+33 (0)1 48 10 09 46
formation@ambient-it.net
140 Avenue Jean Lolive 93500 Pantin - France

Updated 03/15/2024

Sign up

Deep Learning and Neural Networks fundamentals for engineers

3 days (21 hours)

Presentation

Artificial intelligence has revolutionized many scientific fields, and is now beginning to revolutionize a large number of economic sectors (industry, medicine, communications, etc.). Nevertheless, the way it is presented in the mainstream media is often nothing more than a fantasy, far removed from what Machine Learning and Deep Learning actually are.

The aim of this course is to provide engineers who have already mastered IT tools (including a software programming base) with an introduction to Deep Learning and its various fields of specialization, and therefore to the main network architectures available today.

If the mathematical bases are recalled during the course, a level of mathematics of the type BAC+2 is recommended for more comfort. It is possible, in absolute terms, to omit the mathematical axis and retain only a "systems" vision, but this approach will limit your understanding of the subject.

Objectives

- Understanding the fundamental principles of Machine Learning and its evolution towards Deep Learning
- Review of the main tools and applications
- Master simple convolutional and recursive neural networks using Tensorflow examples
- Understanding more advanced models: auto-encoders, gans, reinforcement learning
- Understanding attention and transformation models and reinforcement learning
- Knowledge of the theoretical foundations and practical application of neural network architecture and convergence
- Understanding the limits and benefits of deep learning
- Master the concepts of generative models and distribution approximations

Target audience

Developers, Architects, Big Data Data Analyst / Data Engineer / Data Scientist / Engineers / Analysts

Prerequisites

Knowledge of Python and mathematics.

Software requirements

Python installed.

Further information

- By way of introduction, we offer you an Artificial Intelligence training course.
- Complementing Facebook's Pytorch or Google's TensorFlow technology

Deep Learning and neural networks for engineers training program

[DAY 1]

1. Introduction to AI, Machine Learning & Deep Learning

- The fantasy of artificial intelligence versus today's realities.
- Disappearance of algorithms, new problem modeling.
- Machine learning: learning overview
- Main approaches: supervised learning, unsupervised learning, reinforcement learning, self supervised learning.
- Main actions: classification, regression, clustering, density estimation, reduction of dimensionality, prediction, generation
- Evolutionary algorithms: introduction and current status
- Examples of Machine Learning algorithms: Linear regression, Naive Bayes, Random Tree.

2. Fundamental concepts of a neural network (Application: multi-layer perceptron)

- Reminder of basic mathematics.
- Definition of a neural network: classical architecture, activation and weighting functions for previous activations, network depth.
- Defining neural network learning: cost functions, backpropagation,stochastic gradient descent, maximum likelihood.

- Modeling a neural network: modeling input and output data according to the type of problem (regression, classification, etc.). Curse of dimensionality. Distinction between multi-feature data and signal. Choosing a cost function according to the data.
- Approximating a distribution with a neural network: presentation and examples
- Data Augmentation: how to balance a data set and identify data biases.
- Generalization of neural network results.
- Neural network initialization and regularization: L1/L2 regularization, Batch Normalization, Instance Normalization
- Optimization and convergence algorithms: stochastic gradient descent, batching, Adagrad, AdaDelta, RMSProp,

3. Usual ML / DL tools

- Data management tools: Apache Spark, Apache Hadoop
- Common Machine Learning tools: Numpy, Scipy, Sci-kit
- High-level DL frameworks: PyTorch, Tensorflow, Caffe

4. Deep Learning applications: review of the state of the art and examples of applications

- Data classification
  - Understand data classification in different scenarios: raw data, image, sound, text, etc.
  - Understand the challenges of data classification and the choices implied by a model classification.
  - Presentation of common classification tools, in particular MLP (Multilayer perceptron) or CNN (Convolutional neural network) networks VS Machine Learnig tools (Random Forest, Naive bayes)
  - Presentation of examples of existing solutions (e.g. image classification) customer history, texts written by users, etc.).
  - Clustering: a special case of unsupervised learning. Overview of different algorithms (k-means, Random Forests, etc.).
  - Anomaly detection: tools and limits
- Information prediction and sequential/temporal data
  - Stakes and limits of information prediction. Search for structural rules within the data that can enable predictive logic.
  - Prediction as classification or regression.
  - Common pitfalls of a predictive approach.
  - Presentation of common prediction tools: RNN (Recurrent Neural Networks), LSTM (Long Short Term Memory).
  - Examples: image prediction following a video sequence. Pollution prediction in urban environments, and more.
- Transformation / Data generation
  - What exactly is data transformation? What are the barriers, what's at stake?
  - Re-interpretation of the same data: de-noising, text summary generation, image segmentation.
  - Transformation operation on the same format: translation of text from one language to another other (brief presentation of Google Machine Translation or BERT architecture by Google).
  - Original" data generation operation: neural style, super-resolution, generation images from text presentations.

