DR. D Y PATIL VIDYAPEETH

PUNE – 411018

CENTRE FOR ONLINE LEARNING

PIMPRI, PUNE

SYLLABUS FOR

MASTER OF BUSINESS ADMINISTRATION (M.B.A)

Academic Year 2025

Semester-III Artificial Intelligence & Machine Learning Specialization

Semester	3	**Course Credits**	**4**	**Specialization**	Artificial Intelligence & Machine Learning
Course Code	OMBAIML301			**Type**	Specialization Course
Course Title	Basics of Artificial Intelligence & Machine Learning				

Course Description:

To develop semantic-based and context-aware systems to acquire, organize process, share and use the knowledge embedded in multimedia content. The research will aim to maximize automation of the complete knowledge lifecycle and achieve semantic interoperability between Web resources and services. Artificial Intelligence & Machine learning are pillars on which you can build intelligent applications. Artificial Intelligence and Machine Learning Fundamentals begin by introducing you to Python and discussing AI search algorithms.

Course Objectives:

1) Acquire advanced Data Analysis skills.
2) Stay Industry-relevant and grow in your career.
3) Create AI/ML solutions for various business problems.
4) Build and deploy production-grade AI/ML applications.
5) Apply AI/ML methods, techniques, and tools immediately.

Course Outlines:

Unit 1: Introduction to Data Science and AI & ML: Data Science, AI & ML, Use Cases in Business and Scope, Scientific Method, Modelling Concepts, CRISP-DM Method.

Unit 2: R Essentials Programming: Commands and Syntax, Packages and Libraries, Introduction to Data Types, Data Structures in R - Vectors, Matrices, Arrays, Lists, Factors, Data Frames, Importing and Exporting Data, Control structures, and Functions.

Descriptive Statistics: Data exploration (histograms, bar chart, box plot, line graph, scatter plot), Qualitative and Quantitative Data, Measure of Central Tendency (Mean, Median and Mode), Measure of Positions (Quartiles, Deciles, Percentiles, and Quantiles), Measure of Dispersion (Range, Median, Absolute deviation about median, Variance and Standard deviation), Anscombe's quartet, Other Measures: Quartile and

Percentile, Interquartile Range.

Unit 3: Statistical Analysis Initial Data Analysis: Relationship between attributes: Covariance, Correlation Coefficient, Chi-Square, Measure of Distribution (Skewness and Kurtosis), Box and Whisker Plot (Box Plot and its parts, Using Box Plots to compare distribution) and other statistical graphs.

Probability: Probability (Joint, marginal, and conditional probabilities), Probability distributions (Continuous and Discrete), Density Functions, and Cumulative functions.

Unit 4: Data Acquisition: Gather information from different sources, Internal systems, External systems, Web APIs, Open Data Sources, Data APIs, Web Scraping, and Relational Database access (queries) to process/access data.

Unit 5: Data Pre-processing and Preparation: Data Mining, Wrangling, Plyr packages, Cast/Melt

Unit 6: Data Quality and Transformation: Data imputation, Data Transformation (min-max, log transform, z-score transform, etc.,), Binning, Classing and Standardization, Outlier/Noise & Anomalies.

Unit 7: Handling Text Data: Bag-of-words, Regular Expressions, Sentence Splitting and Tokenization, Punctuations and Stop words, Incorrect spellings, Properties of words and Word cloud, Lemmatization and Term-Document TxD computation, Sentiment Analysis (Case Study).

Unit 8: Principles of Big Data: Introduction to Big Data, Challenges of processing Big Data (Volume, Velocity and Variety perspective), Use Cases.

Unit 9: Big Data Frameworks – Hadoop, Spark, and NoSQL: Processing, Storage and Programming Framework, Hadoop eco-system components and their functions, Essential Algorithms (Word count, Page Rank, IT-IDF), Spark: RDDs, Streaming and Spark ML, NoSQL concepts (CAP, ACID, NoSQL types).

Unit 10: Data Visualization: Science of Visualization, Visualization Periodic Table, Aesthetics and Storytelling, Concepts of measurement - scales of measurement, Design of data collection formats with illustration, Principles of data visualization - different methods of presenting data in business analytics, Concepts of Size, Shape, Color, Various Visualization types, Bubble charts, Geo-maps (Chlorpeths), Gauge charts, Tree map, Heat map, Motion charts, Force Directed Charts etc.

Unit 11: Sampling and Estimation: Sample versus population, Sample techniques (simple, stratified, clustered, random), Sampling Distributions, Parameter Estimation, Unbalanced data treatment.

