



Search



Get unlimited access to the best of Medium for less than \$1/week. [Become a member](#)



Great results on audio classification with fastai library

Ethan Sutin · [Follow](#)

2 min read · Oct 31, 2018

[Listen](#)[Share](#)[More](#)

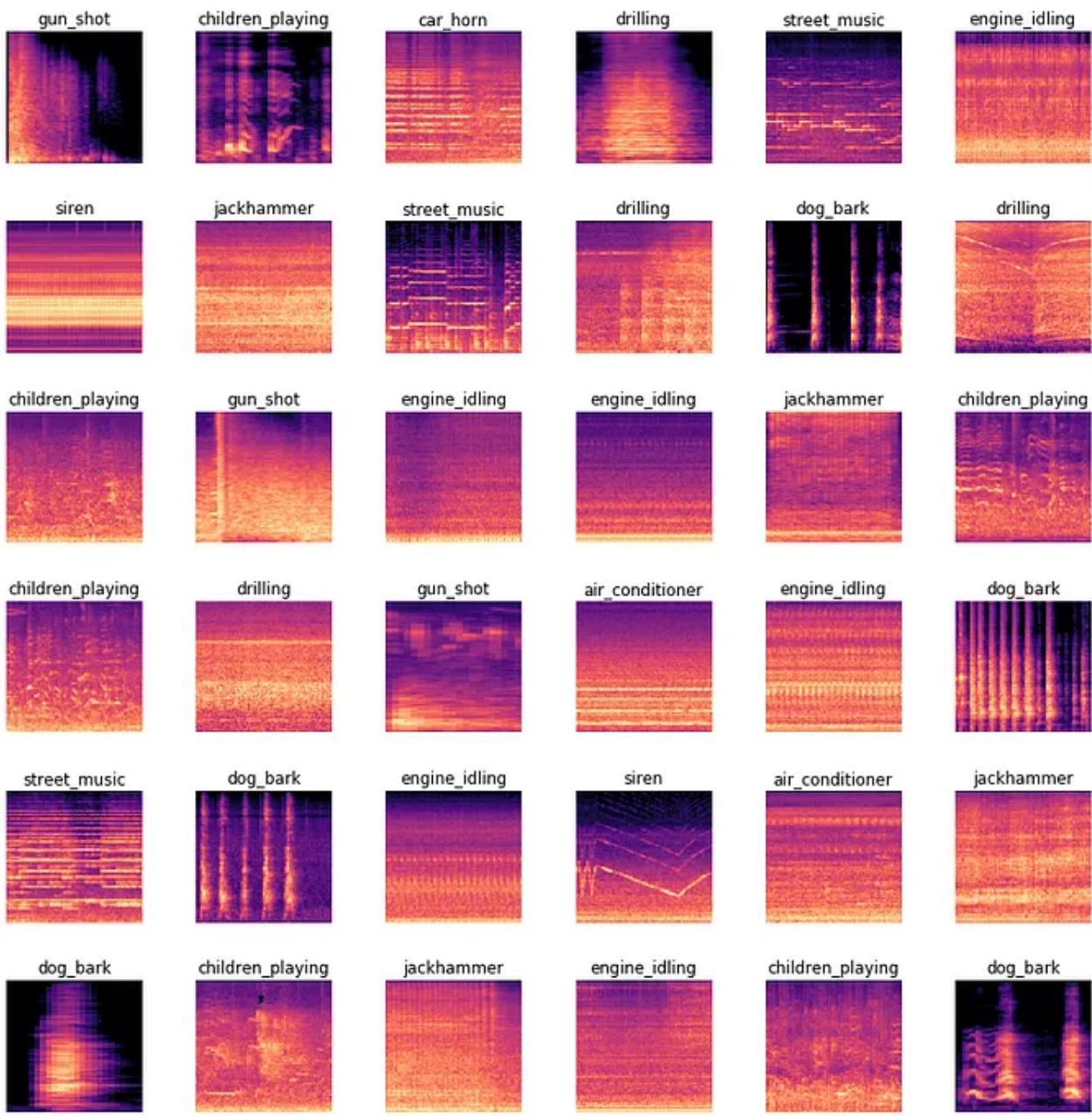
The latest version of Jeremy Howard's fast.ai [deep learning for coders](#) course has just begun. It utilizes the new [fastai library](#) built on top of PyTorch, and it makes it very easy to get great results with very little effort on a range of different tasks.

