

**Hindi Vidya Prachar Smiti’s**

**RAMNIRANJAN JHUNJHUNWALA COLLEGE**

**OF ART, SCIENCE AND COMMERCE**

**(Empowered Autonomous)**

Affiliated to  
**University of Mumbai**

**Syllabus for the M.Sc.**

**Program :** M.Sc. Data Science and Artificial Intelligence

**Program Code :**

***(Choice Based Credit System with effective from the academic year 2023 – 2024 in alignment with the NEP2020 facilitating the inter-and multidisciplinary learning and multiple entry and exit of the students)***

***(CBCS 2023-2024)***

***F.Y. B.Sc. Data Science and Artificial Intelligence Syllabus Semester I & II***

***National Education Policy (NEP 2020)***

| **CURRICULUM FOR MSc DSAI PART- I & II** | | |
| --- | --- | --- |
| **SEMESTERS** | **CREDITS** | **SUBJECTS** |
| **SEM - I** | 3+2 | **Introduction to Supervised Learning** |
| 3+2 | **Data Analysis and Visualization** |
| 2+2 | Database & Warehousing |
| 2+2 | Python for Data Science |
| Mathematics for Data Science |
| 2+2 | Research Methodology - I |
| **Total Credits** | **22** | |
| **SEM - II** | 3+2 | Introduction to Unsupervised Learning |
| 3+2 | Data Engineering |
| 2+2 | Time Series Analysis & Forecasting |
| Social Media Analytics |
| 2+2 | Soft Computing |
| 4 | On Job Training (OJT) |
| **Total Credits** | **22** | |
| **SEM - III** | 4+2 | Generative AI |
| 4+2 | Optimization Technique |
| 4+2 | Advanced Data Analytics |
| 2+2 | Computer Vision |
| Natural Language Processing |
|  |  |
| **Total Credits** | **22** | |
| **SEM - IV** | 4+2 | Introduction to Reinforcement Learning |
| 2 | Machine learning System Design |
| 2+2 | Introduction to Robotics |
| Cloud Native Development |
| 10 | *Research Project and Internship* |
| **Total Credits** | **22** | |

**Semester I**

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Introduction to Supervised Learning** | DSC | 3 | 3 |
| **Course Outcome:**   1. To understand the features of machine learning to apply it to real world problems. 2. To characterize the machine learning algorithms as supervised learning and unsupervised learning and Apply and analyze the various algorithms of supervised and unsupervised learning. 3. To analyze the concept of neural networks for learning linear and non-linear activation functions. 4. To learn the concepts in Bayesian analysis from probability models and methods   **Learning Outcome:**   1. Able to understand the basics of Unsupervised Learning. 2. Understand the informed and uninformed problem types and apply search strategies to solve them. 3. Design and evaluate intelligent expert models for perception and prediction from an intelligent environment. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction to machine learning** How machine learn, Basic structure of Machine Learning Model and comparison to traditional programming, Goal and Applications of Machine learning. Types of Data related to Machine Learning, Introduction to standard machine learning datasets and interpreting the data , Correlation between data, supervised vs Unsupervised learning algorithms. |  |
| **UNIT - II** | **Feature Engineering:**  Feature transformation, feature construction, feature  scaling(Standardization & Normalization), feature extraction: Principal Component Analysis, t-sne, curse of dimensionality, function and power transformation. |  |
| **UNIT - III** | **Regression :** Linear regression and assumptions, Feature Selection using Correlation Coefficient, Polynomial regression, Multivariate Regression, Logistic regression, Cost functions for Regression(Mean Square Error, Mean Absolute Error, Root Mean Square error, R2 Score) and optimization using gradient descent. |  |
| **UNIT - IV** | **Classification :** Linear & Non-Linear Datasets, Decision Tree Algorithm, Role of Entropy and Information Gain(IG) in decision tree algorithm, CART ,Random forest, Classification Metrics - Type I and Type II Errors and Confusion Matrix, Overfitting vs Under fitting, Bias variance trade off, Naive Bayes Classifier, Support Vector Machine (SVM), Kernels in  SVM. |  |
| **References**   1. E. Alpaydin, Introduction to Machine Learning, 3rd Edition, MIT Press, 2015. 2. T Hastie, R Tibshirani and J Friedman, The Elements of Statistical Learning Data Mining, Inference, and Prediction, 2nd Edition, Springer, 2009. 3. C.Bishop, "Pattern Recognition and Machine Learning, Springer. 4. A.F..Vermeulen, "Practical Data Science", APress, 2018. 5. S.Ozdemir, "Principles of Data Science", Packt, 2016. | | |

**Practical List:**

1. Write a python program to import and export data using Pandas library functions
2. Demonstrate various data pre-processing techniques for a given dataset
3. Write a Python program to demonstrate various Data Visualization Techniques.
4. Implement Simple and Multiple Linear Regression Models.
5. Develop Logistic Regression Model for a given dataset
6. Develop Decision Tree Classification model for a given dataset and use it to classify a new sample.
7. Implement Naïve Bayes Classification in Python
8. Build KNN Classification model for a given dataset.
9. Demonstrate Support Vector Machine for a given dataset.
10. Implement Random forest ensemble method on a given dataset.
11. Implement Boosting ensemble method on a given dataset.

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Data Analysis and Visualization** | DSC | 3 | 3 |
| **Course Outcome:**   1. Manipulate large datasets and handle missing or inconsistent values in datasets. 2. Perform statistical analysis using tools. 3. Discover and visualize data sets using PowerBi.   **Learning Outcome:**   1. Use data analysis tools in the pandas library. 2. Load, clean, transform, merge and reshape data. 3. Create informative visualization and summarize data sets. 4. Analyze and manipulate time series data. 5. Solve real world data analysis problems. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction to Data Analysis**: Data Analysis - Exploratory Data Analysis and Data Science Process - Responsibilities of a Data Analyst - Data Analytics vs. Data Analysis - Types of Data - Understanding Different Types of File Formats - Sources of Data - Languages for Data Professionals - Overview of Data Repositories - Data Marts, Data Lakes, ETL, and Data Pipelines - Foundations of Big Data - Identifying Data for Analysis |  |
| **UNIT - II** | **Data Sources** - How to gather and Import Data - Data Loading, Storage and File Formats - Reading and Writing Data in Text Format, Web Scraping, Binary Data Formats, interacting with Web APIs, Interacting with Databases – Data Wrangling - Hierarchical Indexing, Combining and Merging Data Sets Reshaping and Pivoting - Tools for Data Wrangling - Data Cleaning and Preparation - Handling Missing Data, Data Transformation, String Manipulation |  |
| **UNIT - III** | **Intro to data visualization** - Introduction to Visualization and Dashboarding Software - Visualization Tools - Getting started with Tableau Desktop – Connecting to the dataset - Creating charts – Creating common visualizations (bar charts, line charts etc.) - Filtering and sorting data - Adding Titles, Labels, and descriptions - Publish your work to Tableau Cloud - Interactivity with text and visual tooltips - Interactivity with actions (filter, highlight, URL) – Assembling dashboards from multiple charts |  |
| **UNIT - IV** | **Introduction to Power BI** - Understanding Desktop - Understanding Power BI Report Designer - Report Canvas, Report Pages: Creation, Renames - Report Visuals, Fields and UI Options - Experimenting Visual Interactions, Advantages - Reports with Multiple Pages and Advantages - Pages with Multiple Visualizations - PUBLISH Options and Report Verification in Cloud - Adding Report Titles. Report Format Options - Introduction to data storytelling - Creating a data story |  |
| **References**   1. Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython by McKinney, W., 2nd edition. O‟Reilly Media, 2017 2. Doing Data Science: Straight Talk from the Frontline by O‟Neil, C., & Schutt, R, O‟Reilly Media, 2013 3. The Big Book of Dashboards by Steve Wexler, Jeffrey Shaffer, Andy Cotgreave, John Wiley & Sons, 2017 4. Practical Tableau by Ryan Sleeper, O‟Reilly Media, 2018 5. Power BI. Book-1, Business Intelligence Clinic: Create and Learn by Roger F Silva, 2018 6. Introducing Microsoft Power BI by Alberto Ferrari and Marco Russo, Microsoft Press, Washington, 2016 | | |