Unit 12: Inferential Statistics: Develop an intuition of how to understand the data, attributes, distributions, Procedure for statistical testing, etc., Test of Hypothesis (Concept of Hypothesis testing, Null Hypothesis

and Alternative Hypothesis), Cross Tabulations (Contingency table and their use, Chi-Square test, Fisher's exact test), One Sample t test (Concept, Assumptions, Hypothesis, Verification of assumptions, Performing the test and interpretation of results), Independent Samples t test, Paired Samples t test, One way ANOVA (Post hoc tests: Fisher's LSD, Tukey's HSD), z-test and F-test.

Course Outcomes:

On successful completion of the course the learner will be able to:

CO#	Cognitive Abilities	Outcomes
CO301.1	**Remember**	The Basic Concept of Artificial Intelligence & machine Leering
CO301.2	**Apply**	Different Methods of AI and ML
CO301.3	**Analyze**	Analyze the recent changes at the Application Level
CO301.4	**Evaluate**	Evaluate how AI & ML Helps Business to Grow

Suggested Reading:

1. Artificial Intelligence: A Modern Approach- Textbook by Peter Norvig and Stuart J. Russell
2. An Introduction to Statistical Learning: With Applications in R - Gareth M. James, Daniela Witten, Trevor Hastie, Robert Tibshirani
3. Machine Learning: An Artificial Intelligence Approach: Jaime G. Carbonell, Ryszard S. Michalski, Tom M. Mitchell (auth.), Ryszard S. Michalski, Jaime G. Carbonell, Tom M. Mitchell (eds.)

Semester	3	**Course Credits**	**4**	**Specialization**	Artificial Intelligence & Machine Learning
Course Code	OMBAIML302			**Type**	Specialization Course
Course Title	Machine Learning				

Course Description:

Machine Learning is a key to develop intelligent systems and analyze data in science and engineering. Machine learning engines enable intelligent technologies such as Siri, Kinect, or Google self-driving car, to name a few. At the same time, machine learning methods help unlock the information in our DNA and make sense of the flood of information gathered on the web, forming the basis of a new Science of Data. This course provides an introduction to the fundamental methods at the core of modern machine learning. It covers theoretical foundations as well as essential algorithms for supervised and unsupervised learning.

Course Objectives:

1) To have a good understanding of the fundamental issues and challenges of machine learning: data, model selection, model complexity, etc.
2) To be able to design and implement various machine learning algorithms in a view of real-world applications.
3) To describe the Machine Learning approach and its relationship with Data Science.
4) To identify applications of Machine Learning.
5) Have an understanding of the strengths and weaknesses of many popular machine learning approaches.

Course Outlines:

Unit-1 Importance of Machine Learning: What is Machine Learning?, Machine Learning Concept with Definition, why Machine Learning?, Domain and Applications of Machine Learning, Types of Machine Learning, Future of Machine Learning, Machine Learning Support in Python.

Unit-2 Data Pre-Processing: Need of Data Pre-Processing, Steps for Data Pre-Processing in Machine Learning, Techniques of Data Pre-Processing for Machine Learning Using Python: Importing The Dataset, Missing Data, Categorical Data, Splitting the Dataset Into the Training Set and Test Set, Feature Scaling. Creating Our Own Dataset, Why and How to Do Feature Scaling in Machine Learning.

Unit-3 Regression: Simple Linear Regression, Simple Linear Regression Intuition, Simple Linear Regression in Python, Multiple Linear Regression, Multiple Linear Regression in Python, Polynomial Regression Intuition, Polynomial Regression in Python, SVR, SVR Intuition, SVR in Python.

Unit-4 Classification: Logistic Regression, Logistic Regression Intuition, Logistic Regression in Python, K-Nearest Neighbors (K-NN), K-NN Intuition, K-NN in Python, Support Vector Machine (SVM), SVM Intuition, SVM in Python, Kernel SVM, Kernel SVM Intuition, Kernel SVM in Python, Naïve Bayes, Naïve Bayes in Python, Decision Tree Classification, Decision Tree Classification Intuition, Decision Tree Classification in Python, Random Forest Classification, Random Forest Classification Intuition, Random Forest Regression in Python, Evaluating Classification Models Performance.

Unit-5 Clustering: K-Means Clustering, K-Means Clustering Intuition, K-Means Clustering in Python, Hierarchical Clustering, Hierarchical Clustering Intuition, Hierarchical Clustering in Python, Model-Based Clustering, Model-Based Clustering Intuition, Model-Based Clustering in Python, Density-Based Clustering, Density-Based Clustering Intuition, Density-Based Clustering in Python, Fuzzy Clustering, Fuzzy Clustering Intuition, Fuzzy Clustering in Python, Evaluating Clustering Models Performance.