Jeremy shows that fastai is extremely effective at classifying images containing every day things like different breeds of pets, but how about on something less ImageNet-y, such as spectrograms for the purposes of audio classification.

Enter the [UrbanSound8k](#) dataset. The dataset contains 8732 labeled sound excerpts (<=4s) of urban sounds from 10 classes: air_conditioner, car_horn, children_playing, dog_bark, drilling, engine_idling, gun_shot, jackhammer, siren, and street_music.

Can we convert the audio files into spectrograms and then train a CNN to classify them? The short duration of the sounds seems like a good fit for a CNN, so let's give it a shot.

We can use the [librosa](#) python library to generate spectrograms and save them for classification.



Samples of spectrograms generated from urban sounds audio

Now that we have our images, we can train a normal image classifier as usual (except will disable the default fastai image augmentation since it won't make sense on spectrograms).

And the results are pretty impressive. We were quickly able to achieve a 80.5% mean accuracy across the 10 cv folds.

According to the latest [publication](#) on the dataset's [website](#), the state-of-the-art mean accuracy achieved was 79%. It should be noted that is with extensive audio specific augmentation, and without augmentation their top accuracy was 74%.

So even without any audio specific data augmentation, we can very quickly beat the state-of-the-art on an audio classification task. We could increase accuracy further

by introducing data augmentation.

It's pretty cool that fastai out-of-the-box can produce these kind of results even on images distant from the kind found in ImageNet!

And here's a [link](#) to the notebook.

Machine Learning

Artificial Intelligence

Computer Vision

Python

Data Science



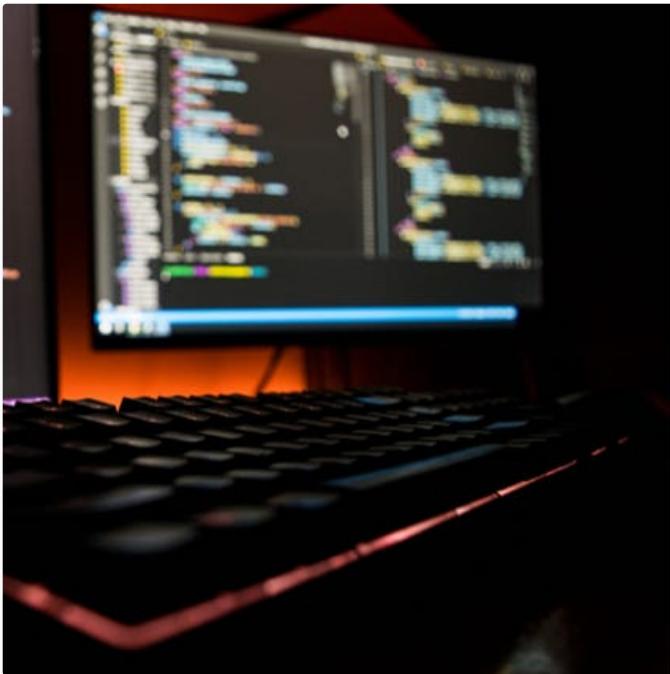
Follow



Written by Ethan Sutin

70 Followers

[More from Ethan Sutin](#)



 Ethan Sutin

Experiences as a first time CTO (Twitter acquisition edition!)

Squad is joining Twitter!