**Practical List:**

1. Show Basic Visualization in Python
2. Show Basic Visualization in R.
3. Connecting to Data and preparing data for visualization in Tableau
4. Use Data aggregation and statistical functions in Tableau.
5. Show Data Visualization using Tableau.
6. Use dashboards of Tableau.
7. Show Data Visualization using PowerBi.
8. Show Data Visualization using DataWrapper.
9. Show Data Visualization using Gnatt Chart.
10. Show Data Visualization using Zoho.
11. Publish visualised data on Cloud in PowerBi

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Database and Warehousing** | DSC | 2 | 2 |
| **Course Outcome:**   * To understand the principles of Data warehousing and Database * To be familiar with the Data warehouse architecture and its Implementation. * To know the Architecture of a Database system. * To understand the various Data preprocessing Methods. * To perform classification and prediction of data.   **Learning Outcome:**   1. Students will be able to describe architecture and methods for storage and provision of enterprise data. 2. Students will develop competency in query development and essential business intelligence reporting. 3. Students will demonstrate competency in data modeling, including dimensional modeling | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Database Concepts:** Why Databases?, Data versus Information, Introducing the Database, Why Database Design Is Important, Overview of MongoDB, Introduction of MongoDB, Nosql Database, Advantage over RDBMS, MongoDB Data Types, Install MongoDB, MongoDB Data Modeling, MongoDB Operators, Query & Projection Operator, MongoDB Update Operator, Aggregation Pipeline Stages , MongoDB limit(), MongoDB sort(), Query Modifiers |  |
| **UNIT - II** | **Database Commands:** Aggregation Commands, **Database:** Create Database, Drop Database Collection, Create Collection, Drop Collection **CRUD :** Documents, Inset Documents, Update Documents, Delete Documents, Query Documents, SQL to MongoDB Mapping , MongoDB text search, Partial Updates & Document Limits, Connectivity, Java MongoDB, PHP MongoDB, Python MongoDB |  |
| **UNIT - III** | **Data Warehousing and Business Analysis**: - Data Warehousing Components, building a Data warehouse, Data Warehouse Architecture, DBMS Schemas for Decision Support, Data Extraction, Cleanup, and Transformation Tools –Metadata – reporting, Query tools and Applications, Online Analytical Processing (OLAP), OLAP and Multidimensional Data Analysis.  **Data Mining:** - Data Mining Functionalities, Data Pre-processing, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation, Architecture of A Typical Data Mining Systems, Classification Of Data Mining Systems.  **Association Rule Mining:** - Efficient and Scalable Frequent Item set Mining Methods, Mining Various Kinds of Association Rules, Association Mining to Correlation Analysis, Constraint-Based Association Mining. |  |
| **UNIT - IV** | **Classification and Prediction:** - Issues Regarding Classification and Prediction, Classification by Decision Tree Introduction, Bayesian Classification, Rule Based Classification, Classification by Back propagation, Support Vector Machines, Associative Classification, Lazy Learners, Other Classification Methods, Prediction, Accuracy and Error Measures, Evaluating the Accuracy of a Classifier or Predictor, Ensemble Methods, Model Section.  **Cluster Analysis**: Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Partitioning Methods, Hierarchical methods, Density-Based Methods, Grid-Based Methods , Model-Based Clustering Methods, Clustering High, Dimensional Data, Constraint-Based Cluster Analysis, Outlier Analysis. |  |
| **Text Book**  Jiawei Han, Micheline Kamber and Jian Pei“Data Mining Concepts and Techniques”, Third Edition, Elsevier, 2011.  **Reference Books**   1. Alex Berson and Stephen J. Smith “Data Warehousing, Data Mining & OLAP”, Tata McGraw – Hill Edition, Tenth Reprint 2007. 2. K.P. Soman, Shyam Diwakar and V. Ajay “Insight into Data mining Theory and Practice”, Easter Economy Edition, Prentice Hall of India, 2006. 3. G. K. Gupta “Introduction to Data Mining with Case Studies”, Easter Economy Edition, Prentice Hall of India, 2006. 4. Pang-Ning Tan, Michael Steinbach and Vipin Kumar “Introduction to Data Mining”, Pearson Education, 2007. 5. Shannon Bradshaw, Eoin Brazil, Kristina Chodorow “MongoDB: The Definitive Guide, 3rd Edition”, O’Reilly, 2019. 6. Kyle Banker Peter Bakkum Shaun Verch Douglas Garrett Tim Hawkins, “MongoDB in Action”, MANNING, 2nd Edition. | | |

| **Practical List:**   1. Demonstrate MongoDB Basic Operations 2. Demonstrate MongoDB Aggregation Operations 3. Demonstrate MongoDB Sorting & limiting & skipping Operations 4. Demonstrate MongoDB Comparison Operators 5. Demonstrate MongoDB Logical Operators 6. Demonstrate MongoDB $abs, $floor, $ceil Operator 7. Demonstrate MongoDB $log, $mod, $divide, $multiply operator. 8. Demonstrate MongoDB $pow, $sqrt, $subtract 9. Demonstrate MongoDB $trunc, $round, $cmp operator 10. Demonstrate MongoDB $concat, $size, $rename operator 11. Implementation of Supervised Learning: 12. Decision Tree 13. Logistic 14. KNN 15. IDK 16. SMO 17. Naive Bayes   12. Implementation of UnSupervised Learning:   1. Clustering    1. EM    2. Hierarchical    3. Density-Based    4. Simple K Means    5. Association 2. Apriori |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Python for Data Science** | DSC | 2 | 2 |
| **Course Outcome:**   1. Learn Programming fundamentals using Python 2. Understand the concepts and usage data types, variables and other basic elements 3. Learn about using operators and control statements in Python 4. Learn about using arrays and strings in Python. 5. Learn about using IPython architecture for Python. 6. Introduce data Science Tools and plot data using appropriate Python visualization libraries   **Learning Outcome:**   1. Upon the successful completion of this course, the student will be able to achieve: 2. Proficiency in using and applying various data types including, string, array list, tuple and dictionary. Ability to use regular expressions to perform complex operations in less code. 3. Learning to make use of date and time in Python for various applications. 4. Proficiency in using IPython architecture for Data Science Applications. 5. Knowledge about use of various data science tools | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction to Python**  Introduction to Python, Python: Keywords, Identifiers, Statements, Comments, Syntax and Indentation, Variables, Basic Data Types: Text (string/str), Numeric (int, float and complex), Boolean (bool), and None (None Type), Type Casting, Exceptions, Python Operators, Python Date and Time, Python Input and Output Statements, Using Python Modules: import Statement, Namespaces and Scoping.  **Python Flow Controls**  If-else Statement, for Loop, while Loop, continue and break statements, pass Statement. |  |
| **UNIT - II** | **Python Functions**  Functions and arguments, Built-in Functions, User Defined Function, Recursive Function, Modules and Packages.  **Database connectivity in Python:** Installing mysql connector, accessing connector module module, using connect, cursor, execute & close functions, reading single & multiple results of query execution, executing different types of statements, executing transactions, understanding exceptions in database connectivity. |  |
| **UNIT - III** | **Python Module: Pandas**  **Pandas objects:** Series, Data frame and Index, Data frame Operations: Descriptive Statistics, Filtering, Sorting, Subletting Merge, Joins, Add and Remove Columns, Summarizing, Grouping Data, Pivot Tables.  **Python Module: Numpy-** Numpy Arrays, Aggregation Functions, Array: Indexing, Slicing, Copying, Shaping, Reshaping, Splitting, Searching, Sorting and Filtering Operations, Array Iterations, Random Functions. |  |
| **UNIT-IV** | **Visualization in Python**  **Introduction to Visualisation Libraries:** Matplotlib, Seaborn and Bokeh, Introduction to Pyplot, Subplots, formatting style of the Plot: Markers, Line, Labels axes, colors and Grid, plotting with keyword Strings, plotting with Categorical Variables, Understanding Charts: Scatter Plot, Heat Map, Box and Whisker Plot, Timeline, Choropleth Map, Formatting Charts, Word Cloud.  **Building and sharing applications using Streamlit**  Streamlit installation, Data Modeling and Data Flow, API Reference: visualize,Creating and deploying app. |  |
| **References**   1. The Python Language Reference Manual, Guido Van Rossum, Fred L. Jr. Drake, Published by Network Theory Ltd. 2. Python: The Complete Reference Book by Martin Brown and Martin C. Brown. 3. Python in nutshell Book by [Alex Martelli](https://www.amazon.in/s/ref=dp_byline_sr_book_1?ie=UTF8&field-author=Alex+Martelli&search-alias=stripbooks), [Anna Ravenscroft](https://www.amazon.in/s/ref=dp_byline_sr_book_2?ie=UTF8&field-author=Anna+Ravenscroft&search-alias=stripbooks), [Steve Hold](https://www.amazon.in/Steve-Holden/e/B001ITX5DG/ref=dp_byline_cont_book_3) 4. Python for Data Analysis: Data Wrangling with pandas, NumPy, and IPython, Wes McKinney 5. “Programming Python, Book by Mark Lutz. ” 6. “Fluent Python, Book by Luciano Ramalho.” 7. “https://www.w3schools.com/python/default.asp” | | |

