

Sheth NKTT College of Commerce and Sheth JTT College of Arts, Thane
(Autonomous)
Syllabus (w.e.f. 2025-26)
M.Sc. (Data Science) Part I

Semester II

Category	Code	Course	Credit
Mandatory	MDS201	Soft Computing	4
	MDSP202	Soft Computing Practical	2
	MDT203	Time Series Analysis and Forecasting	4
	MDTP204	Time Series Analysis and Forecasting Practical	2
	MDE205	Ethical Issues in Data Science	2
Elective (Any One)	MDH206	Human Resource Analytics	2 + 2
	MDHP207	Human Resource Analytics Practical	
	MDP208	Public Healthcare Analytics	
	MDPP209	Public Healthcare Analytics Practical	
	MDS210	Social Media Analytics	
	MDSP211	Social Media Analytics Practical	
	MDOT212	OJT	4
		Total	22

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Major Mandatory		
Name of the Dept: Science and Technology		
Course Title: Soft Computing		
Course Code: MDS201		Course Level:6.0
Type: Theory		
Course Credit: 4 credits		
Hours Allotted: 60 Hours		
Marks Allotted: 100 Marks		
Course Objectives: 1. To understand the fundamentals of artificial neural networks, including supervised, unsupervised, and associative learning models, and their applications in pattern recognition and classification. 2. To explore advanced and special neural networks, such as deep learning, spiking neural networks, and probabilistic neural networks, along with their architectures and functionalities. 3. To learn the principles of fuzzy logic, fuzzy sets, membership functions, fuzzy relations, and defuzzification methods for handling uncertainty and imprecision in data. 4. To apply genetic algorithms for optimization problems by understanding their biological inspiration, operators, and problem-solving capabilities compared to traditional algorithms.		
Course Outcomes: Upon completing this course, the student will be able to: CO1. Understanding Soft Computing Paradigms. CO2. Comprehend the principles and mathematical foundations of Fuzzy logic. CO3. Acquire knowledge of the theory and concepts underlying neural networks, including artificial neurons, activation functions. CO4. Explore the theoretical understanding of Genetic algorithms, including genetic operators (selection, crossover, mutation), encoding schemes, fitness evaluation.		

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	Artificial Neural Network Fundamental concepts, Evolution of neural network, basic model of Artificial Neural Network, Important terminologies, McCulloch Pits neuron, linear separability, Hebb network Supervised Learning Network: Perceptron networks, Adaline, MAdaline, Backpropagation network, Radial Basis Function, Time Delay Network, Functional Link Networks, Tree Neural Network. UnSupervised Learning Networks Fixed weight competitive nets, Kohonen self- organizing feature maps, learning vectors quantization, counter propogation networks, adaptive resonance theory networks. Associative Memory Networks: Training algorithm for pattern Association, Autoassociative memory network, hetroassociative memory network, bi- directional associative memory, Hopfield networks, iterative autoassociative memory networks, temporal associative memory networks networks , extreme learning machine model.	15
II	Special Networks Simulated annealing, Boltzman machine, Gaussian Machine, Cauchy Machine, Probabilistic neural net, cascade correlation network, cognition network, neo-cognition network, cellular neural network, optical neural network Third Generation Neural Networks: Spiking Neural networks, convolutional neural networks, deep learning neural	15
III	Fuzzy Logic Introduction to Fuzzy Logic, 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. Membership Function: features of the membership functions, fuzzification and methods of membership value	15

	assignments. Defuzzification: Lambda-cuts for fuzzy sets, Lambda-cuts for fuzzy relations, Defuzzification methods. Fuzzy Arithmetic and Fuzzy measures: fuzzy arithmetic, fuzzy measures, measures of fuzziness, fuzzy integrals	
IV	Unit 4: Genetic Algorithm Biological Background, Traditional optimization and search techniques, genetic algorithm and search space, genetic algorithm vs. traditional algorithms, basic terminologies, simple genetic algorithm, general genetic algorithm, operators in genetic algorithm, stopping condition for genetic algorithm flow, constraints in genetic algorithm, problem solving using genetic algorithm, the schema theorem, classification of genetic algorithm, Holland classifier systems, genetic programming, advantages and limitations and applications of genetic algorithm	15
	Total Hours	60

References:

1. Artificial Intelligence and Soft Computing by Anandita Das Bhattacharya, SPD, 3rd edition 2018
2. Principles of Soft computing by S.N.Sivanandam S.N.Deepa , Wiley, 3rd edition, 2019
3. Neuro-Fuzzy and Soft Computing by J.S.R.Jang, C.T.Sun and E.Mizutani, Prentice Hall of India, 1st edition, 2004
4. Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis & Applications by S.Rajasekaran, G. A. Vijayalakshami , Prentice Hall of India, 1st edition, 2004
5. Fuzzy Logic with Engineering Applications, Timothy J.Ross, McGrawHill 1st edition, 1997
6. Genetic Algorithms: Search, Optimization and Machine Learning by Davis E.Goldberg, Addison Wesley, 1st edition, 1989
7. Introduction to AI and Expert System by Dan W. Patterson, Prentice Hall of India, 2nd edition, 2009

