

Topics on exam from background lectures

Language models

- ↳ predict next word

- ↳ n-gram model

- ↳ limitations

 - ↳ memory

 - ↳ sparsity

- ↳ neural LM

 - ↳ embeddings

 - ↳ composition fns

 - ↳ RNN

 - ↳ simple element-wise fns

 - ↳ self-attn

 - ↳ multiple heads

 - ↳ parallelizing

- ↳ training

 - ↳ loss fn

 - ↳ gradient descent

- ↳ evaluation
 - ↳ perplexity

↳ Transformers

- ↳ depth (how do you stack multiple Transformer blocks)

↳ positional encodings

- ↳ learned additive emb.
- ↳ length generalization
 - ↳ RoPE
 - ↳ No PE

↳ tokenization

- ↳ word
- ↳ subword
 - ↳ BPE
- ↳ bytes

↳ LLMs

↳ scaling laws

↳ Chinchilla vs. Kaplan et al.

↳ "compute-optimal"

↳ "emergent properties"

↳ in-context learning

↳ zero-shot / few-shot

↳ chain-of-thought

↳ RAG

↳ Training pipeline

↳ pretraining

↳ learns basic linguistic faculties

↳ memorizing facts

↳ post-training

↳ instruction tuning (SFT)

↳ RLHF

Papers

- ↳ understand main ideas
- ↳ discussion Qs