Unit-6 Association Rule Learning: Apriori Algorithm, Apriori Algorithm Intuition, Apriori Algorithm in Python, Eclat Algorithm, Eclat Algorithm Intuition, Eclat Algorithm in Python.

Unit-7 Reinforcement Learning: Important terms used in Deep Reinforcement Learning, Working of Reinforcement Learning, Reinforcement Learning Algorithms, Characteristics of Reinforcement Learning, Types of Reinforcement Learning, Learning Models of Reinforcement, Q-learning Intuition, Q-learning in Python, Reinforcement Learning vs. Supervised Learning, Applications of Reinforcement Learning, why use Reinforcement Learning? And when not to use Reinforcement Learning? Challenges of Reinforcement Learning, Upper Confidence Bound (UCB), UCB Intuition, UCB in Python, Thompson Sampling, Thompson Sampling Intuition, Thompson Sampling in Python.

Unit-8 Deep Learning: Artificial Neural Networks, Artificial Neural Networks Intuition, Artificial Neural Networks in Python, Convolutional Neural Networks, Convolutional Neural Networks Intuition, Convolutional Neural Networks in Python

Unit-9 Dimensionality Reduction: Principal Component Analysis (PCA), PCA Intuition, PCA in Python, Linear Discriminant Analysis (LDA), LDA Intuition, LDA in Python, Kernel PCA (KPCA), Kernel PCA Intuition, Kernel PCA in Python.

Unit-10 Model Selection & Boosting: k-Fold Cross-Validation, k-Fold Cross Validation Intuition, k-Fold Cross-Validation in Python, Grid Search, Grid Search in Python, XGBoost, XGBoost Intuition, XGBoost in Python.

Course Outcomes:

On successful completion of the course the learner will be able to:

CO#	Cognitive Abilities	Outcomes
CO 302.1	**Remember**	Learn the Basics of Machine Learning
CO 302.2	**Understand**	Understand the various terms like Data Pre-Processing, Supervised, Unsupervised Learning, Classification, Regression, Dimensionality Reduction, etc.
CO 302.3	**Analyze**	Analyze and evaluate the various Machine Learning Algorithms
CO 302.4	**Apply**	Design and develop a Machine Learning Algorithms for an organization.

Suggested Reading:

1. Mastering Machine Learning with Python in Six Steps

A Practical Implementation Guide to Predictive Data Analytics Using Python

Manohar Swamynathan

Publication: Apress

2. Machine Learning

Step-by-Step Guide to Implement

Machine Learning Algorithms with Python

Author Rudolph Russell

3. Data Classification Algorithms and Applications

Edited by Charu C. Aggarwal

IBM T. J. Watson Research Center

Yorktown Heights, New York, USA

Semester	3	**Course Credits**	**4**	**Specialization**	Artificial Intelligence & Machine Learning
Course Code	OMBAIML303			**Type**	Specialization Course
Course Title	Performing Analytics with Python				

Course Description:

This course will take you from the basics of Python to exploring many different types of data. This course is designed to teach students how to analyze different types of data using Python. Learner's will know about how to prepare data for analysis, perform simple statistical analysis, create meaningful data visualizations, predict future trends from data, and more.

Course Objectives:

1) To understand the basic concepts of Python & its libraries
2) To understand Data Analytics using Python
3) To understand and use Python data science libraries as a tool for data analytics
4) To understand Data Modelling techniques.

Course Outlines:

Unit-1 - Python Fundamentals for Data Analytics: Python data structures, data types, tuples, operators, Event-driven programming, local and global variables, buttons and input fields - The canvas, static drawing, timers, interactive drawing -Lists, keyboard input, motion, positional/velocity control - Mouse input, more lists, dictionaries. Control statements, Functions, Object Oriented programming concepts using classes, objects and methods, Exception handling, Implementation of user-defined Modules and Package, File handling in python.

Unit-2 - Python Advance Concepts: Images - Classes, tiled images - Acceleration and friction, spaceship class, sprite class, sound - Sets, groups of sprites, collisions, sprite animation - Fundamentals of Data Manipulation with Python, RDBMS, UML.

Unit-3 - Python Libraries: Learn Numpy - a common scientific computation library, Getting familiar with Pandas for structural data processing, Learn the common statistical data analysis techniques in python, Getting familiar with simple linear regression.

Unit -4 : Introduction to Data Understanding and Preprocessing: Knowledge domains of Data Analysis, Understanding structured and unstructured data, Data Analysis process, Dataset generation,

Importing Dataset: Importing and Exporting Data, Basic Insights from Datasets, Cleaning and Preparing the Data: Identify and Handle Missing Values.