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Major Mandatory		
Name of the Dept: Science and Technology		
Course Title: Soft Computing Practical		
Course Code: MDSP202		Course Level:6.0
Type: Practical		
Course Credit: 2 credits		
Hours Allotted: 60 Hours		
Marks Allotted: 50 Marks		
Course Objectives: 1. To implement and experiment with various neural network models, including Perceptron, Backpropagation, and Hopfield networks for solving classification and pattern recognition problems. 2. To explore and apply unsupervised learning techniques, such as Kohonen networks and associative memory models, for clustering and feature mapping tasks. 3. To simulate fuzzy logic systems, including fuzzy set operations, fuzzification, defuzzification, and fuzzy inference systems for decision-making under uncertainty. 4. To apply genetic algorithms for solving optimization problems by designing fitness functions, using selection, crossover, and mutation operations effectively.		
Course Outcomes: Upon completing this course, the student will be able to: CO1. Understanding of the basic principles, concepts, and techniques of Soft Computing. CO2. Acquire the knowledge and skills necessary to apply Soft Computing techniques to solve real-world problems and Soft Computing algorithms.		

Syllabus: NEP 2020 w.e.f 2026-27

Sr. No	Content	Hours
1	Write a program to implement logical gates AND, OR and NOT with McCulloch-Pitts.	
2	Write a program to implement Hebb's rule.	
3	Implement Kohonen Self organizing map.	
4	Solve the Hamming network given the exemplar vectors.	
5	Write a program for implementing BAM network.	
6	Implement a program to find the winning neuron using MaxNet.	
7	Implement De-Morgan's Law.	
8	Implement Union, Intersection, Complement and Difference operations on fuzzy sets.	
9	Create fuzzy relation by Cartesian product of any two fuzzy sets	
10	Perform max-min composition on any two fuzzy relations.	
		60

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Major Mandatory		
Name of the Dept: Science and Technology		
Course Title: Time Series Analysis and Forecasting		
Course Code: MDT203		Course Level:6.0
Type: Theory		
Course Credit: 4 credits		
Hours Allotted: 60 Hours		
Marks Allotted: 100 Marks		
Course Objectives: 1. To understand the fundamentals of time series data, its components, and various methods for trend estimation and decomposition. 2. To analyze seasonal patterns in time series, using methods like moving averages, ratio to trend, and link relatives for seasonal component estimation. 3. To apply forecasting techniques, including exponential smoothing, Box-Jenkins method, and short-term forecasting models for real-world data. 4. To explore cyclic and special processes like AR and MA models, and perform parameter estimation using methods such as the Yule-Walker equations		
Course Outcomes: Upon completing this course, the student will be able to: - CO1. Forecast the trend pattern exhibited by the given data by using various methods. - CO2. Run and interpret time series models and regression models for time series. - CO3. Use the Box-Jenkins approach to model and forecast time series data empirically. - CO4. Analyze and estimate the cyclic components using special processes.		

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
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.	30
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	30
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.	
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) – Yule-Walker equations.	
	Total Hours	60

Reference Books:

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
4. Shumway, R. H., and Stoffer, D. S. (2006). Time Series Analysis and Its Applications With R Examples, 2 ed. Springer, New York, NY
5. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time Series
6. Analysis: Forecasting and Control. Prentice - Hall, Inc., Upper Saddle River, NJ.
7. 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.

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Major Mandatory		
Name of the Dept: Science and Technology		
Course Title: Time Series Analysis and Forecasting Practical		
Course Code: MDTP204		Course Level:6.0
Type: Practical		
Course Credit: 2 credits		
Hours Allotted: 60 Hours		
Marks Allotted: 50 Marks		
Course Objectives: 1. To develop skills in fitting and visualizing growth curves, such as modified exponential, Gompertz, and logistic curves for trend analysis. 2. To apply the moving average method for identifying and plotting trends in time series data. 3. To measure seasonal variations using techniques like Ratio-to-Trend, Ratio-to-Moving Average, and Link Relative methods. 4. To perform forecasting and variance analysis, using exponential smoothing, short-term forecasting methods, and the variate difference method for random components.		
Course Outcomes: Upon completing this course, the student will be able to: CO1. Fit various growth curves, trend and to measure seasonal indices CO2. Understand forecasting by different methods and able to calculate variance of a random component		

Syllabus: NEP 2020 w.e.f 2026-27

Sr. No.	Content	Hours
1	Fitting and plotting of modified exponential curve.	
2	Fitting and plotting of Gompertz curve.	
3	Fitting and plotting of logistic curve.	
4	Fitting of trend by Moving Average Method.	
5	Measurement of Seasonal indices Ratio-to-Trend method.	
6	Measurement of Seasonal indices Ratio-to-Moving Average method.	
7	Measurement of seasonal indices Link Relative method.	
8	Calculation of variance of random component by variate difference method.	
9	Forecasting by exponential smoothing.	
10	Forecasting by short term forecasting methods.	
		60

