VI Semester

MACHINE LEARNING			
Course Code	21AI63	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Course Learning Objectives

CLO 1. Define machine learning and understand the basic theory underlying machine learning.
CLO 2. Differentiate supervised, unsupervised and reinforcement learning
CLO 3. Understand the basic concepts of learning and decision trees.
CLO 4. Understand Bayesian techniques for problems appear in machine learning
CLO 5. Perform statistical analysis of machine learning techniques.

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teacher can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) needs not to be only traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyse information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module-1

Introduction:
Machine learning Landscape: what is ML?, Why, Types of ML, main challenges of ML

Concept learning and Learning Problems – Designing Learning systems, Perspectives and Issues – Concept Learning – Find S-Version Spaces and Candidate Elimination Algorithm –Remarks on VS- Inductive bias.

Text book 2: Chapter 1, Text book 1:Chapter 1 and 2

Teaching-Learning Process	Chalk and board, Active Learning, Problem based learning

Module-2

End to end Machine learning Project: Working with real data, Look at the big picture, Get the data, Discover and visualize the data, Prepare the data, select and train the model, Fine tune your model.

Classification : MNIST, training a Binary classifier, performance measure, multiclass classification, error analysis, multi label classification, multi output classification

Text book 2: Chapter 2, Chapter 3

Teaching-Learning	Chalk and board, Active Learning

Process	
Module-3	
Training Models: Linear regression, gradient descent, polynomial regression, learning curves, regularized linear models, logistic regression **Support Vector Machine:** linear, Nonlinear , SVM regression and under the hood **Text book 2**: **Chapter 4, Chapter 5**	
Teaching-Learning Process	Chalk and board, Problem based learning, Demonstration
Module-4	
Decision Trees Training and Visualizing DT, making prediction, estimating class, the CART training, computational complexity, GINI impurity, Entropy, regularization Hyper parameters, Regression, instability **Ensemble learning and Random Forest**: Voting classifiers, Bagging and pasting, Random patches, Random forests, Boosting, stacking **Text book 2: Chapter 6, Chapter 7**	
Teaching-Learning Process	Chalk& board, Problem based learning
Module-5	
Bayes Theorem – Concept Learning – Maximum Likelihood – Minimum Description Length Principle – Bayes Optimal Classifier – Gibbs Algorithm – Naïve Bayes Classifier– example-Bayesian Belief Network – EM Algorithm **Text book 1: Chapter 6**	
Teaching-Learning Process	Chalk and board, MOOC

Course Outcomes

At the end of the course the student will be able to:

- CO 1. Understand the concept of Machine Learning and Concept Learning.
- CO 2. Apply the concept of ML and various classification methods in a project.
- CO 3. Analyse various training models in ML and the SVM algorithm to be implemented.
- CO 4. Apply the ML concept in a decision tree structure and implementation of Ensemble learning and Random Forest.
- CO 5. Apply Bayes techniques and explore more about the classification in ML.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures not less than 35% (18 Marks out of 50) in the semester-end examination (SEE), and a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

Three Unit Tests each of **20 Marks (duration 01 hour**)

1. First test at the end of 5th week of the semester
2. Second test at the end of the 10th week of the semester
3. Third test at the end of the 15th week of the semester

Two assignments each of **10 Marks**

4. First assignment at the end of 4th week of the semester
5. Second assignment at the end of 9th week of the semester

Group discussion/Seminar/quiz any one of three suitably planned to attain the COs and POs for **20 Marks** (**duration 01 hours**)

6. At the end of the 13th week of the semester

The sum of three tests, two assignments, and quiz/seminar/group discussion will be out of 100 marks and will be **scaled down to 50 marks**

(To have less stressed CIE, the portion of the syllabus should not be common /repeated for any of the methods of the CIE. Each method of CIE should have a different syllabus portion of the course).

CIE methods /question paper has to be designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the subject (**duration 03 hours)**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.

The students have to answer 5 full questions, selecting one full question from each module.

Suggested Learning Resources:

Textbooks

1. Tom M. Mitchell, Machine Learning, McGraw-Hill Education, 2013
2. Aurelien Geron, Hands-on Machine Learning with Scikit-Learn &TensorFlow , O'Reilly, Shroff Publishers and Distributors Pvt. Ltd 2019

Reference:

1. Ethem Alpaydin, Introduction to Machine Learning, PHI Learning Pvt. Ltd, 2nd Ed., 2013
2. T. Hastie, R. Tibshirani, J. H. Friedman, The Elements of Statistical Learning, Springer, 1st edition, 2001
3. Machine Learning using Python, Manaranjan Pradhan, U Dinesh Kumar, Wiley, 2019
4. Machine Learning, Saikat Dutt, Subramanian Chandramouli, Amit Kumar Das, Pearson,2020

