BTS
Training & Consultancy

Analytics & Data Science

Course

Website: www.btsconsultant.com

Email: info@btsconsultant.com

Telephone: 00971-2-6452630

Analytics & Data Science Course

Introduction

If you want to become a data scientist, this is the course to begin with. Using open source tools, it covers all the concepts necessary to move through the entire data science pipeline, and whether you intend to continue working with open source tools, or later opt for proprietary services, this course will give you the foundation you need to assess which options best suit your needs.

Objectives:

You Will Learn How To

- Translate business questions into Machine Learning problems to understand what your data is telling you
- Explore and analyze data from the Web, Word Documents, Email, Twitter feeds, NoSQL stores, Relational Databases and more, for patterns and trends relevant to your business
- Build Decision Tree, Logistic Regression and Naïve Bayes classifiers to make predictions about your customers' future behaviors as well as other business critical events
- Use K-Means and Hierarchical Clustering algorithms to more effectively segment your customer market or to discover outliers in your data
- Discover hidden customer behaviors from Association Rules and Build Recommendation Engines based on behavioral patterns
- Use biologically-inspired Neural Networks to learn from observational data as humans do
- Investigate relationships and flows between people, computers and other connected entities using Social Network Analysis

Who should attend?

- Professionals interested in entering the fields of data analytics, technology, informatics, business intelligence, web analytics or data collection
- Professionals with limited academic and work experience in data analytics and related technology fields
- Anyone interested in data analysis and data collection
- Students interested in prerequisite knowledge prior to pursuing our intermediate or advanced data programs such as Data Science and Big Data Programming & Architecture

Recommended Experience:

There's no expectations regarding specific platforms except basic familiarity with a Windows environment.

Course Outline

- Introduction to R

Exploratory Data Analysis with R

- Loading, querying and manipulating data in R
- Cleaning raw data for modeling
- Reducing dimensions with Principal Component Analysis
- Extending R with user-defined packages

Facilitating good analytical thinking with data visualization

- Investigating characteristics of a data set through visualization
- Charting data distributions with boxplots, histograms and density plots
- Identifying outliers in data
- Working with Unstructured Data

Mining unstructured data for business applications

- Preprocessing unstructured data in preparation for deeper analysis
- Describing a corpus of documents with a term-document matrix
- Make predictions from textual data
- Predicting Outcomes with Regression Techniques

Estimating future values with linear regression

- Modeling the numeric relationship between an output variable and several input variables
- Correctly interpreting coefficients of continuous data
- Assess your regression models for 'goodness of fit'
- Categorizing Data with Classification Techniques

Automating the labelling of new data items

- Predicting target values using Decision Trees
- Constructing training and test data sets for predictive model building
- Dealing with issues of overfitting

Assessing model performance

- Evaluating classifiers with confusion matrices
- Calculating a model's error rate
- Detecting Patterns in Complex Data with Clustering and Social Network Analysis

Identifying previously unknown groupings within a data set

- Segmenting the customer market with the K-Means algorithm
- Defining similarity with appropriate distance measures
- Constructing tree-like clusters with hierarchical clustering
- Clustering text documents and tweets to aid understanding

Discovering connections with Link Analysis

- Capturing important connections with Social Network Analysis
- Exploring how social networks results are used in marketing
- Leveraging Transaction Data to Yield Recommendations and Association Rules

Building and evaluating association rules

- Capturing true customer preferences in transaction data to enhance customer experience
- Calculating support, confidence and lift to distinguish "good" rules from "bad" rules
- Differentiating actionable, trivial and inexplicable rules

Constructing recommendation engines

- Cross-selling, up-selling and substitution as motivations
- Leveraging recommendations based on collaborative filtering
- Learning from Data Examples with Neural Networks

Machine learning with neural networks

- Learning the weight of a neuron
- Learning about how neural networks are being applied to object recognition, image segmentation, human motion and language modeling
- Analyzing labelled data examples to find patterns in those examples that consistently correlate with particular labels for object recognition
- Implementing Analytics within Your Organization

Expanding analytic capabilities

- Breaking down Data Analytics into manageable steps
- Integrating analytics into current business processes
- Reviewing Hadoop, Spark, and Azure services for machine learning

Dissemination and Data Science policies

- Examining ethical questions of privacy in Data Science
- Disseminating results to different types of stakeholders
- Visualizing data to tell a story