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science) Semester: II	
Course Category/Vertical: Major Mandatory	
Name of the Dept: Science and Technology	
Course Title: Ethical Issues in Data Science	
Course Code: MDE205	Course Level:6.0
Type: Theory	
Course Credit: 2 credits	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	
Course Objectives: 1. To understand the ethical principles and challenges related to data collection, storage, sharing, and automated decision-making in data science. 2. To explore emerging ethical issues in technologies like AI, IoT, and machine learning, and to promote responsible practices in data privacy, security, and research.	
Course Outcomes: Upon completing this course, the student will be able to: CO1: Understand the ethical principles, challenges, and decision-making frameworks in Data Science, with particular emphasis on ethical issues arising from data collection, analysis, and automated decision-making processes. CO2: Recognize and address ethical, privacy, and security concerns related to data collection, storage, sharing, and usage, while evaluating emerging ethical issues and future challenges in Data Science.	

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	Ethical foundation in Data Science Introduction to ethical frameworks and principles, Ethics in data science: challenges and importance, Data collection methods, storage, sharing and its ethical considerations, Types of bias in data, Importance of transparency in data science, Ethical considerations in automated decision-making, Data governance frameworks and practices, Ensuring accountability in data science projects.	15
II	Emerging Ethical issues in Data Science Ethical Issues in Data Visualization, Ethical Issues in Machine Learning, Ethical challenges in emerging technologies e.g., AI, IoT, biometrics, blockchain, Ethical challenges in data science research, Ethical considerations in collaborative data science environments, Ethical issues in using the internet, privacy and security, in the context of data science.	15
	Total Hours	30

Reference Books:

1. Data Science Ethics, David Martens ISBN: 9780192847263 Oxford University Press 2023
2. Ethics of Big Data: Balancing Risk and Innovation" by Kord Davis and Doug Patterson O'Reilly 2012.
3. Data Science Ethics Resources - Concepts, Techniques, and Cautionary Tales by David Martens- Oxford University Press-2022.
4. Data Science Association, Data Science Code of Professional Conduct.-
<https://www.datascienceassn.org/code-of-conduct.html>

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Elective		
Name of the Dept: Science and Technology		
Course Title: Human Resource Analytics		
Course Code: MDH206		Course Level:6.0
Type: Theory		
Course Credit: 2 credits		
Hours Allotted: 30 Hours		
Marks Allotted: 50 Marks		
Course Objectives: 1. To understand and analyse various HR problems and issues and evaluate the relevance of HR Analytics in solving them for effective decision-making. 2. To learn and apply different tools, methods, and techniques of HR Analytics to analyse HR data and understand real-world corporate scenarios. 3. To identify and study the applications of HR Analytics in different HR sub-systems such as recruitment, training, performance management, and employee retention.		
Course Outcomes : Upon completing this course, the student will be able to: CO1. Analyze problems and issues in HR and the relevance of HR analytics and logically synthesize the tools, methods and techniques of HR analytics to understand real world corporate scenario. CO2. Identify the application and uses of HR analytics in various HR sub-systems.		

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	HR Measurement Need for HR Measurement, Significance and concept of HR Analytics, HR Analytics and business linkages, Prerequisites of HR Analytics; Models and frameworks of HR Analytics; Measuring intellectual capital, need and rationale for HR Accounting & Audit, Approaches and methods of HR Accounting & Audit HRIS for HR Analytics: What is Human Resource Information System; Role of HRIS in analytics; HRIS development and Implementation, the development process- need analysis, systems design, structure and culture; HRIS Applications making HRIS work.	15
II	Analytics for HR sub-systems HR Analytics for Staffing, Training & Development, Performance Management Systems, Career Planning Systems, Rewards and Compensation Management, Employee Relations Systems. Analytics for HR system HR performance frameworks and measurement systems; Measuring HR Climate and People Management Capabilities; Competency Management Frameworks & Competency Mapping, Integration of competency-based HR System. Measuring HR Effectiveness, The HR Scorecard Trends and Future Challenges Technology and changes in HR Analytics, Role of social media, Big Data and Predictive Analytics in HR, Assessing the effectiveness of HR Analytics, Post analysis steps, Review and monitoring, Issues in HR valuation and measurement; Emerging challenges: Global and Indian Experience	15
	Total Hours	30

Reference Books:

1. Ulrich, D. & Brockbank, W., The HR Value Proposition. Harvard Business School Press 2016
2. How to measure HRM by Jac Fitz-enz 2002
3. Predictive Analytics for Human Resources by Jac Fitz-enz, John Mattox II, Wiley 2014
4. Making Human Capital Analytics Work: Measuring the ROI of Human Capital Processes and Outcomes. By by Jack Phillips, Patricia Pulliam Phillips- 2014

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Elective		
Name of the Dept: Science and Technology		
Course Title: Human Recourse Analytics Practical		
Course Code: MDHP207		Course Level:6.0
Type: Practical		
Course Credit: 2 credits		
Hours Allotted: 60 Hours		
Marks Allotted: 50 Marks		
Course Objectives: 1. To understand and analyse the concept of HR Analytics for improving workforce planning and organizational decision-making. 2. To learn and use various tools and techniques of HR Analytics for collecting, analysing, and interpreting HR data. 3. To identify the applications of HR Analytics in different HR sub-systems such as recruitment, training, performance management, and employee retention.		
Course Outcomes: Upon completing this course, the student will be able to: CO1. Analyze HR analytics and use and understand tools require for HR analytics. CO2. Identify the application and uses of HR analytics in various HR sub-systems.		