Web links and Video Lectures (e-Resources):

1. https://www.youtube.com/playlist?list=PL1xHD4vteKYVpaIiy295pg6_SY5qznc77
2. https://nptel.ac.in/courses/106/106/106106139/

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

VI Semester

DATA SCIENCE AND ITS APPLICATIONS			
Course Code	21AD62	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:2:0	SEE Marks	50
Total Hours of Pedagogy	40 T + 20 P	Total Marks	100
Credits	04	Exam Hours	03

Course Learning Objectives:

CLO 1. Demonstrate the proficiency with statistical analysis of data to derive insight from results and interpret the data findings visually

CLO 2. Utilize the

CLO 3. skills in data management by obtaining, cleaning and transforming the data.

CLO 4. Make use of machine learning models to solve the business-related challenges

CLO 5. Experiment with decision trees, neural network layers and data partition.

CLO 6. Demonstrate how social clustering shape individuals and groups in contemporary society.

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teacher can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) does not mean only traditional lecture method, but different type of teaching methods may be adopted to develop the outcomes.
2. Show Video/animation films to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOTS (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop thinking skills such as the ability to evaluate, generalize, and analyze information rather than simply recall it.
6. Topics will be introduced in a multiple representation.
7. Show the different ways to solve the same problem and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module-1: Introduction

What is Data Science? Visualizing Data, matplotlib, Bar Charts, Line Charts, Scatterplots, **Linear Algebra,** Vectors, Matrices, **Statistics,** Describing a Single Set of Data, Correlation, Simpson's Paradox, Some Other Correlational Caveats, Correlation and Causation, **Probability,** Dependence and Independence, Conditional Probability, Bayes's Theorem, Random Variables, Continuous Distributions, The Normal Distribution, The Central Limit Theorem.

Chapters 1, 3, 4, 5 and 6

Laboratory Component:

1. Installation of Python/R language, Visual Studio code editors can be demonstrated along with Kaggle data set usage.
2. Write programs in Python/R and Execute them in either Visual Studio Code or PyCharm Community Edition or any other suitable environment.
3. A study was conducted to understand the effect of number of hours the students spent studying on their performance in the final exams. Write a code to plot line chart with number of hours spent studying on x-axis and score in final exam on y-axis. Use a red '*' as the point character, label the axes and give the plot a title.

Number of hrs spent studying (x)	10	9	2	15	10	16	11	16
Score in the final exam (0 – 100) (y)	95	80	10	50	45	98	38	93

4. For the given dataset mtcars.csv (www.kaggle.com/ruiromanini/mtcars), plot a histogram to check the frequency distribution of the variable 'mpg' (Miles per gallon)

Teaching-Learning Process	1. Demonstration of different charts 2. PPT Presentation for Theorems and different distributions 3. Live coding and execution for visualization with simple examples

Module-2: Hypothesis and Inference

Statistical Hypothesis Testing, Example: Flipping a Coin, p-Values, Confidence Intervals, p-Hacking, Example: Running an A/B Test, Bayesian Inference, **Gradient Descent,** The Idea Behind Gradient Descent Estimating the Gradient, Using the Gradient, Choosing the Right Step Size, Using Gradient Descent to Fit Models, Minibatch and Stochastic Gradient Descent, **Getting Data,** stdin and stdout, Reading Files, Scraping the Web, Using APIs, Example: Using the Twitter APIs, **Working with Data,** Exploring Your Data, Using NamedTuples, Dataclasses, Cleaning and Munging, Manipulating Data, Rescaling, An Aside: tqdm, Dimensionality Reduction.

Chapters 7, 8, 9 and 10

Laboratory Component:

1. Consider the books dataset BL-Flickr-Images-Book.csv from Kaggle (https://www.kaggle.com/adeyoyintemidayo/publication-of-books) which contains information about books. Write a program to demonstrate the following.

- Import the data into a DataFrame
- Find and drop the columns which are irrelevant for the book information.
- Change the Index of the DataFrame
- Tidy up fields in the data such as date of publication with the help of simple regular expression.
- Combine str methods with NumPy to clean columns

Teaching-Learning Process	1. Demonstration of Hypothesis test. 2. PPT Presentation to explore and manipulate data. 3. Live coding of concepts with simple examples 4. Case Study: Extraction of data from Books dataset

Module-3: Machine Learning

Modeling, What Is Machine Learning?, Overfitting and Underfitting, Correctness, The Bias-Variance Tradeoff, Feature Extraction and Selection, **k-Nearest Neighbors,** The Model, Example: The Iris Dataset, The Curse of Dimensionality, **Naive Bayes,** A Really Dumb Spam Filter, A More Sophisticated Spam Filter, Implementation, Testing Our Model, Using Our Model, **Simple Linear Regression,** The Model, Using

Gradient Descent, Maximum Likelihood Estimation, **Multiple Regression,** The Model, Further Assumptions of the Least Squares Model, Fitting the Model, Interpreting the Model, Goodness of Fit, Digression: The Bootstrap, Standard Errors of Regression Coefficients, Regularization, **Logistic Regression,** The Problem, The Logistic Function, Applying the Model, Goodness of Fit, Support Vector Machines.

