

Persentation:


One project Sentiment analysis: BY IMDB dataset

1. Text preprocessing
2. Tokenizer
3. LSTM architecture
4. Activation SoftMax
5. Adam optimizer
6. testing

SENTIMENT ANALYSIS

Submit

Negative



muril_bilstm_DL

	precision	recall	f1-score	support
-1	0.87	0.89	0.88	750
0	0.78	0.82	0.80	750
1	0.87	0.80	0.83	750
accuracy			0.84	2250
macro avg	0.84	0.84	0.84	2250
weighted avg	0.84	0.84	0.84	2250

LSTM

	precision	recall	f1-score	support
0	0.98	0.83	0.90	4454
1	0.26	0.80	0.39	341
accuracy			0.83	4795
macro avg	0.62	0.81	0.65	4795
weighted avg	0.93	0.83	0.86	4795

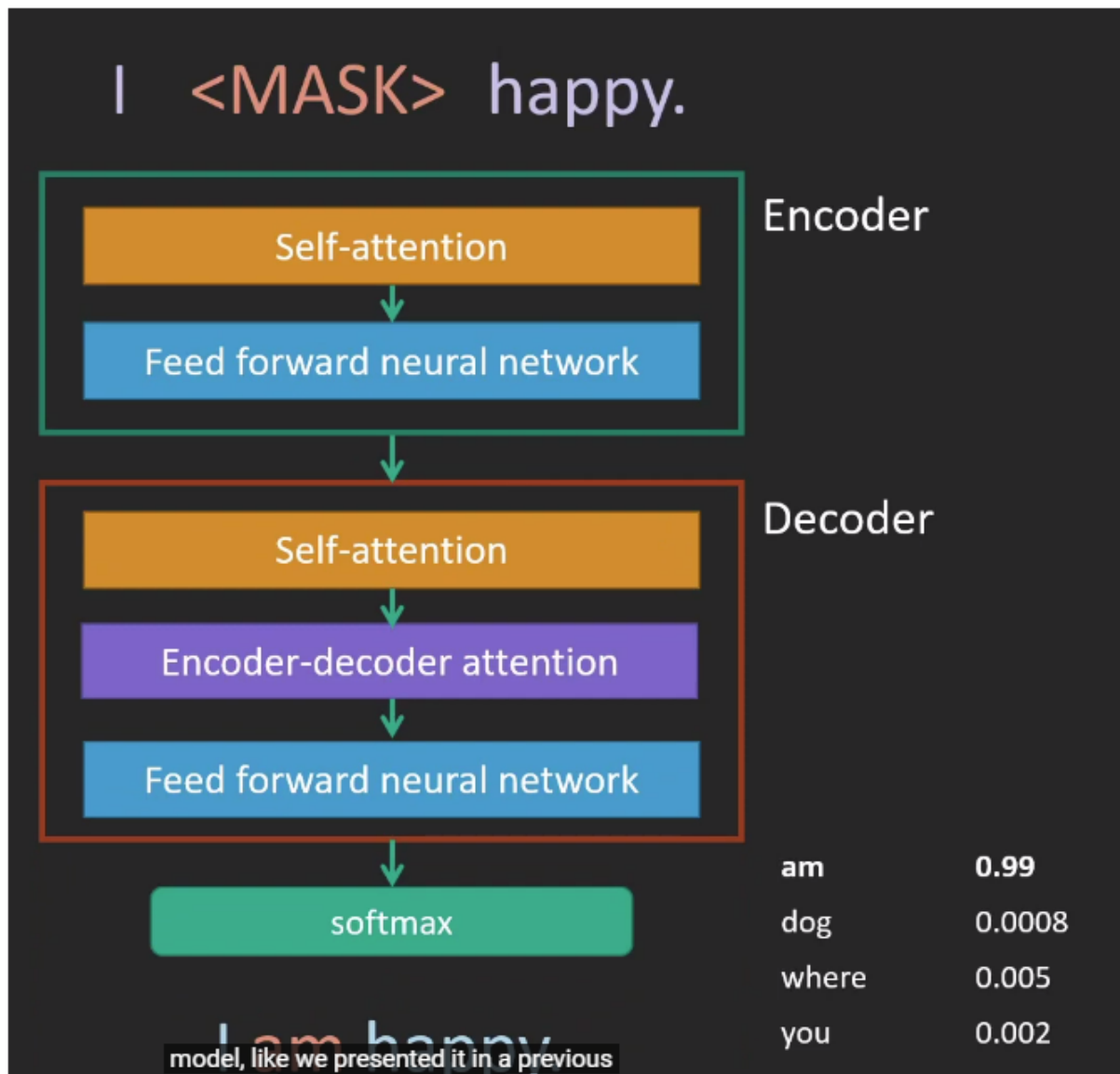
Now I Want to explain some latest architecture