6 min read · Dec 11, 2020

 59  2

PID	CPU	USER	TH	VIRT	RES	FILE	NET	Command
1348	0.50	systemd-tim	1	112M	16M	0	24.00	/usr/sbin/chronyd -f /etc/chrony/chrony.conf
1861	0.50	root	5	4M	2M	0	0.00	csysdig
1477	0.00	root	1	1M	48K	723	0.00	/bin/hupper -pidfile /run/hupper.pid -huppidfile /run/docker.pid -path /Database/branch/master/watch/com.docker.driver
1862	0.00	root	1	1M	4K	0	0.00	/bin/sh -c run-parts /etc/periodic/15min
1864	0.00	root	1	1M	4K	0	0.00	run-parts /etc/periodic/15min
1083	0.00	root	1	1M	52K	0	0.00	/sbin/acpid
1	0.00	root	1	1M	8K	0	0.00	init
1867	0.00	root	1	1M	52K	192	215.50	ntpd -n -q -p pool.ntp.org
1322	0.00	root	14	624M	45M	0	1.31K	/usr/bin/docker daemon --pidfile=/run/docker.pid -H tcp://0.0.0.0:2375 -H unix:///var/run/docker.sock --config-file=/D
767	0.00	root	1	1M	48K	199	43.50	/sbin/syslogd
1865	0.00	root	1	1M	4K	0	0.00	run-parts /etc/periodic/hourly
1627	0.00	root	1	1M	4K	0	0.00	/sbin/getty -L ttyS0 115200 vt100
1441	0.00	root	1	1M	112K	2K	86.50	/usr/sbin/crond -c /etc/crontabs
997	0.00	root	1	1M	52K	0	0.00	udhcpc -R -n -p /var/run/udhcpc.eth0.pid -i eth0 -T 1 -A 3
1866	0.00	root	1	1M	4K	48	0.00	/bin/sh /etc/periodic/15min/ntp
1502	0.00	root	5	6M	2M	0	0.00	/usr/bin/diagnostics-server
1863	0.00	root	1	1M	112K	723	0.00	/bin/sh -c run-parts /etc/periodic/hourly
633	0.00	root	1	1M	48K	0	0.00	/sbin/klogd
1113	0.00	root	5	3M	8K	904	0.00	/sbin/transfused -p /var/run/transfused.pid -l /var/transfused_start.log
1167	0.00	root	11	9M	3M	0	1.31K	/sbin/vsudd -port 2376 -sock /var/run/docker.sock
1807	0.00	root	8	255M	6M	3K	0.00	docker-containerd-shim dc3de59a91bbad2eefc@01840c58d6609573abd46cb0e80a8bf154fbaee02e0d /var/run/docker/libcontainerd/
1830	0.00	root	1	21M	4M	0	0.00	bash
1354	0.00	root	10	346M	17M	0	0.00	docker-containerd -l /var/run/docker/libcontainerd/docker-containerd.sock --runtime docker-runc --debug --metrics-inte

 Ethan Sutin

sysdig running on the new, ‘native’ Docker apps!

Nothing beats the convenience of docker on your laptop for development and exploration. Perhaps you've had experience with docker toolbox...

2 min read · May 7, 2016

👏 4

💬 2



...