Chapters 11, 12, 13, 14, 15 and 16

Laboratory Component:

1. Train a regularized logistic regression classifier on the iris dataset (https://archive.ics.uci.edu/ml/machine-learning-databases/iris/ or the inbuilt iris dataset) using sklearn. Train the model with the following hyper parameter C = 1e4 and report the best classification accuracy.
2. Train an SVM classifier on the iris dataset using sklearn. Try different kernels and the associated hyper parameters. Train model with the following set of hyper parameters RBF-kernel, gamma=0.5, one-vs-rest classifier, no-feature-normalization. Also try C=0.01,1,10C=0.01,1,10. For the above set of hyper parameters, find the best classification accuracy along with total number of support vectors on the test data

Teaching-Learning Process	1. Demonstration of Models 2. PPT Presentation for techniques 3. Live coding of all concepts with simple examples

Module-4: Decision Trees

What Is a Decision Tree?, Entropy, The Entropy of a Partition, Creating a Decision Tree, Putting It All Together, Random Forests, **Neural Networks,** Perceptrons, Feed-Forward Neural Networks, Backpropagation, Example: Fizz Buzz, **Deep Learning,** The Tensor, The Layer Abstraction, The Linear Layer, Neural Networks as a Sequence of Layers, Loss and Optimization, Example: XOR Revisited, Other Activation Functions, Example: Fizz Buzz Revisited, Softmaxes and Cross-Entropy, Dropout, Example: MNIST, Saving and Loading Models, **Clustering,** The Idea, The Model, Example: Meetups, Choosing k, Example: Clustering Colors, Bottom-Up Hierarchical Clustering

Chapters 17, 18, 19 and 20

Laboratory Component:

1. Consider the following dataset. Write a program to demonstrate the working of the decision tree based ID3 algorithm.

Price	Maintenance	Capacity	Airbag	Profitable
Low	Low	2	No	Yes
Low	Med	4	Yes	Yes
Low	Low	4	No	Yes
Low	Med	4	No	No
Low	High	4	No	No
Med	Med	4	No	No
Med	Med	4	Yes	Yes
Med	High	2	Yes	No
Med	High	5	No	Yes
High	Med	4	Yes	Yes
high	Med	2	Yes	Yes
High	High	2	Yes	No
high	High	5	yes	Yes

2. Consider the dataset spiral.txt (https://bit.ly/2Lm75Ly). The first two columns in the dataset corresponds to the co-ordinates of each data point. The third column corresponds to the actual cluster label. Compute the rand index for the following methods:

- K – means Clustering
- Single – link Hierarchical Clustering
- Complete link hierarchical clustering.
- Also visualize the dataset and which algorithm will be able to recover the true clusters.

Teaching-Learning Process	1. Demonstration using Python/ R Language 2. PPT Presentation for decision tree, Neural Network, Deep learning and clustering 3. Live coding for the concepts with simple examples 4. Project Work: Algorithm implementation

Module-5: Natural Language Processing

Word Clouds, n-Gram Language Models, Grammars, An Aside: Gibbs Sampling, Topic Modeling, Word Vectors, Recurrent Neural Networks, Example: Using a Character-Level RNN, **Network Analysis,** Betweenness Centrality, Eigenvector Centrality, Directed Graphs and PageRank, **Recommender Systems,** Manual Curation, Recommending What's Popular, User-Based Collaborative Filtering, Item-Based Collaborative Filtering, Matrix Factorization.

