

## Statistics with R

### Module Overview

The following module comprises of R programming basics and application of several Statistical Techniques using it. The module aims to provide exposure in terms of Statistical Analysis, Hypothesis Testing, Regression and Correlation using R programming language.

### Learning Objectives

The objective of this module to make students exercise the fundamentals of statistical analysis in R environment. They would be able to analysis data for the purpose of exploration using Descriptive and Inferential Statistics. Students will understand Probability and Sampling Distributions and learn the creative application of Linear Regression in multivariate context for predictive purpose.

### Learning Outcomes

After the successful completion of this module, students will be able to:

- Install, Code and Use R Programming Language in R Studio IDE to perform basic tasks on Vectors, Matrices and Data frames.
- Describe key terminologies, concepts and techniques employed in Statistical Analysis.
- Define, Calculate, Implement Probability and Probability Distributions to solve a wide variety of problems.
- Conduct and Interpret a variety of Hypothesis Tests to aid Decision Making.
- Understand, Analyse, Interpret Correlation and Regression to analyse the underlying relationships between different variables.

### Unit I

#### Introduction to R Programming

R and R Studio, Logical Arguments, Missing Values, Characters, Factors and Numeric, Help in R, Vector to Matrix, Matrix Access, Data Frames, Data Frame Access, Basic Data Manipulation Techniques, Usage of various apply functions – apply, lapply, sapply and tapply, Outliers treatment.

## **Unit II**

### **Descriptive Statistics**

Types of Data, Nominal, Ordinal, Scale and Ratio, Measures of Central Tendency, Mean, Mode and Median, Bar Chart, Pie Chart and Box Plot, Measures of Variability, Range, Inter-Quartile-Range, Standard Deviation, Skewness and Kurtosis, Histogram, Stem and Leaf Diagram, Standard Error of Mean and Confidence Intervals.

## **Unit III**

### **Probability, Probability & Sampling Distribution**

Experiment, Sample Space and Events, Classical Probability, General Rules Of Addition, Conditional Probability, General Rules For Multiplication, Independent Events, Bayes' Theorem, Discrete Probability Distributions: Binomial, Poisson, Continuous Probability Distribution, Normal Distribution & t-distribution, Sampling Distribution and Central Limit Theorem.

## **Unit IV**

### **Statistical Inference and Hypothesis Testing**

Population and Sample, Null and Alternate Hypothesis, Level of Significance, Type I and Type II Errors, One Sample t Test, Confidence Intervals, One Sample Proportion Test, Paired Sample t Test, Independent Samples t Test, Two Sample Proportion Tests, One Way Analysis of Variance and Chi Square Test.

## **Unit V**

### **Correlation and Regression**

Analysis of Relationship, Positive and Negative Correlation, Perfect Correlation, Correlation Matrix, Scatter Plots, Simple Linear Regression, R Square, Adjusted R Square, Testing of Slope, Standard Error of Estimate, Overall Model Fitness, Assumptions of Linear Regression, Multiple Regression, Coefficients of Partial Determination, Durbin Watson Statistics, Variance Inflation Factor.

## **References**

1. Ken Black, 2013, *Business Statistics*, New Delhi, Wiley.
2. Lee, Cheng. et al., 2013, *Statistics for Business and Financial Economics*, New York: Heidelberg Dordrecht.

3. Anderson, David R., Thomas A. Williams and Dennis J. Sweeney, 2012, *Statistics for Business and Economics*, New Delhi: South Western.
4. Waller, Derek, 2008, *Statistics for Business*, London: BH Publications.
5. Levin, Richard I. and David S. Rubin, 1994, *Statistics for Management*, New Delhi: Prentice Hall.

# Python Programming

## Module Overview

Python Programming module is intended for students who wish to learn the Python programming language. This module is highly important so as to proceed with this programme. The module comprises of Programming basics with regards to Python Language such as Data Types, Operators, Functions, Classes and Exception Handling.

## Learning Objectives

This module will help students gain much needed knowledge pertaining to Python Programming, so as to prepare them for the advanced modules such as ML. Python scripting is user-friendly and is the most used language in industry when it comes to designing and scripting applications with respect to Emerging Technologies.

## Learning Outcomes

Upon successful completion of this module, students should be able to:

- To understand why Python is a useful scripting language.
- To learn how to use lists, tuples, and dictionaries in Python programs.
- To learn how to write loops and decision statements in Python.
- To learn how to write functions and pass arguments in Python.
- To learn how to design object-oriented programs with Python classes.
- To learn how to use exception handling in Python applications for error handling.