Unit-5: Data Processing and Visualization: Data Formatting, Exploratory Data Analysis, Filtering and hierarchical indexing using Pandas. Data Visualization: Basic Visualization Tools, Specialized Visualization Tools, Seaborn Creating and Plotting Maps.

Unit-6 : Mathematical and Scientific applications for Data Analysis: Numpy and Scipy Package, Understanding and creating N-dimensional arrays, Basic indexing and slicing, Boolean indexing, Fancy indexing, Universal functions, Data processing using arrays, File input and output with arrays.

Unit 7: Analysing Web Data: Data wrangling,Web scrapping, Combing and merging data sets, Reshaping and pivoting, Data transformation, String Manipulation, case study for web scrapping.

Unit-8 : Analytics part - I: Basic nomenclature - Analytics process model - Analytics model requirements - Types of data sources – Sampling - types of data elements - Visual Data Exploration and Exploratory Statistical Analysis - Missing Values - Outlier Detection and Treatment - Standardizing Data – Categorization - weights of evidence coding - Variable selection – Segmentation.

Unit-9: Analytics part - II: Predictive Analytics: Target Definition - Linear Regression - Logistic Regression - Decision Trees - Neural Networks - Support Vector machines - Ensemble Methods - Multiclass Classification Techniques - Evaluating Predictive Models.

Unit-10: - Social Networking: Social Network Analytics: Social Network Definitions - Social Network Metrics - Social Network Learning -Relational Neighbor Classifier - Probabilistic Relational Neighbor Classifier -Relational logistic Regression - Collective Inference. Benchmarking - Data Quality - Software – Privacy - Model Design and Documentation - Corporate Governance. Example applications: Credit Risk Modeling - Fraud Detection - Recommender Systems - Web Analytics.

Unit-11 Value Addition: Web Analytics, Introduction to Machine Learning, Applications of Machine Learning.

Course Outcomes:

On successful completion of the course the learner will be able to:

CO#	Cognitive Abilities	Outcomes
CO303.1	**Remember**	Learn the Basics of Python
CO303.2	**Understand**	Understand the Data Analytics using Python

CO303.3	**Analyze**	Analyze and implement various algorithms using Python
CO303.4	**Apply**	Design and develop an algorithm to solve business problems.

Suggested Reading:

1. Python Data Analytics

With Pandas, NumPy, and Matplotlib

Second Edition

Fabio Nelli

Publication: Apress

2. DATA ANALYSIS FROM SCRATCH WITH PYTHON

Step By Step Guide

Peters Morgan

3. SOCIAL MEDIA ANALYTICS STRATEGY

USING DATA TO OPTIMIZE BUSINESS PERFORMANCE

Alex Gonçalves

Publication: Apress

Semester	3	**Course Credits**	**4**	**Specialization**	Artificial Intelligence & Machine Learning
Course Code	OMBAIML304			**Type**	Specialization Course
Course Title	Statistics & Quantitative Techniques (SQT)				

Course Description:

Resources are always critical in any organization. They are unavailable in an unlimited manner and there are always constraints. Operation research is helpful in the situation of such constraint of resources. Managers have to manage limited available resources in such a way that neither production nor other activities get disturbed in the business. Facility design is a fascinating area for OR. The excitement of operation research lays in the application of quantitative techniques to real-world problems.

Course Objectives:

1) To impart knowledge in concepts and tools of OR and QT; and
2) To help students apply these tools in managerial decision-making.

Course Outlines:

Unit 1: Arranging Data to Convey Meaning: Presenting Data in Tables & Charts: Application Areas for Statistics, Statistical Methods, Understand Data, Organize and Classify Data, Graphical Representation of Data, Good & Bad Data Presentation.

Unit 2: Measures of Central Tendency – Mean, Median, Mode: Numerical Data Properties, Frequency, and Frequency Table, Summary Measures– Central Tendency.

Unit 3: Measures of Dispersion: Summary Measures– Variation.

Unit 4: Correlation Analysis: Correlation Analysis– Scatter Plots, Some Misconceptions About Correlation, Correlation Terminologies.

Unit 5: Simple And Multiple Regressions: Regression Analysis, Simple Regression, Multiple Regressions.

Unit 6: Association of Attributes: Notations, Classes and Class Frequencies, Relationship Between the Class Frequencies, Consistency of the Data, Independence of Attributes, Association of Attributes, Yules' Co-efficient of Association.

Unit 7: Probability & Probability Distribution: Notation and Terminology from Set Theory, Addition Theory of Probability, Conditional Probability, Multiplication Theory of Probability, Applications of Bayes' Theorem, Binomial Distribution, Poisson Distribution, Normal Distribution.

