

BIG DATA ANALYTICS

III SEMESTER

| S. No | Course Code | SUBJECT TITLE |
|--------------|--------------------|---------------------------------------|
| 1 | EBDA-301 | SOCIAL & WEB ANALYTICS |
| 2 | EBDA-302 | HR ANALYTICS |
| 3 | EBDA-303 | OPERATIONS AND SUPPLY CHAIN ANALYTICS |
| 4 | EBDA-304 | MARKETING ANALYTICS – I |
| 5 | EBDA-305 | RETAIL ANALYTICS – I |

IV SEMESTER

| S. No | Course Code | SUBJECT TITLE |
|--------------|--------------------|-----------------------------------|
| 1 | EBDA-401 | BUSINESS INTELLIGENCE |
| 2 | EBDA-402 | MACHINE LEARNING |
| 3 | EBDA403 | DESIGN AND ANALYSIS OF ALGORITHMS |
| 4 | EBDA-404 | SOFTWARE PROJECT MANAGEMENT |
| 5 | EBDA-405 | DATA VISUALIZATION |

SOCIAL & WEB ANALYTICS (EBDA-301)

Course Objective: Familiarize the learners with the concept of social media analytics and understand its significance. • Familiarize the learners with the tools of social media analytics. • Enable the learners to develop skills required for analyzing the effectiveness of social media for business purposes

UNIT-I:

Introduction to Social Media Analytics – Importance, Social Media Audience, Analytics Audience Insights, Social Media Audit & Setting Benchmarks, Making Actionable Recommendations, Brand Lift and Conversion Studies, Identifying Opinions through Sentiment Analysis and Topic Modeling

UNIT-II:

Emerging Analytics: Social Analytics – Data challenge, twitter revolution, analyzing offline customer experiences, analyzing mobile customer experiences, Quantifying the impact of Twitter, Hidden web analytics traps – accuracy or precision, Dealing with data quality, Challenges in Online data mining and Predictive Analytics

UNIT-III:

Introduction to Web Analytics: Concept of web analytics, Importance and benefits of Web Analytics, Web Metrics – Visits and Visitors, Time on page and Time on site, Bounce Rate, Exit Rate, Conversion rate, Engagement, Attributes of metrics,

UNIT-IV:

Clickstream Analysis and KPI's: Understanding the web metrics of a web site, Producing web analytics report, Foundational Analytical strategies – Segmentation, Focus on Customer Behavior, Different Clickstream Analysis, Web analytics challenges

UNIT-V:

Leveraging Qualitative Data, Testing and Experimentation: Lab Usability Studies, Usability Alternatives, Surveys, Web-enabled emerging user research options, Testing – A/B Testing, Multivariate Testing, Actionable Testing ideas, Controlled Experiments, Competitive Intelligent Analysis – CI data sources, types and secrets, web traffic analysis, search and keyword analysis

Reference Books:

1. Clifton B., Advanced Web Metrics with Google Analytics, Wiley Publishing, Inc. 2nd ed. 2. Kaushik A., Web Analytics
2. The Art of Online Accountability and Science of Customer Centricity, Wiley Publishing, Inc. 1st ed.
3. Sterne J., Web Metrics: Proven methods for measuring web site success, John Wiley and Sons

HR ANALYTICS (EBDA-302)

Course Objective: Observe HR issues from an operational and a highly strategic perspective. It will enable you to comprehensively understand the fundamental principles of modern business models and facilitate the creation and implementation of future-ready HR strategies that are in sync with recruitment practices ideal for the future workforce.

UNIT-I:

Meaning of Analytics: Classification; Importance HCMs; Role and Perspective of HCMs.

UNIT-II:

The HCM Model: The Employee Value Proposition; Compensation, Attracting, Motivating and retaining employees now and in the future.

UNIT-III:

The new face of work force planning; The workforce planning; segmentation of skills, The business playbook; the contents and process of creating a playbook.

UNIT-IV:

Quality employee engagement: Employee Engagement Definition and Measurement; Engagement Drivers; Disorder and Disengagement; Behavior Based Signs of Departure, Event based Signs of Departure, Data based Signs of Departure;

UNIT-V:

Meaning of Metrics: The our Human Capital Performance Metrics; The Second Generation and Third Generation Metrics Connecting the Metrics; Predictive Analytics for Human Capital Management.