**Practical List:**

1. Write a python program to demonstrate variables, datatypes, indentation, etc.
2. Implementation of control statements and iterative statements in python.
3. Implementation of exception handling and file handling.
4. Perform CRUD operations on databases using python.
5. Demonstration of Pandas Dataframe and some basic operations.
6. Demonstration of the Aggregation function of NumPy.
7. Demonstration of random function of NumPy.
8. Implementation of charts and graphs using matplotlib.
9. Building basic web applications using streamlit.
10. Implementation of dashboard using streamlit.

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Mathematics for Data Science** | DSC | 2 | 2 |
| **Course Outcome:**  **Learning Outcome:** | | | | |

| **UNIT - I** | **Data Aggregation:** Measures of Central tendency: Mean, Median, mode for raw data, discrete, grouped frequency distribution.  **Measures dispersion:** Variance, standard deviation, coefficient of variation for raw data, discrete and grouped frequency distribution, quartiles, quantiles Real life examples  **Measures of Skewness and Kurtosis:** skewness, kurtosis, types of skewness and kurtosis, interpretation and calculation. |  |
| --- | --- | --- |
| **UNIT - II** | **Standard distributions:** random variable; discrete, continuous, expectation and variance of a random variable, pmf, pdf, cdf, reliability, Introduction and properties without proof for following distributions; binomial, normal, chi-square, t, F. examples  **Hypothesis testing:** one sided, two sided hypotheses, critical region, pvalue, tests based on t, Normal and F, confidence intervals. Analysis of variance : one-way, two-way analysis of variance. |  |
| **UNIT - III** | **Linear Algebra:** Scalars, Vectors, Matrices and their properties, Vector Addition and Multiplication, Norm of a vector, Dot product of two vectors, Cross products , Relation between norm and dot product, Orthogonal and Orthonormal Vectors, Linear Independence of vectors, Linear Dependence and span **Scalars, Vectors and Matrices:** Matrices, Visualizing matrices, Determinants, Properties of Matrices, Matrix multiplication, Types of Matrices, Transpose of matrix, Identity and Inverse of a Matrix, Determinant of a Matrix  **Principal Components Analysis (PCA)**  Eigen Decomposition: Eigenvalues and Eigen Vectors, Covariance Matrix, Correlation Matrix |  |
| **UNIT - IV** | **DIFFERENTIATION:** Implicit Differentiation, Derivatives of Logarithmic Functions, Derivatives of Exponential and Inverse Trigonometric Functions **THE DERIVATIVE IN GRAPHING AND APPLICATIONS:** Analysis of Functions I: Increase, Decrease, and Concavity, Analysis of Functions II: Relative Extrema; Graphing Polynomials, Absolute Maxima and Minima, Applied Maximum and Minimum. **Problems:** Rolle’s Theorem; Mean-Value Theorem |  |
| **References**  1. “Probability and Statistics for Engineers”, Dr. J. Ravichandran,2010  2. “Fundamentals of Data Science: Take the First Step to Become A Data Scientist”, Samuel Burns, Amazon KDP Printing and Publishing.  3.  Essential Math for Data Science by O'reilly.  4. Hands-On Mathematics for Data Scientists: Explore advanced mathematical concepts to prove your abilities to become a successful data scientist by [Dr. Tirthajyoti Sarkar](https://www.amazon.in/s/ref=dp_byline_sr_book_1?ie=UTF8&field-author=Dr.+Tirthajyoti+Sarkar&search-alias=stripbooks) | | |

| **Practical List:**   1. Perform Data Collection, Cleaning, Modelling and Compilation 2. Implementation of the Exploratory Data Visualization Technique 3. Perform Different Statistics Distribution 4. Implementation Hypothesis Testing 5. Perform for Exploring Categorical and Binary Data 6. Perform Anova testing 7. Implementation of Non-parametric tests. 8. Perform Time Series Analysis 9. Implementation of Regression Analysis 10. Computing eigenvalues and eigenvectors for dimensionality reduction. 11. Calculate Maxima and Minima for the given function. 12. Impute the example based on the application of integration 13. Discuss the limit and continuity of the given function. 14. Impute an example based on Rolle’s and Mean Value Theorem. 15. Calculate the dot product of the vector and norm of the vector. 16. Impute Inverse of matrix by adjoint method. 17. Evaluate Definite Integral by substitutions. |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Research Methodology** | DSC | 4 | 4 |
| **Course Outcome:**   1. To understand Research and Research Process 2. To acquaint students with identifying problems for research and develop research strategies 3. To familiarize students with the techniques of data collection, analysis of data and interpretation   **Learning Outcome:**   1. Understand the meaning and objectives of research 2. Familiar with the research in Commerce and Management 3. Classify the research into different types 4. Understand the review of literature and research process 5. Understand the methods of research: case study method and survey method. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Foundations of Research:** Meaning, Objectives, Motivation, Utility. Concept of theory, empiricism, deductive and inductive theory. Characteristics of scientific method – Understanding the language of research – Concept, Construct, Definition, Variable. Research Process  **Problem Identification & Formulation** – Research Question – Investigation Question – Measurement Issues – Hypothesis – Qualities of a good Hypothesis –Null Hypothesis & Alternative Hypothesis. Hypothesis Testing – Logic & Importance |  |
| **UNIT - II** | **Research Design:** Concept and Importance in Research – Features of a good research design – Exploratory Research Design – concept, types and uses, Descriptive Research Designs – concept, types and uses. **Experimental Design:** Concept of Independent & Dependent variables  Qualitative and Quantitative Research: Qualitative research – Quantitative research – Concept of measurement, causality, generalization, replication. Merging the two approaches.  **Measurement:** Concept of measurement– what is measured? Problems in measurement in research – Validity and Reliability. Levels of measurement – Nominal, Ordinal, Interval, Ratio. |  |
| **UNIT - III** | **Sampling:** Concepts of Statistical Population, Sample, Sampling Frame, Sampling Error, Sample Size, Non Response. Characteristics of a good sample. Probability Sample – Simple Random Sample, Systematic Sample, Stratified Random Sample & Multi-stage sampling. Determining size of the sample – Practical considerations in sampling and sample size.  **Data Analysis:** Data Preparation – Univariate analysis (frequency tables, bar charts, pie charts, percentages), Bivariate analysis – Cross tabulations and Chi-square test including testing hypothesis of association. |  |
| **UNIT - IV** | **Interpretation of Data and Paper Writing** – Layout of a Research Paper, Journals in Computer Science, Impact factor of Journals, When and where to publish? Ethical issues related to publishing, Plagiarism and Self-Plagiarism.  Use of Encyclopaedias, Research Guides, Handbook etc., Academic Databases for Computer Science Discipline. |  |
| **References**   1. Business Research Methods – Donald Cooper & Pamela Schindler, TMGH, 9th edition 2. Business Research Methods – Alan Bryman & Emma Bell, Oxford University Press. 3. Research Methodology – C.R.Kothari | | |

