COURSE STRUCTURE AND EVALUATION SCHEME
BS-MS (MATHEMATICS AND DATA SCIENCE)

Year IV, Semester- VII

S. No.	Course Type	Course Title	Subject Code	Credits	Periods			Sessional Marks				ESE	Total Marks
					L	T	P	MSE	TA	Lab	Total		
1	PEC-II			4	3	1	0	30	20	-	50	50	100
2	PEC-III			3	3	0	0	30	20	-	50	50	100
3	PEC-IV			3	3	0	0	30	20	-	50	50	100
4	OEC-II			2	2	0	0	30	20	-	50	50	100
5	BS Project-I		NMA429	8	0	0	16	-	50	-	50	50	100
6	Internship/ Seminar		NMA431	2	0	0	4	-	50	-	50	50	100
	Total Credits: 22												**600**

PEC-II: Mathematical Methods (NMA411), Bayesian Analysis (NMA413), Computer Vision (NMA415**)**

PEC-III: Graph Theory (NMA417), Time Series Analysis (NMA419), Natural Language Processing (NMA421)

PEC-IV: Convex Optimization (NMA423), Stochastic Process & Applications (NMA425), Design and Analysis of Algorithms (NMA427)

Year IV, Semester- VIII (BS (HONORS))

S. No.	Course Type	Course Title	Subject Code	Credits	Periods			Sessional Marks				ESE	Total Marks
					L	T	P	MSE	TA	Lab	Total		
1	PCC	Differential Equations	NMA412	4	3	0	2	15	20	15	50	50	100
2	PCC	Data Mining	NMA414	3	2	0	2	15	20	15	50	50	100
3	PEC-V			3	3	0	0	30	20	-	50	50	100
4	OEC-III			2	2	0	0	30	20	-	50	50	100
5	BS Project-II		NMA422	10	0	0	20		50	-	50	50	100
	Total Credits: 22												**500**

PEC-V: Fluid Mechanics (NMA416), Pattern Recognition (NMA418), Recommender Systems with Python (NMA420)

OR

Year IV, Semester- VIII (BS (HONORS With RESEARCH))

S. No.	Course Type	Course Title	Subject Code	Credits	Periods			Sessional Marks				ESE	Total Marks
					L	T	P	MSE	TA	Lab	Total		
1	PEC-V			4	3	1	0	30	20	-	50	50	100
2	OEC-III			2	2	0	0	30	20	-	50	50	100
3	BS Project-II		NMA422	16	0	0	32		200	-	200	200	400
	Total Credits: 22												**600**

PEC-V**:** Fluid Mechanics (NMA416), Pattern Recognition (NMA418), Recommender Systems with Python (NMA420)

COMPUTER VISION

Course Code: NMA415 **L-T-P-C: 3-1-0-4**

Prerequisites: A basic understanding of optics and images.

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	understand the fundamental concepts, history, and goals of Computer Vision, including image formation, camera models and geometric transformations.
CO 2	apply basic image processing techniques such as filtering, transformations, segmentation and color processing to analyze digital images.
CO 3	extract and analyze image features including texture, color, edges and shape for effective representation and interpretation.
CO 4	implement the concepts of linear filters and shift invariant linear systems to imaging systems.
CO 5	analyze and apply computer vision techniques in real-world applications such as medical imaging, object tracking, face recognition and gesture analysis.

UNIT I Fundamentals **8 Hours**

Introduction and Goals of Computer Vision, A Brief History, Image Formation, Radiometric Quantities, Photometric Stereo, Geometric Transformations, Single Camera Setup of Image Formation, The Digital Camera.

UNIT II Image Processing **9 Hours**

Basics of Image Processing, Point Operators, Geometric Operations, Spatial or Neighbourhood Operations, Operations between Images, Image Transforms, Image Filtering, The Physics of Colour, Human Colour Perception, Colour Image Processing, Mathematical Morphology, Image Segmentation.

UNIT III Image Features **8 Hours**

Texture Descriptors, Texture Representation Methods, Gabor Filter, Colour Features, Gradient-Based Methods for Edge Detection, Object Boundary and Shape Representations, Chain Code and Shape Number, Interest or Corner Point Detectors.

UNIT IV Linear Filters **8 Hours**

Linear filters and Convolution, Shift invariant linear systems, Spatial frequency and Fourier transform, Sampling and aliasing, Scale and image pyramids, Real imaging system vs Shift-Invariant linear systems.

UNIT V Applications of Computer Vision **7 Hours**

Medical Image Segmentation, Motion Estimation and Object Tracking, Face and Facial Expression Recognition, Gesture Recognition, Image Fusion.

Textbooks

1. Richard Szeliski , Computer Vision: Algorithms and Applications, Springer, (1st Edition), 2023.
2. M. K. Bhuyan, Computer Vision and Image Processing – Fundamentals and Applications, CRC press, USA, (1st Edition), 2020.

Reference Books

1. Forsyth and Pance, Computer Vision (A Modern Approach), Pearson Education, (2nd Edition), 2015.
2. Simon J. D. Prince, Computer Vision: Models, Learning, and Inference (1st Edition), 2012.

NATURAL LANGUAGE PROCESSING

Course Code: NMA421 **L-T-P-C: 3-0-0-3**

Prerequisites: An undergraduate-level understanding of learning algorithms, syntax and grammar.

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO 1	understand the fundamental concepts, challenges, components and applications of Natural Language Processing.
CO 2	apply language modeling techniques and perform basic linguistic analysis including POS tagging and named entity recognition.
CO 3	analyze and implement word representation techniques and methods for semantic understanding and disambiguation.
CO 4	apply text analytics and information extraction techniques for classification, sentiment analysis and multilingual information retrieval.
CO 5	understand and compare different machine translation approaches, including statistical and neural methods.

UNIT I Introduction **8 Hours**

Rationalist and Empiricist Approaches to Language, The Ambiguity of Language: Why NLP is Difficult? Lexical resources, Zipf's Law, Advantages and Disadvantages of NLP, Components of NLP, Applications of NLP, NLP pipeline and Phases, NLP APIs and Libraries.

UNIT II Statistical Language Modeling and Linguistic Analysis **8 Hours**

Unigram, Bigram, Trigram, N-gram Models; Advanced Smoothing Techniques and Their Empirical Comparison; Applications of Language Modeling; Natural Language Generation; Part-of-Speech Tagging; Morphology; Named Entity Recognition.

UNIT III Word Representation and Lexical Semantics **8 Hours**

Bag of Words, Skip-gram, Continuous Bag-of-Words (CBOW), Word Embeddings, Lexical Semantics, Word Sense Disambiguation, Knowledge-Based and Supervised WSD.

UNIT IV Text Analytics and Information Extraction **8 Hours**

Sentiment Mining, Text Classification, Information Extraction, Relation Extraction, Question Answering in Multilingual Settings, NLP in Information Retrieval, Cross-Lingual Information Retrieval

UNIT V Machine Translation and Neural Approaches **8 Hours**

Need and Challenges of Machine Translation, MT Approaches, Direct Machine Translation, Rule-Based Machine Translation, Knowledge-Based MT Systems, Statistical Machine Translation (SMT), Parameter Learning in SMT using IBM Models and EM Algorithm, Encoder–Decoder Architecture, Neural Machine Translation.

Textbooks

1. Christopher D. Manning and Hinrich Schutze, Foundations of Statistical Natural Language Processing, Cambridge, MA: MIT Press, (1st Edition), 1999.
2. David Jurafsky and James H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition, Pearson, (2nd Edition) 2014.

Reference Books

1. Steven Bird, Ewan Klein, and Edward Loper, Natural Language Processing with Python-Analyzing Text with the Natural Language Toolkit, O'Reilly Media, Inc., (1st Edition), 2009.
2. Alexander Clark, Chris Fox and Shalom Lappin, The Handbook of Computational Linguistics and Natural Language Processing, Wiley-Blackwell, (1st Edition), 2012

STOCHASTIC PROCESSES AND APPLICATIONS

Course Code: NMA 425 **L-T-P-C: 3-0-0-3**

Prerequisites: A solid foundational background in probability, linear algebra, and calculus.

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to –

CO1	Explain and classify stochastic processes and analyze their statistical properties.
CO2	Model and analyze Poisson and renewal processes in real-world applications.
CO3	Formulate and solve discrete and continuous-time Markov chain models.
CO4	Apply stochastic processes to queueing systems and reliability analysis.
CO5	Implement stochastic simulation techniques for solving data science problems.

Unit I: Fundamentals of Stochastic Processes **8 Hours**
Definition and examples of stochastic processes; *Classification:* discrete/continuous time and state space; Mean, variance, autocorrelation function; Stationary and weakly stationary processes; Random-walk models. *Applications in* Time-series randomness, modelling noisy signals, Monte Carlo intuition.

Unit II: Poisson and Renewal Processes **9 Hours**
Poisson process: definition and properties; Exponential inter-arrival times, Superposition and splitting, Non-homogeneous Poisson process, Renewal process and renewal equation. *Applications in* Event modelling (clickstreams, arrivals, system failures).

Unit III: Markov Chains **10 Hours**
Discrete-Time Markov Chains (DTMC), Transition probability matrix, Chapman-Kolmogorov equations, Classification of states (transient, recurrent, absorbing), Steady-state and limiting distributions, Continuous-Time Markov Chains (CTMC), Birth-death processes. *Applications in* Page ranking models, customer behaviour modelling, queueing transitions.

Unit IV: Queueing Theory and Reliability Models **9 Hours**
Basic queueing structure, Kendall's notation, M/M/1 and M/M/c models, Performance measures: L, W, Lq, Wq; Little's theorem, Reliability concepts, Series and parallel systems. *Applications in* Cloud computing resource allocation, server farms, network traffic.

Unit V: Gaussian Processes and Stochastic Simulation **9 Hours**
Gaussian processes and properties, Brownian motion and Wiener process, Introduction to stochastic differential equations (basic idea only), Monte Carlo simulation, Simulation of Poisson and Markov processes, Error analysis and convergence. *Applications in* Financial modelling, ML kernel methods, simulation-based inference.

Textbooks

1. Cont, R., & Tankov, P., Stochastic Processes with Applications to Finance, CRC Press, UK, 1st Edition, 2004.
2. Castañeda, L. B., Arunachalam, V., & Dharmaraja, S., Introduction to Probability and Stochastic Processes with Applications, Wiley, India, Asian Edition, 2012.
3. Yates, R. D., & Goodman, D. J., Probability and Stochastic Processes: An Indian Adaptation, Wiley, India, 3rd Edition, 2015.

Reference Books

1. Ross, S. M., Stochastic Processes, Wiley, India, 2nd Edition, 1996.
2. Trivedi, K. S., Probability and Statistics with Reliability, Queueing and Computer Science Applications, John Wiley & Sons, USA, 2nd Edition, 2001.
3. Goswami, A., & Rao, B. V., A Course in Applied Stochastic Processes, Hindustan Book Agency, India, 1st Edition, 2006.

Differential Equations

Course Code: NMA412 **L-T-P-C:** 3-0-2-4

Prerequisites: An undergraduate-level understanding of calculus, linear algebra, ordinary differential equations, real, complex and Fourier analysis.

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have the knowledge/skills/competency to:

CO1	Show proficiency in examining the existence and uniqueness of solutions to IVPs
CO2	Demonstrate expertise in the analysis of and methods to solve non-linear differential equations and in examining the stability of linear and non-linear systems of ODEs,
CO3	Explain how the analytical solutions to ODE-BVPs are obtained by different methods
CO4	Formulate and solve PDE-BVs involving the Laplace equation
CO5	Demonstrate understanding of different methods to solve the wave equation -a parabolic PDE.

Detailed Syllabus

Unit-I: Existence and uniqueness of solutions **9 Hours**
Lipschitz condition, Gronwall inequality, successive approximation, Picard's theorem. Existence of solutions in the large systems. Fixed point method, Systems of linear differential equations: nth order equation as a first order System, System of first-order equations, fundamental matrix, non-homogeneous linear systems, Linear Equations with constant coefficients

Unit-II: Non-linear Differential Equations **9 Hours**
Extremal solutions, Upper and lower solutions, Monotone iterative method and method of quasi linearization, Stability of Linear and Nonlinear systems: Critical points, System of equations with constant coefficients, Linear equations with constant coefficients, Lyapunov stability.

Unit-III: **Boundary value problem for ordinary differential equations** **8 Hours**
Sturm-Liouville problem, Eigenvalue, and Eigen functions, Orthogonality of eigen functions, Expansion in Eigen functions, Green's function for boundary value problems.

Unit-IV: The Laplace equation **7 Hours**

Boundary value for Laplace's equation, fundamental solution, Integral representation and mean value formula for harmonic functions, Green's function for Laplace's equation, Solution of the Dirichlet problem for a ball, solution by separation of variables.

Unit-V: The wave equation **8 Hours**

Solution by the method of separation of variables, D'Alembert solution of the wave equation, Solution of the wave equation by the Fourier transform method.

LAB WORK: Implementation of any 2-3 algorithms from each UNIT (selected by the instructor) using MATLAB/MAPLE.

Text Books

1. K. Sankara Rao, Introduction to Partial Differential Equations, PHI Learning Private Limited (3rd Edition), 2011.
2. S.L. Ross: Differential Equations, Wiley (3rd Edition), 2018.

Reference Books

1. Earl A. Coddington, Norman Levinson, The Theory of Ordinary Differential Equations, Krieger Pub Co., (1st Edition), 1984.
2. Tyn Mint-U, Lokenath Debnath, Linear Partial Differential Equations for Scientists and Engineers, Birkhauser Boston Inc., (4th Edition), 2007.
3. Lawrence Evans, Partial Differential Equations, Graduate Studies in Mathematics, AMS, Volume 19 (2nd Edition), 2010
4. S.G. Deo, V. Laxmikantham, and V. Raghvendra, Textbook of Ordinary Differential Equations, McGraw-Hill Education (India) Pvt. Ltd. (2nd Edition), 2002.
5. Shair Ahmad and M.R.M. Rao, Theory of Ordinary Differential Equations with Applications in Biology and Engineering, Affiliated East-West Press Private Ltd., 1999.

Data Mining

Course Code: NMA414 **L-T-P-C: 2-0-2-3**

Prerequisites: Basic concepts of Probability and Statistics, Linear Algebra, Python, Data Structures and Algorithms.

Course Outcomes: On satisfying the requirements of the course and upon its successful completion, students will have knowledge/skills/competency to-

CO1	understand the concepts of data mining, data preprocessing techniques, and the role of data mining in knowledge discovery.
CO2	apply association rule mining techniques and frequent pattern analysis to discover relationships in large datasets.
CO3	analyze and implement classification and prediction models using machine learning algorithms.
CO4	apply clustering algorithms to identify patterns and structures in unlabeled datasets.
CO5	evaluate advanced data mining techniques including outlier detection, text mining, and real-world applications.

UNIT-I: Introduction to Data Mining and Data Preprocessing 9 Hours

Introduction to Data Mining, Data Mining vs Data Warehousing, Knowledge Discovery in Databases (KDD) process, Data Mining Functionalities and Applications. Types of Data and Data Sources. Data Preprocessing: Data Cleaning, Data Integration, Data Transformation, Data Reduction, Data Discretization. Data Warehousing Concepts and OLAP operations.

UNIT-II: Frequent Pattern Mining and Association Rules 9 Hours

Basic Concepts of Frequent Patterns and Association Rules. Market Basket Analysis. Apriori Algorithm and FP-Growth Algorithm. Association Rule Generation and Rule Evaluation Measures: Support, Confidence, Lift. Constraint-based Association Mining.

UNIT-III: Classification and Prediction 8 Hours

Basic Concepts of Classification and Prediction. Decision Tree Induction and Attribute Selection Measures. Bayesian Classification and Naive Bayes Classifier. k-Nearest Neighbor (k-NN) Method. Model Evaluation and Validation Techniques: Confusion Matrix, Cross Validation, Accuracy, Precision, Recall.

UNIT-IV: Clustering Techniques 7 Hours

Introduction to Clustering and Cluster Analysis. Types of Clustering Methods. Partitioning Methods: k-Means, k-Medoids. Hierarchical Clustering: Agglomerative and Divisive Methods. Density-based Clustering (DBSCAN). Cluster Evaluation Measures.

UNIT-V: Advanced Data Mining and Applications 7 Hours

Outlier Detection and Anomaly Detection. Sequential Pattern Mining. Text Mining and Web Mining. Social Network Analysis. Applications of Data Mining in Finance, Healthcare, E-commerce, and Scientific Data.

DATA MINING LAB (Python/R)

Practical exercises using Python (NumPy, Pandas, Scikit-learn) or R for implementing data mining techniques.

1. Installation and introduction to Python/R data science libraries.
2. Data preprocessing: handling missing values, normalization, and transformation.
3. Data visualization and exploratory data analysis.
4. Implementation of the Apriori algorithm for association rule mining.
5. Classification using decision trees.
6. Implementation of Naive Bayes classifier.
7. k-Nearest Neighbor classification.
8. Clustering using the k-Means algorithm.
9. Hierarchical clustering implementation.
10. Density-based clustering (DBSCAN).
11. Evaluation of classification models using a confusion matrix and cross-validation.
12. Mini project using a real-world dataset.

Textbooks

1. Jiawei Han, Micheline Kamber, Jian Pei, Data Mining: Concepts and Techniques, Morgan Kaufmann (3rd Edition), 2012.

2. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, Introduction to Data Mining, Pearson (2nd Edition), 2019.

Reference Books

1. Ian H. Witten, Eibe Frank, Mark A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, (2nd Edition), 2005.

2. Charu C. Aggarwal, Data Mining: The Textbook, Springer, (1st Edition), 2015.

3. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Springer, (2nd Edition), 2009.

Pattern Recognition

Course Code: NMA418 **L-T-P-C:** 3-0-2-4

Pre-requisites: An undergraduate level understanding of probability, statistics and linear algebra is assumed. A basic knowledge of MATLAB will be useful.

Course Outcomes (COs):
On successful completion of the course, the students will be able to:

CO1	Apply fundamental concepts and Bayesian decision theory techniques to design and analyse Pattern Recognition systems.
CO2	Implement and evaluate statistical pattern recognition methods using parametric, non-parametric, and classification techniques.
CO3	Apply feature extraction, feature selection, and clustering techniques for efficient pattern analysis and recognition.
CO4	Analyze and implement structural and syntactic pattern recognition approaches using graph- and sequence-based methods.
CO5	Develop neural and hybrid Pattern Recognition models for solving real-world recognition problems using Python implementations.

Detailed Syllabus

Unit - I: Introduction **10 Hours**
Pattern Recognition: Intro, Basic Concepts and Design Principles, Design Cycle of Pattern Recognition system, Learning Paradigms, Examples, Approaches, Applications.
Bayesian Decision Theory: Bayes Decision Rule, Loss Function, Decision Regions and Boundaries, Discriminant Function for: multivariate Gaussian density and Discrete Features, Minimum Error Rate Classifier.

Unit - II: Statistical Pattern Recognition – I **10 Hours**
Parametric Techniques: ML and MAP Estimation, Bayesian Estimation, Logistic Regression, Naive Bayes, LDA, QDA, Linear SVM, HMMs and Variants.
Non-parametric Techniques: Density Estimation, Parzen Windows, Histogram method, k-NN Density Estimation, PNN, Non-parametric HMM Approaches.
Classification Techniques: k-NN and Fuzzy classification, Reduced Coulomb Energy Networks.

Unit - III: Statistical Pattern Recognition – II **10 Hours**
Feature Extraction: Overview of PCA, LDA, t-SNE, ICA and their use in PR. BoW, TF-IDF, Fourier and Wavelet Transforms, HOG.
Feature Selection: Overviews of Filter, Wrapper and Embedded methods, their Application and Python Implementation in PR.
Clustering Pattern Recognition: Criterion Functions for clustering; Algorithms: Overview and use of K-means, Hierarchical and other Algorithms in PR; GMMs: Expectation-Maximization method for Parameter Estimation, Maximum Entropy Estimation; Cluster validation.
Unit - IV: Structural Pattern Recognition (SPR) **8 Hours**
Introduction, Basic Concepts, Benefits and Example Applications, Components, Learning, Algorithms/Approaches for SPR. Illustration of Workflow of any 3-4 Important/Popular

Algorithms (GED, DTW, GNNs, GCNs etc with their Python Implementations). Syntactic Pattern Recognition.

Unit - V: Neural and Hybrid Pattern Recognition **7 Hours**

Architecture, Illustration of Pattern Recognition mechanism & Application, with Python Implementation of: CNNs, RRNs, LSTMs, GRUs, Autoencoders, and CNN-SVM, HMM-NN, DT-DL, CNN-HMM, SOM-KNN etc.

LAB: Implementation of any 2-3 algorithms from each UNIT (selected by the instructor)

TEXT BOOKS

1. R.C. Gonzalez, R. Woods, Digital Image Processing, Generic, 2018.
2. R.O. Duda, P.E. Hart and D.G. Stork, Pattern Classification, Wiley and Sons, 2001.
3. S. Theodoridis and K. Koutroumbas, Pattern Recognition, Academic Press, 2008.

REFERENCE BOOKS

1. K. Jain, Fundamentals of digital image processing, PHI, 2015.
2. D. Koller, N. Friedman, Probabilistic Graphical Models: Principles and Techniques, MIT Press, 2009.
3. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2009.
4. P.A Devijver and J. Kittler, Pattern Recognition: A Statistical Approach, Prentice-Hall International, Englewood Cliffs, NJ, 1982.
5. K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd Ed. Academic Press, New York, 2013.