

AI/Deep Learning Frameworks

SELECT AN AI FRAMEWORK FOR YOUR AI PROJECTS

“Which AI Framework should I use for my project?”

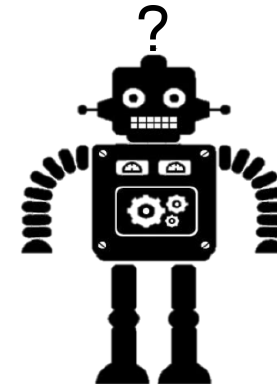


theano



PYTORCH

dmlc
mxnet



DL FRAMEWORKS: THEANO (RIP)

theano

- One of the first DL libraries, from Yoshua Bengio's Lab (MILA) (<http://deeplearning.net/software/theano/>)
- Supported interfaces: Python
- Low scalability, e.g. lack multi-GPU support
- Good documentation (Many tutorials)
- Very flexible
- On Sept. 28, 2017, MILA announced that it will stop developing Theano (RIP Theano)

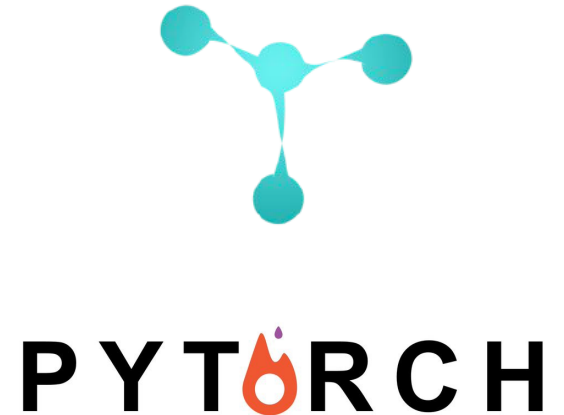
DL FRAMEWORKS: TENSORFLOW

- Open-source software library from Google (<https://www.tensorflow.org/>)
- Supported interfaces: Python, C++, (Java – experimental)
- Capability to run on multiple CPUs/GPUs
- Good documentation
- Flexible
- There is no commercial support



DL FRAMEWORKS: TORCH & PYTORCH

- Torch is scientific computing framework (maintained by Facebook/Twitter/Google (DeepMind))
(<http://torch.ch/>)
- Supported interfaces: C, C++, Lua
- PyTorch (open sourced by Facebook in Jan 2017) is built to integrate Torch in Python (<http://pytorch.org/>)
=> We can use PyTorch naturally in Python like numpy
- PyTorch builds NNs dynamically (build at runtime) => High flexibility for research projects
- Easy to learn and write code



DL FRAMEWORKS: MICROSOFT COGNITIVE TOOLKIT (CNTK)



- **From Microsoft**
- **Supported interfaces: C#, Python, C++, and Command Line**
- **High scalability: Scales across GPUs & machines**
- **Very fast for sequential models, e.g. RNNs, LSTMs**
- **No commercial support**

DL FRAMEWORKS: MXNET



- Apache incubator open source project (<https://mxnet.incubator.apache.org/>)
- Supported interfaces: Python, C++, R, Julia, MatLab
- Scalable: Can run on multiple GPUs and machines
- Amazon's "DL framework of choice"

DL FRAMEWORKS: CAFFE

Caffe

- From Berkeley Vision and Learning Center
(<http://caffe.berkeleyvision.org/>)
- Supported interfaces: Python, MatLab, C++, C, CL
- Initial focus: Computer Vision
- Useful when using Convolutional Neural Networks (CNNs)
- Not appropriate for Recurrent Neural Networks (RNNs)
- Lack of documentation

DL FRAMEWORKS: CAFFE2



- From Facebook, build on top of Caffe (<https://caffe2.ai/>)
- Supported interfaces: Python, C++
- Improvement of Caffe:
 - Higher scalability
 - Better memory optimization
- Better documentation

DL FRAMEWORKS: DEEPLEARNING4J (DL4J)