Chapters 21, 22 and 23

Laboratory Component:

Mini Project – Simple web scrapping in social media

Teaching-Learning Process	1. Demonstration of models 2. PPT Presentation for network analysis and Recommender systems 3. Live coding with simple examples

Course outcome (Course Skill Set)

At the end of the course the student will be able to:

CO 1. Identify and demonstrate data using visualization tools.
CO 2. Make use of Statistical hypothesis tests to choose the properties of data, curate and manipulate data.
CO 3. Utilize the skills of machine learning algorithms and techniques and develop models.
CO 4. Demonstrate the construction of decision tree and data partition using clustering.
CO 5. Experiment with social network analysis and make use of natural language processing skills to develop data driven applications.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures not less than 35% (18 Marks out of 50) in the semester-end examination (SEE), and a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together

Continuous Internal Evaluation:

Three Unit Tests each of **20 Marks (duration 01 hour**)

1. First test at the end of 5th week of the semester
2. Second test at the end of the 10th week of the semester
3. Third test at the end of the 15th week of the semester

Two assignments each of **10 Marks**

4. First assignment at the end of 4th week of the semester
5. Second assignment at the end of 9th week of the semester

Practical Sessions need to be assessed by appropriate rubrics and viva-voce method. This will contribute to **20 marks**.

Note: Minimum of 80% of the laboratory components have to be covered.

- Rubrics for each Experiment taken average for all Lab components – 15 Marks.
- Viva-Voce– 5 Marks (more emphasized on demonstration topics)

The sum of three tests, two assignments, and practical sessions will be out of 100 marks and will be **scaled down to 50 marks**
(to have a less stressed CIE, the portion of the syllabus should not be common /repeated for any of the methods of the CIE. Each method of CIE should have a different syllabus portion of the course).

CIE methods /question paper has to be designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the subject (**duration 03 hours)**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.

The students have to answer 5 full questions, selecting one full question from each module

Suggested Learning Resources:

Text Books

1. Joel Grus, "Data Science from Scratch", 2nd Edition, O'Reilly Publications/Shroff Publishers and Distributors Pvt. Ltd., 2019. ISBN-13: 978-9352138326

Reference Books

1. Emily Robinson and Jacqueline Nolis, "Build a Career in Data Science", 1st Edition, Manning Publications, 2020. ISBN: 978-1617296246.
2. AurélienGéron, "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems", 2nd Edition, O'Reilly Publications/Shroff Publishers and Distributors Pvt. Ltd., 2019. ISBN-13: 978-1492032649.
3. François Chollet, "Deep Learning with Python", 1st Edition, Manning Publications, 2017. ISBN-13: 978-1617294433
4. Jeremy Howard and Sylvain Gugger, "Deep Learning for Coders with fastai and PyTorch", 1st Edition, O'Reilly Publications/Shroff Publishers and Distributors Pvt. Ltd., 2020. ISBN-13: 978-1492045526
5. Sebastian Raschka and Vahid Mirjalili, "Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2", 3rd Edition, Packt Publishing Limited, 2019.ISBN-13: 978-1789955750

Web links and Video Lectures (e-Resources):

1. Using Python : https://www.python.org
2. R Programming : https://www.r-project.org/
3. Python for Natural Language Processing : https://www.nltk.org/book/
4. Data set: https://bit.ly/2Lm75Ly
5. Data set: https://archive.ics.uci.edu/ml/datasets.html

6. Data set : www.kaggle.com/ruiromanini/mtcars
7. Pycharm : https://www.jetbrains.com/pycharm/
8. https://nptel.ac.in/courses/106/106/106106179/
9. https://nptel.ac.in/courses/106/106/106106212/
10. http://nlp-iiith.vlabs.ac.in/List%20of%20experiments.html

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

1. Real world problem solving - Applying the machine learning techniques and developing models

VI Semester

SOFTWARE ENGINEERING & PROJECT MANAGEMENT			
Course Code	**21CS61**	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	2:2:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Course Learning Objectives

CLO 1. Outline software engineering principles and activities involved in building large software programs. Identify ethical and professional issues and explain why they are of concern to Software Engineers.
CLO 2. Describe the process of requirement gathering, requirement classification, requirement specification and requirements validation.
CLO 3. Infer the fundamentals of object oriented concepts, differentiate system models, use UML diagrams and apply design patterns.
CLO 4. Explain the role of DevOps in Agile Implementation.
CLO 5. Discuss various types of software testing practices and software evolution processes.
CLO 6. Recognize the importance Project Management with its methods and methodologies.
CLO 7. Identify software quality parameters and quantify software using measurements and metrics. List software quality standards and outline the practices involved

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module-1

Introduction: The evolving role of software, Software, The changing nature of software, Software engineering, A Process Framework, Process Patterns, Process Assessment, Personal and Team Process Models, Process Technology, Product and Process.

Textbook 1: Chapter 1: 1.1 to 1.3

Process Models: Prescriptive models, Waterfall model, Incremental process models, Evolutionary process models, Specialized process models.

