Updated on 18/07/2023

Sign up

Explainable AI (XAI) training

2 days (14 hours)

Presentation

Expainable AI (XAI) is the set of techniques and best practices that will enable you to interpret your machine learning models.

Year after year, models become more and more powerful. On the other hand, they are more complex than ever, and often described by business teams as "black boxes".

With the arrival of AI applications in ever more critical use cases, the need for interpretation becomes a major issue in order to ensure transparent, ethical and trusted applications.

Expainable AI (XAI) is a set of recent techniques and models for interpreting the most complex machine learning models.

In this course, you'll learn how to implement the most advanced interpretability techniques with various frameworks such as SHAP and LIME, and apply them to a variety of use cases, with a particular focus on image analysis models (CNN).

Objectives

- Understand the importance of interpretability and when to implement it
- Understanding the taxonomy and main families of explainable AI (XAI) techniques
- Implement these techniques on supervised machine learning use cases (PDP, LIME, SHAP, permutation, importance, etc.).
- Understand good implementation practices in a professional context

Target audience

- Data Analyst
- Data Scientist
- Data Engineer
- Machine learning engineer
- Developers
- Big Data Architects
- Lead Developer

Prerequisites

- Knowledge of Python
- General knowledge of supervised machine learning
- Knowledge of general mathematics (probability, statistics, etc.)

Explainable AI (XAI) training program

DAY 1 - Morning

Interpreting machine learning models

- The importance of interpretation
- Assessing the criticality of a use case
- Overview and taxonomy of XAI techniques Global and

agnostic methods

- Partial dependence plots
- Accumulated local effects
- Permutation feature importance
- Advantages and disadvantages
- Lab: application to a concrete use case with scikit-learn and elie5

DAY 1 - Afternoon

Global and agnostic methods

- Local surrogate (LIME)
- Counterfactual explanations
- Shapley and SHAP value frameworks
- Lab: application to a deep learning use case for image analysis.

Practical application

- Extension to other use cases (time series, NLP, etc.)
- Implementation in a professional context
- Review of good and bad practices
- Training summary

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.

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.