References

1. Moore, McCabe, Duckworth, and Alwan. The Practice of Business Statistics: Using Data for Decisions, Second Edition, New York: W.H.Freeman, 2008.
2. Predictive analytics for Human Resources, Jac Fitz-enz, John R. Mattox, II, Wiley, 2014.
3. Human Capital Analytics: Gene Pease Boyce Byerly, Jac Fitz-enz, Wiley, 2013.
4. The HR Scorecard: Linking People, Strategy, and Performance, by Brian E. Becker, Mark A. Huselid, Mark A Huselid, David Ulrich, 2001.
5. HR Analytics: The What, Why and How, by Tracey Smith The New HR Analytics: Predicting the Economic Value of Your Company's Human By Jac FITZ-ENZ, 2010.

OPERATIONS AND SUPPLY CHAIN ANALYTICS (EBDA-303)

Courses Objectives: In present time of intense global competition, customers are demanding more and more variety, with better quality and service at lowest cost. This means that in order to be successful, firms need to develop supply chain strategies and logistical capabilities that serve the needs of their customers whilst maximizing overall profitability. All supply chains, in order to function properly, must focus on the huge opportunity that exists in their analytics.

Unit 1:

Descriptive Analysis using R: Computing an overall summary of a variable and an entire data frame, summary() function, sapply() function, stat.desc() function, Case of missing values, Descriptive statistics by groups, Simple frequency distribution: one categorical variable, Two-way contingency table: Two categorical variables, Multiway tables: More than two categorical variables.

Unit 2:

Basic Concept in R: Data Structure, Import of Data. Graphic Concept in R: Graphic System, Graphic Parameter Settings, Margin Settings for Figures and Graphics, Multiple Charts, More Complex Assembly and Layout, Font Embedding, Output with cairo_pdf, Unicode in figures, Colour settings, R packages and functions related to visualization.

Unit 3:

Visualization of Categorical Data in R: Bar Chart Simple, Bar Chart with Multiple Response Questions, Column Chart with two-line labeling, Column chart with 45o labeling, Profile Plot, Dot Chart for 3 variables, Pie Chart and Radial Diagram, Chart Tables.

Unit 4:

Distributions: Histogram overlay, Box Plots for group, Pyramids with multiple colors, Pyramid: emphasis on the outer and inner area, Pyramid with added line, Aggregated Pyramids, Simple Lorenz curve.

Unit 5:

Shot Time Series, Areas underneath and between time series, presentation of daily, weekly and monthly values, Exceptions and Special cases in Time series, Scatter Plot for Four Quadrants differentiated by colors, Scatter Plot for Outliers Highlighted, Scatter Plot for Areas Highlighted, Exceptions and Special cases in Scatter Plot

Reference Books:

1. Data Visualization with R 100 Examples by Thomas Rahlf, Springer
2. Using R for Introductory Statistics, By John Verzani, CRC Press
2. Davis, Pecar – Business Statistics using Excel, Oxford
3. Ken Black – Business Statistics, 5th ed., Wiley India
4. Chandrasekaran & Umavathi-Statistics for Managers, 1st edition, PHI Learning
5. Big Data Visualization, James D. Miller, Packt Publishing Ltd.

MARKETING ANALYTICS (EBDA-304)

Course objective: this COURSE you will develop in- **Demand - marketing** and analytics that will help you enhance and optimise your organisation's business strategies and benefit from one of the most sought-after professions today

UNIT-I:

Introduction to Metrics: Share of Hearts Minds and Markets Market Share, Relative Market share, Market concentration, Brand Development Index, Category Development Index, Penetration, Share of requirements, Heavy Usage Index, Awareness Attitudes and usage, Customer satisfaction ,Willingness to recommend Net promoter, Willingness to search.

UNIT-II:

Product and Portfolio Management: Trial, Repeat ,Penetration and volume projections, Growth Percentage and CGR, Cannibalization Rate and Fair share draw rate, Brand equity Metrics, Conjoint utilities and consumer preference, Segmentation and conjoint utilities, Conjoint utilities and volume projections.

UNIT-III:

Margins and Profits: Margins ,Selling Price and Channel Margins Average Price per unit and price per statistical unit, Variable costs and fixed costs, Marketing spending-Total, fixed and variable, Breakeven analysis and contribution analysis ,Target Volume.

UNIT-IV:

Advertising Media and web metrics: Advertising Impressions, gross rating, points and opportunities to see, Cost per thousand impressions, Reach net reach and frequency, hare of voice, Impressions page views and hits, Rich Media Display time, Rich Media Interaction rate, Click through rates, Cost per impressions, Cost per click and cost of acquisition, Visits, Visitors and abandonment bounce rate, Friends ,followers and supporters, downloads.

UNIT-V:

Promotion: Baseline sales , Incremental sales, and promotional lift, Redemption rates for coupons/rebates, Percent sales on deal Percent time on deal and average deal depth, Pass through and price waterfall.

