
APPLICATIONS OF DEEP LEARNING

CA5009 - DEEP LEARNING TECHNIQUES AND APPLICATIONS

1. H2O Driverless AI:



H2O Driverless AI is an artificial intelligence (AI) platform for automatic machine learning. Driverless AI automates some of the most difficult data science and machine learning workflows such as feature engineering, model validation, model tuning, model selection, and model deployment. It aims to achieve highest predictive accuracy, comparable to expert data scientists, but in much shorter time thanks to end-to-end automation. Driverless AI also offers automatic visualizations and machine learning interpretability (MLI). Especially in regulated industries, model transparency and explanation are just as important as predictive performance. Modeling pipelines (feature engineering and models) are exported (in full fidelity, without approximations) both as Python modules and as Java standalone scoring artifacts.

2. Huggingface:



Hugging face Transformers provides APIs to easily download and train state-of-the-art pretrained models. Using pretrained models can reduce your compute costs, carbon footprint, and save you time from training a model from scratch. The models can be used across different modalities such as:

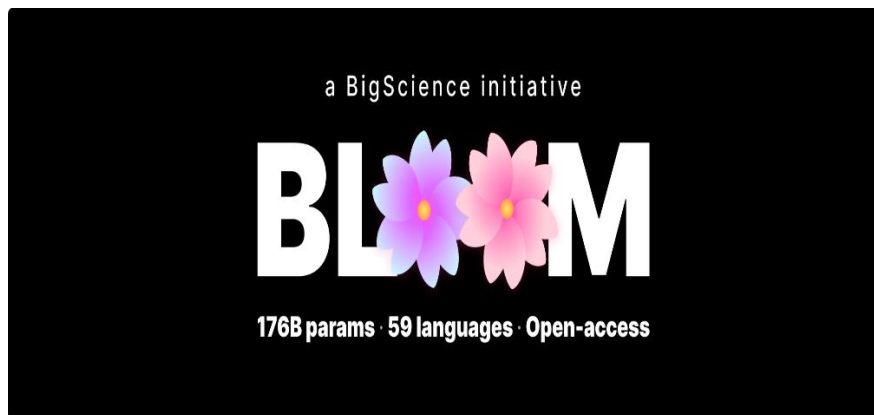
Text: text classification, information extraction, question answering, summarization, translation, and text generation in over 100 languages.

Images: image classification, object detection, and segmentation.

Audio: speech recognition and audio classification.

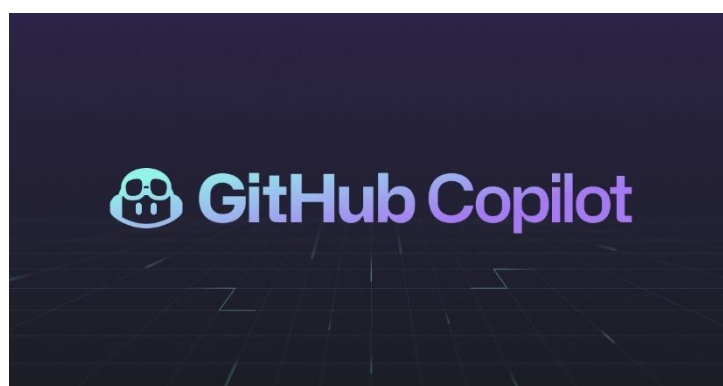
Multimodal: table question answering, optical character recognition, information extraction from scanned documents, video classification, and visual question answering.

3. BLOOM



BLOOM is an autoregressive Large Language Model (LLM), trained to continue text from a prompt on vast amounts of text data using industrial-scale computational resources. As such, it is able to output coherent text in 46 languages and 13 programming languages that is hardly distinguishable from text written by humans. BLOOM can also be instructed to perform text tasks it hasn't been explicitly trained for, by casting them as text generation tasks.

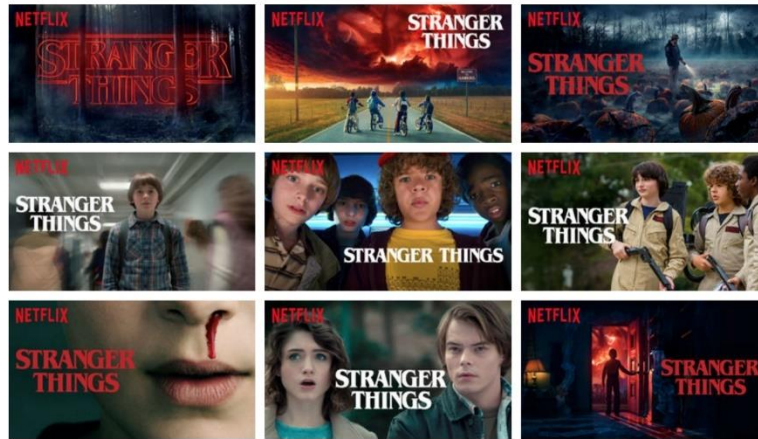
4. GitHub Copilot:



GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in real-time, right from your editor. GitHub Copilot is an AI pair programmer that offers autocomplete-style suggestions as you code. You can receive suggestions from GitHub Copilot either by starting to write the code you want to use, or by writing a natural language comment describing what you want the code to do. GitHub Copilot analyses

the context in the file you are editing, as well as related files, and offers suggestions from within your text editor.

5. Artwork Personalization at Netflix:



Using thousands of video frames from an existing movie or show as a starting point for thumbnail generation, Netflix annotates these images then ranks each image in an effort to identify which thumbnails have the highest likelihood of resulting in your click. These calculations are based on what others who are similar to you have clicked on. One finding could be that users who like certain actors / movie genres are more likely to click thumbnails with certain actors/image attributes.