## Unit I

### Introduction

History of Python, Need of Python Programming, Applications Basics of Python Programming Using the REPL(Shell), Running Python Scripts, Variables, Assignment, Keywords, Input-Output, Indentation.

## Unit II

### Types, Operators and Expressions

Types - Integers, Strings, Booleans; Operators- Arithmetic Operators, Comparison (Relational) Operators, Assignment Operators, Logical Operators, Bitwise Operators, Membership Operators, Identity Operators, Expressions.

## **Unit III**

### **Data Structures and Control Flow**

Lists, Operations, Slicing, Methods, Tuples, Sets, Dictionaries, Sequences, Comprehensions, Conditional blocks using If, Else and El-if, For Loop, For loop using Ranges, String, list and Dictionaries, While Loop, Loop Manipulation using Pass, Continue, Break and Else, Conditional and Loops Block.

## **Unit IV**

### **Functions Modules and Packages**

Defining Functions, Calling Functions, Passing Arguments, Keyword Arguments, Default Arguments, Variable-length arguments, Anonymous Functions, Function Returning Values, Scope of the Variables in a Function - Global and Local Variables. Creating modules, Name Spacing, Introduction to PIP, Installing Packages via PIP, Using Python Packages.

## **Unit V**

### **Object Oriented Programming & Exception Handling**

Classes, Self-Variable, Methods, Constructor Method, Inheritance, Overriding Methods, Data Hiding, Difference between an Error and Exception, Handling Exception, Try Except Block, Raising Exceptions, and User Defined Exceptions.

## **References**

1. R.Nageswara Rao, 2018, *Core Python Programming*, Dreamtech.
2. John Hearty, 2016, *Advanced Machine Learning with Python*, Packt.
3. Jake VanderPlas, 2016, *Python Data Science Handbook: Essential Tools for Working with Data*, O'Reilly.
4. Mark Lutz, 2010, *Programming Python*, O'Reilly.
5. Tim Hall and J-P Stacey, 2009, *Python 3 for Absolute Beginners*, Apress.

# Structured Query Language

## Module Overview

In this course, the students will learn the basics of the SQL/No SQL and the Relational Databases. They will learn about the Relational Model and Relational Model concepts and constraints. The students will get exposure to key concepts with regards to SQL Language and DBMS such as Normalization, Transaction Processing along-side an exposure to No SQL programming.

## Learning Outcomes

This module will help students gain much needed knowledge pertaining to Relational Database Management Systems, Data Models, SQL query processing, Normalization along with an introduction to No SQL Database systems using Mongo DB.

## Learning Objectives

- To understand the basic concepts and the applications of Database Systems.
- To master the basics of SQL and construct queries using SQL.
- To become familiar with the basic issues of Transaction Processing and Concurrency Control.
- To become familiar with NO SQL Programming Language.
- Explain the architecture, define objects, load and query data within No SQL databases.

## Unit I

### Introduction to Database Management Systems

Introduction-Database System Applications, Purpose of Database Systems, Views of Data, Data Abstraction, Instances and Schemas, Data Models, Database Languages, DDL, DML, Database Architecture, Database Users and Administrators, Database Design, ER Diagrams, Entities, Attributes and Entity Sets, Relationships and Relationship sets, Integrity Constraints, Views.

## Unit II

### SQL Operators and Relational Theorems

Relational Algebra and Calculus, Selection and Projection, Set Operations, Renaming, Joins, Division, Relational calculus, Tuple Relational Calculus, Domain Relational Calculus, Forms

of Basic SQL Query, Nested Queries, Comparison Operators, Aggregate Operators, NULL values, Logical connectives, AND, OR and NOT, Outer Joins, Triggers.

## **Unit III**

### **Normalization**

Problems Caused by Redundancy, Decompositions, Functional Dependencies, Normal Forms, First, Second, Third Normal forms, BCNF, Properties of Decompositions, Loss less Join Decomposition, Dependency Preserving Decomposition, Multi Valued Dependencies, Fourth Normal Form, Join Dependencies, Fifth Normal Form.

## **Unit IV**

### **Transactions**

Transaction Management, Transaction Concept, Transaction State, Implementation of Atomicity and Durability, Concurrent, Executions, Serializability, Recoverability, Implementation of Isolation, testing for serializability, Concurrency Control, Lock, Timestamp Based Protocols, Validation Based Protocols, Recovery, Failure Classification, Storage Structure, Atomicity, Log Based Recovery, Remote Backup Systems.