**Semester - II**

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Introduction to Unsupervised Learning** | DSC | 3 | 3 |
| **Course Outcome:**   1. Understand how to evaluate clustering results using a variety of metrics. 2. Understand the tradeoffs and assumptions inherent in different clustering techniques. 3. Understand how unsupervised learning can be used to improve supervised prediction. 4. Perform density estimation using a kernel, with a single random variable. 5. Interpret a biplot result from principal components analysis (PCA).   **Learning Outcome:**   1. Correctly apply and interpret results from clustering methods in scikit-learn, including k-means, agglomerative clustering, hierarchical clustering, and DBSCAN. 2. Correctly apply and interpret results from manifold learning methods, including multidimensional scaling 3. Build awareness of other advanced methods like kernel PCA and spectral clustering. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction:** What Is Unsupervised Learning? Application of Unsupervised Learning, Types of Unsupervised Learning, Advantages and Disadvantages of Unsupervised Learning  **Beyond Supervised Learning:**  Data Augmentation, Transfer Learning,: Fine Tuning, Unsupervised pre-training, Semi-supervised learning, Meta-Learning, Few-shot Learning. |  |
| **UNIT - II** | **Dimensionality Reduction :**Curse of Dimensionality, Principal Component Analysis(PCA),  Computational issues, choosing number of latent dimensions, t-SNE, PCA vs t-SNE, Singular Value Decomposition(SVD) |  |
| **UNIT - III** | **Introduction to Clustering :** Evaluation the output of clustering methods, Partition Based Clustering, K-Means clustering, elbow methods to find number of clusters, K-Medoids clustering, Hierarchical Clustering, Agglomerative and Divisive Approach, Dendrogram, linkage, Density-based spatial clustering of applications with noise(DBSCAN) |  |
| **UNIT - IV** | **Association rule learning :** Working with association rule learning, Support ,Confidence, Lift, Apriori Algorithm, Frequent Itemset, Steps of Apriori Algorithm, Fp-Growth Algorithm, Frequent Pattern(FP) Tree, Difference Between  Fp-Growth and Apriori, Application of Association rule Learning, |  |
| **Reference:**   1. Probabilistic Machine learning An Introduction, Kevi P. Murphy. 2. Machine Learning: Master Supervised and Unsupervised Learning Algorithms with Real Examples by [Dr Ruchi Doshi](https://www.amazon.in/s/ref=dp_byline_sr_ebooks_1?ie=UTF8&field-author=Dr+Ruchi+Doshi&text=Dr+Ruchi+Doshi&sort=relevancerank&search-alias=digital-text), [Dr Kamal Kant Hiran](https://www.amazon.in/s/ref=dp_byline_sr_ebooks_2?ie=UTF8&field-author=Dr+Kamal+Kant+Hiran&text=Dr+Kamal+Kant+Hiran&sort=relevancerank&search-alias=digital-text), [Ritesh Kumar Jain](https://www.amazon.in/s/ref=dp_byline_sr_ebooks_3?ie=UTF8&field-author=Ritesh+Kumar+Jain&text=Ritesh+Kumar+Jain&sort=relevancerank&search-alias=digital-text), [Dr Kamlesh Lakhwani](https://www.amazon.in/s/ref=dp_byline_sr_ebooks_4?ie=UTF8&field-author=Dr+Kamlesh+Lakhwani&text=Dr+Kamlesh+Lakhwani&sort=relevancerank&search-alias=digital-text)**.** | | |

| **Practical List:**   1. Implementation of data augmentation. 2. Build a model using transfer learning. 3. Demonstration of few shot learning. 4. Perform dimensionality reduction using principal component analysis. 5. Implementation of singular value decomposition. 6. Implementation of partition based clustering. 7. Implementation of density based clustering. 8. Implementation of hierarchical clustering. 9. Perform market basket analysis using apriori algorithm. 10. Implementation of FP-Growth algorithm. |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Data Engineering** | DSC | 3 | 3 |
| **Course Outcome:**   1. To develop the skills of managing the data with respect to knowledge generation. 2. Provide the ability to design the data engineering process 3. To propose the data reliability models 4. To define how to use Machine learning model   **Learning Outcome:**   1. Building the storage system with appropriate data technologies 2. Designing the data pipelines and data flow 3 1,4 3. Processing the data infrastructure 4 3 4. Investigate possible diagnostics by designing Databases for Reliability, Scalability, and Availability, Understanding Data Operations for Flexibility 5. Training and measuring the serving Infrastructure for Machine Learning Models | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Selecting Appropriate Storage Technologies:** From Business Requirements to Storage Systems, Technical Aspects of Data, Types Of Structure, Schema Design Consideration Building and **Operationalizing Storage Systems:** Cloud SQL, Cloud Spanner, Cloud Bigtable, Cloud Firestore, BigQuery, Cloud Memorystore, Cloud Storage, Unmanaged Databases  **Designing Data Pipelines:** Overview Of Data Pipelines, GCP Pipeline Components, Migrating Hadoop and Spark To GCP **Designing a Data Processing Solution:** Designing Infrastructure, Designing for Distributed Processing, Migrating a Data Warehouse |  |
| **UNIT - II** | **Building and Operationalizing Processing Infrastructure:** Provisioning and Adjusting Processing Resources, Monitoring Processing Resources  **Designing for Security and Compliance:** Identity and Access Management with Cloud IAM, Using IAM with Storage and Processing Services, Data Security, Ensuring Privacy with the Data Loss Prevention API, Legal Compliance |  |
| **UNIT - III** | **Designing Databases for Reliability, Scalability, and Availability:** Designing Cloud Bigtable Databases for Scalability and Reliability, Designing Cloud Spanner Databases for Scalability and Reliability, Designing BigQuery Databases for Data Warehousing **Understanding Data Operations for Flexibility and Portability:** Cataloging and Discovery with Data Catalog, Data Preprocessing with 12 Dataprep, Visualizing with Data Studio, Exploring Data with Cloud Datalab, Orchestrating Workflows with Cloud Composer  **Deploying Machine Learning Pipelines:** Structure of ML Pipelines, GCP Options for Deploying Machine Learning Pipeline |  |
| **UNIT - IV** | **Choosing Training and Serving Infrastructure:** Hardware Accelerators, Distributed and Single Machine Infrastructure, Edge Computing with GCP Measuring, Monitoring, and **Troubleshooting Machine Learning Models:** Three Types of Machine Learning Algorithms, Deep Learning, Engineering Machine Learning Models, Common Sources of Error in Machine Learning Models  **Leveraging Prebuilt Models as a Service:** Sight, Conversation, Language, Structured Data |  |

| **Reference:**   1. Professional Data Engineer DAN SULLIVAN Sybex - Wiley 3 rd 2020. 2. Data Driven Science and Engineering STEVEN L. BRUNTON, J. NATHAN KUTZ 3. Cambridge University Press 2nd 2019. 4. Data Security in Cloud Computing Vimal Kumar, Sivadon Chaisiri and Ryan Ko The 5. Institution of Engineering and Technology 2020. 6. 4. Data Engineering on Azure Vlad Riscutia Manning Publications 2021. | | |
| --- | --- | --- |