6. DeepFakeLabs:



DeepFaceLab is an open-source project that can be found on GitHub. For a couple of years now, it has been possible to use this software to create deepfakes. The quality of the deepfake that the software makes depends partly on the quality of the

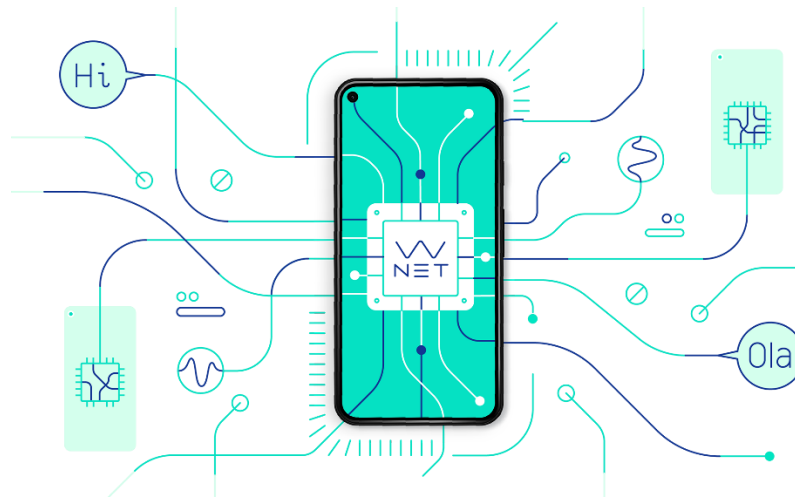
data available, but mainly on the skill of the user. The famous Tom Cruise deepfakes on Tik-Tok are a result of the models and features inside DeepFaceLab.

7. Artificial Intelligence and Autopilot at Tesla:



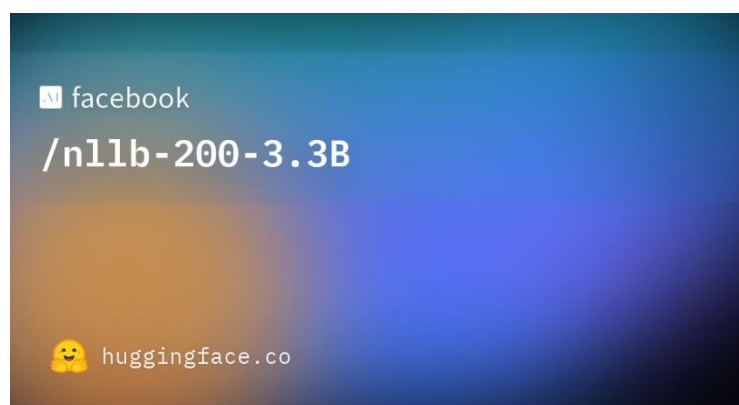
Apply cutting-edge research to train deep neural networks on problems ranging from perception to control. Our per-camera networks analyse raw images to perform semantic segmentation, object detection and monocular depth estimation. Our birds-eye-view networks take video from all cameras to output the road layout, static infrastructure and 3D objects directly in the top-down view. Our networks learn from the most complicated and diverse scenarios in the world, iteratively sourced from our fleet of nearly 1M vehicles in real time. A full build of Autopilot neural networks involves 48 networks that take 70,000 GPU hours to train 🤖. Together, they output 1,000 distinct tensors (predictions) at each timestep.

8. WaveNet:



WaveNet is a generative model that is trained on speech samples. It creates the waveforms of speech patterns by predicting which sounds likely follow each other. Each waveform is built one sample at a time, with up to 24,000 samples per second of sound. And because the model learns from human speech, WaveNet automatically incorporates natural-sounding elements left out of earlier text-to-speech systems, such as lip-smacking and breathing patterns. By including intonation, accents, emotion, and other vital layers of communication overlooked by earlier systems, WaveNet delivers a richness and depth to computer-generated voices. For example, when we first introduced WaveNet, we created American English and Mandarin Chinese voices that narrowed the gap between human and computer-generated voices by 50%.

9. Meta AI:



Meta AI has built a single AI model, [NLLB-200](#), that is the first to translate across 200 different languages with state-of-the-art quality that has been validated through extensive evaluations for each of them. We've also created a new evaluation dataset, FLORES-200, and measured NLLB-200's performance in each language to confirm that the translations are high quality. NLLB-200 exceeds the previous state of the art by an average of 44 percent. We're now using modeling techniques and

learnings from the project to improve and extend translations on Facebook, Instagram, and Wikipedia. We're open-sourcing NLLB-200 models, FLORES-200, model training code, and code for re-creating the training dataset in order to help other researchers improve their translation tools and build on our work.

10.DALL·E 2:

TEXT DESCRIPTION

An astronaut Teddy bears A bowl
of soup

riding a horse lounging in a
tropical resort in space playing
basketball with cats in space

in a photorealistic style in the
style of Andy Warhol as a pencil
drawing



DALL·E 2



DALL·E2 can create original, realistic images and art from a text description. It can combine concepts, attributes, and styles. DALL·E 2 can make realistic edits to existing images from a natural language caption. It can add and remove elements while taking shadows, reflections, and textures into account. DALL·E 2 can take an image and create different variations of it inspired by the original. DALL·E 2 has learned the relationship between images and the text used to describe them. It uses a process called “diffusion,” which starts with a pattern of random dots and gradually alters that pattern towards an image when it recognizes specific aspects of that image.
