APJ Abdul Kalam Technological University

Cluster 6: Ernakulam I

M. Tech Programme in Artificial Intelligence and Data Science

Scheme of Instruction and Syllabus: 2020 Admissions

APJ Abdul Kalam Technological University
(Cluster 6: Ernakulam I)

M. Tech Programme in Artificial Intelligence and Data Science

Scheme of Instruction

Credit requirement: 68 credits (23+19+14+12)

Normal Duration - Regular: 4 semester **Maximum Duration -** Regular: 6 semester

Courses - Core courses: Either 4 or 3 credits courses; **Elective courses:** All of 3 credits

Allotment of credits and examination scheme

Semester I (Credits: 23)

Exam Slot	Course Code	Subjects	L-T-P	Internal Marks	End Semester Exam		Credits
					Marks	Duration	
A	06 DS 6 01 1	Linear Algebra and Optimization	4-0-0	40	60	3 hrs	4
B	06 DS 6 02 1	Stochastic Models and Numerical Optimization	4-0-0	40	60	3 hrs	4
C	06 DS 6 03 1	Principles of Artificial Intelligence and Machine Learning	4-0-0	40	60	3 hrs	4
D	06 DS 6 04 1	Exploration and Statistical Analysis for Data Science	3-0-0	40	60	3 hrs	3
E	06 DS 6 x5 1	Elective I	3-0-0	40	60	3 hrs	3
F	06 DS 6 06 1	Research Methodology	0-2-0	100			2
S	06 DS 6 07 1	Seminar 1	0-0-2	100			2
U	06 DS 6 08 1	Machine Intelligence Lab	0-0-3	100			1
							23 Credits

Elective I	
Course Code	**Subjects**
06 DS 6 15 1	Ensemble Models
06 DS 6 25 1	Convolutional Neural Network
06 DS 6 35 1	Soft Computing
06 DS 6 45 1	Computer Vision

Semester II (Credits: 19)

Exam Slot	Course Code	Subjects	L-T-P	Internal Marks	End Semester Exam		Credits
					Marks	Duration	
A	06 DS 6 01 2	Big Data Analytics *	4-0-0	40	60	3 hrs	4
B	06 DS 6 02 2	Deep Learning and Artificial Neural Network	3-0-0	40	60	3 hrs	3
C	06 DS 6 03 2	Genetic Algorithms	3-0-0	40	60	3 hrs	3
D	06 DS 6 x4 2	Elective II	3-0-0	40	60	3 hrs	3
E	06 DS 6 x5 2	Elective III	3-0-0	40	60	3 hrs	3
F	06 DS 6 06 2	Mini Project	0-0-4	100			2
U	06 DS 6 07 2	Deep Learning Lab	0-0-3	100			1
							19 Credits

Elective II	
Course Code	**Subjects**
06 DS 6 14 2	R for Data Science
06 DS 6 24 2	Data Analytics and Scalable Algorithms
06 DS 6 34 2	Scalable Systems for Data Science
06 DS 6 44 2	Knowledge Engineering and Data Science

Elective III	
Course Code	**Subjects**
06 DS 6 15 2	Big Data for Internet of Things
06 DS 6 25 2	Artificial Intelligence and Robotics
06 DS 6 35 2	Natural Language Processing
06 DS 6 45 2	Machine Learning Models and Storage Management

* - Subject common to M.Tech Data Science/M.Tech Artificial Intelligence and Data Science

Semester III (Credits: 14)

Exam Slot	Course Code	Name	L-T-P	Internal Marks	End Semester Exam		Credits
					Marks	Duration	
A	06 DS 7 x1 1	Elective IV	3-0-0	40	60	3 hrs	3
B	06 DS 7 x2 1	Elective V	3-0-0	40	60	3 hrs	3
S	06 DS 7 03 1	Seminar II	0-0-2	100			2
F1	06 DS 7 04 1	Project Phase I	0-0-8	50			6
							14 Credits

Elective IV	
Course Code	**Subjects**
06 DS 7 11 1	Artificial Intelligence in Cyber Security
06 DS 7 21 1	Game Theory in Artificial Intelligence
06 DS 7 31 1	Image and Video Analytics
06 DS 7 41 1	Cloud Data Management

Elective V	
Course Code	**Subjects**
06 DS 7 12 1	Data Visualisation Techniques
06 DS 7 22 1	Social Network Analysis
06 DS 7 32 1	Text Mining
06 DS 7 42 1	Data Warehouse and Data Lakes

Semester IV (Credits: 12)

Exam Slot	Course Code	Name	L-T-P	Internal Marks	End Semester Exam		Credits
					Marks	Duration	
F2	06 DS 7 01 2	Project Phase II	0-0-21	70	30		12
							12 Credits

SEMESTER I

M. Tech Programme in Artificial Intelligence and Data Science

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6011	**Linear Algebra and Optimization**	**4-0-0 : 4**	**2020**

Course Objectives:

1. To visualise vectors in n-space which is useful in representing data
2. To learn handling of linear system of equations using matrix as a tool
3. To introduce Eigen values and Eigen vectors which are significant in dynamic problems
4. To introduce matrix decompositions methods that reduce a matrix into constituent parts which make it easier to calculate more complex matrix operations
5. To study optimization algorithms with single and multi-variables for large datasets

SYLLABUS:
Vector spaces, subspaces, bases and dimensions, systems of linear equations, Linear transformations, Isomorphism, Inner product, Orthogonality, Eigenvalues and Eigenvectors, Matrix factorisations, Function optimization, Newton's method.

Course Outcomes:

Students should be able to apply
1. Theory of vector space in representing data
2. Matrix operations in solving system of linear equations
3. Matrix decomposition in solving system of equations

Module	Course Content	Hours
I	**Vector Spaces:** Vector Spaces, Subspaces- Definition and Examples, Linear independence of vectors, Bases and dimension, Linear Span, Field-Definition	**6**
	Vector space in R^n: System of linear equations, row space, Column space and null space. Four fundamental spaces, relation between rank and nullity, consistency theorem, basis from a spanning set and independent set.	**7**
	INTERNAL TEST 1 (Module I)	
II	**Linear transformations:** General linear transformation, Matrix of transformation, Kernel and range, properties, Isomorphism, change of basis, invariant subspace, Linear functional.	**7**
	Inner Product: Real and complex inner product spaces, properties of inner product, length and distance, Cauchy-Schwarz inequality, Orthogonality, Orthogonal complement, Orthonormal bases, Gram Schmidt orthogonalisation	**6**
	INTERNAL TEST 2 (Module II)	
III	**EigenSpace:** Properties of Eigen values and Eigen vectors , Eigen values, Eigen vectors, minimal polynomial, Diagonalization, Orthogonal diagonalization, Jordan canonical form	**7**
	Matrix Factorization: LU decomposition, QR Decomposition and singular value decomposition	**6**

IV	**Optimization**: Conditions for local minimization-One dimensional Search methods:Golden search method, Fibonacci method, Newton's Method, Secant Method, Remarks on Line Search Gradient-based methods-introduction, the method of steepest descent, analysis of Gradient Methods, Convergence, Convergence Rate.	**7**
	Analysis of Newton's Method, Levenberg-Marquardt Modification, Newton's Method for Nonlinear Least-Squares. Conjugate direction method, Conjugate Direction Algorithm, Conjugate Gradient Algorithm for Non-Quadratic Quasi Newton method.	**6**
END SEMESTER EXAM (All Modules)		

References:

1. Gilbert Strang Linear Algebra and It's Applications, 4th edition, Cengage Learning, 2006.
2. Stephen Boyd, Lieven Vandenberghe, Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares, Cambridge University Press, 2018
3. W. Keith Nicholson, Linear Algebra with applications, 4th edition, McGraw-Hill, 2002
4. I.N Herstein, Topics in Linear Algebra, Wiley Eastern, 1975.
5. S.Kumaresan, Linear Algebra : A Geometric Approach, Prentice-Hall of India, 2000.
6. Seymour Lipschutz, Marc Lipson, Schaum's outline of linear algebra, 3rd Ed., Mc Graw Hill Edn., 2017
7. Edwin K.P. Chong, Stanislaw H. Zak, An introduction to Optimization, Second edition, Wiley,2013
8. Mohan C. Joshi and Kannan M. Moudgalya, Optimization: Theory and Practice, Narosa Publishing House, New Delhi,2004

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6021	**Stochastic Models and Numerical Optimization**	**4-0-0 : 4**	**2020**

Course Objectives:

1. To study the basic concepts of the theory of stochastic processes
2. To familiarise the most important types of stochastic processes
3. For studying the various properties and characteristics of processes
4. To study the methods for describing and analysing complex stochastic models

SYLLABUS:
Random variables and events, distributions, inequalities and limits, Stochastic processes, Exponential distribution, Markov chains, Discrete Time Markov chains, Continuous Time Markov chains, Mathematical models of optimisation.

Course Outcomes:

Students should be able to
1. Carry out derivations involving conditional probability distributions and conditional expectations
2. Apply Markov chains in real world scenarios.
3. Solve differential equations for distributions and expectations in time continuous processes.
4. Apply optimization processes to generate best results for real world applications.

Module	Course Content	Hours
I	**Events, Measurability, Independence** - Sample Spaces, Events, Measures, Probability, Independence, Conditional probability, Bayes' theorem	**4**
	Random Variables - Functions of random variables, Sequence of random variables, Bernoulli, Binomial, Geometric, Poisson; Uniform, Exponential, Normal, Lognormal, Expectations, Moments and Moment generating functions, Random Vectors - Joint and Marginal distributions, Dependence, Covariance, Transformations of random vectors.	**5**
	Conditioning RVs - Conditional Distribution of a RV, Computing probabilities and expectations by conditioning, IT Application: Time-to-a-pattern for password security	**4**
INTERNAL TEST 1 (Module I)		

II	**Inequalities and Limits of Events, RVs, Distributions** - Inequalities: Markov, Chebyshev, Jensen, Holder, Convergence of RVs: Weak and Strong laws, Central limit theorem, Distributions of extreme.	**4**
	Introduction to Stochastic Processes (SPs): Definition and examples of SPs, classification of random processes according to state space and parameter space, types of SPs, elementary problems. Stationary Processes: Weakly stationary and strongly stationary processes, moving average and auto regressive processes.	**4**
	Exponential Distribution and Poisson Process - Construction of Poisson Process from Exponential Distribution, Conditional Arrival Times. Normal Distribution and Brownian Process - Construction of Brownian Process from Normal Distribution, Hitting Times and Maximum Values, Finance Applications: Option Pricing and Arbitrage Theorem	**6**
INTERNAL TEST 2 (Module II)		
III	**Markov Chains** - Markovian property and Transition probabilities, Irreducibility and Steady-State probabilities.	**3**
	Discrete-time Markov Chains (DTMCs): Definition and examples of MCs, transition probability matrix, Chapman-Kolmogorov equations; calculation of n-step transition probabilities, limiting probabilities, classification of states, ergodicity, stationary distribution, transient MC; random walk and gambler's ruin problem.	**6**
	Continuous-time Markov Chains (CTMCs): Kolmogorov-Feller differential equations, infinitesimal generator, Poisson process, birth-death process, stochastic Petri net, applications to queueing theory and communication networks.	**6**
IV	**Mathematical modeling of optimization** - Objective function, Continuous functions and discrete functions, unimodal, convex, and concave functions, Optimization constraints - internal and external. Hessian matrix, Gradient-free search, Saddle point, Linear programming and simplex method. Case Study : Transportation problem and Assignment problem, Case Study : Genetic Algorithms as optimization problems.	**10**
END SEMESTER EXAM (All Modules)		

References:

1. S.M. Ross, "Introduction to Probability Models", 11th edition, Academic Press, 2014.
2. S.M. Ross, "Stochastic Processes", 2nd Edition, John Wiley & Sons, 1996.
3. S. Resnick, "Adventures in Stochastic Processes", Birkhauser, 1994.
4. A. Müller and D. Stoyan, "Comparison Methods for Stochastic Models and Risks", John Wiley & Sons 2002.
5. R.E. Barlow and F. Proschan, "Mathematical Theory of Reliability", 1965.
6. J. Medhi, Stochastic Processes,3rd Edition,NewAge International, 2009.
7. S Karlin and H M Taylor, A First Course in Stochastic Processes, 2nd edition, Academic Press, 1975.
8. Rao S. S., Optimization Theory and Applications, Wiley Eastern, 1984.