Syllabus: NEP 2020 w.e.f 2026-27

Sr. No.	Content	Hours
1	Analyze employee turnover rates and identify factors contributing to attrition • Collect historical employee data, including tenure, performance ratings, salary, and job satisfaction. • Calculate employee turnover rates for different departments and job roles. • Conduct statistical analysis to identify correlations between turnover and variables such as salary, job satisfaction, and performance. • Generate visualizations (e.g., charts, graphs) to present the findings and propose recommendations to reduce turnover.	
2	Develop a user-friendly HRIS dashboard for monitoring and analyzing HR metrics • Identify key HR metrics to be displayed on the dashboard (e.g., headcount, recruitment pipeline, training hours). • Design the layout and interface of the HRIS dashboard using appropriate programming languages and tools. • Integrate data from various HR systems and databases to populate the dashboard in real-time. • Implement interactive features, such as drill-down capabilities and data filters, to facilitate data exploration and analysis	
3	Analyze training effectiveness and identify skill gaps in the organization • Collect training data, including participant demographics, training modules, pre/post-assessment scores, and performance metrics. • Perform statistical analysis to evaluate the impact of training on employee performance. • Identify areas of improvement and recommend targeted training programs based on identified skill gaps. • Develop a visualization or report summarizing the training needs analysis results	
4	Develop an HR scorecard to measure HR effectiveness and align HR strategies with organizational goals • Identify key HR performance indicators aligned with the organization's strategic objectives. • Collect relevant data for each HR indicator, such as employee satisfaction surveys, training investment data, and performance metrics. • Calculate HR metrics and indicators, such as turnover rate, training ROI, and employee engagement index. • Design a dashboard or report to present the HR scorecard and analyze trends over time.	

5	Use predictive analytics to forecast employee attrition and develop retention strategies • Gather historical HR data, including employee demographics, performance metrics, compensation, and employee exit data. • Build a predictive model (e.g., logistic regression, decision tree) to predict employee attrition. • Validate the model's accuracy and evaluate its performance using appropriate evaluation metrics. • Generate actionable insights and recommendations to proactively address potential attrition risks.	
6	Use predictive analytics to forecast employee attrition and develop retention strategies • Gather historical HR data, including employee demographics, performance metrics, compensation, and employee exit data. • Build a predictive model (e.g., logistic regression, decision tree) to predict employee attrition. • Validate the model's accuracy and evaluate its performance using appropriate evaluation metrics. • Generate actionable insights and recommendations to proactively address potential attrition risks.	
7	Measure and analyze employee engagement levels within the organization • Collect employee engagement survey data, including responses to survey questions related to job satisfaction, work environment, and organizational culture. • Calculate engagement scores and identify key drivers of engagement. • Conduct a sentiment analysis on employee feedback to understand areas of improvement. • Present the findings and propose strategies to enhance employee engagement based on the analysis.	
8	Develop a program to automate repetitive HR processes, such as leave management or performance appraisal • Identify the HR process to be automated and define the required functionalities. • Design and implement a web-based application or script to streamline the process using appropriate programming languages and frameworks. • Integrate the application with relevant HR systems and databases to ensure data accuracy and consistency. • Test and validate the automated process, considering different scenarios and user inputs.	
9	Analyze the effectiveness of the organization's performance management system and provide insights for improvement. • Collect performance evaluation data, including performance ratings, goal achievement metrics, and feedback.	

	• Analyze the distribution of performance ratings across different departments or job roles. • Identify trends and patterns in performance data and assess the fairness and consistency of the evaluation process. • Propose recommendations for enhancing the performance management system based on the analysis results.	
10	Analyze the organization's compensation structure and compare it to industry benchmarks. • Gather salary data for different job roles and levels within the organization. • Perform a salary analysis, including measures like average salary, salary distribution, and salary competitiveness. • Conduct benchmarking by comparing the organization's salary data with industry standards or competitor data.	
	Total	60

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Elective		
Name of the Dept: Science and Technology		
Course Title: Public Health Analytics		
Course Code: MDP208		Course Level:6.0
Type: Theory		
Course Credit: 2 credits		
Hours Allotted: 30 Hours		
Marks Allotted: 50 Marks		
Course Objectives: 1. To understand the evolving landscape of healthcare services and the role of data and Electronic Health Record (EHR) systems in improving healthcare outcomes and decision-making. 2. To learn the importance of data analysis in public health and apply graphical and descriptive techniques to effectively summarize and interpret public health data. 3. To evaluate the reliability, accuracy, and integrity of health and genomic data for meaningful comparison and research in healthcare systems. 4. To develop basic skills in using software tools for data analysis to support healthcare research and decision-making. 5. To design data models that integrate patient information from multiple sources to create comprehensive and patient-centric healthcare insights.		
Course Outcomes (OC): Upon completing this course, the student will be able to: CO1: Understand the evolving healthcare ecosystem, the role of value-based healthcare, and the application of data in public health and electronic health record (EHR) systems, while identifying and evaluating reliable sources of health and genomic data. CO2: Apply data analysis techniques, graphical and descriptive methods, and software tools to analyze public health data and develop integrated data models for creating comprehensive, patient-centric healthcare solutions.		

