. Text feature extraction: Bag of Words, Bag of Ngrams, Tf-Idf. Normalization techniques.
2. Word embeddings. Word2Vec. CBOW, Skip-Gram.
3. Word embeddings: Matrix factorization, PPMI. GloVe algorithm.
4. Metric learning, DSSM. Semantic search.
5. Sequence-to-sequence models. Seq2seq training. Decoding strategies (beam search, greedy, sampling).
6. Attention in seq2seq.
7. Machine translation task. MT metrics, training pipeline.
8. Transformer architecture. Self-attention. Transformer block structure. Positional encoding.
9. Transfer learning idea. Contextual pretrained representations (CoVe, ELMo).
10. Transfer learning with pretrained models. GPT-1, BERT.
11. Generative pretraining. T5, GPT-2/3/4. Zero-shot and few-shot inference. Prompting strategies (chain-of-thought, self-consistency).
12. LLM Alignment. Instruction tuning, RLHF.
13. LLM Alignment. Contrastive learning: SLiC, DPO.
14. PEFT: Sparse methods. Pruning, sparse fine-tuning.
15. PEFT: LoRA, prompt-tuning, multilayer prompt-tuning.
16. PEFT: function composition (adapters, routing).

Reinforcement learning

1. RL problem statement. MDP formalism. Crossentropy method.
2. Model-based RL. Bellman equations. Policy iteration with dynamic programming.
3. Value-based RL. Model-free prediction (Monte-Carlo vs. Temporal Difference).
4. Value-based RL. Model-free control. Q-Learning, SARSA, EV-SARSA.
5. Approximate value-based methods. DQN.
6. Policy-based RL. REINFORCE (with loss derivation).
7. Actor-critic policy gradient. Baselines, Advantage, A2C.
8. Exploration strategies. Eps-greedy, UCB, Thompson sampling.
9. Reinforcement learning for seq2seq. Self-critical Sequence Training.

Theoretical minimum

1. Embedding layer structure. How does backprop through embedding layer work?
2. Word2Vec main idea.
3. Triplet loss, contrastive learning.
4. Metric learning, DSSM main idea.
5. Seq2seq architecture.
6. Autoregressive decoding strategies: greedy, sampling, beam search.
7. Attention mechanism.
8. Transfer learning main idea. BERT pretraining objectives.
9. Generative pretraining, GPT.
10. Why do we need to use human preferences instead of supervised FT in LLM alignment?
11. RL problem statement (environment, agent, state, action). Reward, discounted reward.
12. MDP model and its assumptions.
13. Model-free and model-based RL difference.
14. On-policy and off-policy RL difference.
15. Value function and q-function. Relationship between them.
16. Exploration / exploitation tradeoff. Describe any exploration tactic.
17. Policy-based RL idea. REINFORCE idea (without derivation).
18. How to frame seq2seq decoding as an RL problem?