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6031	**Principles of Artificial Intelligence and Machine Learning**	**4-0-0 : 4**	**2020**

Course Objectives:

1. To familiarise basic principles of Artificial Intelligence
2. To learn and design intelligent agents
3. To familiarise the basic areas of artificial intelligence including problem solving, knowledge representation, reasoning, decision making, perception and action
4. To master the fundamentals of machine learning, mathematical framework and learning algorithm

SYLLABUS:
Intelligent Agents, Problem solving and search, Uninformed search, Knowledge and reasoning, Probabilistic reasoning, Bayesian networks and decision theory, Neural networks, Issues in ANN training, Types of ANN architectures, SVM.

Course Outcomes:

Students should be able to
1. Understand formal methods of knowledge representation, logic and reasoning
2. Understand foundational principles, mathematical tools and program paradigms of artificial intelligence
3. Understand the fundamental issues and challenges of machine learning: data, model selection, model complexity
4. Analyse the underlying mathematical relationships within and across Machine Learning algorithms and the paradigms of supervised and unsupervised learning
5. Apply intelligent agents for Artificial Intelligence programming techniques

Module	Course Content	Hours
I	**Introduction**- Intelligent Agent, Structure of Intelligent Agent and Environment. Problem solving and search strategies - Problem solving agents, Problem-solving through Search - forward and backward, state-space, blind, heuristic, problem-reduction, minimax, constraint propagation, neural and stochastic;	**7**
	Uninformed search strategies-Depth First Search, Breadth First Search, Depth limited search, iterative deepening depth first search, bidirectional search. Informed search strategies-Best First Search, A*, AO*.	**6**
INTERNAL TEST 1 (Module I)		
II	**Knowledge and reasoning**: A knowledge based agent, representation, Propositional Logic-First Order Logic-Soundness and Completeness -Forward and Backward chaining-Resolution-semantic networks- Handling uncertain knowledge– Probabilistic Reasoning –making simple and complex decisions.	**8**
	Bayesian networks; Basics of decision theory, sequential decision problems.	**4**
INTERNAL TEST 2 (Module II)		

III	**Neural Networks**: Introduction, Basic Architecture of Neural Networks, Single Computational Layer:The Perceptron, Choice of Activation functions, Number of Output Nodes and Loss Functions; Multilayer Neural Networks, Training a Neural Network with Backpropagation. **Practical Issues in Neural Network Training**: Problem of Overfitting, Vanishing and Exploding, Gradient Problems, Difficulties in Convergence, Local and Spurious Optima, Computational Challenges.	**7** **6**
IV	**Types of Neural Architectures**: Simulating Basic Machine Learning with Shallow Models, Radial Basis Function Networks, Restricted Boltzmann Machines, Recurrent Neural Network: Architecture, Training, Applications. Convolutional Neural Network: Architecture, Training, Applications-Confusion Matrix, Precision, Recall, F Measure. **Support Vector Machine**: Architecture, Training, Applications. Parameter Estimation Bias -Mean Squared Error -Relative Efficiency – Standard Error - Maximum Likelihood Estimation.	**7** **5**
END SEMESTER EXAM (All Modules)		

References:

1. Stuart Russel, Peter Norvig, Artificial Intelligence ‘A modern Approach, Prentice Hall PTR.
2. Elaine Rich and Kevin Knight. Artificial Intelligence, 3e, Tata McGraw Hill, 2017
3. Charu C. Aggarwal, Neural Networks and Deep Learning, Springer
4. Earl Gose, Richard O Duda, Peter E.Hart, David G.Stork, Pattern Recognition, PHI Learning
5. Richard O Duda, Peter E.Hart, David G.Stock, Pattern Classification, Wiley ,India, Second Edition.

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6041	**Exploration and Statistical Analysis for Data Science**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To obtain a Comprehensive knowledge of various tools and techniques for Data transformation and visualisation
2. To learn the probability and probabilistic models of data science
3. To learn the basic statistics and testing hypothesis for specific problems

SYLLABUS:
Data Science process, Memorization methods, Unsupervised models, Univariate data exploration, Data visualisation, Prediction and filtering, Probability theory and Statistics.

Course Outcomes:

Students should be able to
1. Apply exploratory data analysis and create insightful visualisations to identify patterns
2. Understand the statistical foundations of data science and analyse the degree of certainty of predictions using statistical test and models
3. Understand the basic probability principles and techniques

Module	Course Content	Hours
I	**Data Science process** - Roles and stages in a data science project, Working with files and databases, Exploring and managing data.	**5**
	Exploratory Data Analysis, Selecting models for data science problems, Evaluation and validation of models, Documentation and deployment, Presenting results to stakeholders.	**5**
	INTERNAL TEST 1 (Module I)	
II	**Memorization methods** - single variable and multivariable models, Linear and logistic regression. **Unsupervised methods** - Cluster analysis, K-means algorithm, clusterboot method, Association rules.	**5**
	Exploring Univariate Data - Histograms - Stem-and Leaf Quantile Based Plots - Continuous Distributions -Quantile Plots- QQ Plot- Box Plots.	**5**
	INTERNAL TEST 2 (Module II)	
III	Python based data visualization, Prediction using linear regression - single variable and multi-variable models, Collaborative filtering - user based filtering and item based filtering.	**9**

IV	**Probability Concepts** -Axioms of Probability - Conditional Probability and Independence - Bayes Theorem - Expectation - Mean and Variance Skewness Kurtosis; Common Distributions - Binomial Poisson Uniform - Normal Exponential Gamma-Chi-Square Weibull Beta **Introduction to Statistics** - Sampling, Sample Means and Sample variance sample moments, covariance, correlation, Sampling Distributions - Parameter Estimation Bias - Mean Squared Error - Relative Efficiency - Standard Error - Maximum Likelihood Estimation. Comparing Two Samples - A/B Testing - ANOVA.	**6** **7**
END SEMESTER EXAM (All Modules)		

References:

1. Boris Lublinsky, Kevin T. Smith. Alexcy Yakubovich, "Professional Hadoop Solutions", Wiley, 2015
2. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, "Mining of Massive Datasets". Cambridge University Press, 2014
3. Nathan Yau, "Visualize This: The Flowing Data Guide to Design, Visualization and Statistics", Wiley, 2011
4. Nina Zumel, John Mount "Practical Data Science with R··. Manning Publications. 2014
5. Sameer Madhavan , “Mastering Python for Data Science”, Packt Publishing Limited, 2015
6. Tony Ojeda, Sean Patrick Murphy, Benjarnin Bengfort. Abhijit Dasgupta. "Practical Data Science Cookbook", Packt Publishing Limited, 2014
7. W. N. Venables. D. M. Smith and the R Core Team, "An Introduction to R", 2013
8. Wendy L. Martinez, Angel R. Martinez, Computational Statistics Handbook with MATLAB, Second edition, Chapman Hall/CRC, 2008
9. Douglas C. Montgomery, George C. Runger, Applied Statistics and Probability for Engineers, Sixth Edition, Wiley, 2013
10. Dr.J.Ravichandran, Probability And Statistics For Engineers, First Edition,Wiley, 2010

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6151	**Ensemble Models**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To familiarise the various ensemble algorithms
2. To implement the algorithms and compare performance of each one

SYLLABUS:
Bias-variance tradeoff, Non-generative models, Clustering, Combination methods, Diversity and Pruning, Learning methods, Applications.

Course Outcomes:

Students should be able to
1. Understand the various applications of AdaBoost, Random Forest and other related areas
2. Understand the basic concepts of rule extraction and combination methods

Module	Course Content	Hours
I	Bias, Variance, and the tradeoff, Ensemble learning, Difficulties in ensemble learning, Switching Net models.	**4**
	Non-generative models - Voting, Stacking. Generative models - Boosting, Bagging, Random Forests. AdaBoost algorithm, Multiclass extension. Two ensemble paradigm, Random Tree ensembles, Combination using averaging.	**7**
INTERNAL TEST 1 (Module I)		
II	**Clustering** - Consensus clustering, Use of OpenEnsembles. Similarity based methods and graph based methods. Relabeling based and transformation based methods.	**6**
	Combination methods - Algebraic methods, Behaviour Knowledge space method, Decision template method, Dynamic classifier selection, Mixture of experts.	**4**
INTERNAL TEST 2 (Module II)		
III	Ensemble diversity, Error decomposition, **Diversity measures** - pairwise and non-pairwise, Information theoretic diversity.	**6**
	Ensemble pruning, Categorization of pruning methods, Ordering based and clustering based pruning, Optimization based pruning.	**5**
IV	Semi-supervised learning, Active learning, Cost-sensitive learning, Class-imbalance learning, Reduction of ensemble to single model, Rule extraction and visualization of ensembles.	**7**
	Application : Classifying fraudulent transactions, Predicting Bitcoin prices, Sentiment evaluation in Twitter	**3**
END SEMESTER EXAM (All Modules)		

References:

1. George Kyriakides, Konstantinos Margaritis, “Hands-on Ensemble Learning with Python”, Packet Publishers
2. Alan Gelfand, Crayton Walker, “Ensemble Modeling”, Marcel Decker Inc, 1984.
3. Zhi-Hua Zhou, “Ensemble Models - Foundations and Algorithms”, CRC Press

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6251	**Convolutional Neural Network**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To familiarise how convolution networks are designed
2. To perform feature extractions and classifications

SYLLABUS:
Pattern classification, Feature extraction, Convolutions, Types of layers, Visualisation of CNNs, Applications, Ensembles.

Course Outcomes:

Students should be able to
1. Understand how to build a convolutional neural network, including recent variations such as residual networks
2. Know how to apply convolutional networks to visual detection and recognition tasks

Module	Course Content	Hours
I	**Pattern classification** - Linear classifier, Multiclass classifier, Linear separability, Feature extraction, Artificial Neural Networks - Activation, Bias, Initialization, Convolutions and Pooling activities. Introduction : LexNet, AlexNet	**10**
	INTERNAL TEST 1 (Module I)	
II	**Deriving convolution from fully connected layer**, Role of convolution, Backpropagation of convolution layers and pooling layers, Designing and training ConvNets.	**8**
	INTERNAL TEST 2 (Module II)	
III	Analyzing quantitative results from ConvNets, **Other types of layers** - Local response normalization, Spatial Pyramid pooling, Mixed pooling, Batch normalization.	**6**
	Visualizing Neural Networks - Data oriented techniques, Gradient based techniques, Inverting representation.	**7**
IV	**Classification of traffic signs** - Dataset preparation, Training and validation curves, Using ConvNets for traffic sign classification	**5**
	Ensemble of ConvNets, Stability against Noise, Sliding Window within ConvNets, Sparse Coding.	**5**
	END SEMESTER EXAM (All Modules)	

References:

1. Hamed Habibi Aghdam, Elnaz Jahani Heravi, “Guide to Convolutional Neural Networks : A practical application to Traffic Sign Detection and Classification”, Springer, 2017
2. Ethem Alpaydin, "Introduction to Machine Learning”, MIT Press, Prentice Hall of India,2005.
3. Tom Mitchell, “Machine Learning”, McGraw Hill, 3rd Edition,1997.
4. Christopher M. Bishop, “Pattern Recognition and Machine Learning”, Springer (2006).
5. Kevin P. Murphy, “Machine Learning: A Probabilistic Perspective”, The MIT Press, 2012

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6351	**Soft Computing**	**3-0-0 : 3**	**2020**

Course Objective:

1. To familiarise the salient approaches in soft computing, based on artificial neural networks, fuzzy logic, and genetic algorithms
2. To introduce applications of soft computing in different research areas in Computer Science/ Information Technology

SYLLABUS:
Basic concepts, learning models, ANN architectures, Fuzzy sets and logic, Crisp vs Fuzzy logic, Genetic Algorithms, Hybrid systems.

Course Outcomes:

Students should be able to
1. Learn about soft computing techniques and their applications
2. Analyse various neural network architectures
3. Understand perceptrons and counter propagation networks.
4. Define the fuzzy systems
5. Analyse the genetic algorithms and their applications.