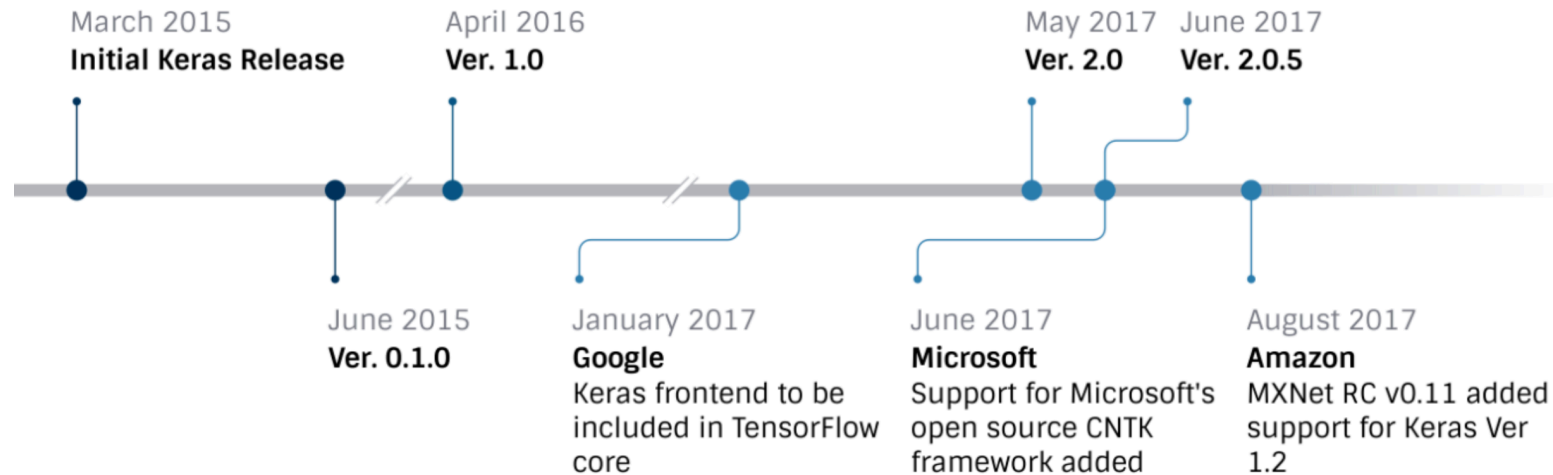
- A Deep Learning framework for Java
(<https://deeplearning4j.org/>)
- Supported interfaces: Java & Scala, compatible with JVM
- Can be implemented on top of Big Data tools, e.g.,
Hadoop, Spark
- Good documentation

DL FRAMEWORKS: HIGH LEVEL NEURAL NETWORK APIs

- An easier way to build DL models:

- Keras (<https://keras.io/>)

- Supported interfaces: Python
 - We can build a DL models in a few line of code
 - Can use Tensorflow, Theano, CNTK , MXNet or DL4J as backend
 - Good documentation



Source: Jasmeet Bhatia

DL FRAMEWORKS: HIGH LEVEL NEURAL NETWORK APIs

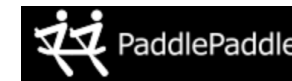
- **An easier way to build DL models:**
 - Lasagne (<http://lasagne.readthedocs.io/en/latest/>)
 - Support interface: Python
 - Can only use Theano as backend
 - Not as good documentation as Keras
- **Note: These APIs make it easy to build DL model, but they might not flexible as using backend DL libraries directly**

DL FRAMEWORKS: FEW MORE

- Chainer (<https://chainer.org/>)
 - From a Tokyo startup (Preferred Network)
 - Support dynamic computation graphs
- DSSTNE (<https://github.com/amzn/amazon-dsstne>)
 - From Amazon, written mainly in C++
 - Library behind Amazon's product recommendations
- DyNet (<https://github.com/clab/dynet>)
 - From CMU
 - Support dynamic computation graphs
- Gluon (<http://gluon.mxnet.io/>)
 - From Microsoft & Amazon: High level API for MXNet [Announced Oct. 20017]
- PaddlePaddle (<http://www.paddlepaddle.org/>)
 - DL framework from Baidu