Reference Books

1. Stephan Sorger, —MarkStrategicng ModelsAnalyticsand Metri 2013.
2. Mark Jeffery, —Data Driven Marketing:should The knowl, Wiley, 2013.
3. Paul W. Farris, Neil T. Bendle, Phillip E. The Definitive Guide to Measuring Marketin

RETAIL ANALYTICS 1 (EBDA-305)

Objective: a retail merchant has many day to day, and long term challenges that must be managed at the same time. As IT technologies have matured, tremendous investments have been made solving both the tactical and strategic business operations problems. For the retailer recently considering the retail analytics landscape, it appears quite fragmented and difficult for to determine where to start and which solutions they should consider

Unit-1

Retail Analytics Survey This report details the market survey and recommendation for choosing a retail analytics solution with a landscape analysis and four examples of budget and benefit analysis. Research was conducted using in person interviews, internet research from the retail leader point of view using internet search engines, and a synthesis of findings within the team discussions.

Unit-2

Analysis Framework Thirty five companies were investigated for analytics capabilities useful to retail businesses and their offerings categorized according to where they contributed in the retail business model-The Tickto model has 5 categories strategy and planning, store operations, marketing, supply chain, and merchandising

Unit -3

spend level The original project scoping proposed three target spend levels for analysis \$10K, \$100K, and \$1M. In the course of this investigation it was discovered that there were some companies that were positioning “free”. Typically, this was an entry path to paid services so a Freemium category was added.- The Lokad pricing model is quite interesting in that they claim all features are included in every plan

Unit-4

Merchandising. Merchandise financial planning. Plan and manage sales, margin and inventory turns across all categories and channels. Merchandise inseason management. Set datadriven financial performance goals and match inventory to inseason demand.- Pricing Regular price optimization. Set the best price for every item you sell based on multiple factors – the competition, your goals, business rules, ad placement, etc. – at a customer, market or store level. React quickly – and correctly – to changes in the market

Unit -5

SAS At the very high end, companies such as SAS and SAP are offering solutions that cover all aspects of business (in this case retail business), not just analytics. This leads to a more integrated analytics offering that can provide a richer capability for the end customer. In return, due to complexity of the offering and wide range of capabilities, companies who are in this budget range will need more specialized workforce, such as data scientists, sophisticated IT and security infrastructure, a development team, and business analysts.

Reference Books:

1. Supply chain management by Sunil Chopra, and Peter Meindl, Pearson
2. Jeremy F. Shapiro. Modeling the Supply Chain. Duxbury Thomson Learning
3. D. Simchi-Levi, P. Kaminsky, E. Simchi-Levi, and Ravi Shankar, Designing and Managing the Supply Chain concepts, Strategies and Case studies, Third Edition, Tata McGraw Hill, New Delhi, 2008.
4. Rahul Saxena • Anand Srinivasan, Business Analytic

BUSINESS INTELLIGENCE (EBDA-401)

Course objective; this COURSE you will develop in- **BUSINESS INTELLIGENCE** and analytics that will help you enhance and optimise your organisation's business strategies and benefit from one of the most sought-after professions today.

UNIT-I:

Introduction to Business Analytics and Data warehouse: Concept of Business Analytics and Business Intelligence Systems, concept of data warehouse, characteristics of data warehouse, data warehouse architecture, data pre processing, tools for extraction, clean up and transformation

UNIT-II:

Online Analytical Processing: Concepts of OLTP and OLAP, multidimensional analysis - MOLAP, ROLAP, Data Warehouse development - Data Cubes, Fact tables, Dimension Tables, Dimension Schema, Star Schema and Snow flake Schema.

UNIT-III:

Introduction to Data Mining: Concept, KDD process, benefits of data mining, steps in data mining, data mining for business problems.

UNIT-IV:

Data Mining Tasks: Market Basket Analysis, predictive models, cluster analysis, text mining, web mining, selecting and using the right technique.

UNIT-V:

Big Data: Concept of big data, significance, business applications of big data, introduction to Apache Hadoop, business performance management - performance measurement, metrics, KPIs and business activity monitoring(BAM).