Module	Course Content	Hours
I	**Basic concept of Soft Computing**; Basic concept of neural networks, Mathematical model, Properties of neural network, Typical architectures: single layer, multilayer, competitive layer	**5**
	Different learning methods: Supervised, Unsupervised & reinforced; Common activation functions; Feed forward, Feedback & recurrent N.N; Application of N.N	**5**
INTERNAL TEST 1 (Module I)		
II	**Architecture, Algorithm & Application** of McCulloh-Pitts, Hebb Net, Perceptron (with limitations & Perceptron learning rule Convergence theorem), Back propagation Neural Network, ADALINE, MADALINE, Discrete Hopfield net, Bidirectional Associative Memory, Maxnet	**11**
INTERNAL TEST 2 (Module II)		
III	**Fuzzy Sets & Logic** : Fuzzy versus Crisp; Fuzzy sets—membership function, linguistic variable, basic operators, properties; Fuzzy relations—Cartesian product, Operations on relations;	**6**
	Crisp logic—Laws of propositional logic, Inference; Predicate logic—Interpretations, Inference; **Fuzzy logic**—Quantifiers, Inference; Defuzzification methods.	**5**

IV	**Genetic Algorithm (GA)** Basic concept; role of GA in optimization, Fitness function, Selection of initial population, Cross over(different types), Mutation, Inversion, Deletion, Constraints Handling; Evolutionary Computation.	**5**
	Hybrid Systems : GA based BPNN(Weight determination); Neuro Fuzzy Systems—Fuzzy BPNN--fuzzy Neuron, architecture, learning; Fuzzy Logic controlled G.A.	**5**
END SEMESTER EXAM (All Modules)		

References:

1. Neural Networks- A Comprehensive foundation, Simon Haykin, 2nd Ed; Pearson
2. Principles of Softcomputing, S.N. Sivanandam, S.N.Deepa, Wiley India.
3. Neural Networks, Fuzzy Logic & Genetic Algorithms – Synthesis & applications, T.S. Rajasekaran & G.A. Vijaylakshmi Pai, PHI
4. Genetic Algorithm & fuzzy Logic Systems - Sanchez, Takanori, Zadeh; World Scientific
5. Genetic Algorithm, Goldberg David E.; Pearson
6. Fuzzy Set Theory & Its Applications, Zimmermann H. J, Allied Publishers Ltd, 1991.
7. Neuro-Fuzzy and Soft Computing, A Computational Approach to Learning and Machine Intelligence, Jyh-Shing Roger Jang, Chuen-Tsai Sun, Eiji Mizutani, Prentice-Hall of India Pvt. Ltd., 2004. ISBN:978-0-13261-066-7
8. Fuzzy Logic with Engineering Applications (3rd Edn.), Timothy J. Ross, Willey, 2010.

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6451	**Computer Vision**	**3-0-0 : 3**	**2020**

Course Objectives:

1. Introduce the standard computer vision problems and identify the solution methodologies.
2. To introduce concepts of Linear discriminant based and tree based classifiers

SYLLABUS:
Image formation and modelling, Affine structures, Bayesian decision theory, Linear discriminants and Trees, Pattern recognition.

Course Outcomes:

Students should be able to
1. Understand and implement the algorithms for 3D reconstruction from various cues.
2. Understand and implement the various segmentation, pattern analysis, objection detection/recognition methods.

Module	Course Content	Hours
I	**Image formation and Image model**- Components of a vision system- Cameras- camera model and camera calibration- Radiometry- Light in space- Light in surface - Sources, shadows and shading, Multiple images-The Geometry of multiple views- Stereopsis	**8**
	INTERNAL TEST 1 (Module I)	
II	**Affine structure from motion**- Elements of Affine Geometry, Affine structure and motion from two images- Affine structure and motion from multiple images- From Affine to Euclidean images. High level vision- Geometric methods- Model based vision- Obtaining hypothesis by pose consistency, pose clustering and using Invariants, Verification.	**6** **6**
	INTERNAL TEST 2 (Module II)	
III	**Bayesian Decision Theory**- Minimum error rate classification Classifiers, discriminant functions, decision surfaces- The normal density and discriminant-functions for the Normal density.	**7**
IV	**Linear discriminant based classifiers and tree classifiers** - Linear discriminant function based classifiers- Perceptron- Minimum Mean Squared Error (MME) method, Support Vector machine, Decision Trees: CART, ID3. **Recent Advances in Pattern Recognition** - Neural network structures for pattern recognition, Pattern classification using Genetic Algorithms.	**9** **6**
	END SEMESTER EXAM (All Modules)	

References:

1. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
2. R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification, John Wiley, 2001.
3. Richard Hartley and Andrew Zisserman, Multiple View Geometry in Computer Vision, Second Edition, Cambridge University Press, 2004.
4. S. Theodoridis and K. Koutroumbas, Pattern Recognition, 4th Ed., Academic Press, 2009.

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6061	**Research Methodology**	**0-2-0 : 2**	**2020**

Course Objective:

1. To prepare the student to do the M. Tech project works with a research bias.
2. To formulate a viable research question.
3. To develop skill in the critical analysis of research articles and reports.
4. To analyze the benefits and drawbacks of different methodologies.
5. To understand how to write a technical paper based on research findings.

SYLLABUS:
Introduction to Research Methodology - Types of research - Ethical issues - Copy right - royalty - Intellectual property rights and patent law - Copyleft - Open access - Analysis of sample research papers to understand various aspects of research methodology : Defining and formulating the research problem - Literature review - Development of working hypothesis - Research design and methods - Data Collection and analysis - Technical writing - Project work on a simple research problem

Course Outcomes:

Students should be able to
1. Understand research concepts in terms of identifying the research problem
2. Propose possible solutions based on research
3. Write a technical paper based on the findings
4. Get a good exposure to a domain of interest
5. Get a good domain and experience to pursue future research activities

Module	Course Content	Hours
I	**Introduction to Research Methodology**: Motivation towards research - Types of research: Find examples from literature. **Professional ethics in research** - Ethical issues-ethical committees. Copy right - royalty - Intellectual property rights and patent law - Copy left - Open access - Reproduction of published material - Plagiarism - Citation and acknowledgement. **Impact factor**. Identifying major conferences and important journals in the concerned area. Collection of at least 4 papers in the area.	**6**
INTERNAL TEST 1 (Module I)		
II	**Defining and formulating the research problem** - Literature Survey - Analyze the chosen papers and understand how the authors have undertaken literature review, identified the research gaps, arrived at their objectives, formulated their problem and developed a hypothesis.	**8**
INTERNAL TEST 2 (Module II)		

III	**Research design and methods**: Analyze the chosen papers to understand formulation of research methods and analytical and experimental methods used. Study of how different it is from previous works. **Data Collection and analysis**. Analyze the chosen papers and study the methods of data collection used. - Data Processing and Analysis strategies used– Study the tools used for analyzing the data.	**7**
IV	**Technical writing** - Structure and components, contents of a typical technical paper, difference between abstract and conclusion, layout, illustrations and tables, bibliography, referencing and footnotes-use of tools like Latex. Identification of a simple research problem – Literature survey-Research design- Methodology –paper writing based on a hypothetical result.	**7**

References:

1. C. R. Kothari, Research Methodology, New Age International, 2004
2. Panneerselvam, Research Methodology, Prentice Hall of India, New Delhi, 2012.
3. J. W. Bames, Statistical Analysis for Engineers and Scientists, Tata McGraw-Hill, New York.
4. Donald Cooper, Business Research Methods, Tata McGraw-Hill, New Delhi.
5. Leedy P. D., Practical Research: Planning and Design, McMillan Publishing Co.
6. Day R. A., How to Write and Publish a Scientific Paper, Cambridge University Press, 1989.
7. Manna, Chakraborti, Values and Ethics in Business Profession, Prentice Hall of India, New Delhi, 2012.
8. Sople, Managing Intellectual Property: The Strategic Imperative, Prentice Hall ofIndia, New Delhi, 2012.
9. Vinod Chandra S. S., Anand H. S. Research Methodology, Pearson Education, ISBN: 978-93-528-6351-8, 2017

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6071	**Seminar I**	**0-0-2 : 2**	**2020**

Course Objectives:

1. To introduce the students to research, make them understand research papers and prepare presentation material
2. To understand cutting edge technology in the chosen area
3. To improve oral communication skills through presentation
4. To prepare original technical write up on the presentation

Course Outcomes:

After completion of course, students will be able to:
1. Develop skills in doing literature survey, technical presentation and report preparation
2. Improve the proficiency in English
3. Improve presentation skills
4. Improve analytical and reasoning ability
5. Improve technical writing skills

Syllabus:

The aim of this course is to introduce the student to research, and to acquaint him with the process of presenting his work through seminars and technical reports. Students have to register for the seminar and select a topic in consultation with any faculty member offering courses for the programme. The student is expected to do an extensive literature survey and analysis in an area related to computer science (other than the area of specialisation). The study should preferably result in design ideas, designs, algorithms, and theoretical contributions in the form of theorems and proofs, new methods of proof, new techniques or heuristics with analytical studies, implementations and analysis of results.

The presentation shall be of 30 minutes duration and a committee with the Head of the Department as the chairman and two faculty members from the department as members shall evaluate the seminar based on the coverage of the topic, presentation and ability to answer the questions put forward by the committee.

Students shall individually prepare and submit a seminar report based on experimental study / industrial training on the corresponding topic, in the prescribed format given by the Department. The reference shall include standard journals (ACM/IEEE), conference proceedings and equivalent documents, reputed magazines and textbooks, technical reports and web based material, approved by the supervisor. The references shall be incorporated in the report following IEEE standards reflecting the state-of-the-art in the topic selected.

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6081	**Machine Intelligence Lab**	**0-0-3 : 1**	**2020**

Course Objectives:

1. Implement basic algorithm in AI
2. Make use of Data sets in implementing the machine learning algorithms
3. Implement the machine learning concepts and algorithms in any suitable language of choice

Course Outcomes:

Students should be able to
1. Apply AI algorithms to solve real world problems
2. Understand the implementation procedures for the machine learning algorithms
3. Design Java/Python programs for various Learning algorithms
4. Apply appropriate data sets to the Machine Learning algorithms
5. Identify and apply Machine Learning algorithms to solve real world problems

Sl No	List of Experiments
1	Introduction to Python-based notebook environments
2	Implement A* algorithm for one the following problems: i) 8 puzzle ii) Missionaries and Cannibals
3	Implement and test hill climbing based search algorithms to solve Travelling Salesman Problem
4	Solve and implement map coloring problem by backtracking and constraint propagation
5	Solve and implement the game of tic-tac-toe using mini-max
6	Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and use it to classify a new sample
7	Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets
8	Write a program to implement the naive Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets. Calculate the accuracy, precision, and recall for your data set
9	Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using any standard Heart Disease Data Set
10	Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering
11	Write a program to implement k-Nearest Neighbour algorithm to classify the iris data set. Print both correct and wrong predictions

12	Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw the corresponding graphs
13	Write a program to implement 5-fold cross validation on a given dataset. Compare the accuracy, precision, recall, and F-score for your data set for different folds
Ten experiments to complete mandatory	

APJ Abdul Kalam Technological University
Master of Technology – Course Plan

SEMESTER II

M. Tech Programme in
Artificial Intelligence and Data Science

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6012	**Big Data Analytics**	**4-0-0 : 4**	**2020**

Course Objectives:

1. To familiarise the Big Data Platform and its use cases
2. To provide an overview of Apache Hadoop
3. To provide HDFS Concepts and Interfacing with HDFS
4. To familiarise Map Reduce analytics using Hadoop and related tools like Pig, Hive etc.

SYLLABUS:

Introduction, Nature of data, Mining data streams, Estimations, Predictive analysis, Visualizations, Hadoop, HDFS, MapReduce, Case Study.