 Ethan Sutin

A Bot To Link All People and Parks

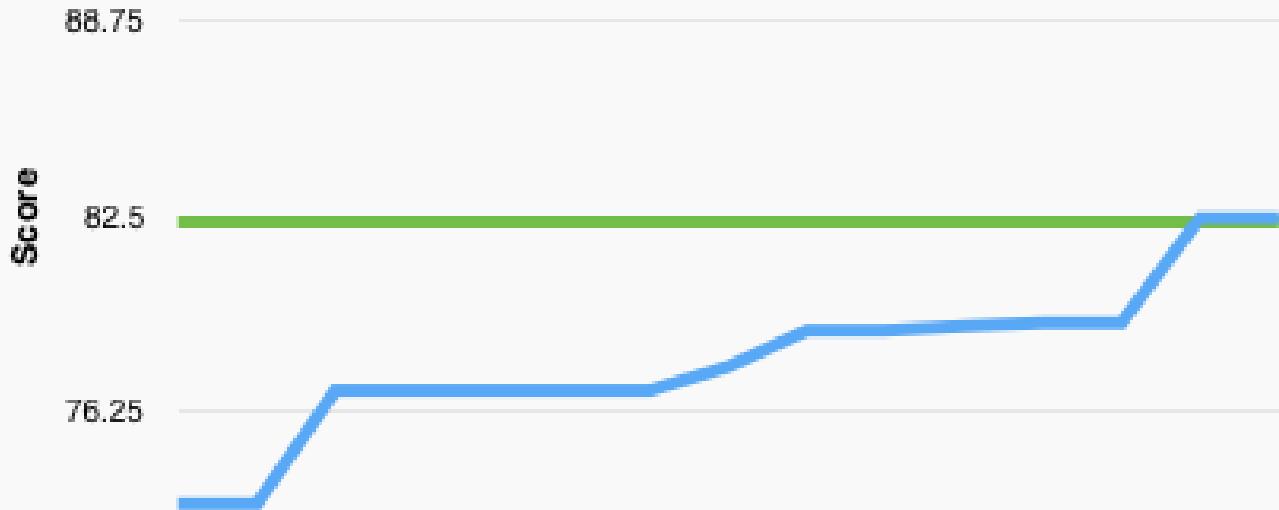
4 min read · May 19, 2016

👏 5

💬



...



 Ethan Sutin

Competition Works: Consistent Machine Comprehension Gains

NLP and NLU is hard.

2 min read · Jan 19, 2018

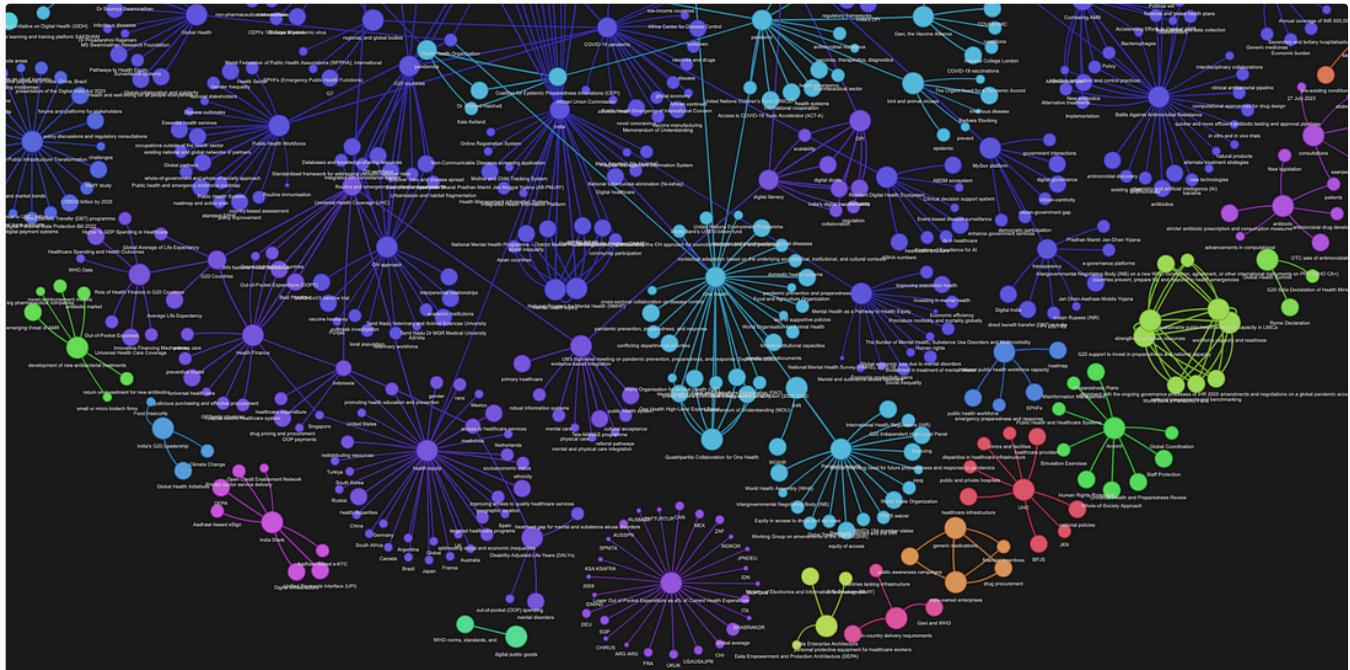
 33



...

