

dim\_videos (1)

dimensions for marts & gold layer

Results - dim\_videos

sk_video_id	video_id	title	description
7914e056cfc5654cee2a859d2dd074b1	rFwQDDbYTm4	[Classic] Playing Atari with Deep Reinforcement Learning (Paper Explained)	▸ #ai #dqn #deepmind After the initial success of deep neural networks, the next step was to develop algorithms that can learn to play Atari games. This video explains the Deep Q-Network (DQN) architecture, which was the first to achieve human-level performance on a variety of Atari games.
9eb13106ada3c80e714ff6381c3fcd54	Nq3auVtvd9Q	[Classic] ImageNet Classification with Deep Convolutional Neural Networks (Paper Explained)	▸ #ai #research #alexnet AlexNet was the start of the deep learning revolution. This video explains the architecture of AlexNet, which was the first to achieve human-level performance on the ImageNet classification task.
0b807d5dd2698333e13648cf08f9e098	w3knicSHx5s	Dynamic Inference with Neural Interpreters (w/ author interview)	▸ #deeplearning #neuralinterpreter #ai This video includes an interview with the author of the paper, who discusses the challenges of dynamic inference and the role of neural interpreters.
74795196bd9b1e41afa69f3482e466ce	_PyusGsbBPY	Stochastic RNNs without Teacher-Forcing	▸ We present a stochastic non-autoregressive RNN that does not require teacher forcing during training. This allows for more efficient training and inference, and can be applied to a variety of tasks.
49a67b81ac24ab6cac973a9a72729712	iDulhoQ2pro	Attention Is All You Need	▸ https://arxiv.org/abs/1706.03762 Abstract: The dominant sequence-to-sequence models today are based on the encoder-decoder framework. In this paper, we propose a simple architecture called the Transformer, which is based on the self-attention mechanism.
68d15710bd1c593c72b1cedfad4d69ee	H5vpBCLo74U	XLNet: Generalized Autoregressive Pretraining for Language Understanding	▸ Abstract: With the capability of modeling bidirectional contexts, deep bidirectional models have achieved state-of-the-art performance on a variety of natural language understanding tasks. In this paper, we propose a new pretraining task called XLNet, which is a generalization of the autoregressive pretraining task.
13a6fe7ee9f413531239c80002639a3d	-MCYbmU9kfg	RoBERTa: A Robustly Optimized BERT Pretraining Approach	▸ This paper shows that the original BERT model, if trained correctly, can achieve state-of-the-art performance on a variety of natural language understanding tasks. We propose a new pretraining task called RoBERTa, which is a robustly optimized BERT pretraining approach.
20a9e366bf36db3255456a5e648b419d	69ljnZaoeao	LeDeepChef Deep Reinforcement Learning Agent for Families of Text-Based Games	▸ The AI cook is here! This agent learns to play a text-based game called LeDeepChef, which is a family of text-based games. The agent is trained using deep reinforcement learning and can play the game at a level comparable to human players.
462cff86a95282f4644b8f2cfa55edd9	i4H0kxrias	Reformer: The Efficient Transformer	▸ The Transformer for the masses! Reformer solves the biggest problem of the Transformer: it is too slow. Reformer is a new Transformer architecture that is much faster than the original Transformer, while maintaining the same performance.
8458e2a044e2a7b4bbeec27e80a1a0ae	tC01FRB0M7w	Turing-NLG, DeepSpeed and the ZeRO optimizer	▸ Microsoft has trained a 17-billion parameter language model that can generate human-like text. This model is trained using Turing-NLG, DeepSpeed, and the ZeRO optimizer. This video discusses the challenges of training such a large model and the solutions that were used.
50f04d3c4bfec49e4bf45a807c5be773	p3sAF3gVMMA	Deep Learning for Symbolic Mathematics	▸ This model solves integrals and ODEs by doing seq2seq! https://arxiv.org/abs/1908.09462 Abstract: We propose a deep learning architecture for symbolic mathematics, which can solve a variety of tasks including integration and ordinary differential equations.
6f1469e53f010794ffb7c028c37c625	qeEO2GECQk0	Evaluating NLP Models via Contrast Sets	▸ Current NLP models are often "cheating" on supervised learning tasks by memorizing the training data. We propose a new evaluation task called Contrast Sets, which is designed to evaluate the ability of NLP models to generalize to new tasks.