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	Healthcare data management • What is Health Data Management? Benefits and challenges of health data management, how to store all that data. • Electronic Health Records– Components of EHR- Coding Systems- Benefits of EHR- Barrier to Adopting EHR Challenges • Phenotyping Algorithms • Statistical analysis of healthcare data- Measures of Central Tendency and Dispersion, Confidence Limits and Hypothesis Testing, Statistical Tests for Categorical Data, T-Tests for Related and Unrelated Data, Analysis of Variance. • Data Quality and Governance	15
II	Healthcare data Analysis: • Biomedical Image and Signal Analysis, Genomic Data Analysis for Personalized Medicine. • Natural Language Processing and Data Mining for Clinical Text, Mining the Biomedical- Social Media Analytics for Healthcare. • Predictive Models for Integrating Clinical and Genomic Data, Privacy Preserving Data, Publishing Methods in Healthcare. • Mobile Imaging and Analytics for Biomedical Data, Data Analytics for Pharmaceutical Discoveries- Clinical Decision Support Systems.	
	Total	30

Reference Books:

1. Healthcare Business Intelligence + Website - A Guide to Empowering Successful Data Reporting and Analytics Hardcover by LB Madsen (Author),2012
2. Practical Text Analytics: Interpreting Text and Unstructured Data for Business Intelligence (Marketing Science) 1st Edition by Dr. Steven Struhl,2016

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Elective		
Name of the Dept: Science and Technology		
Course Title: Public Health Analytics Practical		
Course Code: MDPP209		Course Level:6.0
Type: Practical		
Course Credit: 2 credits		
Hours Allotted: 60 Hours		
Marks Allotted: 50 Marks		
Course Objectives: 1. To develop practical skills in analyzing public health data using statistical tools and software for identifying health trends and patterns. 2. To apply data visualization and predictive techniques for evaluating health programs and supporting evidence-based public health decisions		
Course Outcomes: Upon completing this course, the student will be able to: CO1: Analyze, visualize, and interpret healthcare data using descriptive, diagnostic, and inferential statistical techniques with tools such as Python, R, or Excel to support evidence-based decision-making in healthcare. CO2: Apply regression, forecasting, machine learning, deep learning, and natural language processing (NLP) techniques to perform predictive analytics and extract meaningful insights from healthcare data.		

Syllabus: NEP 2020 w.e.f 2026-27

Sr. No.	Content	Hours
1	EHR data modeling, data mining, exploratory data analysis using tableau or power BI.	
2	Medical expenditure prediction	
3	A Twitter Healthcare data extraction, processing and sentiment analysis.	
4	Use keywords related to health (e.g., COVID-19 vaccine)	
5	Using classification algorithms in prediction of diseases	
6	Clinical Text Data Categorization and Feature Extraction	
7	Medical Image classification (x-rays)	
8	Image object detection using deep learning.	
9	Health forecasting using time-series.	
10	Heart sound classification using signal processing.	
	Total Hours	30

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Elective		
Name of the Dept: Science and Technology		
Course Title: Social Media Analytics		
Course Code: MDS210		Course Level:6.0
Type: Theory		
Course Credit: 2 credits		
Hours Allotted: 30 Hours		
Marks Allotted: 50 Marks		
Course Objectives: 1. To understand the fundamentals of social media platforms and their role in digital communication, marketing, and user engagement. 2. To explore tools and techniques for collecting, analyzing, and interpreting social media data from various channels. 3. To apply analytical methods to measure campaign performance, sentiment analysis, and audience behavior. 4. To use social media insights for strategic decision-making in business, branding, and public communication.		
Course Outcomes: Upon completing this course, the student will be able to: CO1: Understand the concepts, significance, and applications of social media analytics, and develop the ability to collect, analyse, and evaluate social media data for effective decision-making. CO2: Apply appropriate analytical and visualization techniques to assess the effectiveness of social media strategies while addressing ethical and legal issues related to the use of social media data.		