Unit 8: Linear Programming: Formulation & Graphical Solutions to LPP: Variables, Constraints, Objective, Phases of an Operations Research Project, Linear Programming– Formulation Graphical Solutions to LPP.

Unit 9: Transportation: Mathematical Formulation of Transportation Problem, North-West Corner Rule, Lowest Cost Entry Method, Vogel's Approximation Method, Test for Optimization.

Unit 10: Assignment Problems: Mathematical Statement of Assignment Problem, Solution Method for Assignment Problem, Travelling Salesman Problem.

Unit 11: Queuing Theory: Single Server & Multi-Server: Analysing Queuing Process, Constituents of Queuing System, Service Facility, Queuing Discipline, Kendall Notations, Single Server Models, Multi-Server Models

Unit 12: Markov Chain: Monte Carle Simulation: Simulation Procedure, Application of Simulation.

Unit 13: Games Theory: Zero-Sum Games, Fundamental Principles of Game Theory, Reducing by Dominance, Saddle Point, Strictly Determined Game, Mixing Strategies, Flow of Solution, Assumptions for Games Theory.

Unit 14: Decision Theory: Criteria for Decision Making: Decision Tables, Decision Making Process, Decision Criteria for Certainty, Decision Criteria for Uncertainty [5 Criteria], Decision Criteria for Risk.

Course Outcomes:

On successful completion of the course the learner will be able to:

CO#	Cognitive Abilities	Outcomes
CO 304.1	**Remember**	Learn the basic concepts of operational research.
CO 304.2	**Apply**	Easily apply the tools in managerial decision-making.
CO 304.3	**Analyze**	Grab the opportunities in operation as a career.
CO 304.4	**Creating**	Easily do operational research for better growth.

Suggested Reading:

1. Taylor III. Bernard W., Introduction to Management Science, Dorling Kindersley (India) Pvt. Ltd., licenses of Pearson Education in South Asia, 9th Edition, 2008.

2. Vohra N. D., Quantitative Techniques in Management, Tata McGraw-Hill Publishing Co. Ltd., New Delhi, 3rd Edition, 2007.

Semester	3	**Course Credits**	**4**	**Specialization**	Artificial Intelligence & Machine Learning
Course Code	OMBAIML305			**Type**	Specialization Course
Course Title	Business Intelligence				

Course Description:

Business Intelligence (BI) refers to technologies, applications, and practices for the collection, integration, analysis, and presentation of business information. The purpose of business intelligence is to support better business decision-making. This course provides an overview of the technology of BI and the application of BI to an organization's strategies and goals.

Course Objectives:

1) Introduce the concepts and components of Business Intelligence (BI)
2) Evaluate the technologies that make up BI (data warehousing, OLAP)
3) Define how BI will help an organization and whether it will help yours
4) Identify the technological architecture that makes up BI systems
5) Plan the implementation of a BI system

Course Outline:

Unit 1: Business Intelligence: Introduction, Meaning, Purpose, and Structure of Business Intelligence Systems. Understanding Multidimensional Analysis Concepts: Attributes, Hierarchies, and Dimensions in Data Analysis. Understanding Dimensional Data Warehouse: Fact Table, Dimension Tables, Surrogate Keys, and alternative Table Structure. What is multi-dimension OLAP?

Unit 2: On-Line Analytical Processing (OLAP): What Is OLAP? (OLAP and OLTP, Operational Data Stores, Variations in Data and Approach), OLAP Applications and Functionality, Multi-Dimensions (Thinking in More Than Two Dimensions, What Are the Possibilities? Drilling and Pivoting), OLAP Architecture (Cubism, Tools, ROLAP, MOLAP, HOLAP).

Unit 3: Creating your First Business Intelligence Project: Creating Data source, Creating Data view. Modifying the Data view. Creating Dimensions, Time, and Modifying dimensions. Parent-Child Dimension.

Unit 4: Data Mining: Meaning and purpose. Creating data for data mining. Data mining model creation. Selecting data mining algorithm. Understanding data mining tools. Mapping Mining Structure to Source Data columns. Using Cube Sources. Configuring Algorithm parameters.

Unit 5: The Spectrum of Business Intelligence: Enterprise and Departmental Business Intelligence, Strategic and Tactical Business Intelligence, Power and Usability in Business Intelligence, Finding the Right Spot on the Continuum, Business Intelligence: Art or Science?

Unit 6: Business Intelligence User Interfaces: Querying and Reporting, Reporting and Querying Toolkits, Basic Approaches (Building Ad-Hoc Queries, Building On-Demand Self-Service Reports, Enhancing and Modifying), Data Access (Pull-Oriented Data Access, Push-Oriented Data Access), Dashboards (EIS Is the Engine, Metric System and KPIs, Business Intelligence Dashboards).