## **Unit V**

### **No SQL**

Overview of No SQL, Types of No SQL Databases, No SQL Storage Architecture, CRUD Operations in MongoDB, Querying, Modifying and Managing No SQL Databases, Indexing and Ordering, Migrating from RDBMS to No SQL, No SQL in Cloud, Database Administration.

## **References**

1. Guy Harrison, 2015, *Next Generation Databases: No SQL and Big Data*, Apress.
2. Ramez Elmasri, ShamkatB.Navathe, 2013, *Database Systems*, Pearson.
3. Pramod J. Sadalage, Martin Fowler, 2012, *No SQL Distilled*, Addison Wesley.
4. A.Silberschatz, H.F. Korth, S.Sudarshan, 2006, *Database System Concepts*, McGraw Hill.
5. Raghurama Krishnan, Johannes Gehrke, 2003, *Database Management Systems*, McGraw Hill.

## **SaS and Tableau**

### **Module Overview**

This module forms an introduction as well as an in-depth study in the discipline of SaS programming with regards to Statistical Analysis along with Data Visualization through Tableau. The module comprises of basic introduction and programming in SaS, using Procedures within SaS and Data Visualization using Tableau application.

### **Learning objectives**

This course will provide students and exposure towards SaS, Tableau and its usability in the field of analytics. The course comprises of an introduction to SaS, its procedures, visualizations along with Tableau application usage and visualization basics.

### **Learning outcomes**

Upon successful completion of this module, students should be able to:

- Deploy SaS in a virtual environment and import data for analysis.
- Prepare and manipulate datasets for analysis in SaS.
- Perform exploratory data analysis within SaS environment using various procedures and functions.
- Understand Tableau Interface, Panes and Implement Visualization Techniques.
- Prepare, Deploy and Publish Stories, Dashboards based on Analytical Cases.

## **UNIT I**

### **Introduction to SaS**

Overview of SaS university edition, Deploying SaS Studio on virtual platform, File Management, SaS libraries, importing data, Structure of Data and Data Types, Program Syntax, saving data, PROC IMPORT and PROC CONTENTS, Displaying Data and Generating Logs, List Input.

## **UNIT II**

### **SaS Programming**

Variables and Syntax Rules, Data Set Options, Operators, In-File Statement, Input Styles, Select Statements, Leave and Continue, Decision Making via SaS, Where Statement, Looping Constructs, SaS Functions, Arrays and Array Processing, Modifying and Combining Data Sets.

## **UNIT III**

### **SaS Procedures**

Proc Data, Proc Sort, Proc Means Sort, Proc Means, Proc Univariate, Proc Freq, Proc Plot, Proc Sgplot, Proc Summary, Proc Contents, Proc Append, Proc Copy, Proc SQL, Proc Delete, Proc Format, Proc Import, Proc Export, Proc Transpose, Proc GChart, Proc GPlot, Proc Report.

## **UNIT IV**

### **Visualization with Tableau -I**

Tableau Software Ecosystem, Toolbar Icons, Data Window and Aggregation, Tableau Data Source, Data Extract, Connect to Data, Measure Names, Number of Records & Measures, Heat Maps, Tree maps, Bar Chart, Line Chart, Area Fill Charts, Pie Chart, Scatter Plot, Circle View, Bullet Graph, Packed Bubble, Histogram, Boxplot and Gantt Chart, Sorting Data, Enhancing Views with Filters, Sets, Groups & Hierarchies.

## **UNIT V**

### **Visualization with Tableau - II**

Cross-tabulation, Dashboard Designing, Dashboard Actions, Joining Database, Functions in Tableau, Aggregate Functions, Numeric Functions, Date Functions, Stories, Advanced Mapping, Advanced Parameters, Tableau Best Practices, Combining Multiple Dashboards into Stories, Publishing Stories and Dashboards.

## **References**

1. Ron Cody, 2018, *An Introduction to SaS University Edition*, SaS Institute.
2. Ron Cody, 2018, *Learning SaS by Example*, SaS Institute.
3. Deepti Gupta. 2018, *Applied Analytics through Case Studies Using SaS*, Apress.
4. Joshua N. Milligan, 2015, *Learning Tableau*, Packt.
5. Ben Jones, 2014, *Communication Data with Tableau: Designing, Developing and Delivering Data Visualization*, O'Reilly.