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	Unit 1: Social Media Analytics: An Overview Core Characteristics of social media, Types of social media, social media landscape, Need for Social Media Analytics (SMA), Seven Layers of Social Media Analytics, Types of Social Media Analytics, Social Media Analytics Cycle, Location Analytics - Sources of Location Data, Categories of Location Analytics, Social Information Filtering: Social Sharing and filtering, Automated Recommendation systems, Traditional Vs social Recommendation Systems, understanding social media and Business Alignment, social media KPI, formulating a Social Media Strategy, Managing Social Media Risks CaseStudy: Online Behavior on Twitter	15
II	Unit 2: Social Network Structure, Measures & Visualization Basics of Social Network Structure - Nodes, Edges & Tie Describing the Networks Measures - Degree Distribution, Density, Connectivity, Centralization, Tie Strength & Trust. Network Visualization - Graph Layout, Visualizing Network features, Scale Issues. Capturing Correlations: Triangles, Clustering, and Assortative. Social Media Network Analytics - Common Network Terms, Common Social Media Network Types, Types of Networks, Common Network Terminologies, Network Analytics Tools. Case Study: LinkedIn	15
	Total Hours	30

Reference Books:

1. Seven Layers of Social Media Analytics_ Mining Business Insights from Social Media Text, Actions, Networks, Hyperlinks, Apps, Search Engine, and Location Data, Gohar F. Khan, 2015
2. Analyzing the Social Web 1st Edition by Jennifer Golbeck, 2013
3. Mining the Social Web_ Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites, Matthew A Russell, O.,Reilly, 2019
4. Charu Aggarwal (ed.), Social Network Data Analytics, Springer, 2011

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical: Elective		
Name of the Dept: Science and Technology		
Course Title: Social Media Analytics Practical		
Course Code: MDSP211		Course Level:6.0
Type: Practical		
Course Credit: 2 credits		
Hours Allotted: 60 Hours		
Marks Allotted: 50 Marks		
Course Objectives: 1. To understand, visualize, and interpret healthcare data for meaningful insights and informed decision-making. 2. To apply statistical techniques such as linear regression and forecasting methods to predict growth rates, healthcare expenditure, and other numerical variables in the medical field. 3. To develop the ability to perform descriptive, diagnostic, and inferential statistical analysis using tools such as Python, R, or Excel. 4. To learn and apply machine learning and deep learning techniques for predictive analysis in healthcare data. 5. To understand and perform Natural Language Processing (NLP) and sentiment analysis for analyzing healthcare-related textual data.		
Course Outcomes: Upon completing this course, the student will be able to: CO1: Analyze, visualize, and interpret healthcare data using descriptive, diagnostic, and inferential statistical techniques, along with regression and forecasting methods, using tools such as Python, R, or Excel. CO2: Apply machine learning, deep learning, and natural language processing (NLP) techniques to perform predictive analytics and extract meaningful insights from healthcare data.		

Sr. No.	Content	Hours
1	Study Various • Social Media platforms (Facebook, twitter, YouTube etc) • Social Media analytics tools (Facebook insights, google analytics netlytic etc.) • Social Media Analytics techniques and engagement metrics (page level, post level, member level) using Gephi Tool.	
2	Scrape an online Social Media Site for Data. Use python to scrape information from twitter. Exploratory Data Analysis and visualization of Social Media Data.	
3	Create sociograms for the persons-by-persons network and the community- by- community network for a given relevant problem. Create a one-mode network and two-node network for the same. Datasets: les-Misérables, Airlines, Internet Core Routers.	
4	Develop Content (text, emoticons, image, audio, video) based social media analytics model for business. (e.g., Content Based Analysis: Topic, Issue, Trend, sentiment/opinion analysis, audio, video, image analytics)	
5	Develop Structure based social media analytics model for any business. (e.g., Structure Based Models -community detection, influence analysis)	
6	Develop a dashboard and reporting tool based on real time social media data Using Power BI.	
7	Use Google Visualization Charts to analyze social media data.	
8	Analyze social media data Network Analysis with Orange Software	
9	Use Graph Neural Networks on the datasets (Planetoid Cora Dataset) Jazz Musicians Network.	
10	Analyze Twitter conversations to identify the most active and influential users using Machine Learning Algorithms with Gephi Tool.	

Syllabus: NEP 2020 w.e.f 2026-27

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous)
Thane (W)

Programme Name: M.Sc (Data Science)		Semester: II
Course Category/Vertical:		
Name of the Dept: Science and Technology		
Course Title: On Job Training		
Course Code:		Course Level:6.0
Type:		
Course Credit: 4 credits		
Hours Allotted: 120 Hours		
Marks Allotted: 100 Marks		
Course Objectives: 1. To provide practical industry exposure by allowing students to apply academic knowledge in real-world work environments and develop job-specific skills. 2. To enhance professional competencies such as communication, teamwork, problem-solving, and adaptability through hands-on experience in a professional setting.		
Course Outcomes: Upon completing this course, the student will be able to: CO1. Apply data science theories, tools, and techniques to collect, clean, analyse real-world datasets and develop basic predictive models for solving industry-based problems. CO2. Demonstrate professional competence through effective teamwork, communication, ethical practices, and the ability to prepare technical reports and presentations in a real work environment.		