Unit 7: Visualization: The Basics, Unconstrained Views, Guided Analysis (The Business Intelligence Two-Step), Handling Unstructured Data, Identifying Photographs.

Unit 8: Building the BI Project: Planning the BI project, Project Resources; Project Tasks, Risk Management and Mitigation, Cost-justifying BI solutions and measuring success, Collecting User Requirements, Requirements-Gathering Techniques; Prioritizing & Validating BI Requirements, Changing Requirements; BI Design and Development, Best Practices for BI Design; Post-Implementation Evaluations, Maintaining Your BI Environment.

Unit 9: Reporting authoring: Building reports with relational vs Multidimensional data models; Types of Reports – List, crosstabs, Statistics, Chart, Map, financial, etc; Data Grouping & Sorting, Filtering Reports, Adding Calculations to Reports, Conditional formatting, Adding Summary Lines to Reports. Drill up, drill-down, drill-through capabilities. Run or schedule reports, different output forms – PDF, Excel, CSV, XML, etc.

Unit 10: BI Deployment, Administration & Security: Centralized Versus Decentralized Architecture, BI Architecture Alternatives, phased & incremental BI roadmap, System Sizing, Measurements and Dependencies, System Sizing, Measurements, and Dependencies. Setting Early Expectations and Measuring the Results. End-User Provisos. OLAP Implementations. Expanding BI Authentication Authorization, Access Permissions, Groups and Roles, Single-sign on Server Administration, Manage Status & Monitoring, Audit, Mail server & Portal integration, Back-Up and Restore.

Course Outcome:

On successful completion of the course the learner will be able to:

CO#	Cognitive Abilities	Outcomes
CO305.1	**Remember**	Different Components Business Intelligence
CO305.2	**Apply**	Business User Interface to Different Modules of the Business
CO305.3	**Analyze**	Plan the implementation of a BI system
CO305.4	**Creating**	Evaluate the technologies that make up BI (data warehousing, OLAP)

Suggested Reading:

1. Artificial Intelligence for Business Optimization: Research and Applications - Bhuvan Unhelkar, Tad Gonsalves.
2. Business Intelligence_ A Managerial Approach- [Pearson Education [distributor], Prentice-Hall (2010_2011) - Turban, Efraim_Sharda, Ramesh_Delen, Dursun_King, David
3. Business intelligence and analytics: systems for decision support - Turban, Efraim
4. Business intelligence guidebook: from data integration to analytics - Sherman, Rick - Elsevier Science; Morgan Kaufman, Year: 2015;2014

Semester	3	**Course Credits**	**4**	**Specialization**	Artificial Intelligence & Machine Learning
Course Code	OMBAIML306			**Type**	Specialization Course
Course Title	R Programming for Data Analysis and Visualization				

Course Description:

The dependency of Business firms on data and information systems is continuously growing, a lot of raw data is generated on daily basis, which is meaningless unless analyzed and interpreted by using some sophisticated techniques. Data visualization provides a good, organized pictorial representation of the data which makes it easier to understand, observe, and analyze. In this course, learners will learn how to visualize data using R and its libraries to facilitate decision-making. Data visualization allows managers to gain insight into their vast amounts of data. It benefits them to recognize new patterns and errors in the data.

Course Objectives:

1) To understand the importance of data analysis and visualization techniques for managers
2) To learn to analyze the descriptive statistics using R
3) To learn to test hypotheses to interpret data by using R package
4) To Understand what plots are suitable for a type of data managers encounter with while making decisions.
5) To Visualize data by creating various graphs using R base package, lattice, and ggplot2 packages.

Course Outline:

Unit 1: Preliminaries and Introduction to R: What is R, Installing R and R studio, R studio overview, Working in the console, Arithmetic Operators, Logical Operations, Using Functions, Getting Help in R and Quitting RStudio

Types of variables, Creating variables, Using Variables, Numeric, character, and Logical variables and operators, Sorting Numeric, Character, and Factor Vectors, Special Values, Using the console, Creating an object in R, Data types in R - Integers and doubles, Objects and Data Types, Coercion rules in R, Functions in R, Functions, and arguments, Objects and Functions, Packages in R,

What is a Vector? create vectors, Naming a vector in R, Vector recycling, Vectorized operations, Slicing and indexing a vector in R, Extracting elements from a vector, Using the [] brackets, The power of vectorized operations, Changing the dimensions of an object in R.