| **Practical List:**   1. Collection of raw data from various sources 2. Implement of data preprocessing pipeline 3. Perform ETL operations using tool 4. Build GCP machine learning pipeline 5. Integration of data into different file formats 6. Designing Cloud Bigtable Databases for Scalability and Reliability 7. Troubleshooting machine learning model 8. Perform data visualization with tools. |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Time Series Analysis and Forecasting** | DSC | 2 | 2 |
| **Course Outcome:**   1. Forecast the trend pattern exhibited by the given data by using various methods 2. Run and interpret time series models and regression models for time series 3. Use the Box-Jenkins approach to model and forecast time series data empirically 4. Analyze and estimate the cyclic components using special processes   **Learning Outcome:**   1. Fit various growth curves, trend and to measure seasonal indices 2. Understand forecasting by different methods 3. Able to calculate variance of a random component | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | Introduction to Trend Introduction to times series data, application of time series from various fields, Components of a time series, Decomposition of time series.  **Trend:** Estimation of trend by free hand curve method, method of semi averages, fitting a various mathematical curve and growth curves. |  |
| **UNIT - II** | Trend and Seasonal Component Method of moving averages, Detrending, Effect of elimination of trend on other components of the time series.  **Seasonal Component:** Estimation of seasonal component by Method of simple averages, Ratio to Trend, Ratio to moving average and Link relatives |  |
| **UNIT - III** | **Forecasting Variate component method** - Stationary Time series: Weak stationary, autocorrelation function and correlogram of moving average  Forecasting: Exponential smoothing methods, short term forecasting methods: Brown‟s discounted regression, Box-Jenkins Method. |  |
| **UNIT - IV** | **Cyclic Component Deseasonalization** - Cyclic Component: Harmonic Analysis. **Some Special Processes**: Moving-average (MA) process and Autoregressive (AR) process of orders one and two, Estimation of the parameters of AR (1) and AR (2) – YuleWalker equations. |  |
| **References**   1. Kendall, M. (1976) Time Series. 2nd Edition, Charles Griffin and Co Ltd., London and High Wycombe. 2. Chatfield C. (1980). The Analysis of Time Series –An Introduction, 6th Edition, Chapman & Hall. 3. Mukhopadhyay P. (2011). Applied Statistics, 2nd ed. Revised reprint, Books and Allied Shumway, R. H., and Stoffer, D. S. (2006). Time Series Analysis and Its Applications With R Examples, 2 ed. Springer, New York, NY 4. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time Series Analysis: Forecasting and Control. Prentice - Hall, Inc., Upper Saddle River, NJ. 5. Yaffee, R. and McGee, M. (2000). Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS. Academic Press, Inc., San Diego, CA. | | |

| **Practical List:**   1. Decompose time series data to find trend, seasonality, cyclic and irregularity.. 2. Data conversion of non-stationary to stationary. 3. Perform a duckey-fuller test to check stationarity of data. 4. Implementation of moving averages models. 5. Demonstration of autocorrelation functions and partial autocorrelation functions. 6. Implementation of Autoregressive models. 7. Implementation of ARIMA model. 8. Implementation of SARIMA model. 9. Time series forecasting using exponential smoothing. 10. Implementation of LSTM for time series forecasting. |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Social Media Analytics** | DSC | 2 | 2 |
| **Course Outcome:**   1. To understand and familiarize the learners with the concept of social media. 2. Social media analytics integrates with the learners to understand the significance. 3. Enable the learners to develop skills required for analyzing the effectiveness of social media. 4. Familiarize the learner with different visualization techniques for social media decisions. 5. Examine the ethical and legal implications of leveraging social media data.   **Learning Outcome:**   1. To understand the fundamental concepts of social media networks. 2. To Collect, monitor, store and track social media data 3. To analyze and visualize social media data 4. To design and develop social media analytics models. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Users:** The Who of social media. Measuring Variations in User Behavior in Wikipedia, **Long Tails Everywhere:** The 80/20 Rule (p/q Rule), Online Behavior on Twitter. **Networks:** The How of Social Media. Types and Properties of Social Networks, Visualizing Networks, **Degrees:** The Winner Takes All, Capturing Correlations: Triangles, Clustering, and Assortativity. |  |
| **UNIT - II** | **Temporal Processes:** The When of Social Media. What Traditional Models Tell You About Events in Time, Inter-Event, Bursty Activities of Individuals, Forecasting Metrics in Time.  **Content:** The What of Social Media. Defining Content: Focus on Text and Unstructured Data, Using Content Features to Identify Topics, Extracting Low-Dimensional Information from High-Dimensional Text. |  |
| **UNIT - III** | **Processing Large Datasets.** MapReduce: Structuring Parallel and Sequential Operations, Multi-Stage MapReduce Flows, Patterns in MapReduce Programming, Sampling and Approximations: Getting Results with Less Computation, **Sampling and Approximations:** Getting Results with Less Computation, Bloom Filter, Count-Min Sketch, Executing on a Hadoop Cluster (Amazon EC2). |  |
| **UNIT - IV** | **Learn, Map, and Recommend.** Social Media Services Online, Problem Formulation, Learning and Mapping, Prediction and Recommendation. Social Media Data, From Data to Insights, Luis Madureira, Analytics in Social Media, Dedicated vs. Hybrid Tools.  Reports, Milan Veverka, Strategy, Tactics , Michael Wu, Prescriptive Analytics, The Future of Social Media Analytics. |  |
| **References**   1. Gabor Szabo, Gungor Polatkan, Oscar Boykin, Antonios Chalkiopolos “Social Media Data Minning and Analytics”, John Wiley , & Sons, 2019. 2. Alex Goncalves “Social Media Analytics Strategy”, Apress, 2017. | | |

| **Practical List:**   1. Perform Analysis on Youtube, Instagram and Twitch using tools. 2. Perform data collection from any of the social media platform. 3. Perform Preprocessing of collected data and store it. 4. Analyze and Visualize Social Media data collected. 5. Perform Hyperlink Analytics. 6. Perform Text Analytics 7. Perform Action Analytics. 8. Analyze how Individual/Organization use Social Media and Social Media privacy policy. |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Soft Computing** | DSC | 2 | 2 |
| **Course Outcome:**   1. To Provide the knowledge of soft computing concepts like fuzzy logic, neural networks and genetic 2. algorithms, where Artificial Intelligence is the mother branch of all. 3. To learn effective techniques and their roles in building intelligent systems. 4. To learn how to use neural networks for classification and regression problems.   **Learning Outcome:**   1. Identify and describe soft computing techniques and their roles in building intelligent machines. 2. Select soft computing methodology to solve a particular problem. 3. Apply fuzzy logic and reasoning to solve engineering problems. 4. Apply genetic algorithms to combinatorial optimization problems. 5. Apply neural networks for classification and regression problems. 6. Evaluate and compare solutions by various soft computing approaches for a given problem. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction**  Introduction of soft computing, soft computing vs. hard computing, applications of soft computing.  **Types of soft computing techniques**  Fuzzy Computing, Neural Computing, Genetic Algorithms.  **Other Soft Computing Techniques**  Associative Memory, Adaptive Resonance Theory, Classification, Clustering, Bayesian Networks, Probabilistic reasoning  **Associative Memory Networks**  Training algorithm for pattern Association, Autoassociative memory network, Hetroassociative memory network, bi-directional associative memory, Hopfield networks. |  |
| **UNIT - II** | **Special Networks**  Simulated annealing, Boltzmann machine, Gaussian Machine, Cauchy Machine, Probabilistic neural net, cognition network, cellular neural network, optical neural network.  **Introduction to Fuzzy Logic, Classical Sets and Fuzzy sets** Classical sets, Fuzzy sets.  **Classical Relations and Fuzzy Relations**  Cartesian Product of relation, classical relation, fuzzy relations, tolerance and equivalence relations, non-iterative fuzzy sets, Fuzzification, Defuzzification.  **Genetic Algorithm**  Biological Background, Traditional optimization and search techniques, genetic algorithm vs. traditional algorithms, basic terminologies, operators in genetic algorithm, stopping condition for genetic algorithm flow, problem solving using genetic algorithm. |  |

| **Practical List:**   1. Implementation of single layer perceptron 2. Implementation of multi-layer perceptron 3. Perform backpropagation in neural network 4. Implementation of fuzzy logic. 5. Implementation of heb rule learning 6. Implementation of self organizing map 7. Implementation of delta rule learning 8. Implementation of genetic algorithm. |
| --- |