Textbook 1: Chapter 2: 2.1, 2.2, 2.4 to 2.7

Requirements Engineering: Requirements Engineering Task, Initiating the Requirements Engineering process, Eliciting Requirements, Developing use cases, Building the analysis model, Negotiating Requirements, Validating Requirements, Software Requirement Document **(Sec 4.2)**

Textbook 1: Chapter 3: 3.1 to 3.6, Textbook 5: Chapter 4: 4.2

Teaching-Learning Process	Chalk and board, Active Learning, Problem based learning

Module-2

Introduction, Modelling Concepts and Class Modelling: What is Object orientation? What is OO development? OO Themes; Evidence for usefulness of OO development; OO modelling history. Modelling as Design technique: Modelling, abstraction, The Three models. Class Modelling: Object and Class Concept, Link and associations concepts, Generalization and Inheritance, A sample class model, Navigation of class models, Introduction to RUP**(Textbook: 5 Sec 2.4)** and UML diagrams

Textbook 2: Chapter 1,2,3

Building the Analysis Models: Requirement Analysis, Analysis Model Approaches, Data modeling Concepts, Object Oriented Analysis, Scenario-Based Modeling, Flow-Oriented Modeling, class Based Modeling, Creating a Behavioral Model.

Textbook 1: Chapter 8: 8.1 to 8.8

Teaching-Learning Process	Chalk and board, Active Learning, Demonstration

Module-3

Software Testing: A Strategic Approach to Software Testing, Strategic Issues, Test Strategies for Conventional Software, Test Strategies for Object -Oriented Software, Validation Testing, System Testing, The Art of Debugging.

Textbook 1: Chapter 13: 13.1 to 13.7

Agile Methodology & DevOps: Before Agile – Waterfall, Agile Development,

Self-Learning Section:
What is DevOps?, DevOps Importance and Benefits, DevOps Principles and Practices, 7 C's of DevOps Lifecycle for Business Agility, DevOps and Continuous Testing, How to Choose Right DevOps Tools?, Challenges with DevOps Implementation.

Textbook 4: Chapter 2: 2.1 to 2.9

Teaching-Learning Process	Chalk and board, Active Learning, Demonstration

Module-4

Introduction to Project Management:
Introduction, Project and Importance of Project Management, Contract Management, Activities Covered by Software Project Management, Plans, Methods and Methodologies, Some ways of categorizing Software Projects, Stakeholders, Setting Objectives, Business Case, Project Success and Failure, Management and Management Control, Project Management life cycle, Traditional versus Modern Project Management Practices.

Textbook 3: Chapter 1: 1.1 to 1.17

Teaching-Learning Process	Chalk and board, Active Learning, Demonstration

Module-5

Activity Planning:
Objectives of Activity Planning, When to Plan, Project Schedules, Sequencing and Scheduling Activities, Network Planning Models, Forward Pass– Backward Pass, Identifying critical path, Activity Float, Shortening Project Duration, Activity on Arrow Networks.

Textbook 3: Chapter 6: 6.1 to 6.16

Software Quality:
Introduction, The place of software quality in project planning, Importance of software quality, software quality models, ISO 9126, quality management systems, process capability models, techniques to enhance software quality, quality plans.

Textbook 3: Chapter 13: (13.1 to 13.6 , 13.9, 13.11, 13.14),

Teaching-Learning Process	Chalk and board, Active Learning, Demonstration

Course Outcomes

At the end of the course the student will be able to:

CO 1. Understand the activities involved in software engineering and analyze the role of various process models

CO 2. Explain the basics of object-oriented concepts and build a suitable class model using modelling techniques

CO 3. Describe various software testing methods and to understand the importance of agile methodology and DevOps

CO 4. Illustrate the role of project planning and quality management in software development

CO 5. Understand the importance of activity planning and different planning models

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures not less than 35% (18 Marks out of 50) in the semester-end examination (SEE), and a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together

Continuous Internal Evaluation:

Three Unit Tests each of **20 Marks (duration 01 hour**)

1. First test at the end of 5th week of the semester
2. Second test at the end of the 10th week of the semester
3. Third test at the end of the 15th week of the semester

Two assignments each of **10 Marks**

4. First assignment at the end of 4th week of the semester
5. Second assignment at the end of 9th week of the semester

Group discussion/Seminar/quiz any one of three suitably planned to attain the COs and POs for **20 Marks** (**duration 01 hours**)

6. At the end of the 13th week of the semester

The sum of three tests, two assignments, and quiz/seminar/group discussion will be out of 100 marks and will be **scaled down to 50 marks**

(to have less stressed CIE, the portion of the syllabus should not be common /repeated for any of the methods of the CIE. Each method of CIE should have a different syllabus portion of the course).

CIE methods /question paper has to be designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the subject (**duration 03 hours)**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.