Amazon DSSTNE



DL FRAMEWORKS: COMMUNITY & ACTIVITY

| DL Frameworks | No. GitHub Stars | No. GitHub Forks | No. GitHub Commits last month | No. SO Questions |
|---------------|------------------|------------------|-------------------------------|------------------|
| Theano | 7,403 | 2,354 | 38 | 2,350 |
| Tensorflow | 81,729 | 40,069 | 1,032 | 17,597 |
| Torch | 7,540 | 2,210 | 0 | 1,280 |
| PyTorch | 10,147 | 2,132 | 187 | 381 |
| CNTK | 13,366 | 3,487 | 63 | 345 |
| MXNET | 12,393 | 4,568 | 110 | 308 |
| CAFFE | 21,745 | 13,358 | 2 | 2,240 |
| CAFFE2 | 6,523 | 1,472 | 159 | 47 |
| DL4J | 7,896 | 3,876 | 108 | 316 |
| Keras | 22,888 | 8,347 | 78 | 5,209 |

- **GitHub: Code hosting service**
- **SO (Stack Overflow): online community for developers to learn and share knowledge**

DL FRAMEWORKS: FRAMEWORK INTEROPERABILITY

- **Can different DL frameworks work interchangeably?**

- **Yes: Attempting**
- **Facebook & Microsoft recently introduce open ecosystem for interchangeable AI frameworks – ONNX (Open Neural Network Exchange Format)**
- **Currently supported: Caffe2, CNTK, MXNet, and PyTorch**
- **Project site: <https://onnx.ai/>**



DL FRAMEWORKS

| Software | Creator | Software license ^[a] | Open source | Platform | Written in | Interface | OpenMP support | OpenCL support | CUDA support | Automatic differentiation ^[1] | Has pretrained models | Recurrent nets | Convolutional nets | RBM/DBNs | Parallel execution (multi node) |
|-----------------------------|---|---------------------------------|-------------|---|------------------------|---|--|--|-------------------------|--|---|---------------------|---------------------|--------------------|---------------------------------|
| Deeplearning4j | Skyrmind engineering team; Deeplearning4j community; originally Adam Gibson | Apache 2.0 | Yes | Linux, Mac OS X, Windows, Android (Cross-platform) | C++, Java | Java, Scala, Clojure, Python (Keras), Kotlin | Yes | On roadmap ^[8] | Yes ^{[9][10]} | Computational Graph | Yes ^[11] | Yes | Yes | Yes | Yes ^[12] |
| Dlib | Davis King | Boost Software License | Yes | Cross-Platform | C++ | C++ | Yes | No | Yes | Yes | Yes | No | Yes | Yes | Yes |
| Keras | François Chollet | MIT license | Yes | Linux, Mac OS X, Windows | Python | Python, R | Only if using Theano or MXNet as backend | Under development for the Theano backend (and on roadmap for the TensorFlow backend) | Yes | Yes | Yes ^[13] | Yes | Yes | Yes | Yes ^[14] |
| MXNet | Distributed (Deep) Machine Learning Community | Apache 2.0 | Yes | Linux, Mac OS X, Windows, ^{[25][26]} AWS, Android, ^[27] iOS, JavaScript ^[28] | Small C++ core library | C++, Python, Julia, Matlab, JavaScript, Go, R, Scala, Perl | Yes | On roadmap ^[29] | Yes | Yes ^[30] | Yes ^[31] | Yes | Yes | Yes | Yes ^[32] |
| Apache SINGA | Apache Incubator | Apache 2.0 | Yes | Linux, Mac OS X, Windows | C++ | Python, C++, Java | No | Yes | Yes | ? | Yes | Yes | Yes | Yes | Yes |
| TensorFlow | Google Brain team | Apache 2.0 | Yes | Linux, Mac OS X, Windows ^[33] | C++, Python | Python (Keras), C/C++, Java, Go, R ^[34] | No | On roadmap ^[35] but already with SYCL ^[36] support | Yes | Yes ^[37] | Yes ^[38] | Yes | Yes | Yes | Yes |
| Theano | Université de Montréal | BSD license | Yes | Cross-platform | Python | Python (Keras) | Yes | Under development ^[39] | Yes | Yes ^{[40][41]} | Through Lasagne's model zoo ^[42] | Yes | Yes | Yes | Yes ^[43] |
| Torch | Ronan Collobert, Koray Kavukcuoglu, Clement Farabet | BSD license | Yes | Linux, Mac OS X, Windows, ^[44] Android, ^[45] iOS | C, Lua | Lua, LuaJIT, ^[46] C, utility library for C++/OpenCL ^[47] | Yes | Third party implementations ^{[48][49]} | Yes ^{[50][51]} | Through Twitter's Autograd ^[52] | Yes ^[53] | Yes | Yes | Yes | Yes ^[54] |
| Wolfram Mathematica | Wolfram Research | Proprietary | No | Windows, Mac OS X, Linux, Cloud computing | C++ | Wolfram Language | No | No | Yes | Yes | Yes ^[55] | Yes | Yes | Yes | Yes |
| Microsoft Cognitive Toolkit | Microsoft Research | MIT license ^[16] | Yes | Windows, Linux ^[15] (OSX via Docker on roadmap) | C++ | Python (Keras), C++, Command line, ^[17] BrainScript ^[18] (.NET on roadmap ^[19]) | Yes ^[20] | No | Yes | Yes | Yes ^[21] | Yes ^[22] | Yes ^[22] | No ^[23] | Yes ^[24] |
| Caffe | Berkeley Vision and Learning Center | BSD license | Yes | Linux, Mac OS X, Windows ^[2] | C++ | Python, MATLAB | Yes | Under development ^[3] | Yes | Yes | Yes ^[4] | Yes | Yes | No | ? |
| Caffe2 | Facebook | Apache 2.0 | Yes | Linux, Mac OS X, Windows ^[5] | C++, Python | Python, MATLAB | Yes | Under development ^[6] | Yes | Yes | Yes ^[7] | Yes | Yes | No | Yes |
| MatConvNet | Andrea Vedaldi, Karel Lenc | BSD license | Yes | Windows, Linux ^[15] (OSX via Docker on roadmap) | C++ | MATLAB, C++ | No | No | Yes | Yes | Yes | Yes | Yes | No | Yes |
| Neural Designer | Artelnics | Proprietary | No | Linux, Mac OS X, Windows | C++ | Graphical user interface | Yes | No | No | ? | ? | No | No | No | ? |
| OpenNN | Artelnics | GNU LGPL | Yes | Cross-platform | C++ | C++ | Yes | No | No | ? | ? | No | No | No | ? |
| Gensim | | | | | | | | | | | | | | | |
| Paddle | | | | | | | | | | | | | | | |
| Pytorch | | | | | | | | | | | | | | | |