**Semester - III**

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Deep Learning & Generative AI** | DSC | 4 | 4 |
| **Course Outcome:**   1. To understand generative AI deeply, including its historical development. 2. To understand the dynamics of reinforcement learning and the power of data search in Generative AI. 3. To understand potential digital transformation opportunities enabled by generative AI for your organization. 4. To understand what it will take – from both technology and culture - to make AI work in your organization.   **Learning Outcome:**   1. Discover how diverse domains like art, biology, emotional support, and learning apply Generative AI. 2. Comprehend and implement prompt engineering to enhance productivity. 3. Learn CNN for image translation using Generative AI. 4. Learn Auto Encoders and Generative Adversarial Network with Generative AI. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction to Deep Learning:**  History of Deep Learning, Learning XOR, Perceptron, Gradient Based Learning, Deep Neural Network, Architecture Design of ANN and Backpropagation, Activation & Loss Functions, Drop out & Regularization  **Introduction to Generative Modelling:**  What is Generative Modelling, Generative vs Discriminative Modelling, The Rise of Generative Modelling and Framework, Probabilistic Generative Models, Challenges of Generative Models. |  |
| **UNIT - II** | **Sequence Modelling :** Recurrent Neural Network(RNN), Bidirectional RNN, Encoder- Decoder Sequence to sequence Architecture, LSTM, Gated recurrent unit, Attention Mechanism, Introduction of Transformer, What is self-attention. |  |
| **UNIT - III** | **Convolutional Neural Network:**  CNN and visual cortex system, Convolution Operation, Pooling & padding, CNN vs ANN, Image to Image Translation, ResNET and VGG16 |  |
| **UNIT - IV** | **Auto encoders(AE):**  Architecture of auto encoder, Types: Under complete AE, Sparse AE, Denoising AE, Variation AE, Generating New Faces using Variational Auto encoders  **Introduction to Generative Adversarial Network:**  Generator, Discriminator, Training the GAN, GAN Challenges: Oscillating Loss,DCGAN, Mode Collapse, Uniormative Loss, Hyperparameter. |  |
| **Reference:**   1. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play by David Foster 2. Ian Goodfellow, Yoshua Bengio, Aaron Courville - Deep Learning (2017, MIT) | | |

| **Practical Examiner :**   1. Implementation of gradient descent optimizer. 2. Implementation of drop out and regularization. 3. Implementation of generative models. 4. Design recurrent neural network. 5. Implementation of LSTM. 6. Demonstration of pre-train models. 7. Design convolutional neural network for image classification. 8. Implementation of generative adversarial network. |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Optimization Techniques** | DSC | 4 | 4 |
| **Course Outcome:**   1. Comprehend the techniques and applications of Engineering optimization. 2. Analyze characteristics of a general linear programming problem 3. Apply basic concepts of mathematics to formulate an optimization problem 4. Analyse various methods of solving the unconstrained minimization problem 5. Analyze and appreciate variety of performance measures for various optimization problems   **Learning Outcome:**   1. Apply operations research techniques like linear programming problems in industrial optimization problems. 2. Solve allocation problems using various OR methods. 3. Understand the characteristics of different types of decision making environments and the appropriate decision making approaches and tools to be used in each type. 4. Recognize competitive forces in the marketplace and develop appropriate reactions based on existing constraints and resources. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction to optimization** Introduction to Classical Methods & Linear Programming Problems Terminology, Design Variables, Constraints, Objective Function, Problem Formulation. Calculus method, Kuhn Tucker conditions, Method of Multipliers. |  |
| **UNIT - II** | **Linear Programming** Problem Linear Programming Problem, Simplex method, Two-phase method, Big-M method, duality, Integer linear Programming, Dynamic Programming, Sensitivity analysis. |  |
| **UNIT - III** | **Single Variable Optimization** Problems Optimality Criterion, Bracketing Methods, Region Elimination Methods, Interval Halving Method, Fibonacci Search Method, Golden Section Method. Gradient Based Methods: Newton-Raphson Method, Bisection Method, Secant Method, Cubic search method. |  |
| **UNIT - IV** | **Intelligent Optimization Techniques** Introduction to Intelligent Optimization, Genetic Algorithm: Types of reproduction operators, crossover & mutation, Simulated Annealing Algorithm, Particle Swarm Optimization (PSO), Genetic Programming (GP): Principles of genetic programming, terminal sets, functional sets, differences between GA & GP, random population generation, solving differential equations using GP. |  |
| **References**   1. Reinforcement Learning: An Introduction Second edition, in progress Richard S. Sutton and Andrew G. Barto c 2014, 2015 | | |

| **Practical List:**   1. Introduction to Optimization 2. Introduction to MATLAB 3. Classical Optimization Techniques 4. Unconstrained Optimization: Elimination Methods 5. Unconstrained Optimization: Interpolation Method 6. Unconstrained Optimization: Direct Root Methods 7. Constrained Optimization: Equality Constraints 8. Constrained Optimization: Inequality Constraints |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Advanced Data Analytics** | DSC | 2 | 2 |
| **Course Outcome:**   1. To develop skills to both design and critique visualizations. 2. To introduce visual perception and core skills for visual analysis. 3. To understand R programming and Spreadsheet for analysis. 4. To understand issues and best practices in information dashboard design.   **Learning Outcome:**   1. Explain principles of visual perception 2. Apply core skills for visual analysis 3. Apply visualization techniques for various data analysis tasks 4. Design information dashboard | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT I** | **Introduction To The R Language**, Data Manipulation Techniques Using R Programming: Data In R, Reading And Writing Data, R And Databases, Dates, Factors, Subscribing, Character Manipulation, Data Aggregation, Reshaping Data |  |
| **UNIT II** | **Statistical Applications Using R Programming: Basics**, The R Environment, Probability And Distributions, Descriptive Statistics And Graphics, One- And Two-Sample Tests, Regression And Correlation, Analysis Of Variance And The Kruskal–Wallis Test, Tabular Data, Power And The Computation Of Sample Size, Advanced Data Handling, Multiple Regression, Linear Models, Logistic Regression, Survival Analysis, Rates And Poisson Regression, Nonlinear Curve Fitting |  |
| **UNIT - III** | **Functionality Using Ranges:** Using Ranges, Selecting Ranges, Entering Information Into a Range, Using AutoFill  **Creating Formulas:** Using Formulas, Formula Functions – Sum, Average, if, Count, max, min, Proper, Upper, Lower, Using AutoSum,  **Advance Formulas**: Concatenate, Vlookup, Hlookup, Match, Countif, Text, Trim  **Spreadsheet Charts:**  Creating Charts, Different types of chart, Formatting Chart Objects, Changing the Chart Type, Showing and Hiding the Legend, Showing and Hiding the Data Table |  |
| **UNIT - IV** | **Data Analysis:**  Sorting, Filter, Text to Column, Data Validation  **PivotTables**: Creating PivotTables, Manipulating a PivotTable, Using the PivotTable Toolbar, Changing Data Field, Properties, Displaying a PivotChart, Setting PivotTable Options, Adding Subtotals to PivotTables  **Spreadsheet Tools:** Moving between Spreadsheets, Selecting Multiple Spreadsheets, Inserting and Deleting Spreadsheets Renaming Spreadsheets, Splitting the Screen, Freezing Panes, Copying and Pasting Data between Spreadsheets, Hiding , Protecting worksheets |  |
| **References** | | |