See all from Ethan Sutin

Recommended from Medium

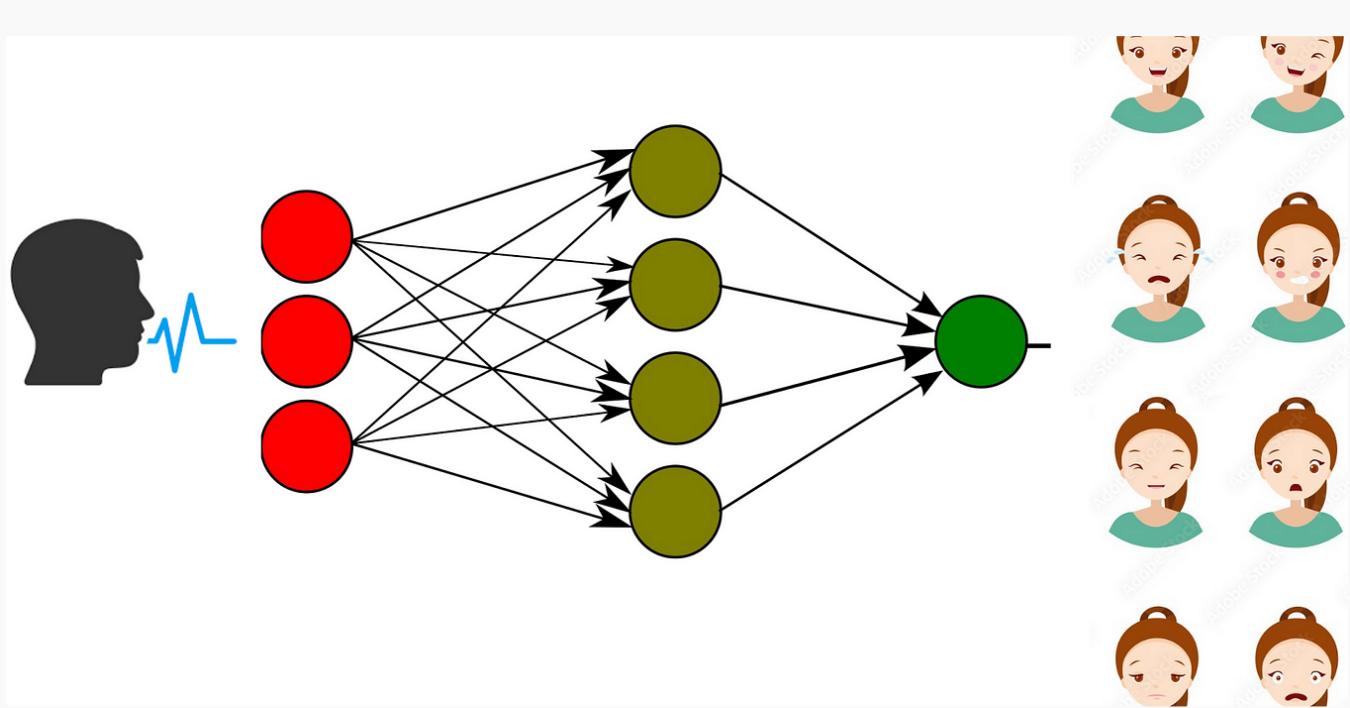


Rahul Nayak in Towards Data Science

How to Convert Any Text Into a Graph of Concepts

A method to convert any text corpus into a Knowledge Graph using Mistral 7B.

12 min read · Nov 10



Soukaina Alaoui

Building a Speech Emotion Analyzer in Python

Introduction

3 min read · Nov 11

24



...

