Machine Learning with Apache Mahout

Description

This one-day course is designed to help Software Engineers and Data Scientists understand the high-level concepts and classifications of machine learning systems, with a strong focus on building Recommender Systems.

You will gain an understanding of the tools and high-level conceptual ideas needed to understand what a machine learning solution is (and is not) capable of, and how to identify a suitable use case. You will learn how to construct an example solution at the conceptual level using pre-provided building blocks in order to get a feel for the general design patterns.

You will learn hands-on how to build a scalable hybrid real-time Recommender System based on Apache Hadoop, Apache Mahout, and Apache Solr, and how to optimise the system to deliver real business value.

Delegates will learn how to

- Classes and categories of machine learning systems
- Capabilities and limitations of end solutions, in business terms
- Capabilities and limitations of technology, in solution capability terms
- How to use case identification and structure
- How to structure and plan a machine learning project for your business

Audience

Software Engineers, Data Scientists, or Technologists with a background in Java programming or a similar modern programming language.

Prerequisites

- Programming skills in Java (or similar modern programming language)
- Basic understanding of Hadoop architecture
- Basic understanding of Hadoop MapReduce for data processing

---

Outline

Concepts

- Machine learning system classifications
- Capabilities and limitations

Use Cases

- Top level use case categorisations
- Identifying and categorising your own use case
- Deep-dive use case example

Technology

- Technology landscape
- Capabilities and limitations
- Selecting the right tools for the job
- Implementation choices
- Optimisation
- Performance and scalability
- Integration