

Recurrent neural networks – once again Consider a very simple network (one layer only) x Consider a very simple network (one layer only) tets remember we DO have biases and sums every time we pass a vector into an activation function... And lets rewrite a bit weirdly: y t-1 y t y t+1 Model state of time t x t x t+1 Time flow







1



























So what is attention?

ZD.

A mechanism in which the model can look at tokens in a sequence inferring which of them **update** others (refer to others) and in what way.

If this process is layered and iterated over, it allows for building inference

(token A is updated by token B changing its meaning, but it then updates token C, so actually token B also changes meaning of C, etc.)









What ChatGPT is aiming to do? "Under the hood": prediction of the next word ChatGPT has no memory, intent, logical reasoning... "On the surface": providing expected output 1 thas seen so many inferences, that it can pretend to do it well 1 thas seen so many conversations that it tries to mimic conversation style – adjusting to your inputs 1 thas seen so many facts that it has an "internal world model" 5 Fine-tuning allows it to closely follow commands

16





















E.g.:	My interaction with Monday
 ✓ 1. ✓ 2. X 3. X 4. 	Full introduction of context Review of my paper, discussion about merit, etc. Lets now add something based on Monday knowledge Can it reflect on its mistake?
	And this is a common pattern! LLMs are usually very good mirrors – but they reflect only what we put into them.
	They feel competent (and they truly are, up to a point). And its dangerous, because it lulls us into a sense of security

Interacting with ChatGPT

ZDu

(You want to tailor the model for you)

- You want to "warm up" (prime the latent space)
- You want to personalize the interaction (Who you both are) and explicitly say what are the goals and context
- You want to organize response preparation (highlight > text) So it is YOU who is responsible for the text
- You want to adjust (modify) the response
- You want to check merit

28

Working with ChatGPT creatively...

ZD.

+ It behaves like a mirror

- (You can watch your ideas grow and multiply) + Its great at fast reiterating, proposing different viewpoints, exploring options
- + Its great at formulation of ideas
- ("I feel like A somehow connects to B. Why do I think like that?")
- + It can wrap ideas in tunable stories and descriptions and propose examples + Rapid style transfer and cross-domain work
- ("change my book section into a movie script")

- Answers tend to drift – unless tightly guided - High fluency -> false confidence (Answers feel wise but are often shallow and derivative) - You need to explicitly ask for criticism (some models offer critique by default (Monday), others can be asked to roleplay a critic) - It behaves like a mirror

(You will have your biases and mistakes multiplied as well)



ZD. Specialized models Lots of "general purpose models" with skew towards particular functionalities: OpenAl GPT series (3.5, 4.0, 4.5, o3, o4 Monday, etc.) Grok, DeepSeek, Gemini, LlaMA, Amazon Titan, and so on... Its hard to pick "best candidates" - some will be a bit more creative, some faster, some will keep context a bit better, some will code a bit better, etc ... Its better to first learn to use one model well, than to jump from model to model Lots of "content generation models" aimed at different contents: Sora, Synthesia (Video), Midjourney, Dalle2, (Images), ElevenLabs (voice), Suno.ai, (music), Synthea (datasets), AlphaFold (proteins), Point-E (point clouds), Lots of "specialized text models": Perplexity, ScholarGPT, Consensus, (scientific text processing), Robin, Harvey (legal texts, contracts), BloombergGPT (finance, reports), GitHub Copilot,

31

Things to remember:

ZDu

- 1. Show sequential 1-layer RNN scheme
- 2. Explain pros and cons of LSTM nets (scheme not required!)
- 3. Explain what a transformer is and what are main types of transformers 4. Explain what are the system prompts for ChatGPT and what they do
- 5. Explain what an attention mechanism is
- 6. Explain steps to properly configure and prompt ChatGPT
- Explain what are the benefits and risks of using ChatGPT creatively
 Explain what a hallucination is