| **Practical List:**   1. Learn all the basics of R-Programming (Data types ,Variables Operators etc.) 2. Implement R-Loops with different examples. 3. Learn the basics of functions in R and implement them with examples. 4. Implement data frames in R. Write a program to join columns and rows in a data frame using c bind()and r bind() in R. 5. Implement different String Manipulation functions in R 6. Implement different data structures in R (Vectors, Lists, Data Frames) 7. Write a program to read a csv file and analyze the data in the file in R 8. Create pie charts and bar charts using R. 9. Create a data set and do statistical analysis on the data using R. 10. Calculate Summary Statistics in Excel 11. Generate Comparative Statistics in Excel 12. Create Graphs in Excel 13. Advanced Data Analysis using PivotTables and Pivot Charts 14. Tabulation, bar diagram, Multiple Bar diagram, Pie diagram, Measure of central tendency: Mean, median, mode, Measure of dispersion: variance, standard deviation, Coefficient of variation. Correlation, regression lines. 15. t-test , F-test, ANOVA one way classification, chi square test, independence of attributes. 16. Time series: forecasting Method of least squares, moving average method. Inference and discussion of results. |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Computer Vision** | DSC | 2 | 2 |
| **Course Outcome:**   1. To review image processing techniques for computer vision. 2. To understand shape and region analysis. 3. To understand Hough Transform and its applications to detect lines, circles, ellipses. 4. To understand three-dimensional image analysis techniques. 5. To understand motion analysis. To study some applications of computer vision algorithms.   **Learning Outcome:**   1. Implement fundamental image processing techniques required for computer vision. 2. Perform shape analysis. 3. Implement boundary tracking techniques. 4. Apply Hough Transform for line, circle, and ellipse detections. 5. Apply 3D vision techniques. 6. Implement motion related techniques. 7. Develop applications using computer vision techniques. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction:** What is computer vision?, A brief history, Image formation, Geometric primitives and transformations, 2D transformations, 3D transformations, 3D rotations, 2D to 3D projections, Lens distortions, Photometric image formation, Lighting, Reflectance and shading, Optics, The digital camera, Sampling and aliasing, Colour ,Compression |  |
| **UNIT - II** | **Image Processing:** Point operators, Linear filtering, More neighbourhood operators, Fourier transformation,  **Feature detection:** edge detection, corner detection, line and curve detection, active contours, SIFT and HOG descriptors, shape context descriptors, Morphological operations,Object recognition, R-CNN, YOLO. |  |
| **UNIT III** | **Segmentation:** Active contours, split & merge, watershed, region splitting, region merging, graph-based segmentation, mean shift and model finding, normalized cut.  **Recognition :** Object detection, face recognition, Instance recognition, VGGFace2, FaceNet. |  |
| **UNIT IV** | **Deep Learning Networks :**  Convolutional Neural Network, CCN Architecture,LeNet Architecture, AlexNet Architecture, DeconvNet Architecture, VGGNet, SegNet Architecture. |  |
| **References:**   1. Computer Vision: Algorithms and Applications, R. Szeliski, Springer, 2011. 2. Introductory techniques for 3D computer vision, E. Trucco and A. Verri, Prentice Hall, 1998 3. Computer Vision: Principles, Algorithms, Applications, Learning, [E. R. Davies](https://www.google.co.in/search?tbo=p&tbm=bks&q=inauthor:%22E.+R.+Davies%22) | | |

| **Practical List:**   1. Windows and Plots 2. Program to change the Brightness of Image 3. To Flip the image around the vertical and horizontal line 4. Display the color components of the image 5. Red Green Blue Components of Image 6. To find the negative of an image 7. Calculate the Histogram of a given image 8. Histogram Equalization of an image 9. Program for Image Filtering(low pass filter)   1)Average filter  2)Weighted Average filter  3)Median filter   1. High pass filters using   1)Sobel operator  2) Laplacian operator   1. Edge detection with gradient and convolution of an Image 2. Program to find threshold of grayscale image 3. Program to find threshold of RGB image 4. Program to estimate and subtract the background of an image 5. Program to convert color image to gray and hsv |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Natural Language Processing** | DSC | 2 | 2 |
| **Course Objectives**  1. To learn sentence structure.  2. To learn Morphological analysis, Lexical analysis, Syntactic and Semantic analysis.  3. To learn feature engineering concepts and rule-based systems for NLP.  4. Using Machine learning and deep learning for NLP.    **Learning Outcomes**  **Upon completion of this course, the student should be able to**  **1.**  Analyse corpus and corpora of NL.  2. Learn language modeling, formal grammars, statistical parsing, machine translation, and dialog processing.  3. Understanding statistical sequence labeling, information extraction, question answering and summarization, advanced topics in speech recognition, speech synthesis. | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| **UNIT - I** | **Introduction:** What is copus and its types, Components of NLP, Regular expressions, Coporpora, Word Tokenization, Stemming & Lemmatizing, Sentence Segmentation and Minimum Edit Distance.  **Text Representation:** OneHotEncoder for Textual data, Bag of Words, Ngram, Bi-Gram and Trigrams, TFIDF, Word2Vec. |  |
| **UNIT - II** | **Naive Bayes, Text Classification, and Sentiment:** Naive Bayes Classifiers, Training the Naive Bayes Classifier, Optimizing for Sentiment Analysis, Naive Bayes as a Language Model, Evaluation: Precision, Recall, F-measure. |  |
| **UNIT - III** | **Hidden Markov Models and Speech Recognition:** Speech Recognition Architecture, Overview of Hidden Markov Models, The Viterbi Algorithm Revisited, Advanced Methods for Decoding, Acoustic Processing of Speech, Computing Acoustic Probabilities, Training a Speech Recognizer, Waveform Generation for Speech Synthesis, Human Speech Recognition , Word Classes and Part-of-Speech Tagging: Tagsets for English, Part of Speech Tagging, Rule-based Part-of-speech Tagging, Stochastic Part-of-speech Tagging, Transformation-Based Tagging. |  |
| **UNIT - IV** | **Transformers and Large Language Models**  The Transformer: A Self-Attention Network,  Multihead Attention, Transformer Blocks,  The Residual Stream view of the Transformer Block, The Language Modelling Head, Large Language Models with Transformers,  Large Language Models: Training Transformers |  |
| **Reference Books:**   1. Daniel Jurafsky and James Martin Speech and Language Processing (2nd Edition), Prentice Hall:2 edition,2008. 2. “Python Natural Language Processing”, Jalaj Thanaki, Packt. 3. “Natural Language Processing with Python: Analyzing Text with the Natural Launguage Toolkit”, By Steven Bird, Ewan Klein, and Edward Loper, NLTK. 4. “Speech and Language Processing”, Daniel Jurafskey and James H. Martin, Prentice Hall, 2009. | | |

| **Practical**   1. Web Scraping: Scrape data from a webpage & store it into csv format. 2. Implementation of Sentimental Analysis 3. Implementation of Text Preprocessing 4. Demonstrate the Parser in NLP 5. Perform Feature Extraction technique in NLP task 6. One Hot Encoding 7. Bag-Of-Words(BOW) 8. N-Grams 9. Term Frequency Inverse Document Frequency 10. Demonstrate the working of Word Embedding in Natural Language Processing |
| --- |

**Semester - IV**

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Introduction to Reinforcement Learning** | DSC | 4 | 4 |
| **Course Outcome**   1. To understand the basics of Reinforcement Learning. 2. The student should be able to a) model a control task in the framework of MDPs. 3. To identify the model based on the model free methods. 4. To identify stability/convergence and approximation properties of RL algorithms. 5. Use deep learning methods to RL problems in practice.   **Learning Outcomes**   1. Define the key features of reinforcement learning 2. Given an application problem (e.g. from computer vision, robotics, etc), decide if it should be formulated as a RL problem 3. Implement in code common RL algorithms. 4. Describe (list and define) multiple criteria for analyzing RL algorithms and evaluate algorithms 5. Describe the exploration vs exploitation challenge and compare and contrast at least two approaches for addressing this challenge | | | | |

| **UNIT** | **TOPICS** | **HRS** |
| --- | --- | --- |
| UNIT - I | **Introduction Reinforcement Learning:** Overview of Reinforcement Learning concepts and terminology, Difference between supervised, Semi-Supervised, unsupervised, and reinforcement learning. |  |
| UNIT - II | **Markov Decision Processes(MDPS):**  Introduction to MDPs, Formal definition of an MDPs and its components, Value Functions (V-function and Q-Learning), Bellman quotations and optimality principle. |  |
| UNIT - III | **Monte-Carlo Simulation:**  Monte Carlo Prediction, Monte Carlo Estimation of Action Values, Monte Carlo Control, Monte Carlo Control without Exploring Starts, Off-policy Prediction via Importance Sampling Importance Sampling on Truncated Returns |  |
| UNIT - IV | **Eligibility Traces:**  n-Step TD Prediction, The Forward View of TD(λ), The Backward View of TD(λ),  Equivalences of Forward and Backward Views,  Sarsa(λ),  Watkins’s Q(λ),  Off-policy Eligibility Traces using Importance Sampling, Implementation Issues, Variable λ |  |