Claude 3: consists of a family of three models (Claude 3 Haiku, Claude 3 Sonnet, and Claude 3 Opus); Claude 3 Opus (the strongest model) seems to outperform GPT-4 on common benchmarks like MMLU and HumanEval; Claude 3 capabilities include analysis, forecasting, content creation, code generation, and converting in non-English languages like Spanish, Japanese, and French; 200K context windows supported but can be extended to 1M token to select customers; the models also have strong vision capabilities for processing formats like photos, charts, and graphs; Anthropic claims these models have a more nuanced understanding of requests and make fewer refusals.

Mistral Large : a new LLM with strong multilingual, reasoning, maths, and code generation capabilities; features include: 1) 32K tokens context window, 2) native multilingual capacities, 3) strong abilities in reasoning, knowledge, maths, and coding benchmarks, and 4) function calling and JSON format natively supported.

StarCoder 2: a family of open LLMs for code with three different sizes (3B, 7B, and 15B); the 15B model was trained on 14 trillion tokens and 600+ programming languages with a context window of 16K token and employing a fill-in-the-middle objective; it matches 33B+ models on many evaluation like code completion, code reasoning, and math reasoning aided through PAL.

BERT:



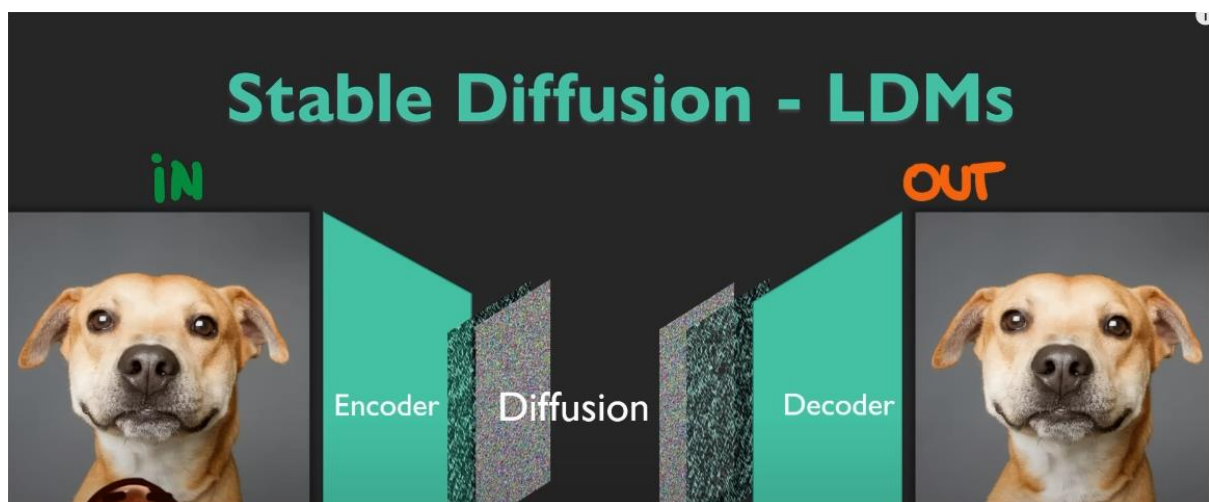
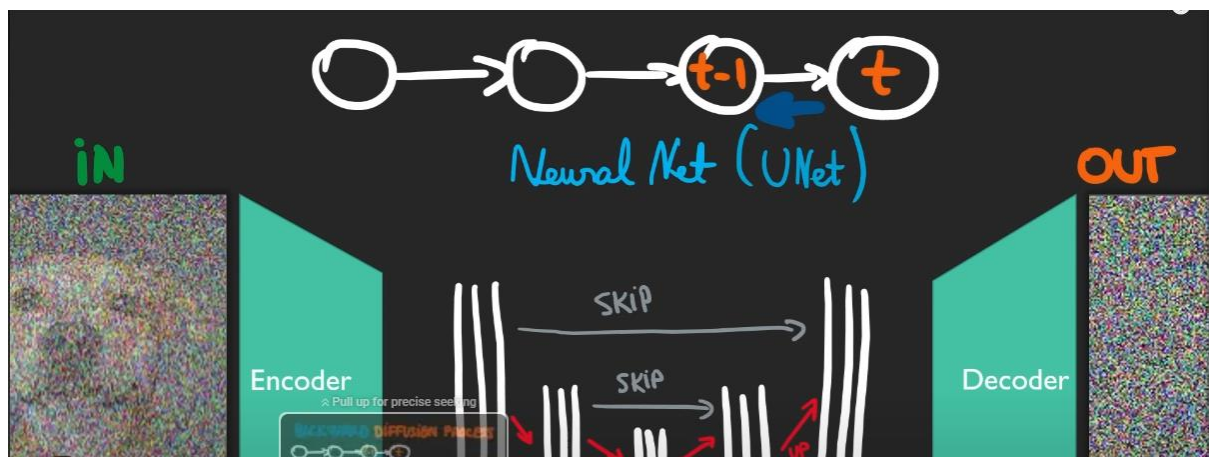
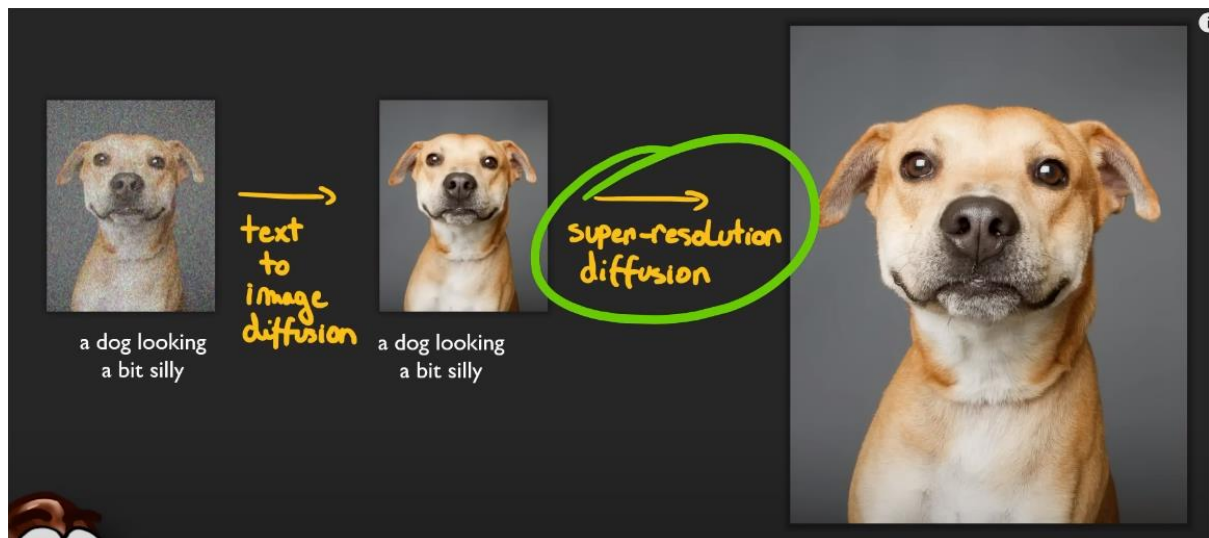
BERTology meets Biology | Solving biological problems with Transformers

PlanGPT: shows how to leverage LLMs and combine multiple approaches like retrieval augmentation, fine-tuning, tool usage, and more; the proposed framework is applied to urban and spatial planning but there are a lot of insights and practical tips that apply to other domains.

	CHINCHILLA	SPARROW
plausibility	61%	78%
rule breaking	20%	8%
Winogender	↑	↓

	Chat GPT	SPARROW
deliver evidence	X	✓
fine-tune on human feedback	✓	✓
follow rules	X	✓
paper	X	✓
model demo	✓	X

Stable Diffusion 3: a suite of image generation models ranging from 800M to 8B parameters; combines diffusion transformer architecture and flow matching for improved performance in multi-subject prompts, image quality, and spelling abilities; technical report to be published soon and linked here.



Thank you for hearing