Reference Books:

1. Digital Image Processing – by Rafael.C.Gonzalez & Richard E.Woods, 3rd edition, Pearson Education, 2008
2. Fundamentals of Digital Image Processing – by A.K. Jain, PHI
3. Digital Image Processing – William K, Part I - John Wiley edition.
4. Digital Image Processing using MATLAB – by Rafael.C.Gonzalez, Richard E.Woods, & Steven 6. L.Eddins, Pearson Education, 2006

MACHINE LEARNING (EBDA-402)

OBJECTIVE : • To introduce students to the basic concepts and techniques of Machine Learning. • To have a thorough understanding of the Supervised and Unsupervised learning techniques • To study the various probability based learning techniques • To understand graphical models of machine learning algorithms

UNIT I INTRODUCTION - Learning – Types of Machine Learning – Supervised Learning – The Brain and the Neuron – Design a Learning System – Perspectives and Issues in Machine Learning – Concept Learning Task – Concept Learning as Search – Finding a Maximally Specific Hypothesis – Version Spaces and the Candidate Elimination Algorithm – Linear Discriminants – Perceptron – Linear Separability – Linear Regression.

UNIT II LINEAR MODELS - Multi-layer Perceptron – Going Forwards – Going Backwards: Back Propagation Error – Multi-layer Perceptron in Practice – Examples of using the MLP – Overview – Deriving Back-Propagation – Radial Basis Functions and Splines – Concepts – RBF Network – Curse of Dimensionality – Interpolations and Basis Functions – Support Vector Machines

UNIT III TREE AND PROBABILISTIC MODELS - Learning with Trees – Decision Trees – Constructing Decision Trees – Classification and Regression Trees – Ensemble Learning – Boosting – Bagging – Different ways to Combine Classifiers – Probability and Learning – Data into Probabilities – Basic Statistics – Gaussian Mixture Models – Nearest Neighbor Methods – Unsupervised Learning – K means Algorithms – Vector Quantization – Self Organizing Feature Map

UNIT IV DIMENSIONALITY REDUCTION AND EVOLUTIONARY MODELS - Dimensionality Reduction – Linear Discriminant Analysis – Principal Component Analysis – Factor Analysis – Independent Component Analysis – Locally Linear Embedding – Isomap – Least Squares Optimization – Evolutionary Learning – Genetic algorithms – Genetic Offspring: - Genetic Operators – Using Genetic Algorithms – Reinforcement Learning – Overview – Getting Lost Example – Markov Decision Process

UNIT V GRAPHICAL MODELS - Markov Chain Monte Carlo Methods – Sampling – Proposal Distribution – Markov Chain Monte Carlo – Graphical Models – Bayesian Networks – Markov Random Fields – Hidden Markov Models – Tracking Methods

REFERENCES:

1. Stephen Marsland, —Machine Learning – An Algorithmic Perspective, Second Edition, Chapman and Hall/CRC Machine Learning and Pattern Recognition Series, 2014.
2. Tom M Mitchell, —Machine Learning, First Edition, McGraw Hill Education, 2013.
3. Peter Flach, —Machine Learning: The Art and Science of Algorithms that Make Sense of Data, First Edition, Cambridge University Press, 2012.
4. Jason Bell, —Machine learning – Hands on for Developers and Technical Professionals, First Edition, Wiley, 2014
5. Ethem Alpaydin, —Introduction to Machine Learning 3e (Adaptive Computation and Machine Learning Series), Third Edition, MIT Press, 2014

DESIGN AND ANALYSIS OF ALGORITHMS (EBDA-403)

OBJECTIVES: • To study the various ways of analyzing algorithms • To understand the need for asymptotic notations • To understand the various algorithm design techniques • To understand string matching algorithms • To learn about NP class of problems and their variations

UNIT I ANALYSING ALGORITHMS The Role of Algorithms in Computing - Growth of Functions – Recurrences - The Substitution Method - The Recurrence Tree Method - The Master Method - Probabilistic Analysis and Randomized Algorithms – Amortized Analysis – Aggregate Analysis – Accounting Method

UNIT II- DIVIDE AND CONQUER & GREEDY DESIGN STRATEGIES 9 Analysis of Quick Sort, Merge Sort – Quick Sort Randomized Version – Sorting in Linear Time - Lower Bounds for Sorting - Selection in Expected Linear Time - Selection in Worst case Linear Time – Greedy Algorithms - Elements of Greedy Strategy - Huffman Code, Dijkstra's Shortest Path Algorithm.

UNIT III DYNAMIC PROGRAMMING AND OTHER DESIGN STRATEGIES Dynamic Programming – Matrix Chain Multiplication - Elements of Dynamic programming – Longest Common Sequences – Warshall's and Floyd's Algorithm – Transitive Closure - All Pairs Shortest Path Algorithm – Analysis – Backtracking – Graph Coloring Problem - Branch and Bound Strategy - Knapsack Problem.

UNIT IV FLOW NETWORKS AND STRING MATCHING Flow Networks – Ford Fulkerson Method - String Matching - Naive String Matching Algorithm – Knuth Morris Pratt Algorithm - Analysis.