Course Outcomes:

Students should be able to
1. Describe big data and use cases from selected business domains
2. Explain the components of Hadoop and Hadoop Eco-System
3. Install, configure, and run Hadoop and HDFS
4. Perform map-reduce analytics using Hadoop
5. Use Hadoop related tools such as HBase, Pig, and Hive for big data analytics

Module	Course Content	Hours
I	**Introduction to big data:** Introduction to Big Data Platform, Challenges of Conventional Systems - Intelligent data analysis, Nature of Data - Analytic Processes and Tools - Analysis vs Reporting.	**12**
	INTERNAL TEST 1 (Module I)	
II	**Mining data streams:** Introduction To Streams Concepts, Stream Data Model and Architecture - Stream Computing - Sampling Data in a Stream, Filtering Streams, Counting Distinct Elements in a Stream	**7**
	Estimating Moments, Counting Oneness in a Window, Decaying Window - Real time Analytics Platform (RTAP) Applications - Case Studies - Real Time Sentiment Analysis- Stock Market Predictions.	**7**
	INTERNAL TEST 2 (Module II)	
III	**Components of Hadoop** - Analysing the Data with Hadoop- Scaling Out- Hadoop Streaming- Design of HDFS-Java interfaces to HDFS Basics	**6**
	Developing a Map Reduce Application-Anatomy of a Map Reduce Job, Scheduling-Shuffle and Sort - Task execution. Case Study: IBM InfoSphere BigInsights and Streams.	**6**

IV	**Introduction to HBase**, Filesystems for HBase, Client API - The Basics, Hbase clients – REST, Shell Commands, Map Reduce Integration **Introduction to Pig**, Grunt, pig data model, Pig Latin, Advanced pig latin, developing and testing Pig Latin scripts, Map Reduce Integration **Hive**, data types and file formats, HiveQL data definition, HiveQL data manipulation, HiveQL queries, HiveQL views, HiveQL Indexes, functions.	**7** **7**
END SEMESTER EXAM (All Modules)		

References:

1. Michael Minelli, Michelle Chambers, and AmbigaDhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley,2013.
2. P. J. Sadalage and M. Fowler, "NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence", Addison-Wesley Professional,2012.
3. Tom White, "Hadoop: The Definitive Guide", Third Edition, O'Reilley,2012. Analytics for Enterprise Class Hadoop and Streaming Data", McGrawHill Publishing, 2012.
4. Eric Sammer, "Hadoop Operations", O'Reilley,2012.
5. E. Capriolo, D. Wampler, and J. Rutherglen, "Programming Hive", O'Reilley,2012.
6. Lars George, "HBase: The Definitive Guide", O'Reilley,2011.
7. Alan Gates, "Programming Pig", O'Reilley,2011.
8. Chris Eaton, Dirk De Roos, Tom Deutsch, George Lapis, Paul Zikopoulos,"Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data", McGrawHill Publishing, 2012.
9. Anand Rajaraman and Jeffrey D. Ullman,"Mining of Massive Datasets", Cambridge University Press, 2012.
10. Arshdeep Bahga, Vijay Madisetti, "Big Data Science & Analytics: A Hands - On Approach",VPT, 2016

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6022	**Deep Learning & Artificial Neural Network**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To familiarise and master the tools of Artificial Intelligence
2. To explore in depth deep neural architectures for learning and inference
3. To evaluate the performance of neural architectures in comparison to other machine learning methods

SYLLABUS:
Neural networks, Perceptrons, Backpropagation, Deep Networks, Deep Reinforcement learning, RNN, LSTM, Deep Unsupervised learning, Adversarial learning, Deep Generative Models.

Course Outcomes:

Students should be able to
1. Understand basic Neural Network architectures
2. Apply fundamental principles, theory and approaches for learning with deep neural networks
3. Analyse main variants of deep learning and their typical applications
4. Analyse how deep learning fits within the context of other Machine Learning approaches

Module	Course Content	Hours
I	**Neural networks**- Perceptrons, sigmoid units; Learning in neural networks - output vs hidden layers; linear vs nonlinear networks; linear models (regression) - LMS algorithm.	**6**
	Perceptrons classification - limitations of linear nets and perceptrons - multi-Layer Perceptrons (MLP)- activation functions - linear, softmax, tanh, ReLU; error functions - feed-forward networks.	**6**
INTERNAL TEST 1 (Module I)		
II	**Backpropagation** - recursive chain rule - Learning weights of a logistic output neuron - loss functions - learning via gradient descent - optimization momentum method; Adaptive learning rates RmsProp - mini-batch gradient descent - bias-variance trade off, regularization - overfitting - inductive bias regularization - drop out - generalization.	**7**
	Deep neural networks - convolutional nets case studies using Keras/ Tensorflow.	**2**
INTERNAL TEST 2 (Module II)		
III	**Introduction to deep reinforcement learning** - neural nets for sequences - Recurrent Nets, LSTM	**5**
	Introduction to Deep unsupervised learning autoencoders - PCA to autoencoders - Deep Generative Models - Generative Models and Variational Inference - Autoregressive Models and Invertible Transformations	**6**

IV	**Adversarial Learning** - Unifying Variational Autoencoders and Generative Adversarial Networks - Adverserial Autoencoders - Evaluation of Generative Models **Geometry of Deep Generative Models** - Application - Model based Reinforcement Learning.	**5** **5**
END SEMESTER EXAM (All Modules)		
References: 1. Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning, Second edition, MITPress, 2016 2. Duda R.O., Hart P.E., Stork D.G., Pattern Classification, Second edition, Wiley - Interscience, 2001 3. Theodoridis, S., Koutroumbas, K. Pattern Recognition, Fourth edition, Academic Press,2008 4. Russell S., Norvig N., Artificial Intelligence: A Modern Approach, Prentice Hall Series in Artificial Intelligence,2003 5. Bishop C.M. Neural Networks for Pattern Recognition, Oxford University Press,1995 6. HastieT.,Tibshirani R.and FriedmanJ .,The Elements of Statistical Learning, Springer,2001 7. Koller D.and Friedman N. Probabilistic Graphical Models, MITPress, 2009		

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6032	**Genetic Algorithms**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To familiarise the basic background of genetic algorithm
2. To explain how NP problems can be tried solving using genetic algorithm strategies

SYLLABUS:

Evolutionary computation, Genetic Algorithms, Encoding, Steady state algorithms, Genetic programming, GA in engineering, GA in optimization, GA in scientific models and theoretical foundations, GBML.

Course Outcomes:

Students should be able to

1. Explain the of the principles underlying Evolutionary Computation in general and Genetic Algorithms in particular.
2. Apply Evolutionary Computation Methods to find solutions to complex problems
3. Summarise current research in Genetic Algorithms and Evolutionary Computing

Module	Course Content	Hours
I	**A brief history of evolutionary computation**-biological terminology-search space -encoding, reproduction-elements of genetic algorithm-genetic modeling- Comparison of GA and traditional search methods.	**8**
	INTERNAL TEST 1 (Module I)	
II	Steady state algorithm - fitness scaling - inversion. **Genetic programming** - Genetic Algorithm in problem solving, Implementing GP.	**6** **6**
	INTERNAL TEST 2 (Module II)	
III	**Genetic Algorithm in engineering and optimization**- natural evolution - simulated annealing and Tabu search. **Genetic Algorithm in scientific models and theoretical foundations** - computer implementation - low level operator and knowledge based techniques in Genetic Algorithm.	**7** **8**
IV	**Applications of Genetic based machine learning**-Genetic Algorithm and parallel processors, constraint optimization, uses of GA in solving NP hard problems, multilevel optimization, real life problem.	**7**
	END SEMESTER EXAM (All Modules)	

References:

1. Melanie Mitchell, “An introduction to Genetic Algorithm”, Prentice-Hall of India, New Delhi, Edition: 2004.
2. David.E.Golberg, “Genetic algorithms in search, optimization and machine learning”, Addition-Wesley-1999.
3. S.Rajasekaran G.A Vijayalakshmi Pai, “Neural Networks, Fuzzy logic and Genetic Algorithms Synthesis and Applications”, Prentice Hall of India, New Delhi-2003.
4. Nils.J.Nilsson, “Artificial Intelligence- A new synthesis”, Original edition-1999.

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6142	**R for Data Science**	**3-0-0 : 3**	**2020**

Course Objectives:

1. Introduction to data science life cycle
2. In depth knowledge of most popular machine learning techniques
3. Supervised and unsupervised learning techniques
4. Real life case studies and simulated projects to sharpen your skill sets
5. Assistance in creating a portfolio which will allow you to showcase your newly acquired skills

SYLLABUS:
R environment, Data loading and organization, factors, Data exploration and cleaning, Matrices, Machine Learning models, Statistical models, Documentation and plotting.

Course Outcomes:

Students should be able to
1. Analyse data and find relative patterns to predict outcomes
2. Analyse continuous data in varying scenarios
3. Perform Confirmatory Data analysis
4. Demonstrate expert knowledge in outcome predictions

Module	Course Content	Hours
I	**Introduction**, Reading and getting data into R, Vectors and assignment, Logical and Index vectors, Generating regular sequences, Missing values, Ordered and Unordered Factors, The function tapply() and ragged arrays, Ordered factors, Reading data from files.	10
INTERNAL TEST 1 (Module I)		
II	**Exploring and cleaning data for analysis**, Data organization, Arrays and Matrices, Basics of Arrays in R, Matrix operations, Advanced Matrix operations, Additional Matrix facilities, Lists and Data frames.	10
INTERNAL TEST 2 (Module II)		
III	**Mapping models to Machine Learning**, Evaluating and Validating models, Probability distributions in R, **Statistical models in R**, Building linear models, Generalized linear models, Nonlinear least squares and maximum likelihood models.	11
IV	**Documentation**, Graphical analysis, plot() function, Displaying multivariate data, Using graphics parameters, Matrix plots, Exporting graphs, ggplot package.	11
END SEMESTER EXAM (All Modules)		

References:

1. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, "Mining of Massive Datasets". Cambridge University Press, 2014
2. Nathan Yau, "Visualize This: The Flowing Data Guide to Design, Visualization and Statistics", Wiley, 2011
3. Nina Zumel, John Mount "Practical Data Science with RÂ·Â·. Manning Publications. 2014
4. Tony Ojeda, Sean Patrick Murphy, Benjarnin Bengfort. Abhijit Dasgupta. "Practical Data Science Cookbook", Packt Publishing Limited, 2014
5. W. N. Venables. D. M. Smith and the R Core Team, "An Introduction to R", 2013

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6242	**Data Analytics & Scalable Algorithms**	**3-0-0 : 3**	**2020**

Course Objectives:

1. Familiarise how data affects the performance of analysis algorithms and their scalability.
2. Identify the role of attributes in parallelisation of analytics algorithms.
3. Explore the use of pre-trained models in data analytics.

SYLLABUS:
Data and Relations, Correlation, Analytical models, Clustering, ReLU, Augmentation, Convolution and Pooling, Parallelization, Hashing, Batch processing, Available networks.

Course Outcomes:

Students should be able to
1. Identify data errors and dependencies among attributes by modelling them as sets & relations
2. Apply regression, classification, and clustering models on a given dataset
3. Analyse data and processes for opportunities on parallelisation
4. Apply data on pertained networks and perform classification

Module	Course Content	Hours
I	**Data and Relations** - Data scales, Set and Matrix representations, Relations, Similarity and dissimilarity measures, Sequence relations. Data preprocessing - Error types, error handling, filtering, transformation, merging. Data visualisation.	**6**
	Correlation - Linear, Causality, Chi-Square tests. Regression - Linear regression, Robust regression, RBF networks, Cross validation and feature selection.	**6**
	INTERNAL TEST 1 (Module I)	
II	Finite state machines, Recurrent models, Autoregressive models, Naive Bayes classifier, LDA, SVM, Learning Vector Quantization.	**5**
	Cluster partitions, Sequential clustering, Prototype based clustering, Fuzzy clustering, Relational clustering, Cluster tendency assessment, Cluster validity, Self organising map.	**5**
	INTERNAL TEST 2 (Module II)	
III	ReLU nonlinearity, Data Augmentation, MLP Convolutional Layer, Global Average Pooling, Dimensionality Reduction, Cascading, CNN-based feature extraction.	**6**
	Scalability through parallelization - Data parallelization, Process parallelization, Scaling using feature engineering, Feature reduction through spatial transforms.	**5**
IV	**Use of hashing**, Multiple feature hashing, Multimodal fusion for classification, Batch processing frameworks.	**6**
	Case Studies : AlexNet, VGG, GoogLeNet, ResNet	**3**
	END SEMESTER EXAM (All Modules)	

References:

1. Thomas A. Runkler, “Data Analytics - Models and Algorithms for Intelligent Data Analysis”, Springer 2012.
2. Stefanos Vrochidis, Benoit Huet, Edward Chang, Ioannis Kompatsiaris, “Big Data Analysis for Large-Scale Multimedia Search”, Wiley 2019.
3. J. O. Moreira, Andre Carvalho, Tomas Horvath, “A General Introduction to Data Analytics”, Wiley 2019.