- Reinforcement Learning: controlling an environment
  - Introducing Deep Reinforcement Learning.
  - Applications: control of numerical simulations, automatic cars, robotics.

[DAY 2]

5. Convolutional Neural Networks (CNN)

- Introduction to CNNs: basic principles and applications.
- Fundamental operation of a CNN: convolutional layer, use of a kernel, padding & stride, feature map generation, pooling layers. 1D, 2D and 3D extensions.
- Presentation of the different CNN architectures that have taken the state of the art in image classification to a new level.
  LeNet, VGG Networks, Network in Network, Inception, Resnet. Presentation of the innovations brought by each architecture and their more global applications (Convolution 1x1 or residual connections).
- Using an attention model.
- Application to a typical classification scenario (text or image).
- U Networks for multi-class generation or segmentation.
- Exploiting the internal representations of a CNN for unsupervised secondary learning.

6. Recurrent Neural Networks (RNN)

- Introduction to RNNs: fundamental principles and applications.
- Fundamental operation of recursive neural networks: hidden activation, back propagation through time, unfolded version.
- Developments towards GRU (Gated Recurrent Units) and LSTM (Long Short Term Memory). Presentation of the different states and evolutions brought by these architectures.
- Convergence problems and vanishing gradient.
- Types of classic architectures: prediction of a time series, classification...
- RNN Encoder Decoder architecture.
- NLP applications: word/character encoding, translation.
- Video applications: prediction of the next generated frame of a video sequence.

7. Attention patterns, CNN vs RNN

- Definition of a fundamental attention model in image analysis, Deep Reinforcement Learning approach.
- Attention model upstream of a CNN to target a particular area.
- CNN Sequence to sequence with a model of attention, architectural detail.
- Caution in RNN Encoder-decoder architecture.
- Attention is all you need" model, state of the art, Cellule transformer.

8. Debugging / analysis of network operation

- The current state of mathematical understanding of convergence in Deep Learning.

- Analysis of "dead" neurons or fundamental kernels in a CNN.
- GradCam and Saliency maps.
- Analysis of an attention model.
- Hierarchical Contextual Decompositions.

[DAY 3]

9. Generational models: Variational AutoEncoder (VAE) and Generative Adversarial Networks (GAN)

- Presentation of generative models, link with CNNs/RNNs.
- Auto-encoder: dimensionality reduction and generation of a compressed version. Limits, collapse mode.
- Variational Auto-encoder: generational model and data distribution approximation. Defining and using latent space. Reparameterization trick. Applications and observed limits.
    - Generative Adversarial Networks: fundamental principles. Two-network architecture (generator and discriminator) with alternate learning, cost functions available.
- GAN convergence and difficulties encountered.
- Improved convergence: Wasserstein GAN, Earth Moving Distance.
- CycleGAN, progressive growing GANs, BigGAN.
- Applications for image or photo generation, text generation, super-resolution.

10. Deep Reinforcement Learning.

- Introducing reinforcement learning: controlling an agent in an environment defined by a state and possible actions.
- Fundamental approach: Markov decision process, Q-Learning VS Policy Gradient, Monte Carlo, SARSA, TD Learning.
- Deep Q Learning: experience replay, and application to video game control. Double Q Learning & Dueling Q Networks. Rainbow.
- Policy gradients. On-policy && off-policy. Actor critic architecture and A3C asynchronous approach.
- Proximal Policy Optimization (OpenAI).
- Entropy parameter to encourage exploration.
- World models, imagination augmented agents.
- Introducing AlphaGo and AlphaGo Zero.

Companies concerned

This training 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 enrolment 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 objectives.

This questionnaire also enables us to anticipate any connection or internal security problems (intra-company or virtual classroom) that could be problematic for the follow-up and smooth running of the training session. This questionnaire also enables us to anticipate any connection or internal security difficulties within the company (intra-company or virtual classroom) that 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, with 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.