Unit 2: Matrices: Matrices, Creating Matrices in R, Naming Dimensions, Colnames() and Rownames(), Matrix Operations, Indexing an element from a matrix, Slicing a matrix in R, Matrix operations in R, Visualizing With Matplot(), Subsetting, Visualizing Subsets Creating a factor in R, Factors in R

Unit 3: Fundamentals of Programming in R: Relational operators in R, Logical operators in R, Vectors and logical operators, If, else, else if statements in R, If, else, else if statements - Keep-In-Mind's, For loops in R, While loops in R, Repeat loops in R, Boolean logical operators, Building a function in R 2.0, Building a function in R 2.0 – Scoping, Creating functions, Calling functions, Exercise Scoping.

Unit 4: R packages and scripts: Installing and loading packages, setting up your working directory, Downloading and importing data, working with missing data, Extracting a subset of a data frame, Writing R scripts, Adding comments and documentation, Creating reports

Unit 5: Data Preparation: What are factors, gsub() and sub(), Dealing with Missing Data, What is an NA?, Way To Locate Missing Data, Data Filters: which() for Non-Missing Data, Data Filters: is.na() for Missing Data, Removing records with missing data, Reseting the dataframe index, Replacing Missing Data: Factual Analysis Method, Replacing Missing Data: Median Imputation Method (Part 1), Replacing Missing Data: Deriving Values Method, Visualizing results.

Unit 6: Data frames: Creating a data frame in R, Importing data into R, The Tidyverse package, Exploring your dataset, Using the $ sign, Basic operations with a Data Frame, Filtering a Data Frame, Introduction to qplot, Visualizing with Qplot, Building Dataframes, Merging Data Frames, Messy Data, Renaming Columns (Variable Names), Attaching / Detaching, Tabulating Data: Constructing Simple Frequency Tables, Ordering Factor Variables

Unit 7: Lists in R: Import Data Into R, Handling Date-Times in R, what is list, Naming components of a list, Extracting components lists: [] vs [[]] vs $, Adding and deleting components, Subsetting a list, Creating A Timeseries Plot, Apply Family of Functions, Import Data into R.

Unit 8: What is Apply Family: Using apply(), Recreating the apply function with loops, Using lapply(), Combining lapply() with [], Adding your own functions, Using sapply(), Nesting apply() functions, which.max() and which.min:

Unit 9: Advanced Visualization using R with GGplot2: Using the ggplot2package to visualize data, Grammar of Graphics- GGplot2, what is a Factor? Aesthetics, Plotting With Layers, Overriding Aesthetics,

Mapping vs Setting, Line chart, Bar charts, Histograms, and Density Charts, Pie chart Stacked Area Chart, Scatter plot and Trendline, Starting Layer Tips, Statistical Transformations, Using Facets, Coordinates, Perfecting by Adding Themes, Applying themes from ggthemes to refine and customize charts and graphs, Building data graphics for dynamic reporting

Unit 10: Data querying: SQL and R: Writing SQL statements in R, Using the Select, From, Where, Is, Like, Order By, Limit, Max, Min SQL functions, Querying the database with the dplyr, filter(), select()

Unit 11: Statistical data analysis using R: Measures of central Tendency (Mean, Median, Mode) using R, Measurement of Variation - Range, IQR and Standard Deviation (Using R), Descriptive Statistics Using psych Package, Central Limit Theorem

Unit 12: Hypotheses Testing using R: One Sample z Test Using R, One-Sample z Test using BSDA Package, One-Sample t-Test Using R, Visualizing One-Sample t-Test Results using Visualize Package, One-Sample Variance Test Using Envstats Package, Two-Sample Z Test Using R, Visualizing Two Sample Z Test Using Visualize Package, Two-Sample t-Test (Equal Variance) Using R, Two-Sample t-Test (Unequal Variance) Using R, Paired t-Test Using R, Two-Sample Variance Test (F Distribution) Using R, Visualizing Two Sample Variance Test Results using Visualize Package, Analysis of Variance (ANOVA) Using R, Goodness of Fit Test Using R, Contingency Table Using R

Unit 13: Introduction to Plotly, lattice, Esquisse, and colourpicker packages: Draw multiple-axes, scatter plots, line plots, histograms, contour plots, heatmaps, network graphs, 3D charts, and time series using plotly and lattice, drag-drop using Esquisse, exporting to visualization powerpoint and word documents, changing color themes using colourpicker.