| **References**  Reinforcement Learning: An Introduction Second edition, in progress Richard S. Sutton and Andrew G. Barto c 2014, 2015 | | |
| --- | --- | --- |

| **Practical List:**   1. Implementation of markov decision process. 2. Perform CartPole Balancing with Q-Learning. 3. Implement Mountain CAr with Deep Q-Networks(DQN). 4. Solve grid-world navigation problem. 5. Implement a solution to the multi-armed Bandit problem. 6. Solving Atari games with advanced DNs. 7. Implementation of monte carlo simulation. 8. Implement TD(0) for value estimation. |
| --- |

| **Course Code** | **Course Name** | | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Machine Learning System Design** | | DSC | 4 | 4 |
| **Course Outcome:**   1. To understand how to Design, Deploy and Monitor Systems using Machine Algorithms. 2. To understand Time Series Forecasting and Graph Neural Network. 3. To understand ML Infrastructure and its Platform.   **Learning Outcome:**   * Develop an appreciation for what is involved in Learning models from data * Understand a wide variety of learning algorithms * Understand how to evaluate models generated from data * Apply the algorithms to a real problem, optimize the models learned and report on the expected accuracy that can be achieved by applying the models. | | | | | |
| **UNIT** | | **TOPICS** | | | **HRS** |
| **UNIT - I** | | Understanding machine learning production, ML and Data Systems Fundamentals, Training Data, Feature engineering, Model selection, development, and training | | |  |
| **UNIT - II** | | Model evaluation, Deployment, Diagnosis of ML system failures & data distribution shifts & monitoring, Monitoring & Continual Learning | | |  |
| **UNIT - III** | | Model Deployment, Experiment tracking & versioning with Weights & Biases | | |  |
| **Unit IV** | | Deploying time series forecasting and graph neural networks, ML Infrastructure and Platform | | |  |
| **References:**   1. Designing Machine Learning Systems by Chip Huyen Released May 2022 Publisher(s): O'Reilly Media, Inc.  Machine Learning System Design With end-to-end examples by Valerii Babushkin and Arseny Kravchenko, MEAP began April 2023 | | | | | |

| **Practical List :**   1. Perform project setup of machine learning. 2. Implement a data pipeline of machine learning systems. 3. Implement system modeling. 4. Perform selecting of machine learning system. 5. Perform training of model. 6. Perform debugging of machine learning models. 7. Demonstrate serving of machine learning system 8. Implement testing of machine learning systems on production. 9. Perform deployment of machine learning model on production. 10. Implement maintenance of ,machine learning model. |
| --- |

| **Course Code** | **Course Name** | | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Introduction to Robotics** | | DSC | 2 | 2 |
| **Course Outcome:**  The objectives of this course are Identify robots and its peripherals for satisfactory operation and control of robots for industrial and non-industrial applications.  **Learning Outcome:**  On completion of the course the student will be able to:  1. list and explain the basic elements of industrial robots  2. analyse robot kinematics and its control methods.  3. classify the various sensors used in robots for better performance.  4. summarize various industrial and non-industrial applications of robots. | | | | | |
| **UNIT** | | **TOPICS** | | | **HRS** |
| **UNIT - I** | | Understanding machine learning production, ML and Data Systems Fundamentals, Training Data, Feature engineering, Model selection, development, and training | | |  |
| **UNIT - II** | | Model evaluation, Deployment, Diagnosis of ML system failures & data distribution shifts & monitoring, Monitoring & Continual Learning | | |  |
| **UNIT - III** | | Model Deployment, Experiment tracking & versioning with Weights & Biases | | |  |
| **Unit IV** | | Deploying time series forecasting and graph neural networks, ML Infrastructure and Platform | | |  |
| **References:**  Designing Machine Learning Systems by Chip Huyen Released May 2022 Publisher(s): O'Reilly Media, Inc. Machine Learning System Design With end-to-end examples by Valerii Babushkin and Arseny Kravchenko, MEAP began April 2023 | | | | | |

| **Practical List :**   1. Perform project setup of machine learning. 2. Implement a data pipeline of machine learning systems. 3. Implement system modeling. 4. Perform selecting of machine learning system. 5. Perform training of model. 6. Perform debugging of machine learning models. 7. Demonstrate serving of machine learning system 8. Implement testing of machine learning systems on production. 9. Perform deployment of machine learning model on production. 10. Implement maintenance of ,machine learning model. |
| --- |

| **Course Code** | **Course Name** | **Group** | **Teaching Scheme** | **Credits** |
| --- | --- | --- | --- | --- |
| **Lectures** |
|  | **Cloud Native Development** | DSC | 2 | 2 |
| **Course Outcome:**   1. The evolution of cloud-native computing and security. 2. How to secure cloud and cloud-native applications. 3. To understand Automation and DevOps. 4. To understand Docker and CI/CD pipelines 5. DevSecOps and the future of security.   **Learning Outcome:**   1. Explain cloud-native architectures and discuss benefits for using this approach to a range of audiences. 2. Review cloud-native application designs and implementations, and propose strategies to improve their performance. 3. Able to create Docker Images 4. Understands Microservices and Monolithic Application. | | | | |
| **UNIT I** | **Cloud-Native Development:** What Is Cloud-Native? Traditional Development vs. Cloud-Native, Agile Practices - Feature-Driven Development, Value-Driven Development, Agile Teams, Writing User Stories, Activity: Analyzing the Case Study Requirements  **Automation and DevOps:** Infrastructure as Code (IaC), Continuous Integration and Continuous Delivery (CI/CD) | | |  |
| **UNIT II** | **Microservices: Introduction to Microservices -** Monolithic vs. Microservice Applications, Recognizing Microservice Boundaries, Stateful vs. Stateless Services, Managing Databases, Activity: Architecting Microservice Applications. **Twelve-Factor Apps -** The Twelve Factors, Implementing Twelve-Factor Apps. **Microservice Architecture -** Designing Loosely-Coupled Services, Communicating Between Microservices, REST, gRPC, OpenAPI, Activity: Running the Case Study | | |  |
| **UNIT III** | **Docker: Understanding Docker -** Containers, Advantages of Containers, Images. **Using Docker -** Basic Docker Commands, Building Docker Images, Dockerfile, Starting Containers, Stopping Containers, Deleting Containers and Images, Activity: Containerizing the Case Study. **Deploying Docker Containers:** Container Registries, Push and Pull, Activity: Managing Docker Containers in a Container Registry | | |  |
| **UNIT IV** | **CI/CD Pipelines -** Cloud-Based (Azure, AWS, Google Cloud) CI/CD Tools, Kubernetes Jenkins, Spinnaker, Terraform, GitHub Actions, Activity: Building a CI Pipeline | | |  |
| **References**   1. Boris Scholl, Trent Swanson, Peter Jausovec “ Cloud Native “  , O'Reilly Media, Inc., 2019 2. Michael Wittig and Andreas Wittig “ Cloud Native Applications“, MANNING, 2016 3. Matthew A. Titmus “Cloud Native Go “, O'Reilly Media, Inc., 2021 | | | | |

| **Practical List:**   1. Perform basic commands of Docker. 2. Perform Dockerﬁle Commands 3. Perform Docker Compose Commands 4. Perform Docker Networking 5. Perform Docker Volumes 6. Perform Docker Object Commands 7. Perform Docker Advanced Commands    1. docker history image    2. docker save image > ﬁle    3. docker load < ﬁle    4. docker commit container image 8. Perform Docker System Commands    1. docker info    2. docker version    3. docker system df    4. docker system events    5. docker system prune 9. Perform Docker Swarm Commands 10. Perform Container Orchestration with Docker Swarm |
| --- |