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6342	**Scalable Systems for Data Science**	**3-0-0 : 3**	**2020**

Course Objectives:

1. Teach the fundamental systems aspects of designing using Big Data platforms
2. Explore distributed program models and abstractions

SYLLABUS:
Distributed systems, Distributed File System, Network protocols, Hadoop, Parallel Data Mining agents, MapReduce, Grid technology.

Course Outcomes:

Students should be able to
1. Distinguish distributed programming models for Big Data like Map Reduce, Stream processing and Graph processing.
2. Design and develop applications on Big Data platforms and their optimisations on commodity clusters and Clouds.
3. Scale data science algorithms and analytics using Big Data platforms.

Module	Course Content	Hours
I	**Introduction to Distributed Systems**, evolution, characteristics, design issues, user requirements, Distributed computing models-workstation model, workstation-server model, processor–pool model. Protocols for distributed systems -VMTP and FLIP.	**8**
INTERNAL TEST 1 (Module I)		
II	**Distributed file system**: Components of DFS, design issues, interfaces, implementation, File Caching and Replication. Sun Network File System – architecture and implementation, Google File System. Naming- Namespace and contexts and name resolution.	**6**
	Network Protocols, Naming, RPC, RMI, Web Services, CORBA, A message passing model for Inter-Process Communication, Coordination algorithms, Leader Election, Bully Algorithm, Maxima Finding on a Ring	**7**
INTERNAL TEST 2 (Module II)		
III	**Hadoop Architecture** - Clusters, HDFS, YARN, Basic file system operations in HDFS, File permissions in HDFS, Functional Programming Model of MapReduce, Job Chaining, Submitting MapReduce job to YARN	**10**
IV	**Parallel Data Mining Agents**, Parallel Data Access, Parallel Data Analysis, Parallel GA in Big Data Analysis, Evolutionary Algorithm Based Techniques to Handle Big Data, Statistical and Evolutionary Feature Selection Techniques Parallelized Using MapReduce Programming Model, **The Role of Grid Technologies**: A Next Level Combat with Big Data	**11**
END SEMESTER EXAM (All Modules)		

References:

1. Sunita Mahajan, Seema shah, Distributed Computing ,Oxford University Press, first edition, 2010
2. George Coulouris, Jean Dellimore and Tim Kindberg, Distributed Systems – Concepts and designing, Pearson Education Asia, Fifth Edition 2006, New Delhi.
3. Pradeep. K, Sinha, Distributed Operating Systems ,PHI Edition, first Edition,1997.
4. Andrew S Tenenbaum, Distributed Operating Systems, Pearson Education Asia
5. Distributed Systems An Algorithmic Approach, Sukumar Ghosh, CRC Press, 2007
6. Techniques and Environments for Big Data Analysis : Parallel, Cloud, and Grid Computing, Studies in Big Data Vol 17, 2016
7. Web based Parallel / Distributed Medical Data Mining Using Software Agents - Hillol Kargupta, Brian Stafford, Ilker Hamzaoglu, Los Alamos National Labs, 1997
8. Data Analytics with Hadoop: An Introduction for Data Scientists

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6442	**Knowledge Engineering and Data Science**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To explore the practical application of intelligent technologies into the different domains
2. To give students insight and experience in key issues of data and knowledge processing

SYLLABUS:

Formalisms, Items and Objects, Schema and normalization, Analysis models, Evidence and knowledge, Analysis and synthesis, Ontology.

Course Outcomes:

Students should be able to
1. Understand and describe the concepts central to the creation of knowledge bases and expert systems.
2. Conduct an in-depth examination of an existing expert system with an emphasis on basic methods of creating a knowledge base.
3. Demonstrate proficiency with statistical analysis of data.
4. Build and assess data-based models.

Module	Course Content	Hours
I	**Formalisms** - Logic as a programming language, Logic as a knowledge language, Logic as a database language, lambda calculus, Data, information and knowledge, Knowledge based systems.	**5**
	Items and Objects - unified representation, structure of data, information, and knowledge items, structure of object, data, information, and knowledge objects. Algebra of objects.	**5**
	INTERNAL TEST 1 (Module I)	
II	**Schema and normalization** - r-schema and i-schema, o-schema, t-schema, Classical normal forms.	**5**
	Analysis - conceptual view of objects, c-coupling map, constraints. Functional model - functional view, f-coupling map, constraints. Layout - internal view, i-coupling.	**6**
	INTERNAL TEST 2 (Module II)	
III	**Evidence and Knowledge**, Abductive Reasoning, Probabilistic Reasoning, Belief functions, Baconian and Fuzzy probability, Evidence based reasoning. Ontology of problem solving tasks, Building knowledge based agents. Agent Design and Development using Learning Technology.	**6**
	Problem solving through analysis and synthesis, Inquiry driven analysis and synthesis for Evidence-based reasoning, Believability assessment.	**5**

IV	**Ontology Design and Development**, Reasoning with ontologies and rules - Reduction and synthesis rules, Rule and ontology matching, Partially learned knowledge, Reasoning with partially learned knowledge. Generalization and specialization for knowledge based agents, Rule learning - Analogy-based generalization, Hypothesis learning.	**10**
	END SEMESTER EXAM (All Modules)	

References:

1. John Debenham, “Knowledge Engineering - Unifying Knowledge Base and Database Design”, Springer 2007.
2. Jude Hemnath, Madhulika Bhatia, Oana Geman, “Data Visualization and Knowledge Engineering”, Springer 2020.
3. Gheorghe Tecuci, Dorin Marcu, Mihai Boicu, David A. Schum, “Knowledge Engineering - Building Cognitive Assistants for Evidence Based Reasoning”, Cambridge University Press, 2016.

Course Code	Course Name	L-T-P : C	Year of Introduction
06DS6152	**Big Data for Internet of Things**	**3-0-0 : 3**	**2020**

Course Objectives:

1. Familiarise the scale of data generated from Internet of Things
2. Identify semantic relationships between different pieces of data generated from IoT
3. Explore the possibility of integrating cloud infrastructure to analyse data from IoT

SYLLABUS:

Big Data platforms for IoT, Challenges in IoT environment, Spatial dimensions of data, Fog Computing, Automation with Web, Data and Analysis in cloud

Course Outcomes:

Students should be able to
1. Understand aspects of volume, velocity, variety, veracity, and variability in IoT-based data
2. Understand fog computing as an infrastructure for load balancing in IoT-based data analysis
3. Apply concepts in IoT and data analysis to design smart systems
4. Apply cloud computing technologies to manage high-scale data from IoT and social networks

Module	Course Content	Hours
I	**Big Data Platforms for the Internet of Things**: network protocol- data dissemination – current state of art- Improving Data and Service Interoperability with Structure, Compliance, **Conformance and Context Awareness**: interoperability problem in the IoT context- Big Data Management Systems for the Exploitation of Pervasive Environments - Big Data challenges and requirements coming from different Smart City applications. Adaptive Pipelined Neural Network Structure in Self- aware Internet of Things: self-healing systems- Role of adaptive neural network.	**11**
	INTERNAL TEST 1 (Module I)	
II	**Spatial Dimensions of Big Data**: Application of Geographical Concepts and Spatial Technology to the Internet of Things- Applying spatial relationships, functions, and models. **Fog Computing**: A Platform for Internet of Things and Analytics: a massively distributed number of sources - Big Data Metadata Management in Smart Grids: semantic inconsistencies – role of metadata.	**10**
	INTERNAL TEST 2 (Module II)	
III	**Toward Web Enhanced Building Automation Systems**: heterogeneity between existing installations and native IP devices - loosely-coupled Web protocol stack –energy saving in smart building- Intelligent Transportation Systems and Wireless Access in Vehicular Environment Technology for Developing Smart Cities: advantages and achievements- Emerging Technologies in Health Information Systems: Genomics Driven Wellness Tracking and Management System (GO-WELL) – predictive care – personalized medicine.	**11**

IV	**Data and Analytics in Cloud-Based M2M Systems** - potential stakeholders and their complex relationships to data and analytics applications - Social Networking Analysis - Building a useful understanding of a social network - Leveraging Social Media and IoT to Bootstrap Smart Environments : lightweight Cyber Physical Social Systems - citizen actuation.	**10**
END SEMESTER EXAM (All Modules)		

References:

1. Stackowiak, R., Licht, A., Mantha, V., Nagode, L.," Big Data and The Internet of Things Enterprise Information Architecture for A New Age", Apress, 2015.
2. Dr. John Bates , "Thingalytics - Smart Big Data Analytics for the Internet of Things", john Bates, 2015.

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD6252	**Artificial Intelligence and Robotics**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To familiarise the principles of reinforcement learning which is one of the key learning techniques for robots
2. To familiarise uncertainty handling in robotics through probabilistic approaches
3. To learn how measurements work for robots

SYLLABUS:

Overview of robotics, Dynamic programming, Approximate solutions, Recursive state estimation, Filters, Measurement.

Course Outcomes:

Students should be able to

1. Learn the foundations of reinforcement learning for robotics
2. Understand basic probabilistic principles behind Robotics intelligence
3. Learn different measurement techniques for robotics
4. Understand POMDP and its significance for robotics
5. Implement principles of robotics intelligence for solving real world problems

Module	Course Content	Hours
I	**Overview**: Robotics introduction, historical perspective on AI and Robotics, Uncertainty in Robotics Reinforcement Learning: Basic overview, examples, elements, Tabular Solution Methods - Multi-armed bandits, Finite Markov decision process, **Dynamic programming** (Policy Evaluation, Pol- icy Iteration, Value Iteration), Monte Carlo Methods, Temporal-Difference Learning (Q-learning, SARSA)	11
	INTERNAL TEST 1 (Module I)	
II	**Approximate Solution Methods** - On-policy Prediction with Approximation, Value function approximation, Non-linear function approximation, Reinforcement Learning in robotics	10
	INTERNAL TEST 2 (Module II)	
III	**Recursive state estimation:** Robot Environment Interaction, Bayes filters, Gaussian filters The Kalman filter, The Extended Kalman Filter, The information filter, The particle filter Robot motion: Velocity Motion Model, Odometry Motion Model, Motion and maps	11
IV	**Measurement:** Beam Models of Range Finders, Likelihood Fields for Range Finders, Correlation- Based Sensor Models, Feature-Based Sensor Models, Overview of POMDP	10
	END SEMESTER EXAM (All Modules)	

References: 1. Sebastian Thrun, Wolfram Burgard, Dieter Fox, Probabilistic Robotics, MIT Press, 2005 2. Richard S. Sutton, Andrew G Barto, Reinforcement Learning: An Introduction, 2nd edition, MIT Press, 2018 3. Jens Kober, Jan Peters, Learning Motor Skills: From Algorithms to Springer, 2014 4. Francis X. Govers, Artificial Intelligence for Robotics, Packt, 2018

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD6352	**Natural Language Processing**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To introduce the fundamental concepts and theory of Natural Language Processing (NLP) and its practical applications
2. To explore Linguistic and statistical approaches to language processing in the three major subfields of NLP

SYLLABUS:
Introduction to NLP, N-grams, Neural networks in NLP, Vector semantics and embeddings, Sentiment classification, POS tagging, Sequence processing with RNN.