Course Outcome:

On successful completion of the course the learner will be able to:

CO#	Cognitive Abilities	Outcomes
CO 306.1	**Understand**	The Use of R Programming for Data Visualization
CO 306.2	**Apply**	Apply Different Operations on Data Frames
CO 306.3	**Analyze**	The Efficiency of Tools for Data Visualization
CO 306.4	**Evaluate**	Evaluate the Different Outcomes of the R Programming for a Data Set

Suggested Reading:

1. Artificial Intelligence for Business Optimization: Research and Applications - Bhuvan Unhelkar, Tad Gonsalves.

2. Business Intelligence_ A Managerial Approach- [Pearson Education [distributor], Prentice-Hall (2010_2011) - Turban, Efraim_Sharda, Ramesh_Delen, Dursun_King, David
3. Business intelligence and analytics: systems for decision support - Turban, Efraim
4. Business intelligence guidebook: from data integration to analytics - Sherman, Rick - Elsevier Science; Morgan Kaufman, Year: 2015;2014

Semester-IV Artificial Intelligence & Machine Learning Specialization

Semester	4	**Course Credits**	**4**	**Specialization**	Artificial Intelligence & Machine Learning
Course Code	OMBAIML401			**Type**	Specialization Course
Course Title	Application of AI in Business				

Course Description:

Artificial intelligence has a wide range of uses in businesses, including streamlining job processes and aggregating business data. Researchers aren't exactly sure what artificial intelligence means for the future of business, specifically as it relates to blue-collar jobs. AI is expected to take digital technology out of the two-dimensional screen and bring it into the three-dimensional physical environment surrounding an individual. "Artificial intelligence" is a broad term that refers to any type of computer software that engages in humanlike activities – including learning, planning, and problem-solving. Calling specific applications "artificial intelligence" is like calling a car a "vehicle" – it's technically correct, but it doesn't cover any of the specifics.

Course Objectives:

1) Understand how AI integrated with Business.
2) Learn IOT for Business.
3) Understand How AI Framework and Its Application works in real Business Word.
4) Learn about the different AI-Based Business Software.

Course Outline:

Unit 1 – Introduction to AI Framework: Defining AI, History of AI, impact of AI on jobs, a technology overview, role of Big Data, Cloud AI, Fourth industrial revolution, What Is Machine Learning? barriers to AI, AI Advantages and Disadvantages, AI cases studies.

Unit 2 – AI Impact on Digital Marketing: AI Application for Digital Marketing, Microsoft Translator, Google Voice Typing and recognition, Search Engine Searching Algorithm, Ai Base Search Engine & Digital Marketing Tools, SEO Optimization.

Unit 3 - IoT (Internet of Things): IoT Architecture and protocols, Various Platforms for IoT, Real-time Examples of IoT, Overview of IoT components and IoT Communication Technologies, Challenges in IoT.

Unit 4 - Cloud Platforms for IoT: Virtualization concepts and Cloud Architecture, Cloud computing, benefits Cloud services (SaaS, PaaS, IaaS), Cloud providers & offerings, Study of IoT Cloud platforms, ThingSpeak API and MQTT, Interfacing ESP8266 with Web services.

Unit 5 - Business Intelligence Technology Counterparts: Data Warehousing (What Is a Data Warehouse?, Data Marts and Analytical Data, Organization of the Data Warehouse), Enterprise Resource Planning, Distributing the Enterprise (First ERP, then Business Intelligence, The Current State of Affairs), Customer Relationship Management(CRM, ERP, and Business Intelligence, Customer Decisions, Decisions About Customers), Business Intelligence and Financial Information.

Unit 6 - The Spectrum of Business Intelligence: Enterprise and Departmental Business Intelligence, Strategic and Tactical Business Intelligence, Power and Usability in Business Intelligence, Finding the Right Spot on the Continuum, Business Intelligence: Art or Science?.

Unit 7 -Artificial Intelligence and Expert Systems: Concepts and Definitions of Artificial Intelligence, Artificial Intelligence Versus Natural Intelligence, Basic Concepts of Expert Systems, Applications of Expert Systems, Structure of Expert Systems, Knowledge Engineering, Development of Expert Systems.

Unit 8 - Business intelligence Applications: Marketing models: Relational marketing, Sales force management, Logistic and production models: Supply chain optimization, Optimization models for logistics planning, Revenue management systems.

Data envelopment analysis: Efficiency measures, Efficient frontier, The CCR model, Identification of good operating practices

Unit 9 - How AI Transforms Business: Machine Learning, Cyber Security, Customer relationship management, Internet and Data Research, Digital personal assistants, HR and Recruitment, Marketing and sales management, and Optimizing finance and operations.

Unit 10 - AI Software Tools for Business: SAS Business Analytics (SAS BA), QlikView, Board, Splunk, Sisense, Microstrategy, KNIME, Dundas BI, TIBCO Spotfire, Tableau Big Data Analytics, SAp Business Object, LOOKER.

Course Outcome:

On successful completion of the course the learner will be able to:

CO#	Cognitive Abilities	Outcomes
CO401.1	**Remember**	How AI Integrated with the Business
CO401.2	**Apply**	Apply AI in Different Business Sectors