INFOSOFT IT SOLUTIONS

Training | Projects | Placements

Revathi Apartments, Ameerpet, 1st Floor, Opposite Annapurna Block, Infosoft It solutions, Software Training & Development Institute, +91-9059683947|+91-9182540872

Apache Mahout Course Content

Key features

This course will prepare you to:

- Explain the architecture of the APACHE Mahout component.
- Configure and use new functionalities in APACHE Mahout.
- Use the standard APACHE Mahout Sub Modules.
- Explain the APACHE Mahout Controlling Configuration and Customization option.

Introduction to Machine Learning And Mahout

- In Mahout Training, you will know what machine learning is, what Apache mahout is and what is clustering.
- Machine Learning Fundamentals
- Apache Mahout Basics
- History of Mahout
- Supervised and Unsupervised Learning techniques
- Mahout and Hadoop
- Introduction to Clustering and Classification.

Apache Mahout And Hadoop

- Myrrix is a recommendation engine based on mahout, therefore this module is designed for mahout training and myrrix.
- Mahout on Apache Hadoop
- Setup Mahout and Myrrix.

Recommendation Engine In Mahout Training

- This module will focus on Recommendation algorithm and Mahout optimizations.
- Recommendations using Apache Mahout
- Introduction to Recommendation systems
- Content Based Mahout Optimizations.

Implementing A Recommender And Recommendation Platform

- Understanding the various recommendations, implementing Recommendors, different types of similarities in Apache mahout.
- User based recommendation
- User Neighbourhood
- Item based Recommendation
- Implementing a Recommender using MapReduce Platforms
- Similarity Measures
- Manhattan Distance
- Euclidean Distance
- Cosine Similarity

- Pearson’s Correlation Similarity
- Log likelihood Similarity
- Tanimoto Evaluating
- Recommendation Engines (Online and Offline)
- Recommendors in Production.

Clustering

- This module is designed to give you thoroughly over the clustering concepts.
- Clustering
- Common Clustering Algorithms in Apache mahout training
- K-means Canopy Clustering
- Fuzzy K-means and Mean Shift etc.
- Representing Data Feature Selection
- Vectorization in Apache Mahout training
- Representing Vectors
- Clustering documents through example TF-IDF and Implementing clustering in Hadoop

Classification

- By the end of this training module, you will be able to develop a classifier on your own.
- Examples
- Basic Predictor variables and Target variables
- Common Algorithms
- SGD
- SVM
- Navie Bayes

- Random Forests
- Training and evaluating a Classifier
- Developing a Classifier

Apache Mahout And Amazon EMR

- We’ll focus on Apache Mahout and Amazon EMR, have an overview on Weka, Octave Matlab and SAS.
- Mahout on Amazon
- EMR Mahout Vs R
- Introduction to tools like Weka, Octave, Matlab and SAS