The students have to answer 5 full questions, selecting one full question from each module

Suggested Learning Resources:

Textbooks

1. Roger S. Pressman: Software Engineering-A Practitioners approach, 7th Edition, Tata McGraw Hill.
2. Michael Blaha, James Rumbaugh: Object Oriented Modelling and Design with UML, 2nd Edition, Pearson Education, 2005.
3. Bob Hughes, Mike Cotterell, Rajib Mall: Software Project Management, 6th Edition, McGraw Hill Education, 2018.

4. Deepak Gaikwad, Viral Thakkar, DevOps Tools From Practitioner's Viewpoint, Wiley.
5. Ian Sommerville: Software Engineering, 9th Edition, Pearson Education, 2012.

Reference:

1. Pankaj Jalote: An Integrated Approach to Software Engineering, Wiley India.

Weblinks and Video Lectures (e-Resources):

1. https://onlinecourses.nptel.ac.in/noc20_cs68/preview
2. https://www.youtube.com/watch?v=WxkP5KR_Emk&list=PLrjkTql3jnm9b5nr-ggx7Pt1G4UAHeFlJ
3. http://elearning.vtu.ac.in/econtent/CSE.php
4. http://elearning.vtu.ac.in/econtent/courses/video/CSE/15CS42.html
5. https://nptel.ac.in/courses/128/106/128106012/ (DevOps)

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

Case study, Field visit

VI Semester

NATURAL LANGUAGE PROCESSING			
Course Code	21AI643	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Course Learning Objectives

CLO 1. Analyse the natural language text.
CLO 2. Define the importance of natural language.
CLO 3. Understand the concepts Text mining.
CLO 4. Illustrate information retrieval techniques.

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same program
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module-1

Overview and language modeling: Overview: Origins and challenges of NLP-Language and Grammar-Processing Indian Languages- NLP Applications-Information Retrieval. Language Modeling: Various Grammar- based Language Models-Statistical Language Model.

Textbook 1: Ch. 1,2

Teaching-Learning Process	Chalk and board, Online demonstration, Problem based learning

Module-2

Word level and syntactic analysis: Word Level Analysis: Regular Expressions-Finite-State Automata-Morphological Parsing-Spelling Error Detection and correction-Words and Word classes-Part-of Speech Tagging. Syntactic Analysis: Context-free Grammar-Constituency- Parsing-Probabilistic Parsing.

Textbook 1: Ch. 3,4

Teaching-Learning Process	Chalk and board, Online Demonstration

Module-3

Extracting Relations from Text: From Word Sequences to Dependency Paths:
Introduction, Subsequence Kernels for Relation Extraction, A Dependency-Path Kernel for Relation Extraction and Experimental Evaluation.

Mining Diagnostic Text Reports by Learning to Annotate Knowledge Roles: Introduction, Domain Knowledge and Knowledge Roles, Frame Semantics and Semantic Role Labeling, Learning to Annotate Cases with Knowledge Roles and Evaluations.

A Case Study in Natural Language Based Web Search: InFact System Overview, The GlobalSecurity.org Experience.

Textbook 2: Ch. 3,4,5

Teaching-Learning Process	Chalk and board, Online Demonstration

Module-4

Evaluating Self-Explanations in iSTART: Word Matching, Latent Semantic Analysis, and Topic Models: Introduction, iSTART: Feedback Systems, iSTART: Evaluation of Feedback Systems,

Textual Signatures: Identifying Text-Types Using Latent Semantic Analysis to Measure the Cohesion of Text Structures: Introduction, Cohesion, Coh-Metrix, Approaches to Analyzing Texts, Latent Semantic Analysis, Predictions, Results of Experiments.

Automatic Document Separation: A Combination of Probabilistic Classification and Finite-State Sequence Modeling: Introduction, Related Work, Data Preparation, Document Separation as a Sequence Mapping Problem, Results.

Evolving Explanatory Novel Patterns for Semantically-Based Text Mining: Related Work, A Semantically Guided Model for Effective Text Mining.

Textbook 2: Ch. 6,7,8,9

Teaching-Learning Process	Chalk and board, Online Demonstration

Module-5

INFORMATION RETRIEVAL AND LEXICAL RESOURCES: Information Retrieval: Design features of Information Retrieval Systems-Classical, Non classical, Alternative Models of Information Retrieval – valuation Lexical Resources: World Net-Frame Net- Stemmers-POS Tagger- Research Corpora.