Course Outcomes:

Students should be able to
1. Understand approaches to syntax and semantics in NLP
2. Understand approaches to generate, dialogue and summarisation within NLP
3. Understand current methods for statistical approaches to machine translation
4. Understand machine learning techniques used in NLP, including hidden Markov models and unsupervised methods

Module	Course Content	Hours
I	**Introduction** – What is Natural Language Processing (NLP) - Syntax, semantics, pragmatics, and ambiguity in NLP, Regular Expressions, Text Normalisation, Edit Distance.	5
	N-gram Language Models-N-Grams, Evaluating Language Models, Generalisation and Zeros, Smoothing, Kneser-Ney Smoothing, The Web and Stupid Backoff, Perplexity's Relation to Entropy.	5
INTERNAL TEST 1 (Module I)		
II	**Neural Networks and Neural Language Models**-Units, Feed-Forward Neural Networks, Training Neural Nets, Neural Language Models.	4
	Vector Semantics and Embeddings-Lexical Semantics, Vector Semantics, Words and Vectors, Cosine for measuring similarity, TF-IDF: Weighing terms in the vector, Applications of the tf-idf vector model,Word2vec, Visualizing Embeddings ,Semantic properties of embeddings, Bias and Embeddings, Evaluating Vector Models.	7
INTERNAL TEST 2 (Module II)		

III	**Sentiment Classification** – What is sentiment classification. Machine Learning for Sentiment Classification - Training the Classifier (Naive Bayes, Logistic Regression, Support Vector Machine, Decision Tree, Random Forest), Optimising for Sentiment Analysis - Other text classification tasks – Evaluation of classification models: Precision, Recall, F-measure, Test sets and Cross-validation, Statistical Significance Testing. **Part-of-Speech Tagging**-English Word Classes, The Penn Treebank Part-of-Speech Tagset, Part-of-Speech Tagging, HMM Part-of-Speech Tagging, Maximum Entropy Markov Models, Bi-directionality, Part-of-Speech Tagging for Morphological Rich Languages. Information Extraction-Named Entity Recognition, Relation Extraction, Extracting Times, Extracting Events and their Times, Template Filling.	**6** **6**
IV	**Sequence Processing with Recurrent Networks**-Simple Recurrent Neural Networks, Applications of Recurrent Neural Networks, Deep Networks: Stacked and Bidirectional RNNs, Managing Context in RNNs: LSTMs and GRUs, Words, Subwords and Characters Neural Language Models and Generation Revisited, Encoder-Decoder Networks, Attention, Applications of Encoder-Decoder Networks. Case study: Machine translation, Question Answering	**5** **4**
	END SEMESTER EXAM (All Modules)	

References:

1. Dan Jurafsky and James H. Martin. Speech and Language Processing (3rd ed)
2. Manning C, Schuetze H. Foundations of Statistical Natural Language Processing, MIT Press
3. James Allen, "Natural Language Understanding", 2/E, Addison-Wesley, 1994
4. Steven Bird, Natural Language Processing with Python, 1st Edition, O'Reilly, 2009
5. Jacob Perkins, Python Text Processing with NLTK 2.0 Cookbook, Packt Publishing, 2010

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD6452	**Machine Learning Models and Storage Management**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To familiarise the conventional and non-conventional methods of Machine Learning.
2. To familiarise and manage data organisation and access for different scales of processing.
3. Familiarise the lifecycle of machine learning process in normal and parallelised scenarios.

SYLLABUS:

Machine Learning through queries, Query optimisation, Execution strategies, Hardware accelerators, Data Access, Resource management, ML life cycle, Parallel machines.

Course Outcomes:

Students should be able to
1. Apply database query languages to perform Machine Learning tasks.
2. Understand performance improvement of machine learning processes using hardware, data, and resource acceleration.
3. Perform ML tasks by following standardised life cycle stages.
4. Understand how machine learning can be parallelised using different data structures.

Module	Course Content	Hours
I	**Machine Learning through database queries and UDFs**, Sampling based methods, Multi-table ML, Learning over joins, Statistical relational learning, Deeper integration and specialised DBMSs. Scope of optimization, planning rewrites, automatic operator fusion. Case Study : Google Big Query	**10**
	INTERNAL TEST 1 (Module I)	
II	**Execution strategies** - Data parallel and task parallel execution, Model-parallel execution, Hybrid strategies, Hardware accelerators. Case Study : GPU, TPU **Data Access** - Caching and Buffer pool management, Data compression, NUMA awareness, Indexing.	**6** **5**
	INTERNAL TEST 2 (Module II)	
III	**Resource provisioning, scheduling, and configuration**. Failure management and transient resources. Case Study : AutoML **Managing ML life cycle**, Data sourcing and cleaning, Feature engineering and Deep learning, Model selection, management, and deployment. Benchmarking ML systems.	**5** **5**
IV	**Parallel machine models and pseudocode**, Parallel algorithm analysis, Processing in parallel - arrays, linked lists, queue-like structures. Unbounded arrays. Hash tables and associative arrays, Universal hashing, probing, perfect hashing, parallel hashing. Case Study : IBM Parallel Machine Learning Toolbox.	**9**
	END SEMESTER EXAM (All Modules)	

References:
1. Matthias Boehm, Arun Kumar, Jun Yang, “Data Management in Machine Learning Systems”, Morgan and Claypool
2. Peter Sanders, Kurt Mehlhorn, Martin Dietzfelbinger, Roman Dementiev, “Sequential and Parallel Algorithms and Data Structures: The Basic Toolbox”, Springer 2019.
3. Ron Bekkerman, Mihkail Bilenko, John Langford, “Scaling up Machine Learning - Parallel and Distributed Approaches”, Cambridge University Press.

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD6062	**Mini Project**	**0-0-4 : 2**	**2020**

Course Objectives:

1. To develop practical ability and knowledge about tools/techniques in order to solve the actual problems related to the industry, academic institutions or similar area.

Course Outcomes:

Student should be able to
1. Identify and solve various problems associated with designing and implementing a intelligent system or application.
2. Test the designed system or application.

Syllabus

Students can take up any application level/system level experimental design / implementation tasks of relatively minor intensity and scope as compared to the major-project, pertaining to a relevant domain of study. Projects can be chosen either from the list provided by the faculty or in the field of interest of the student. At the end of each phase, presentation and demonstration of the project should be conducted, which will be evaluated by a panel of examiners. A detailed project report duly approved by the guide in the prescribed format should be submitted by the student for final evaluation.

Publishing the work in Conference Proceedings/ Journals with National/ International status with the consent of the guide will carry an additional weightage in the review process.

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD6072	**Deep Learning Lab**	**0-0-3 : 1**	**2020**

Course Objectives:

1. Implement the various deep learning algorithms in Python.
2. Learn to work with different deep learning frameworks like Keras, Tensor flow, PyTorch, Caffe etc.

Course Outcomes:

Student should attain
1. Expert knowledge in solving real world problems using state of art deep learning techniques

References:

1. Francois Chollet, "Deep learning with Python" – Manning Publications.

SI No	List of Experiments
1	Basic image processing operations : Histogram equalization, thresholding, edge detection, data augmentation, morphological operations
2	Implement SVM/Softmax classifier for CIFAR-10 dataset: (i) using KNN, (ii) using 3 layer neural network
3	Study the effect of batch normalisation and dropout in neural network classifier
4	Familiarisation of image labelling tools for object detection, segmentation
5	Image segmentation using Mask RCNN, UNet, SegNet
6	Object detection with single-stage and two-stage detectors (Yolo, SSD, FRCNN, etc.)
7	Image Captioning with Vanilla RNNs
8	Image Captioning with LSTMs
9	Network Visualisation: Saliency maps, Class Visualisation
10	Generative Adversarial Networks
11	Chatbot using bi-directional LSTMs
12	Familiarisation of cloud based computing like Google colab
Nine experiments to complete mandatory	

SEMESTER III

M. Tech Programme in
Artificial Intelligence and Data Science

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7111	**Artificial Intelligence in Cyber Security**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To equip students realise the scope of artificial intelligence in preventing security threats
2. To automate the process of detection using artificial intelligence tools
3. To give an overview to the intrusion techniques

SYLLABUS:
Time Series analysis, Time series trends, Anomaly detection, Statistical and machine learning approaches, Heuristics, Intrusion management.

Course Outcomes:

Student should be able to
1. Deploy artificial intelligence based solutions for preventing cyber attacks
2. Understand the basic underlying architecture used for intrusion detection
3. Understand the heuristic methods used for cyber security

Module	Course Content	Hours
I	**Time series analysis**, Stochastic time series model, ANN time series model, Support Vector time series models, Time series decomposition, Time series analysis in cybersecurity.	5
	Time series trends and seasonal spikes, Predicting DDoS attacks - ARMA, ARIMA, ARFIMA. Voting ensemble	5
INTERNAL TEST 1 (Module I)		
II	Using data science to catch email fraud and spam, **Anomaly detection** using K-means, Using windows logs and active directory data. Decision tree and Context-based malicious event detection.	4
	Statistical and machine learning approaches to detection of attacks on computers - Techniques for studying the Internet and estimating the number and severity of attacks, network based attacks, host based attacks. Statistical pattern recognition for detection and classification of attacks, and techniques for visualizing network data, etc.	6
INTERNAL TEST 2 (Module II)		
III	Using **heuristics** to detect malicious pages, Using machine learning, logistic regression, and SVM to detect malicious URLs. Multiclass classification to detect malicious URLs.	5
	Levenshtein distance to differentiate malicious URLs from others. Using TensorFlow for intrusion detection. Machine learning to detect financial fraud - imbalanced data and credit card frauds, managing under-sampled data for logistic regression. Adam gradient optimiser for deep learning. Feature extraction and cosine similarity to quantify bad passwords.	5

IV	**Overview of intrusions**, system intrusion process, dangers of system intrusions, history and state of the art of intrusion detection systems (IDSs): anomaly detection, misuse detection, types of IDS: Network-Based IDS. Host-Based IDS, Hybrid IDS, Intrusion Prevention Systems (IPS): Network-Based IPS, Host-Based IPS, Intrusion Detection Tools, the limitations and open problems of intrusion detection systems, advanced persistent threats, case studies of intrusion detection systems against real-world threats and malware.	**10**
END SEMESTER EXAM (All Modules)		

References:

1. Soma Halder, Sinan Ozdemir, "Hands-on Machine Learning for Cybersecurity", Packt Publishing.
2. Roberto Di Pietro, Luigi V. Mancini, Intrusion Detection System, Springer ,2008
3. Anderson, Ross (2001). Security Engineering: A Guide to Building Dependable Distributed Systems. New York: John Wiley & Sons. pp. 387–388. ISBN 978-0-471-38922-4.
4. Anderson, James P., "Computer Security Threat Monitoring and Surveillance," Washing, PA, James P. Anderson Co., 1980.

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7211	**Game Theory in Artificial Intelligence**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To introduce how human behaviour can be modelled using game theory principles for artificial intelligence
2. To learn various ways game theory helps in different learning mechanisms
3. To introduce how game theory can be used to produce novel and accurate data for data science problems

SYLLABUS:
Game theory, Nash Equilibrium, Cooperative games, Multi-agent AI, Imitation and reinforcement learning, Adversary training.

Course Outcomes:

Student should be able to
1. Understand behavioural game theory for artificial intelligence domain
2. Learn the concepts of game theory for learning techniques in artificial intelligence
3. Apply game theoretic principles for dealing data for data science
4. Model modern problems in AI and DS using game theory
5. Implement game-theoretic solutions for AI and DS

Module	Course Content	Hours
I	**Introduction to Game Theory** - Cooperative vs Non-Cooperative Games, Symmetric vs Asymmetric Games, Perfect vs Imperfect Information Games, Simultaneous vs Sequential Games, Zero-Sum vs Non-Zero Sum Games, Nash Equilibrium, Inverse Game Theory	**6**
	Two-person cooperative games without transferable payoffs, N-person cooperative games, Decisions under risk and uncertainty, Decisions in conflicts against p-intelligent players, Utility theory	**5**
INTERNAL TEST 1 (Module I)		
II	**Multi-agent AI systems**, Agent Architectures and Hierarchical control, Multiagent framework, Representation of Games, Computing strategies with perfect information, Planning under certainty, Partially observable multi agent reasoning, Reasoning under uncertainty, Group decision making, Mechanism design, Learning Belief networks, Ontologies and Knowledge-based systems.	**10**
INTERNAL TEST 2 (Module II)		
III	**Imitation and Reinforcement learning**, Multi-agent Reinforcement learning, Markov Decision process, Deep Q-learning, Imitation learning with Dagger algorithm, Multi-arm bandits, Monte Carlo methods, Temporal Difference Learning, Policy Gradient methods. Case Study : Personalised Web Services.	**10**
IV	**Adversary training**, GANs, Generative Models - HMM, RBM, Discriminative model - SVM, Attacks on Machine Learning, Conditional GAN, DCGAN, InfoGAN, Stack GAN, Wasserstein GAN, Virtual batch normalization.	**10**

END SEMESTER EXAM (All Modules)
References: 1. Maňas, M., “Games and Economic Decisions”, SNTL, Praha, 1998 2. Morris, P., “Introduction to Game Theory”, Springer Verlag, New York, 1994 3. David L. Poole, Alan K. Mackworth, “Artificial Intelligence: Foundations of Computational Agents”, Cambridge University Press 4. Andrea Lonza, “Reinforcement Learning Algorithms with Python”, Packt Publishing 5. Richard S. Sutton, Andrew G. Barto, “Reinforcement Learning: An Introduction”, MIT Press 6. Yevgeniy Vorobeychik, Murat Kantarcioglu, “Adversarial Machine Learning”, Morgan & Claypool. 7. Navin K Manaswi, “Generative Adversarial Networks with Industrial Use Cases”, BPB.