A. Introduction

- On Job training (OJT) is an integral component of the M.Sc. Data Science program that provides students with a unique opportunity to bridge the gap between theoretical knowledge gained in the classroom and practical application in a real-world environment. This training aims to equip students with both technical and non-technical skills that are essential for success in the industry.
- By participating in OJT, students are able to apply the concepts and theories learned during their coursework to real-world scenarios. They gain hands-on experience, problem-solving skills, and a deeper understanding of how the industry operates. This practical exposure enhances their competence and confidence, preparing them to tackle the challenges they may encounter in their professional careers.
- From an organizational perspective, hosting OJT programs allows companies to gain insights into the curriculum and content of the M.Sc. Data Science program. They can provide valuable feedback on the relevance of the coursework and industry requirements, enabling academic institutions to continually improve the program's alignment with industry needs. This collaboration between academia and industry fosters a mutually beneficial relationship, ensuring that graduates are well-prepared for the job market.
- Moreover, OJT benefits the faculty members involved in the program. They have the opportunity to gain firsthand exposure to the industry and observe the type of work being performed. This experience enables them to enhance their teaching methodologies and delivery techniques, ensuring that they remain up-to-date with the latest industry practices. The insights gained from OJT also enable faculty members to provide relevant guidance and mentorship to students, preparing them for successful careers in the field of data science.

B. Enhancing Practical Skills through OJT

- The On Job Training (OJT) program spans 4-6 weeks, requiring a minimum of 120 hours of physical presence at the organization.
- Students are expected to find their own OJT placements, although the institution provides support and guidance in securing positions with reputable organizations.
- OJT must be conducted outside the home institution to expose students to real-world work environments.
- OJT covers any subject within the syllabus, allowing students to align their experience with their academic interests.

- In recognition of changing dynamics, some OJT sessions can be conducted online to accommodate virtual work environments.
- OJT will offer students the opportunity to apply classroom learning in a real-world setting, fostering the development of technical and non-technical skills.
- Mutual Benefits: Organizations gain insights into the program's curriculum and industry requirements, enabling them to provide constructive feedback and enhance course relevance.
- OJT bridges the gap between theoretical knowledge and practical application, preparing students for successful careers in data science
-

C. Interning organization:

Students have the flexibility to pursue their OJT in various types of organizations, including but not limited to:

- Software Development Firms: Gain practical experience in software development and programming.
- Hardware/Manufacturing Firms: Learn about hardware design, manufacturing processes, and quality assurance.
- Small-Scale Industries/Service Providers: Explore opportunities in diverse sectors such as banking, clinics, NGOs, and professional institutions like CA firms or law firms.
- Civic Departments: Engage with local civic departments such as ward offices, post offices, police stations, or panchayats to understand their functioning and contribute to their activities.
- Research Centre"s/University Departments/Colleges: Contribute as research assistants or in similar roles for research projects or initiatives, fostering collaboration between academia and industry.

Note: The listed options provide a range of possible OJT placements, offering students valuable exposure to different sectors and professional settings.

D. OJT mentors:

To enhance the learning experience and ensure the quality of the MSc program, each student participating in the OJT will be assigned two mentors: a faculty mentor from the institution and an industry mentor from the organization where the student is interning.

- **Industry Mentor Role:** The industry mentor plays a crucial role in guiding the student during the internship. They ensure that the internee fulfills the requirements of the organization and successfully meets the demands of the assigned project. Through their expertise and experience, industry mentors provide valuable insights into real-world practices and industry expectations.
- **Faculty Mentor Role:** The faculty mentor serves as the overall coordinator of the OJT program. They oversee the entire internship

process and evaluate the quality of the OJT in a consistent manner across all students. The faculty mentor ensures that the OJT aligns with the program's objectives and provides valuable learning opportunities. They also facilitate communication between the institution, industry mentor, and student to ensure a fruitful OJT experience.

By having both an industry mentor and a faculty mentor, students benefit from a Comprehensive guidance system that combines industry expertise and academic support. This dual mentoring approach ensures a well-rounded and rigorous OJT experience for every student in the program

E. Submission of documentation for OJT

The student will make two documents as part of the OJT

- **Online diary:** This ensures that the student updates daily activity, which could be accessed by both the mentors. Daily entry can be of 3- 4 sentences giving a very brief account of the learning/activities/interaction taken place. The faculty mentor will be monitoring the entries in the diary regularly as shown in **Appendix-I**
- **OJT report:** A student is expected to make a report based on the OJT he or she has done in an organization. It should contain the following:
 - **Certificate:** A certificate in the prescribed Performa (given in **Appendix II** and **Appendix III**) from the organization where the OJT was done.
 - **Title:** A suitable title giving the idea about what work the student has performed during the OJT.
 - **Description of the organization:** A small description of the organization where the student has interned
 - **Description of the activities** done by the section where the intern has worked: A description of the section or cell of the organization where the intern worked. This should give an idea about the type of activity a new employee is expected to do in that section of the organization.
 - **Description of work allotted and done by the intern:** A detailed description of the work allotted, and actual work performed by the intern during the OJT period. It shall be the condensed and structured version of the daily report mentioned in the online diary.
 - **Self-assessment:** A self-assessment by the intern on what he or she has learned during the OJT period. It shall contain both technical as well as interpersonal skills learned in the process.