NOW WHAT?!

- From wiki: https://en.wikipedia.org/wiki/Comparison_of_deep_learning_software

DL FRAMEWORKS: WHICH TO CHOOSE?

- **You work in industry**
 - **Speed & Scalability**
 - **Stability**

DL4J

CAFFE

TensorFlow

CNTK

MXNET

DL FRAMEWORKS: WHICH TO CHOOSE?

- **You work in a research organization**
 - **Flexibility**
 - **Easy debugging**

PyTorch & Torch

Theano

TensorFlow

MXNET

CNTK

DL FRAMEWORKS: WHICH TO CHOOSE?

- **You are DL beginner/practitioner**
 - **Easiness**
 - **Quick prototyping**

Keras

Lasagne

TensorFlow

DL FRAMEWORKS: WHICH TO CHOOSE?

- **You are Uni. Prof.**
 - Use an easy to learn framework

PyTorch

TensorFlow

MXNet

DL FRAMEWORKS: WHICH TO CHOOSE?

- If you are working on computer vision

Caffe

Caffe2

MXNet

Torch

DL FRAMEWORKS: WHICH TO CHOOSE?

- If you are using RNNs for sequence data

PyTorch

CNTK

Hot AI Trend/Technologies in 2017

- **Natural Language Generation**
 - Generate text from existing data
 - Use cases: report generation, summarizing business info.
- **Speech Recognition**
 - Make computer understand human speech
 - Use cases: Voice control systems, Mobile applications
- **Virtual Agents**
 - Bots to interact with human
 - Use cases: Smart home manager, chatbots
- **Machine Learning Platforms**
 - Platforms providing APIs, toolkit to create models
 - Use cases: Used in enterprise applications
- **Deep Learning Platforms**

Hot AI Trend/Technologies in 2017

- **AI-Optimized Hardware**
 - Provide hardware for AI jobs
- **Biometrics**
 - Allows more natural interactions between humans and machines
 - Use cases: Identification, physical measurement and analysis
- **Robotics**
 - More intelligent robots
- **Decision Management**
 - Systems to automate decision-making
- **Text Analysis and NLP**
 - Computers understand human languages

Discussion!