



## Deep Learning News

10 10

[All Sections](#)

We share news, articles, interesting repos relating to deep learning in this channel.

[Dave DeBarr](#)

Apr 13, 2021



<https://twitter.com/paperswithcode/status/1381957151078416391?s=20>

... which reminds me ...

This is a decent place to find state-of-the-art results for ML: <https://paperswithcode.com/sota>

[the organization by task is nice]

Also: [https://www.reddit.com/r/machinelearningmemes/comments/frt4qg/oc/?utm\\_source=share&utm\\_medium=web2x&context=3](https://www.reddit.com/r/machinelearningmemes/comments/frt4qg/oc/?utm_source=share&utm_medium=web2x&context=3)

↩ Reply (2 likes)

[Dave DeBarr](#)

Apr 16, 2021

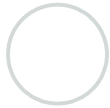


Machine Learning Street Talk podcast: a chat with François Chollet:

[https://open.spotify.com/episode/5dLR29kIUtMkT5Z9hO6NHM?si=IX\\_7LrzDQn2yTeEn-MVAbA&utm\\_source=copy-link](https://open.spotify.com/episode/5dLR29kIUtMkT5Z9hO6NHM?si=IX_7LrzDQn2yTeEn-MVAbA&utm_source=copy-link)

↩ Reply

○

[Dave DeBarr](#)

May 1, 2021

⋮

In case you're into vision:

[DINO: Emerging Properties in Self-Supervised Vision Transformers \(Facebook AI Research Explained\) - YouTube](#)



Note: may make more sense after we've covered transformers.

← [Reply](#) 👍 (1 like)

○

[Dave DeBarr](#)

May 13, 2021

⋮

Hugging Face released Vision Transformers this week ...

[Vision Transformer \(ViT\) — transformers 4.5.0.dev0 documentation \(huggingface.co\)](#)

```
from transformers import ViTFeatureExtractor, ViTForImageClassification
from PIL import Image
import requests
url = 'http://images.cocodataset.org/val2017/000000039769.jpg'
image = Image.open(requests.get(url, stream=True).raw)
feature_extractor = ViTFeatureExtractor.from_pretrained('google/vit-base-patch16-224')
model = ViTForImageClassification.from_pretrained('google/vit-base-patch16-224')
inputs = feature_extractor(images=image, return_tensors="pt")
outputs = model(**inputs)
logits = outputs.logits
# model predicts one of the 1000 ImageNet classes
predicted_class_idx = logits.argmax(-1).item()
print("Predicted class:", model.config.id2label[predicted_class_idx])
```

← [Reply](#) 👍

○

[Dave DeBarr](#)

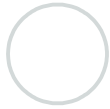
May 15, 2021

⋮

tensorflow 2.5.0 was released this week: [Release TensorFlow 2.5.0 · tensorflow/tensorflow · GitHub](#)

← Reply 👍

○

[Dave DeBarr](#)

May 18, 2021

⋮

<https://arxiv.org/abs/2105.05991>

Improving Code Autocompletion with Transfer Learning

> But what if limited examples of IDE autocompletion in the target programming language are available for model training?

> We find that these unsupervised pretrainings improve model accuracy by over 50% on very small fine-tuning datasets and over 10% on 50k labeled examples.

> We confirm the real-world impact of these pretrainings in an online setting through A/B testing on thousands of IDE autocompletion users, finding that pretraining is responsible for increases of up to 6.63% autocompletion usage.

← Reply 👍

○

[Dave DeBarr](#)

May 18, 2021

⋮

<https://www.blog.google/technology/ai/lamda>

LaMDA — short for "Language Model for Dialogue Applications" — can engage in a free-flowing way about a seemingly endless number of topics, an ability we think could unlock more natural ways of interacting with technology and entirely new categories of helpful applications.

LaMDA's conversational skills have been years in the making. Like many recent language models, including BERT and GPT-3, it's built on Transformer, a neural network architecture that Google Research invented and open-sourced in 2017.

← Reply 👍

○

[Dave DeBarr](#)

⋮



May 18, 2021

<https://twitter.com/quocleix/status/1394832822624415745?s=20>

> Transformers have become one of the most important architectural innovations in deep learning and have enabled many breakthroughs over the past few years. Here we propose a simple attention-free network architecture, gMLP, based solely on MLPs with gating, and show that it can perform as well as Transformers in key language and vision applications. Our comparisons show that self-attention is not critical for Vision Transformers, as gMLP can achieve the same accuracy. For BERT, our model achieves parity with Transformers on pretraining perplexity and is better on some downstream tasks. On finetuning tasks where gMLP performs worse, making the gMLP model substantially larger can close the gap with Transformers. In general, our experiments show that gMLP can scale as well as Transformers over increased data and compute.

← Reply



○

[Dave DeBarr](#)

May 27, 2021



New edition soon-ish ...

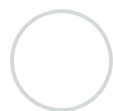
<https://www.manning.com/books/deep-learning-with-python-second-edition>

Edited by [Dave DeBarr](#) on May 27 at 10:46pm

← Reply



○

[Dave DeBarr](#)

Jun 2, 2021



"Tokenizer-free" Hugging Face model:

<https://twitter.com/huggingface/status/1399793113850421255?s=19>

← Reply