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7311	**Image and Video Analytics**	**3-0-0 : 3**	**2020**

Course Objectives:

1. Familiarise the principles behind processing of video, audio, and image signals
2. Use different analytical models and loss functions for multimedia analytics

SYLLABUS:
Audio sampling, Image color spaces, Loss functions, Digital Image Segmentation, Object classification, Object tracking, Speech and handwriting recognition.

Course Outcomes:

Student should be able to
1. Represent multimedia content using appropriate coding methods and quantisation.
2. Understand loss functions involved in quantised signals.
3. Apply segmentation, object detection, and filtering methods for real world applications.
4. Apply multimedia analytics in domains of speech recognition, handwriting recognition, and object detection.

Module	Course Content	Hours
I	**Audio acquisition** - Sampling and aliasing, Sampling theorem, Linear quantization, Nonuniform scalar quantization, Time-domain audio processing, Linear predictive coding. **Image color spaces**, image representation, formats, and descriptors. Video principles and MPEG standard.	**7**
	Loss function, Zero-one loss function, Quantization Error Minimization, Vector quantisation, Neural Gas and Topology Representing Network.	**4**
INTERNAL TEST 1 (Module I)		
II	**Digital Image Segmentation** - Classification of segmentation techniques, Edge detection, Edge linking, Thresholding, Region growing, Region splitting and merging, Watershed based segmentation. Shadow detection and removal. Image processing using OpenCV - blending, smoothing, and reshaping.	**7**
	Object Classification - Shape based object classification, Motion based object classification, Viola Jones Object Detection Framework, Object classification using CNNs, use of RCNN for object classification.	**4**
INTERNAL TEST 2 (Module II)		
III	**Video Object tracking**, Temporal models, Kalman filter, Region based tracking, Contour based tracking, Feature based tracking, Model based tracking, Particle filtering, Models for shape, style, and identity.	**9**
IV	**Speech and Handwriting recognition** - HMM, Lexicon selection, N-gram model performance. Video segmentation - Shot boundary detection, Keyframe extraction, Hand pose colour-based recognition. Baggage exchange detection - Use of GMM, Tracking using Kalman filter, Object labelling. Classification of building images in video sequences - Edge based recognition, changing region detection.	**10**
END SEMESTER EXAM (All Modules)		

References:

1. Francesco Camastra, Alessandro Vinciarelli, “Machine Learning for Audio, Image and Video Analysis: Theory and Applications”, Springer 2015.
2. Maheshkumar H Kolekar, “Intelligent Video Surveillance Systems: An Algorithmic Approach”, CRC Press.
3. Vesna Zeljkovic, “Video Surveillance Techniques and Technologies”, IGI Global
4. Himanshu Singh, “Practical Machine Learning and Image Processing”, APress, 2019
5. Simon J. D. Prince, “Computer Vision: Models, Learning, and Inference”, Cambridge University Press

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7411	**Cloud Data Management**	**3-0-0 : 3**	**2020**

Course Objectives:

1. Familiarise the different types of cloud infrastructures.
2. Manage cloud infrastructure in terms of organisation, scale, and security.
3. Appraise different cloud offerings based on replication and availability.

SYLLABUS:
Cloud infrastructures, Data security in cloud, Data location and control, Storage management locations, Data security during mobility.

Course Outcomes:

Student should be able to
1. Demonstrate the concepts and technologies of Cloud Computing
2. Understand the security aspects associated with Cloud Computing
3. Demonstrate the virtual server component of Cloud Computing
4. Understand Cloud storage and usage monitoring along with security mechanism

Module	Course Content	Hours
I	**Cloud infrastructures**; public, private, hybrid. Service provider interfaces; Saas, Paas, Iaas. VDC environments; concept, planning and design, business continuity and disaster recovery principles. Managing VDC and cloud environments and infrastructures. Scalability and Cloud Services- Large Scale Data Processing- Databases and Data Stores- Data Archival.	**11**
INTERNAL TEST 1 (Module I)		
II	**Data Security** - Storage strategy and governance; security and regulations. Designing secure solutions; the considerations and implementations involved. Securing storage in virtualized and cloud environments. Monitoring and management; security auditing and SIEM.	**10**
INTERNAL TEST 2 (Module II)		
III	**Data Location and Control** - Architecture of storage, analysis and planning. Storage network design considerations; NAS and FC SANs, hybrid storage networking technologies (iSCSI, FCIP, FCoE), design for storage virtualization in cloud computing, host system design considerations.	**6**
	Global storage management locations, scalability, operational efficiency. Global storage distribution; terabytes to petabytes and greater. Policy based information management; metadata attitudes; file systems or object storage.	**6**
IV	**Securing data for transport**, Designing backup/recovery solutions to guarantee data availability in a virtualized environment. Design a replication solution, local remote and advanced. Investigate Replication in NAS and SAN environments. Data archiving solutions; analyzing compliance and archiving design considerations.	**9**
END SEMESTER EXAM (All Modules)		

References:

1. Greg Schulz, “Cloud and Virtual Data Storage Networking”, Auerbach Publications, 2011.
2. Marty Poniatowski, “Foundations of Green IT” Prentice Hall; 1 edition, 2009.
3. EMC, “Information Storage and Management” Wiley; 2 edition, 2012.
4. Volker Herminghaus, Albrecht Scriba, “Storage Management in Data Centers” Springer 2009.
5. Klaus Schmidt, “High Availability and Disaster Recovery” Springer 2006.

<table>
<tr><td>Course Code</td><td>Course Name</td><td>L-T-P : C</td><td colspan="2">Year of Introduction</td></tr>
<tr><td>06AD7121</td><td>Data Visualisation Techniques</td><td>3-0-0 : 3</td><td colspan="2">2020</td></tr>
<tr><td colspan="5">Course Objectives:

1. Familiarise how data can be presented to various stakeholders.
2. Identify peculiarities in data with the help of visualisation.
3. Design dashboards for easy understanding of underlying data.</td></tr>
<tr><td colspan="5">SYLLABUS:
Visualization, Plotting in R, Visual analytics, Validation, Presentation to stakeholders, ggplot library, Dashboards.</td></tr>
<tr><td colspan="5">Course Outcomes:

Student should be able to
1. Understand the necessity of visualisation in data management.
2. Apply visual analytics principles to appropriately preprocess data for visualisation.
3. Use R functions to generate plots for given data.
4. Perform validation of visualisations based on type and purpose of data.
5. Create dashboards and drill-down methods for data visualisation.</td></tr>
<tr><td>Module</td><td colspan="3">Course Content</td><td>Hours</td></tr>
<tr><td>I</td><td colspan="3">Introduction to visualization - the visualization pipeline, The Value of Visualization, Data - Why Do Data Semantics and Types Matter, Data Types, Dataset Types, Attribute Types, Semantics

Plotting in R - plot() function, Displaying multivariate data, Using graphics parameters, Matrix plots, Exporting graphs.</td><td>5

6</td></tr>
<tr><td colspan="5">INTERNAL TEST 1 (Module I)</td></tr>
<tr><td>II</td><td colspan="3">Visual Analytics - Optimal visualization types, Binning values, Calculated fields, Table calculations, Level of Detail calculations.

Validation - Four Levels of Design, Angles of Attack, Threats and Validation Approaches, Validation Examples, Defining Marks and Channels, Using Marks and Channels, Channel Effectiveness, Relative vs. Absolute Judgments.</td><td>5

6</td></tr>
<tr><td colspan="5">INTERNAL TEST 2 (Module II)</td></tr>
<tr><td>III</td><td colspan="3">Presenting results to stakeholders, ggplot library in R - layers, geoms, stats, positioning, annotations, scales, axes and legends, facetting, autoplot and fortify (atleast one example of each case to be done).</td><td>10</td></tr>
<tr><td>IV</td><td colspan="3">Dashboard development - Dashboard design principles, Dashboard interactivity, Connected “drill-down” dashboards. Visualization case studies - Textual data, Temporal data.</td><td>9</td></tr>
<tr><td colspan="5">END SEMESTER EXAM (All Modules)</td></tr>
</table>

References:

1. Few, Stephen, “Show Me the Numbers: Designing Tables and Graphs to Enlighten.” 2nd Edition. Analytics Press 2012
2. Tamara Munzner, Visualization Analysis and Design (VAD), CRC press, 2014
3. Complete ggplot reference manual at https://ggplot2.tidyverse.org/reference/
4. Nina Zumel, John Mount "Practical Data Science with R··. Manning Publications. 2014

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7221	**Social Network Analysis**	**3-0-0 : 3**	**2020**

Course Objectives:

1. To familiarise the components of the social network.
2. To model and visualise the social network.
3. To mine the users in the social network.
4. To familiarise the evolution of the social network.
5. To know the applications in real time systems.

SYLLABUS:

Web architecture, Visualizing online social networks, Modelling social networks, Aggregation and reasoning, Evolution in social networks, Case Study.

Course Outcomes:

Student should be able to
1. Work on the internal components of the social network
2. Model and visualise the social network
3. Mine the behaviour of the users in the social network
4. Predict the possible next outcome of the social network

Module	Course Content	Hours
I	**Introduction to Web** - Limitations of current Web – Development of Semantic Web – Emergence of the Social Web – Statistical Properties of Social Networks -Network analysis - Development of Social Network Analysis - Key concepts and measures in network analysis - Discussion networks - Blogs and online communities - Web-based networks.	**10**
INTERNAL TEST 1 (Module I)		
II	**Visualizing Online Social Networks** - A Taxonomy of Visualizations - Graph Representation - Centrality- Clustering - Node-Edge Diagrams - Visualizing Social Networks with Matrix- Based Representations- Node-Link Diagrams - Hybrid Representations - Modelling and aggregating social network data – Random Walks and their Applications –Use of Hadoop and Map Reduce - Ontological representation of social individuals and relationships.	**11**
INTERNAL TEST 2 (Module II)		
III	**Aggregating and reasoning with social network data**, Advanced Representations – Extracting evolution of Web Community from a Series of Web Archive - Detecting Communities in Social Networks - Evaluating Communities – Core Methods for Community Detection & Mining - Applications of Community Mining Algorithms - Node Classification in Social Networks.	**10**

IV	**Evolution in Social Networks** – Framework - Tracing Smoothly Evolving Communities - Models and Algorithms for Social Influence Analysis - Influence Related Statistics - Social Similarity and Influence - Influence Maximization in Viral Marketing - Algorithms and Systems for Expert Location in Social Networks - Expert Location without Graph Constraints - with Score Propagation – Expert Team Formation - Link Prediction in Social Networks - Feature based Link Prediction – Bayesian Probabilistic Models - Probabilistic Relational Models. Case Study of Google Page Rank and Facebook Graphs.	**11**
END SEMESTER EXAM (All Modules)		

References:

1. Ajith Abraham, Aboul Ella Hassanien, Václav Snášel, Computational Social Network Analysis: Trends, Tools and Research Advances, Springer, 2012
2. Borko Furht, Handbook of Social Network Technologies and Applications, Springer, 1st edition, 2011
3. Charu C. Aggarwal, Social Network Data Analyticsll, Springer; 2014
4. Giles, Mark Smith, John Yen, Advances in Social Network Mining and Analysis, Springer, 2010.
5. Guandong Xu , Yanchun Zhang and Lin Li, Web Mining and Social Networking – Techniques and applications, Springer, 1st edition, 2012
6. Peter Mika, Social Networks and the Semantic Web, Springer, 1st edition, 2007.
7. Przemyslaw Kazienko, Nitesh Chawla, Applications of Social Media and Social Network Analysis, Springer, 2015

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7321	**Text Mining**	**3-0-0 : 3**	**2020**

Course Objectives:

1. Apply text processing techniques to prepare documents for statistical modelling
2. Apply relevant machine learning models for analysing textual data and correctly interpreting the results
3. Use machine learning models for text prediction
4. Evaluate the performance of machine learning models for textual data

SYLLABUS:
Basic NLP processes, Document representation, Text categorization and clustering, Topic modeling, Document summarization, Sentiment analysis, Text visualization.

