AMBIENT-IT

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

Updated on 14/05/2024

Sign up

Ray.io training

2 days (14 hours)

Presentation

Our Ray.io training course will help you simplify the scaling of your compute-intensive Python workloads. Whether you're a data scientist or an AI engineer, this framework will enable you to set up compute clusters to create large-scale machine learning workflows.

Our program will enable you to build and deploy your first distributed applications, and integrate them with machine learning libraries such as PyTorch or TensorFlow. You will also learn how to optimize your applications to improve performance.

Our training course will also teach you how to deploy machine learning models and implement efficient autoscaling and resource allocation scenarios for your production services. Cluster monitoring and security will also be covered.

Like all our training courses, it will run on the latest version of the tool: Ray 2.21

Objectives

- Creating a distributed application
- Handle large volumes of data
- How to deploy a machine learning model
- Monitoring your clusters

Target audience

- **Data scientist**
- IA Engineer

Prerequisites

- Knowledge of Python
- Knowledge of Machine Learning

RAY.IO TRAINING PROGRAM

INTRODUCTION TO RAY

- Overview and features
- Ray installation and initial environment setup
- Key components and architecture
- Use Cases
- Ray community and available resources

FUNDAMENTAL CONCEPTS

- Exploring Ray Core:
  - tasks
  - actors
- Distributed execution engine
- Using the API to create distributed applications
- Error and exception handling in a distributed environment
- Best practices for managing reports in applications

DISTRIBUTED PROGRAMMING

- Building and deploying a first simple distributed application
- Strategies for deploying applications on local clusters or in the cloud
- Debugging and performance optimization
- Practical examples of Ray's integration with machine learning libraries such as PyTorch and TensorFlow
- Ray Tune for hyperparameter optimization in machine learning projects

DATA PROCESSING AND DISTRIBUTED TRAINING

- Ray Data for distributed processing of large data sets
- Introduction to Ray Train for distributed training of deep learning models
- Practical implementation of a data processing pipeline
- Configuring and running a distributed drive on multiple nodes
- Scalability techniques and resource management in distributed environments

ONLINE INFERENCE AND MODEL MANAGEMENT WITH RAY SERVE

- Deploying and managing machine learning models in production with Ray Serve

- Model composition and multiplexing techniques for efficient deployments
- Configuring and monitoring Ray Serve services in production
- Practical scenarios: Deploying machine learning applications using FastAPI and Ray Serve
- Autoscaling and resource allocation for production services.

MONITORING, SECURITY AND OPTIMIZATION OF RAY CLUSTERS

- Ray cluster configuration, deployment and management
- Using the Ray dashboard to monitor and debug applications and clusters
- Good security practices for distributed applications
- Advanced techniques for optimizing Ray cluster performance
- Solving common problems and troubleshooting in Ray environments

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 registration is finalized, the learner receives a self-assessment questionnaire which enables 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 for 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, 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.