Lists



Predictive Modeling w/ Python

20 stories · 618 saves



Practical Guides to Machine Learning

10 stories · 706 saves



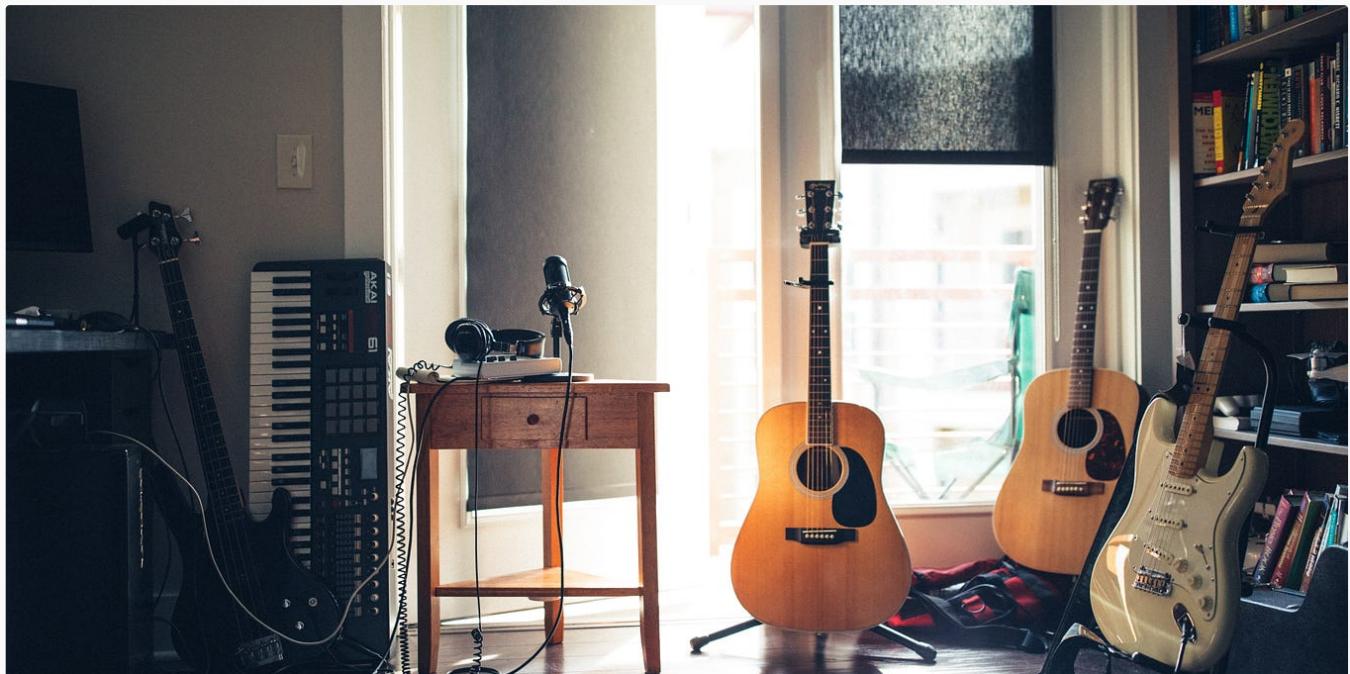
Natural Language Processing

880 stories · 410 saves



ChatGPT

22 stories · 275 saves



Technocrat in CoderHack.com

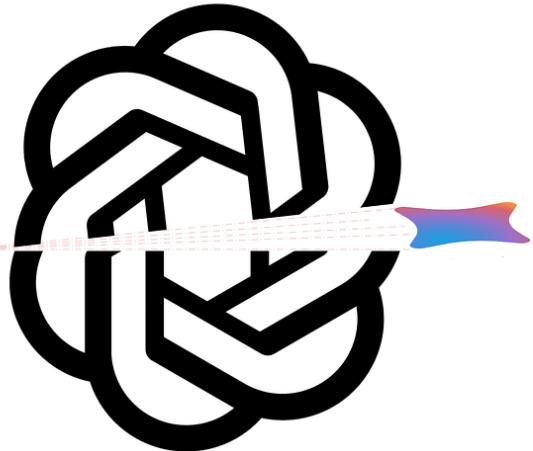
Introduction to LibROSA

LibROSA is a Python package for audio and music analysis. It provides various functions to quickly extract key audio features and metrics...

3 min read · Sep 15



Bard



 AL Anany 

The ChatGPT Hype Is Over—Now Watch How Google Will Kill ChatGPT.

It never happens instantly. The business game is longer than you know.

★ · 6 min read · Sep 2







13 hours ago

nds too easy

2 1 Reply

Paul Rose

I Found A Very Profitable AI Side Hustle

And it's perfect for beginners

6 min read · Oct 19

9.7K

180



...



Datadrifters in GoPenAI

Whisper Large v2 and Distil Whisper: Transcribe Audio Files with Insane Speed!

You can now push the boundaries of what's possible with ASR and translation with Whisper Large V2 and Distil Whisper Large V2!

★ · 6 min read · Nov 2



62



...

[See more recommendations](#)