3df819e65b0361c42e14e43a2ea8ccf6	AU30czb4iQA	Imputer: Sequence Modelling via Imputation and Dynamic Programming	▸ The imputer is a sequence-to-sequence model that strikes a balance between accuracy and efficiency. It is trained using a combination of imputation and dynamic programming, and can be applied to a variety of tasks including missing data imputation and sequence classification.
b0008887012d8c7c82947e3f51bc9351	1HEdXwEYrGM	Predicting the rules behind - Deep Symbolic Regression for Recurrent Sequences (w/ author interview)	▸ #deeplearning #symbolic #research This video includes an interview with the author of the paper, who discusses the challenges of predicting the rules behind recurrent sequences and the role of deep symbolic regression.
ac81f1375f3c0237800664d9d72fbaf0	_8KNb5iqbIE	Longformer: The Long-Document Transformer	▸ The Longformer extends the Transformer by introducing sliding window self-attention. This allows the model to process documents of arbitrary length, while maintaining the same performance as the original Transformer.
e55eb62354120232ad15860ffe8d2a0f	wTIPGoHLw_8	I talk to the new Facebook Blender Chatbot	▸ This is what a 9 Billion parameter transformer can do. I take a look at the new Facebook Blender Chatbot, which is a large language model trained using a combination of supervised and unsupervised learning.
a0a34214915fa08fb4a01b8235923efa	cIUtRNhY6Rw	TAPAS: Weakly Supervised Table Parsing via Pre-training (Paper Explained)	▸ Answering complex questions about tabular information is hard. We propose a new task called TAPAS (Table Answering via Pre-training), which is designed to evaluate the ability of NLP models to answer questions about tabular information.
b53678117e01a69c459309110d1ad7d3	G3pOvrKkFuk	[Code] PyTorch sentiment classifier from scratch with Huggingface NLP Library (Full Tutorial)	▸ Huggingface released its newest library called NLP, which gives you a simple and easy way to use state-of-the-art NLP models. This video is a full tutorial on how to use the library to build a sentiment classifier from scratch.
690ae85f81bf1c332fe20420ad1fa0ba	utuz7wBGjKM	[News] OpenAI Model Generates Python Code	▸ This code completion engine can write an entire function from just a few lines of code. This is a new OpenAI model that is trained to generate Python code, and can be used for a variety of tasks including code completion and code generation.
b2f595cbb53b6b491346cd2fb26c183b	lIebBjbBevs	When BERT Plays the Lottery, All Tickets Are Winning (Paper Explained)	▸ BERT is a giant model. Turns out you can prune away many of its layers without losing performance. This paper discusses the challenges of pruning BERT and the solutions that were used.
d63c56cc711ac96c8a19cc52bbd79ad5	SY5PvZrJhLE	GPT-3: Language Models are Few-Shot Learners (Paper Explained)	▸ #gpt3 #openai #gpt-3 How far can you go with ONLY language models? GPT-3 is a new OpenAI model that is trained to generate text, and can be used for a variety of tasks including text generation and text classification.
b3a75398ea46c2e615d5e397796e97e5	q7QP_lfqnQM	Synthesizer: Rethinking Self-Attention in Transformer Models (Paper Explained)	▸ Do we really need dot-product attention? The attention mechanism in Transformers is often criticized for being slow and inefficient. This paper proposes a new attention mechanism called Synthesizer, which is designed to be faster and more efficient.
2345f78298d1524079dc665822d0535d	nxEr4VNgYOE	Movement Pruning: Adaptive Sparsity by Fine-Tuning (Paper Explained)	▸ Deep neural networks are large models and pruning has become a necessary step for deployment. This paper discusses the challenges of pruning and the solutions that were used.
f82c461053aa5c80f63509f4ff9513ef	rl4nUngiR2k	BLEURT: Learning Robust Metrics for Text Generation (Paper Explained)	▸ Proper evaluation of text generation models, such as machine translation and text summarization, is a challenging task. This paper proposes a new metric called BLEURT, which is designed to be more robust and accurate than existing metrics.
e9bb9eca38282dd3bbdf4c4c56826b79	AJwnbSP_rq8	▸ GPT-NeoX-20B - Open-Source huge language model by EleutherAI (Interview w/ co-founder Connor L...	▸ #eleuther #gptneo #gptj EleutherAI announces GPT-NeoX-20B, a new open-source language model that is trained using a combination of supervised and unsupervised learning.