UNIT V NP PROBLEMS NP-Completeness – Polynomial Time Verification – Theory of Reducibility - Circuit Satisfiability – NP - Completeness Proofs – NP Complete Problems: Vertex Cover, Hamiltonian Cycle and Traveling Salesman Problems – Approximation Algorithms – Approximation Algorithms to Vertex - Cover and Traveling Salesman Problems

REFERENCES:

1. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, —Introduction to Algorithms, Third Edition, Prentice Hall, 2010.
2. Ellis Horowitz, Sartaj Sahni and Sanguthevar Rajasekaran, —Fundamentals of Computer Algorithms, Second Edition, Universities Press, 2008.
3. Kenneth A. Berman and Jerome L. Paul, —Algorithms, Cengage Learning India, 2010
4. Alfred V Aho, John E Hopcroft and Jeffrey D Ullman, —The Design and Analysis of Computer Algorithms, First Edition, Pearson Education, 2006.

SOFTWARE PROJECT MANAGEMENT (EBDA-404)

Course Objective: Understand the process groups and nine knowledge areas of the Understand approaches for managing and optimizing the software development process o Understand efficient techniques for managing each phase of the systems development lifecycle o Use and application of tools to facilitate the software project management process (e.g. Microsoft Project)

UNIT-I:

Introduction to Software Project Management: Software project features, problems with software projects, need for software project management, evaluation of different projects - technical evaluation, cost-benefit analysis and evaluation techniques, risk evaluation, project selection, project planning - an overview of stepwise project planning.

UNIT-II:

Managing Human Resources: Role of project manager, building a project team, dealing with issues, project development models - Waterfall model, V- process, Spiral Model and Agile Development Model, software prototyping, incremental models, object oriented model, selecting the appropriate model.

UNIT-III:

Software Effort Estimation: Software effort estimation technique, function point analysis, COCOMO model, activity planning, project scheduling, network planning model, creating activity network, identifying critical activities, identifying critical path.

UNIT-IV:

Risk Management: Risk identification, risk assessment, risk planning, risk management, resource allocation - identifying resources requirements, scheduling resources.

UNIT-V:

Project Monitoring and Control: Evaluate progress of project, cost monitoring, project control, software quality assessment, significance of software quality, software quality metrics, quality management, software testing - introduction to testing tools.

References:

1. A Practitioner's Guide to Test Case Design by LEE Copland, Artech House Publishers, Boston - London.
2. Software Testing – A Craft's man Approach, Paul C. Jorgensen, A CRC Press LLC.
3. Software Quality Theory and Management by Alan C. Gillies, Chapman & Hall.
4. Software Quality by Galrry S. Marliss , Thomson.
5. Metrics and Models in Software Quality Engineering by Stephen H. Kan , Pearson Education.
6. Handbook of Software Quality Assurance by G. Gordon Sculmeyer, Artech House Publishers, Boston –London

DATA VISUALIZATION (EBDA-405)

Course Objective: Business has simply become more **digital**, more **global** and more **sustainable**. An international and forward-looking mindset, as well as the ability to lead people effectively through change and technological innovation are highly sought-after leadership traits.

UNIT–I: Introduction to Visualization: Concept and importance of data visualization, Choosing appropriate visual encodings – ordering of items, number of distinct values, structure of visualization, Positioning - Placement and Proximity, Graphs and Layouts, Colors, Size, Text and Typography, Shape, Lines

UNIT–II: Charts in Tableau: Introduction to Tableau, Connecting to Data Source: Text Files, Excel, Access, other databases, merging multiple data sources, Univariate Charts, Bivariate Charts, Multivariate Charts and Maps

UNIT–III: User defined fields and Customization: Using predefined fields, calculating percentages, applying if-then logic, applying logical functions, showing totals and percentages, discretizing data, manipulating text, aggregate data, Customization in Tableau

UNIT–IV: Data Visualization with Power BI: Introduction to Power BI, Primary tools of Power BI, Reports in BI, Charts in BI, Slicers, Map Visualizations

UNIT–V: Dashboards and Customization with Power BI: Dashboard Vs reports, Creating a dashboard, Dashboard Tiles, Pinning Tiles, Custom Visualization

Reference Books:

1. Software design, David Budgen, second edition, Pearson education, 2003.
2. Software Engineering: A practitioner's Approach, Roger S Pressman, seventh edition Mc-Graw Hill International Edition, 2009.
3. Software Engineering, Ian Sommerville, seventh edition, Pearson education, 2004.
4. Software Project Management, Bob Hughes & Mike Cotterell, fourth edition, Tata Mc-Graw Hill, 2006
5. The art of Project management, Scott Berkun, O'Reilly, 2005.