Textbook 1: Ch. 9,12

Teaching-Learning Process	Chalk and board, Online Demonstration

Course Outcomes

At the end of the course the student will be able to:

CO 1. Analyse the natural language text.
CO 2. Define the importance of natural language.
CO 3. Understand the concepts Text mining.
CO 4. Illustrate information retrieval techniques.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures not less than 35% (18 Marks out of 50) in the semester-end examination (SEE), and a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together

Continuous Internal Evaluation:

Three Unit Tests each of **20 Marks (duration 01 hour**)

1. First test at the end of 5th week of the semester
2. Second test at the end of the 10th week of the semester
3. Third test at the end of the 15th week of the semester

Two assignments each of **10 Marks**

4. First assignment at the end of 4th week of the semester
5. Second assignment at the end of 9th week of the semester

Group discussion/Seminar/quiz any one of three suitably planned to attain the COs and POs for **20 Marks** (**duration 01 hours**)

6. At the end of the 13th week of the semester

The sum of three tests, two assignments, and quiz/seminar/group discussion will be out of 100 marks and will be **scaled down to 50 marks**

(to have less stressed CIE, the portion of the syllabus should not be common /repeated for any of the methods of the CIE. Each method of CIE should have a different syllabus portion of the course).

CIE methods /question paper has to be designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the subject (**duration 03 hours)**

1. The question paper will have ten questions. Each question is set for 20 marks
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.

The students have to answer 5 full questions, selecting one full question from each module

Suggested Learning Resources:

Textbooks

1. Tanveer Siddiqui, U.S. Tiwary, "Natural Language Processing and Information Retrieval", Oxford University Press, 2008.
2. Anne Kao and Stephen R. Poteet (Eds), "Natural LanguageProcessing and Text Mining", Springer-Verlag London Limited 2007.

Reference Books:

1. Daniel Jurafsky and James H Martin, "Speech and Language Processing: Anintroduction to Natural Language Processing, Computational Linguistics and SpeechRecognition", 2nd Edition, Prentice Hall, 2008.
2. James Allen, "Natural Language Understanding", 2nd edition, Benjamin/Cummingspublishing company, 1995.
3. Gerald J. Kowalski and Mark.T. Maybury, "Information Storage and Retrieval systems", Kluwer academic Publishers, 2000.

Weblinks and Video Lectures (e-Resources):

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

VII Semester

DATA VISUALIZATION			
Course Code	21AD71	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Course Learning Objectives

CLO 1. Understand and use various plot types with Python
CLO 2. Explore and work with different plotting libraries
CLO 3. Create effective visualizations
CLO 4. Implement exemplary applications related to Network Programming and Web Service
CLO 5. Exhibit the awareness of the importance and limitation of the exploratory data analysis paradigm

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teacher can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) does not mean only traditional lecture method, but different type of teaching methods may be adopted to develop the outcomes.
2. Show Video/animation films to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOTS (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop thinking skills such as the ability to evaluate, generalize, and analyze information rather than simply recall it.
6. Topics will be introduced in a multiple representation.
7. Show the different ways to solve the same problem and encourage the students to come up with their own creative ways to solve them.

Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module-1: Data Visualization and Data Exploration

Introduction: Data Visualization, Importance of Data Visualization, Data Wrangling, Tools and Libraries for Visualization

Overview of Statistics: Measures of Central Tendency, Measures of Dispersion, Correlation, Types od Data, Summary Statistics

Numpy: Numpy Operations - Indexing, Slicing, Splitting, Iterating, Filtering, Sorting, Combining, and Reshaping

Pandas: Advantages of pandas over numpy, Disadvantages of pandas, Pandas operation - Indexing, Slicing, Iterating, Filtering, Sorting and Reshaping using Pandas

Text Book 1: Chapter 1

Teaching-Learning Process	5. PPT – Visualization tools 6. Demonstration of operations on data

Module-2: Plots

Comparison Plots: Line Chart, Bar Chart and Radar Chart; **Relation Plots:** Scatter Plot, Bubble Plot ,

Correlogram and Heatmap; **Composition Plots:** Pie Chart, Stacked Bar Chart, Stacked Area Chart, Venn Diagram; **Distribution Plots:** Histogram, Density Plot, Box Plot, Violin Plot; **Geo Plots:** Dot Map, Choropleth Map, Connection Map; What Makes a Good Visualization?