Course Outcomes:

Student should be able to
1. Describe basic concepts and methods in text mining, for example text representation, text classification and clustering, and topic modelling
2. Use the text mining concepts and methods to model real-world problems into text mining tasks, develop technical solutions, and evaluate the effectiveness of the solutions.
3. Communicate text mining process, result, and major findings to various audience including both experts and laypersons.

Module	Course Content	Hours
I	**Basic techniques in natural language processing** - tokenziation, part-of-speech tagging, chunking, syntax parsing, named entity recognition. Case study : Public NLP toolkits.	**6**
	Document representation - representing unstructured text documents with appropriate format and structure, automated text mining algorithms.	**5**
	INTERNAL TEST 1 (Module I)	
II	**Text categorization** - supervised text categorization algorithms, Naive Bayes, kNN, Logistic Regression, SVM, Decision Trees.	**6**
	Text clustering - connectivity-based (or hierarchical) clustering, centroid-based (k-means) clustering.	**5**
	INTERNAL TEST 2 (Module II)	
III	**Topic modeling** - general idea of topic modeling, basic topic models, Probabilistic Latent Semantic Indexing, Latent Dirichlet Allocation (LDA). Applications - classification, imagine annotation, collaborative filtering, and hierarchical topical structure modeling.	**7**
	Document summarization - Extraction- based summarization methods.	**3**

IV	**Sentiment analysis** - concept, sentiment polarity prediction, review mining, aspect identification. **Text visualization** - introduction to mathematical and programming tools.	**6** **3**
	END SEMESTER EXAM (All Modules)	

References:

1. Charu C. Aggarwal and Cheng Xiang Zhai, “Mining Text Data”, Springer, 2012.
2. Daniel Jurafsky and James H Martin, “Speech & Language Processing”, Pearson Education India, 2000.
3. Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schuetze, “Introduction to Information Retrieval”. Cambridge University Press, 2007.

<table>
<tr><th>Course Code</th><th colspan="2">Course Name</th><th>L-T-P : C</th><th>Year of Introduction</th></tr>
<tr><td>06AD7421</td><td colspan="2">Data Warehouse and Data Lakes</td><td>3-0-0 : 3</td><td>2020</td></tr>
<tr><td colspan="5">Course Objectives:

1. Familiarise the need of data warehouse and data lake in the context of data mining.
2. Familiarise the use of OLAP and Data cubes for data mining.
3. Identify data management options for data lakes.
4. Integrate data lake with existing data APIs.</td></tr>
<tr><td colspan="5">SYLLABUS:
Data mining, Data Warehouse, OLAP, Data cubes, Data lakes, Visual analytics, Integrations with existing analytical frameworks.</td></tr>
<tr><td colspan="5">Course Outcomes:

Student should be able to
1. Understand the process of mining insights from large data.
2. Apply data processing and structuring to organise data in warehouses.
3. Create dashboards for illustrating data and insights from data lakes.
4. Design APIs to integrate data lakes with existing data delivery methods.</td></tr>
<tr><th>Module</th><th colspan="3">Course Content</th><th>Hours</th></tr>
<tr><td>I</td><td colspan="3">Introduction: Fundamentals of data mining, Data Mining Functionalities, Classification of Data Mining systems, Data Mining Task Primitives, Integration of a Data Mining System with a Database or a Data Warehouse System, Major issues in Data Mining. Data Preprocessing: Need for Preprocessing the Data, Data Cleaning, Data Integration and Transformation, Data Reduction, Discretization and Concept Hierarchy Generation.</td><td>10</td></tr>
<tr><td colspan="5">INTERNAL TEST 1 (Module I)</td></tr>
<tr><td>II</td><td colspan="3">Data Warehouse and OLAP Technology for Data Mining: Data Warehouse, Multidimensional Data Model, Data Warehouse Architecture, Data Warehouse Implementation, Further Development of Data Cube Technology, From Data Warehousing to Data Mining Data Cube Computation and Data Generalization: Efficient Methods for Data Cube Computation, Further Development of Data Cube and OLAP Technology, Attribute-Oriented Induction.</td><td>11</td></tr>
<tr><td colspan="5">INTERNAL TEST 2 (Module II)</td></tr>
<tr><td>III</td><td colspan="3">Data Lake, Data lake maturity, Data puddles, Data ponds, Data lake organisation - landing zone, production zone, work zone, sensitive zone. Differences between Data lake and Data Warehouse. Data lake Business Reporting, Visual Analytics: Definition, concepts, Different types of charts and graphs, Emergence of data visualization and visual analytics. Lambda architecture driven data lake, Applied lambda for data lake.</td><td>11</td></tr>
<tr><td>IV</td><td colspan="3">Integrations - Acquisition of batch data using Apache Sqoop, Acquisition of stream data using Apache Flume, Messaging layer using Apache Kafka, Data Store using Apache Hadoop, Indexed datastore using ElasticSearch.</td><td>10</td></tr>
<tr><td colspan="5">END SEMESTER EXAM (All Modules)</td></tr>
</table>

References:

1. Sam Aanhory, Dennis Murray , "Data Warehousing in the Real World”, Pearson Edn Asia.
2. Paulraj Ponnaiah, "Data Warehousing Fundamentals”, Wiley student Edition
3. Alex Gorelik, “The Enterprise Big Data Lake”, O’Reilly
4. Tomcy John, Pankaj Misra, “Data Lake for Enterprises”, Packt Publishing
5. Jiawei Han & Micheline Kamber, “Data Mining – Concepts and Techniques”, Morgan Kaufmann Publishers, Elsevier, 2nd Edition, 2006.
6. U. Dinesh Kumar, “Business Analytics – The Science of Data Driven Decision Making”, Wiley 2017.
7. Ramesh Sharda, DursunDelen, Efraim Turban, “Business Intelligence: A Managerial Perspective on Analytics”, Pearson, 3e.

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7031	**Seminar II**	**0-0-2 : 2**	**2020**

Course Objectives:

1. To introduce the students to research, make them understand research papers and prepare presentation material
2. To understand cutting edge technology in the chosen area
3. To improve oral communication skills through presentation
4. To prepare original technical write up on the presentation

Course Outcomes:

After completion of course, students will be able to:
1. Develop skills in doing literature survey, technical presentation and report preparation
2. Improve the proficiency in English
3. Improve presentation skills
4. Improve analytical and reasoning ability
5. Improve technical writing skills

Syllabus:

The aim of this course is to introduce the student to research, and to acquaint him with the process of presenting his work through seminars and technical reports. Students have to register for the seminar and select a topic in consultation with any faculty member offering courses for the programme. The student is expected to do an extensive literature survey and analysis in an area related to the area of specialisation. The study should preferably result in design ideas, designs, algorithms, and theoretical contributions in the form of theorems and proofs, new methods of proof, new techniques or heuristics with analytical studies, implementations and analysis of results.

The presentation shall be of 30 minutes duration and a committee with the Head of the Department as the chairman and two faculty members from the department as members shall evaluate the seminar based on the coverage of the topic, presentation and ability to answer the questions put forward by the committee.

Students shall individually prepare and submit a seminar report based on experimental study / industrial training on the corresponding topic, in the prescribed format given by the Department. The reference shall include standard journals (ACM/IEEE), conference proceedings and equivalent documents, reputed magazines and textbooks, technical reports and web based material, approved by the supervisor. The references shall be incorporated in the report following IEEE standards reflecting the state-of-the-art in the topic selected.

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7041	**Project Phase I**	**0-0-8 : 6**	**2020**

Course Objectives:

1.To undertake research in an area related to the program of study.
2. To acquaint students to literature survey and design of a project.

Course Outcomes:

Student should be able to
1. Identify the topic, objectives and methodology to carry out the project.
2. Finalise the project plan for their course project.

Syllabus:

Every student should carry out project, related to areas of Data Sciences, under the supervision of a Supervisor(s). The project work shall commence in the third semester and shall be completed by the end of fourth semester. Candidates are required to undertake a suitable research project work; the topic shall be approved by a committee constituted by the Head of the concerned Department. Every student will be required to present the topic at the beginning of the Phase-I to illustrate the scope of the work and to finalize the topic. The third semester includes the design phase and the fourth semester includes the implementation and final thesis submission.

The student should report the status of their progress weekly to the concerned supervisor. Students should submit the project report at the end of the respective semesters, on dates announced by the college/department. Project evaluation will be based on presentations, viva voce, demonstration, review reports, design reports and final thesis. Progress of the project work is to be evaluated at the end of the third semester. For this a committee headed by the head of the department with two other faculty members in the area of the project, of which one shall be the project supervisor. If the project is done outside the college, the external supervisor associated with the student will also be a member of the committee.

Normally students are expected to do the project within the college. However they are permitted to do the project in an industry or in a government research institute under a qualified supervisor from that organization. This is only possible in the fourth semester and the topic of investigation should be in line with the project part planned in the 3rd semester. Student should apply for this through the project supervisor indicating the reason for this well in advance, preferably at the beginning of the 3rd semester.

Project evaluation marks shall be as follows:-
Total marks for the Project: 150
In the 3rd Semester: Marks:50

Project Progress evaluation:
Progress evaluation by the Project Supervisor : 20 Marks
Presentation and evaluation by the committee : 30 Marks

SEMESTER IV

M. Tech Programme in
Artificial Intelligence and Data Science

Course Code	Course Name	L-T-P : C	Year of Introduction
06AD7012	**Project Phase II**	**0-0-21 : 12**	**2020**

Course Objectives:

1. To undertake research in an area related to the program of study.
2. To enable students to implement and deploy a system and carry out performance analysis.

Course Outcomes:

Student should be able to

1. Get a good exposure to a domain of interest.
2. Get a good domain and experience to pursue future research activities.

Syllabus:

The Phase II work shall be based on the work in Phase I. Normally students are expected to do the project within the college. However they are permitted to do the project in an industry or in a government research institute under a qualified supervisor from that organization; the topic of investigation should be in line with the project part planned in the 3rd semester. Student should apply for this through the project supervisor indicating the reason for this well in advance, preferably at the beginning of the 3rd semester. This application is to be vetted by a departmental committee constituted for the same by the Principal and based on the recommendation of the committee the student is permitted to do the project outside the college. The same committee should ensure the progress of the work periodically and keep a record of this. The application for this shall include the following:-

Topic of the Project, Project work plan in the 3rdSemester, Reason for doing the project outside, Institution/Organization where the project is to be done, External Supervisor Name, Designation, Qualification and Experience, Letter of consent of the External Supervisor as well as from the organization.

Final evaluation of the project will be taken up only on completion of the project in the fourth semester. This shall be done by a committee constituted for the purpose by the principal of the college. The concerned head of the department shall be the chairman of this committee. It shall have two senior faculty members from the same department, project supervisor and the external supervisor, if any, of the student and an external expert either from an academic/ R&D organization or from Industry as members. Final project grading shall take into account the progress evaluation done in the third semester and the project evaluation in the fourth semester. If the quantum of work done by the candidate is found to be unsatisfactory, the committee may extend the duration of the project up to one more semester, giving reasons for this in writing to the student. Normally further extension will not be granted and there shall be no provision to register again for the project.

Project work is to be evaluated both in the third and the fourth semesters. Based on these evaluations the grade is finalized in the fourth semester.
Project evaluation marks shall be as follows:-
Total marks for the Project: 150
In the 4th Semester: Marks:100

Project evaluation by the supervisor/s : 30 Marks
Presentation& evaluation by the Committee : 40 Marks
Evaluation by the External expert : 30 Marks

Students are required to publish their work in reputed national/ International Journals/ Conference Proceedings etc which will carry weightage in final marks.