# Big Data Analytics

## Module Overview

This module forms an introduction as well an in-depth study in the field of Big Data and Hadoop. It comprises of the fundamentals as well as advanced topics needed to progress in this technology. The students will learn about the applications, usage and several use case scenarios pertaining to Big Data- Hadoop where they can use the knowledge and progress ahead.

## Learning objectives

This course will help students gain knowledge and understanding about Big Data Technology, Hadoop Ecosystem and various tools related to it. The students will learn about the HDFS File System, Map Reduce Framework, Analysing data using Hbase and Hive along with the Integration of R with Hadoop.

## Learning outcomes

Upon successful completion of this module, students should be able to:

- Understand the fundamentals of Big Data and its Applications in various Domains.
- Conceptualize and Incorporate the Technologies behind Big Data.
- Understand HDFS File Structure, Map Reduce Framework, the architectures related to them and to use them to solve complex problems.
- Integrate R with Hadoop and solve analytical problems.
- Understand and Use Hive/Hbase shell pertaining to relational data handling under Hadoop.

## UNIT I

### Introduction to Big Data

What Is Big Data? History of Data Management, Evolution of Big Data, Structuring of Big Data, Elements of Big Data, Application of Big Data in the Business Context, Careers in Big Data. Business Applications of Big Data: The Significance of Social Network Data, Financial Fraud and Big Data, Fraud Detection in Insurance, Use of Big Data in the Retail Industry.

## UNIT II

### Technologies for Handling Big Data

Distributed and Parallel Computing for Big Data, Understanding Hadoop, Cloud Computing, Grid Computing and In-Memory Technology for Big Data. VMWare Installation of Hadoop, Linux and its Shell Commands, Different Hadoop Distributions and their advantages, Hortonworks, Cloudera, MapR.

## UNIT III

### Understanding the Hadoop Ecosystem

The Hadoop Ecosystem, Storing Data with HDFS, Design of HDFS, HDFS Concepts, Command Line Interface to HDFS, Hadoop File Systems, Java Interface to Hadoop, Anatomy of a file read, Anatomy of a file write, Replica placement and Coherency Model. Parallel Copying with distcp, keeping an HDFS Cluster Balanced.

## Unit IV

### Map Reduce Fundamentals

Origins of Map Reduce, How Map Reduce Works, Optimization Techniques for Map Reduce Jobs, Applications of Map Reduce, Java Map Reduce classes (new API), Data flow, combiner functions, running a distributed Map Reduce Job. Configuration API, setting up the development environment, Managing Configuration.

## Unit V

### Integrating R with Hadoop, Understanding Hive & Hbase

Understanding R-Hadoop, Integration Procedure, Packages needed for R under Hadoop Ecosystem, Text Mining for Deriving Useful Information using R within Hadoop, Introduction to Hive & Hbase, Hive and Hbase Architecture, Understanding Queries, Mining Big Data with Hive & Hbase.

## References

1. Arshdeep Bahga, 2016, *Big Data Science & Analytics: A Hands-On Approach*, VPT.
2. Tom White, 2012, *Hadoop: The Definitive Guide*, O'Reilly.
3. Adam Shook and Donald Miner, 2012, *Map Reduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems*, O'Reilly.
4. Dean Wampler, Edward Capriolo & Jason Rutherglen, 2012, *Programming Hive*, O'Reilly.

5. Lars George, 2011, *HBase - The Definitive Guide: Random Access to Your Planet-Size Data*, O'Reilly.

## **Social Media Analytics**

### **Module Overview**

This module comprises of advanced disciplines and units pertaining to Analytics, Data Acquisition via web, Understanding and analysing Unstructured Data along with the future implementations and growth in the discipline of Analytics.

### **Learning objectives**

This course aims at giving exposure on the advanced aspects with regards to Analytics. The course comprises of Social Media, Mobile, Text Analytics along with Web Scraping and the future advancements in the field of Analytics.

### **Learning outcomes**

Upon successful completion of this module, students should be able to:

- Apply and use Social Media Analytics for the betterment of the business.
- Use Mobile Analytics for solving complex business problems and to stop churn.
- Evaluate the business problem and apply analytics techniques for better output.
- Analyse and understand patterns and techniques in Social Media & Mobile Analytics to solve complex problems.

## **UNIT I**

### **Overview**

Social Media, On-Line Social Network, Off-Line Social Network, Metrics and Measurement, Dashboard, Target Audience, Desired Action, Content, Market Research Online Communities, Cluster Analysis, Conjoint Analysis, Multidimensional Scaling, Social Media Listening, Social Media Scoring, Social Media Modelling.

## **UNIT II**

### **Mobile Analytics**

Understanding Mobile Analytics Concepts, difference between Mobile Analytics and Site Analytics, Natural language Processing with Mobile Analytics, Text Mining for Mobile Analytics, Mobile Analytics Tools, Churn Analytics.

## **UNIT III**

### **Text Analytics**

Text Data, Sources of Text Data, Information Clusters, Patterns, Trends, Tagging, Natural Learning Process, Lexical Analysis, Social Network Nodes, Linkage Structure, Node Labelling, Content-Based Classification, Word Stemming, Stemming Algorithms, Polarity of the Attitude, Psychological Profiling, Sentiment Analysis.

## **UNIT IV**

### **Web Scraping**

Web Scraping of unstructured data, Gathering data from HTTP and HTTPS format, Web Scraping from XML and JSON file, Regular expressions, Extraction Strategies, Term Document Matrix, Data Cleansing, Data Manipulation and Data Transformation after Scrapping.

## **UNIT V**

### **Future of Analytics**

Introduction to Big Data, Predictive Analysis for Business, Social Information Processing and Distributed Computing, Advances in Machine Learning, Traditional Data Models Evolve, Analytics to Solve Social Problems, Location Based Data Explosion, Data Privacy Backlash, Internet of Things, Artificial Intelligence.

### **References**

1. Galit Shamuelli, 2017, *Data Mining for Business Analytics: Concepts, Techniques and Applications with R*, Wiley.
2. Luis Torgo, 2017, *Data Mining with R: Learning Case Studies*, Chapman.
3. Zaki & Meira, 2014, *Data Mining and Analysis Fundamental Concepts and Algorithms*, Cambridge.
4. Han, Kamber & Pei, 2013, *Data Mining: Concepts and Techniques*, Morgan Kaufmann.
5. Han, Jiawei and Kamber, Micheline, 2012, *Data Mining: Concepts and Techniques*, Morgan Kaufman.

# Natural Language Processing

## Module Overview

The following module aims to provide an exposure to students based on Natural Language Processing Techniques, Algorithms and Analytics based on Textual Data. The module comprises of NLP basics, Feature Extraction, Relationship Extraction, Text Categorization, Clustering and Topic Modelling Algorithms and Practices.

## Learning Objectives

The course introduces the concepts of Text Analytics, Unstructured Information Analysis for better decision making by deriving valuable insights. The course will help the students understand the roots behind Text Mining which evolved from Machine Learning, Natural Language Processing and Statistics. Upon completion, students are expected to be able to describe basic concepts and methods of Text Mining, Information Extraction, Text Classification and Clustering, Topic Modelling.

## Learning Outcomes

After the successful completion of this module, students will be able to:

- Understand approaches to Syntax and Semantics in NLP.
- Understand various methods for Statistical approaches to Machine Translation.
- Build Models which extract information from Textual Unstructured Data.
- Understand and implement Topic Modelling and Probabilistic Models for Information Extraction.
- Implement and deploy programs based on Relationship Extraction, POS Tagging and Clustering Algorithms based on NLP.

## Unit I

### Introduction to Text Mining

Basics of Text Mining, Natural Language Content Analysis, Core Text Mining Operations, Associations, Using Background Knowledge for Text Mining, Domain Ontologies, Domain Lexicons. Text Mining Pre-processing Techniques, Task Oriented Approaches, NLP Tasks, Tokenization, Part-of-Speech Tagging, Syntactical Parsing and Shallow Parsing.

## **Unit II**

### **Extracting Features, Relations from Text**

Finding Implicit Features, Finding Opinion Phrases and their Polarity, Context-Specific Word Semantic Orientation, Analysis of Word and Document Frequency, tf-idf, Zipf's Law, bind tf\_idf Function, Subsequence Kernels for Relation Extraction, Capturing Relation Patterns with a String Kernel.

## **Unit III**

### **Text Categorization and Clustering**

Applications of Text Categorization, Document Representation, Knowledge Engineering Approach to Text Categorization, Machine Learning Approach to Text Categorization, Evaluation of Text Classifiers. Clustering Tasks in Text Analysis, Clustering Algorithms and Clustering of Textual Data.

## **Unit IV**

### **Relationships between Words**

Tokenizing by N-gram, Counting and Filtering N-gram, Analysing Bigrams to provide Context in Sentiment Analysis, visualizing a Network of Bigrams using ggraph, Counting and Correlating Pairs of Words with the widyr Package, Counting and Correlating among Sections, Examining Pairwise Correlation.

## **Unit V**

### **Topic Modelling and Probabilistic Models for Information Extraction**

Latent Dirichlet Allocation, Word Topic Probabilities, Per-Document Classification, By-words Assignments, Alternative LDA Implementations. Hidden Markov models, Stochastic Context Free Grammar, Conditional Random fields, Parallel Learning Algorithms.

## **References**

1. Julia Silge, David Robinson, 2018, *Text Mining with R-A Tidy Approach*, O'Reilly
2. Matthew L. Jockers, 2014, *Text Analysis with R for Students of Literature*, Springer.
3. James Pustejovsky, Amber Stubbs, 2012, *Natural Language Annotation for Machine Learning*, O'Reilly.
4. Steve R. Poteet, 2007, *Natural Language Processing with Text Mining*, Springer.
5. James Sanger, Ronen Feldman, 2002, *The Text Mining Handbook: Advanced Approaches in Analysing Unstructured Data*, Cambridge.

# **Machine Learning and Artificial Intelligence**

## **Module Overview**

This module comprises of conceptual, statistical and Machine Learning algorithms. It covers all the necessary models pertaining to Machine Learning such as Regression, Decision Trees, Support Vectors, Clustering, Association Analysis and Dimensionality Reduction.

## **Learning Objectives**

After this course students will gain critical knowledge and understanding about major Data Mining procedures like Decision Tree, Cluster Analysis, Neural Networks, Support Vector Machine, Bayesian Networks and Machine Learning fundamentals. Students will be able to apply and practice this gained knowledge in variety of Business Scenarios.

## **Learning Outcomes**

Upon successful completion of this module, students should be able to:

- Understand and employ a wide variety of Statistical and Machine Learning Algorithms.
- Identify the characteristics of Datasets, Problem Statement and develop Machine Learning programs with reference to known Computing Techniques.
- Understand the Model Performance Evaluation and select the best one based on the solution.
- Implement Machine Learning techniques and the Programming Framework to obtain acceptable decisions for the Real-World problems.

## **Unit I**

### **Classification and Regression Tree**

Classification & Regression, working of a Decision Tree, Attribute Selection Measures, Information Gain, Gain Ratio, Gini Index, Building Decision Trees, CART, C5.0, and CHAID Trees, Prediction by Decision Tree, Advantages and Disadvantages of Decision Trees, Model Overfitting, Building Decision Trees in R.

## **Unit II**

### **Clustering**

Cluster Analysis versus Factor Analysis, Overview of Basic Clustering Methods, Agglomerative Hierarchical Clustering, Within-Group Linkage, Nearest Neighbour or Single

Linkage, Furthest Neighbour or Complete Linkage, Centroid Clustering, Ward's Method, K-Means Algorithm, Dendrogram, Profiling of Cluster, Cluster Evaluation.

## **Unit III**

### **Support Vector Machine**

Decision Boundaries for Support Vector Machine, Maximum Margin Hyperplanes, Structural Risk Minimization, Linear SVM-Separable Case, Linear SVM-Non-Separable Case, Kernel Function, Kernel Trick, Kernel Hilbert Space, Model Evaluation.

## **Unit IV**

### **Market Basket Analysis**

Market Basket Analysis and Association Analysis, Market Basket Data, Stores, Customers, Orders, Items, Order Characteristics, Product Popularity, Tracking Marketing Interventions, Association Rules, Support, Confidence, Lift, Chi-Square Value, Sequential Pattern Analysis.

## **Unit V**

### **Introduction to Artificial Intelligence**

Current Trends in AI, Intelligent Agents, Environments, Problem Solving Agents, Searching Techniques, Knowledge and Reasoning in AI, Forms of Learning, Structure of a Neural Network, Analogy with Biological Neural Network, Activation Functions, Gradient Descent, Model Accuracy.

## **References**

1. Kevin Knight, Elaine Rich, B.Nair, 2017, *Artificial Intelligence*, McGraw.
2. Han, Jiawei and Kamber, Micheline, 2012, *Data Mining: Concepts and Techniques*, Morgan Kaufman Publishers.
3. AnandRajaraman, 2011, *Mining of Massive Datasets*, Cambridge University Press.
4. Mitchell, 2013, *Machine Learning*, McGraw Hill.
5. Stuart Russell, Peter Norvig, 2004, *Artificial Intelligence – A Modern Approach*, Pearson.