A Deep Dive into Matplotlib
Introduction, Overview of Plots in Matplotlib, **Pyplot Basics:** Creating Figures, Closing Figures, Format Strings, Plotting, Plotting Using pandas DataFrames, Displaying Figures**,** Saving Figures; **Basic Text and Legend**

Functions: Labels, Titles, Text, Annotations, Legends**; Basic Plots:**Bar Chart, Pie Chart, Stacked Bar Chart, Stacked Area Chart, Histogram, Box Plot, Scatter Plot, Bubble Plot; **Layouts:** Subplots, Tight Layout, Radar Charts, GridSpec; **Images:** Basic Image Operations, Writing Mathematical Expressions

Text Book 1: Chapter 2, Chapter 3

Teaching-Learning Process	3. PPT - Visualization techniques 4. Demonstration of operations on plots using Matplotlib

Module-3: Simplifying Visualizations using Seaborn

Introduction, Advantages of Seaborn **Controlling Figure Aesthetics:** Seaborn Figure Styles, Removing Axes Spines, Contexts; **Color Palettes:** Categorical Color Palettes, Sequential Color Palettes, Diverging Color Palettes; **Interesting Plots in Seaborn:** Bar Plots, Kernel Density Estimation, Plotting Bivariate Distributions, Visualizing Pairwise Relationships, Violin Plots;

Text Book 1: Chapter 4

Teaching-Learning Process	1. PPT - Visualization techniques 2. Demonstration of operations on plots using Seaborn

Module-4: Plotting Geospatial Data

Introduction, Geoplotlib, The Design Principles of Geoplotlib, Geospatial Visualizations, Tile Providers, Custom Layers, Introduction to Folium

Visualizing Data: Building a Google map from geocoded data, Visualizing networks and interconnection and Visualizing mail data

Making Things Interactive with Bokeh
Introduction, Bokeh, Concepts of Bokeh, Interfaces in Bokeh, Output, Bokeh Server, Presentation, Integrating, Adding Widgets

Text Book 1: Chapter5, Chapter 6

Teaching-Learning Process	5. PPT - Visualization techniques 6. Demonstration of operations using Geoplotlib

Module-5: Networked Programs

HyperText Transfer Protocol – HTTP, The World's Simplest Web Browser, Retrieving an image over HTTP, Retrieving web pages with urllib, Parsing HTML and scraping the web, Parsing HTML using regular expressions, Parsing HTML using BeautifulSoup, Reading binary files using urllib

Using Web Services
eXtensibleMarkup Language – XML, Parsing XML, Looping through nodes, JavaScript Object Notation – JSON, Parsing JSON

Text Book 2: Chapters 12 and Chapter 13

Teaching-Learning Process	7. PPT – On web services, browsers, HTTP, HTML 8. Demonstration of parsing and looping - XML,JSON

Course Outcomes

At the end of the course the student will be able to:

CO 1. Demonstrate the data visualization techniques.
CO 2. Analyze data represented in the form of graphs & charts
CO 3. Experiment with different visualization tools
CO 4. Identify geospatial data and interconnection of data.
CO 5. Make use of the web for data extraction

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures not less than 35% (18 Marks out of 50)in the semester-end examination(SEE), and a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together

Continuous Internal Evaluation:

Three Unit Tests each of **20 Marks (duration 01 hour**)

1. First test at the end of 5th week of the semester
2. Second test at the end of the 10th week of the semester
3. Third test at the end of the 15th week of the semester

Two assignments each of **10 Marks**

4. First assignment at the end of 4th week of the semester
5. Second assignment at the end of 9th week of the semester

Group discussion/Seminar/quiz any one of three suitably planned to attain the COs and POs for **20 Marks** (**duration 01 hours**)

6. At the end of the 13th week of the semester

The sum of three tests, two assignments, and quiz/seminar/group discussion will be out of 100 marks and will be **scaled down to 50 marks**

(to have less stressed CIE, the portion of the syllabus should not be common /repeated for any of the methods of the CIE. Each method of CIE should have a different syllabus portion of the course).

CIE methods /question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the subject (**duration 03 hours)**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.

The students have to answer 5 full questions, selecting one full question from each module

Suggested Learning Resources:

Text Books

1. Data Visualization workshop, Tim Grobmann and Mario Dobler, Packt Publishing.
2. Python for Everybody: Exploring Data Using Python 3, Charles R. Severance, Create Space Independent Publishing Platform, 1st Edition, 2016

Reference:

1. "Data Visualization": A Successful Design Process, Kirk, Andy, Packt Publishing Ltd,2012
2. Think Python: How to Think Like a Computer Scientist, Allen B. Downey, Green Tea Press, 2nd Edition, 2015
3. Interactive Data visualization for the Web, Murray, Scott, O'Reilly Media, Inc., 2013
4. Visualizing Data: Exploring and Explaining Data with The Processing Environment, Fry, Ben, O'Reilly

Media, Inc., 2007

Web links and Video Lectures (e-Resources):

1. https://www.youtube.com/watch?v=eFByJkA3ti4
2. https://www.youtube.com/watch?v=JhK2qVi5dC4
3. https://www.youtube.com/watch?v=UjYzNhBVIvY
4. http://book.visualisingdata.com/
5. https://matplotlib.org/
6. https://docs.python.org/3/tutorial/
7. https://www.tableau.com/

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning