

PyTorch Fast.AI training: Deep Learning

3 days (21 hours)

Presentation

[PyTorch](#) and the Fastai library, respectively supported by Facebook Research (FAIR) and a community of developers around Jeremy Howard (co-founder of Kaggle), form a powerful duo for rapidly discovering the power of Deep Learning algorithms.

PyTorch release 1 released in December 2018 and Fastai version 2 are available seamlessly on Google's Kaggle or colab kernel environment.

After a brief review of the Machine Learning paradigm, and the various tasks where Deep Learning outperforms traditional methods, we will implement classic tasks such as image and text classification, word vectors and auto-encoders. Using the Anaconda environment, we'll be using CUDA NVIDIA to speed up our processing.

Depending on the progress of the course and practical work, participation in one or more data challenges on Kaggle (or other) will be organized. The comparison with the competing framework, Google's [TensorFlow](#) with the Keras library (now integrated with TF v2.0), will be discussed during the course.

By the end of the course, participants will have a thorough understanding of PyTorch's features and capabilities, as well as its role and contribution in AI compared to other frameworks and libraries. Participants will also have received the practice they need to implement PyTorch in their own projects.

Like all our training courses, this one will introduce you to the latest stable version and its new features: [PyTorch 2.6](#) and [Fastai 2.7](#).

Objectives

- Understand the concept of Deep Learning and be able to use it with pyTorch with or without fast.ai
- Use web resources (github, arxiv) to build a case library ready-to-use
- Be able to install and use Pytorch/ Fastai independently
- Understand the concept of Machine Learning and be able to use it in PyTorch

Target audience

- Developers
- Architects
- Big Data Data analyst
- Data scientist

Prerequisites

- Knowledge of Python
- Knowledge of numpy is a plus

Technical requirements

- Have Anaconda installed or access to Google Collab

Further information

- As an introduction to [Artificial Intelligence](#), we offer you the following training course
- Complementing Google's [TensorFlow](#) technology

PyTorch training program

Day 1 - morning: Deep Learning: The Artificial Intelligence Revolution

- Defining machine learning as an optimization problem
- Types machine learning algorithms: supervised (classification, regression), unsupervised, semi-supervised, self-supervised, reinforcement learning
- Underlearning, overlearning, generalization (learning curve), metric and evaluation
- The mathematics behind Machine Learning: linear algebra, statistics/probability, gradient descent applied to linear regression
- The advantages and limitations of Machine Learning explained on logistic regression and SVM
- Advantages of Deep Learning over Machine Learning explained on MLP and CNNs
- Reasons to switch to Deep Learning
- Real-life examples and applicable scenarios

- Introducing Pytorch
- Introducing fastai and pytorch lightning overlays

Day 1 - afternoon: Setting up and getting to grips with Pytorch: manipulating vectors and tensors with Pytorch

- Installation: installation of Pytorch, anaconda, drivers; Google Collab
- Using Pytorch like Numpy: tensors, operations and variables
- Use cases
- Creation of a neural network module
- Loading a dataset and splitting it into train and test datasets with Pytorch and Sklearn
- Using Sklearn metrics and Tensorboard to visualize results
- Hyperparameter optimization and the AutoML principle
- Saving and restoring a model
- Using callbacks

Day 2 - morning: Introduction to CNN and its application to image processing

- Implementation and architecture of a CNN (convolution, pooling, data augmentation)
- Understanding and visualizing a CNN and the convolution product
- Convolution and pooling (classification) layers in a CNN
- Transfer learning and development of convolutional neural networks
- Explanation of backpropagation on an example and implementation in Pytorch
- End-to-end image : downloading a learning database (dog/cat, bear: teddy, grizzly, black bear), using a state-of-the-art architecture (ResNet) for transfer learning.
- Comparison of CNN creation with Pytorch and FastAI

Day 2 - afternoon: Multi Layer Perceptron (MLP) on tabular data

- Explanation of the embedding concept and its importance for semantic dimension reduction
- Loading csv data with the Pandas library
- Creation of quantitative variables from qualitative/categorical variables, treatment missing data (nan values), normalization with Pandas then FastAI
- Creating an auto-encoder
- Application to visualization (PCA, T-SNE, UMAP) and clustering embeddings (application to item recommendation)

Day 3 - morning: Text processing

- Explanation of word embeddings
- Use with the Spacy library (word2vec, CBOW, Glove, TF-IDF)
- Explanation of embedding sentences
- Use with the Universal Sentence Encoder library
- Use of Flair and HuggingFace libraries for classic NLP tasks (English and French sentiment analysis, prediction, ...)

Day 3 - afternoon: Introduction to LSTMs and application to time series analysis and prediction

- LSTM implementation and architecture (sequence-to-sequence, prediction/many-to-one, ...)
- RNN vs LSTM vs GRU
- Explaining the attention mechanism (vs. Transform)
- Creating an LSTM
- Application for time series prediction (stock prices or sales)

Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire enabling us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives with regard to the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.