F. Interaction between mentors:

To ensure the smooth conduct of the OJT a meet-up involving the intern, industry mentor, and the faculty mentor will be scheduled as a mid-term review. The meeting can preferably be online to save time and resources. The meeting ensures the synergy between all stakeholders of the OJT. A typical meeting can be of around 15 minutes where at the initial stage the intern brief about the work and interaction goes for about 10 minutes. This can be followed by the interaction of the mentors in the absence of the intern. This ensures that issues between the intern and the organization, if any, are resolved amicably.

G. OJT workload for the faculty:

Every student is provided with a faculty member as a mentor. So, a faculty mentor will have a few students under him/her. A faculty mentor is the overall in charge of the OJT of the student. He/she constantly monitors the progress of the OJT by regularly overseeing the diary, interacting with the industry mentor, and guiding on the report writing etc. Considering the time and effort involved, a faculty mentor who is in-charge of 20 students shall be provided by a workload of 3 hours

Evaluation Scheme

Theory courses of 4 credits: Total marks 100. Out of the total, 50% each for internal and external evaluation.

A. Internal Evaluation (30m + 10m + 10m = 50 Marks)

The internal assessment marks shall be awarded as follows:

1. 30 marks (Any one of the following):

- a. Written Test of 30 Marks
- b. SWAYAM (Advanced Course) of minimum 20 hours and certification exam completed or
- c. NPTEL (Advanced Course) of minimum 20 hours and certification exam completed or
- d. Valid International Certifications (Prometric, Pearson, Certiport, Coursera, Udemy and the like)
- e. Certification marks of one completed exam shall be awarded to one course only. For four courses, the students will have to complete four certifications.

(Note: Only those certification/courses suggested by the department shall be deemed valid, Student cannot do any certification on their own)

2. 10 marks

10 marks from every course (Two 4 credits mandatory courses, one 2 credits mandatory course, one 4 credits elective course) coming to a total of 40 marks, shall be awarded on publishing of research paper in UGC approved / Other Journal with plagiarism less than 15%. The marks can be awarded as per the impact factor of the journal, quality of the paper, importance of the contents published, social value.

3. 10 marks

Open Book examination based on problem solving related to the respective subject.

Suggested format of Question paper of 20 marks for the written test.

Q1. Attempt any two of the following: 16 marks

- a.
- b.
- c.
- d.

Q2. Attempt any two of the following: 14 marks

- a.
- b.
- c.

d.

B. External Examination: (50 marks) Duration : 2 hrs

All questions are compulsory

Q1 (Based on all units) Attempt any two of the following: 10 marks

- a. Unit 1
- b. Unit 2
- c. Unit 3
- d. Unit 4

Q2 (Based on Unit 1) Attempt any two of the following: 10 marks

Q3 (Based on Unit 2) Attempt any two of the following: 10 marks

Q4 (Based on Unit 3) Attempt any two of the following: 10 marks

Q5 (Based on Unit 4) Attempt any two of the following: 10 marks

Theory courses of 2 credits: Total marks 50. Out of the total, 50 % each for internal and external evaluation.

A. Internal Evaluation (25 Marks)

The internal assessment marks shall be awarded as follows:

1. 10 marks from every course (Two 4 credits mandatory courses, One 2 credits mandatory course, One 4 credits elective course) coming to a total of 40 marks, shall be awarded on publishing of research paper in UGC approved / Other Journal with plagiarism less than 15%.
The marks can be awarded as per the impact factor of the journal, quality of the paper, importance of the contents published, social value.
2. 10 marks - Open Book examination based on problem solving related to the respective subject.
3. 5 marks - Assignment/Group discussion.

B. External Examination: (25 marks) Duration : 1 hr

All questions are compulsory

Q1 (Based on Unit 1) Attempt any two of the following: 13 marks

Q2 (Based on Unit 2) Attempt any two of the following: 12 marks

Practical courses of 2 credits: Total marks 50. Out of the total, 50 % each for internal and external evaluation.

A. Practical Evaluation Internal (25 marks)

- | | |
|---|----|
| 1. Performance during all practical sessions | 10 |
| 2. Problem solving with the acquired programming skills | 10 |
| 3. Viva Voce | 05 |

B. Practical Evaluation External (25 marks)

A Certified copy of hard-bound journal is essential to appear for the practical examination.

- | | |
|-----------------------|----|
| 1. Practical Question | 15 |
| 2. Journal | 05 |
| 3. Viva Voce | 05 |

Practical courses of 4 credits: Total marks 100. Out of the total, 50 % each for internal and external evaluation.

A. Practical Evaluation Internal (50 marks)

- | | |
|---|----|
| 1. Performance during all practical sessions | 20 |
| 2. Problem solving with the acquired programming skills | 25 |
| 3. Viva Voce | 05 |

B. Practical Evaluation External (50 marks)

A Certified copy of hard-bound journal is essential to appear for the practical examination.

- | | |
|-----------------------|----|
| 1. Practical Question | 20 |
| 2. Practical Question | 20 |
| 3. Journal | 05 |
| 4. Viva Voce | 05 |