Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Scheme and Detailed syllabus (DJS23)

Third Year B.Tech

in

Computer Science and Engineering (Data Science)

(Semester VI)

Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

Scheme of Semester VI (DJS23) for Department of Computer Science and Engineering (Data Science)

Sr. No.	Course Code	Course	Teaching Scheme				Semester End Examination (SEE) - A						Continuous Assessment (CA) - B						Aggregate (A+B)	Credits Earned	
			Theory (Hrs)	Practical (Hrs)	Tutorial (Hrs)	Credits	Duration (Hrs)	Theory	Oral	Pract	Oral & Pract	SEE Total (A)	Term Test 1 (TT1)	Term Test 2 (TT2)	Term Test 3 (TT3)	Term Test Total (TT1 + TT2+ TT3)	Term Work	CA Total (B)			
1	DJS23DCPC601	Machine Learning – III (Reinforcement Learning)	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	3	**4**
	DJS23DLPC601	Machine Learning – III (Reinforcement Learning) Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25	1	
2	DJS23DCPC602	Natural Language Text Processing	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	3	**4**
	DJS23DLPC602	Natural Language Text Processing Laboratory	--	2	--	1	2	--	25	--	--	25	--	--	--	--	25	25	50	1	
3	DJS23DLPC603	Advanced Statistics Laboratory	--	4	--	2	--	--	--	--	--	--	--	--	--	--	50	50	50	2	**2**
4@	DJS23DCPE611	Time Series Analysis	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	4	**4**
	DJS23DLPE611	Time Series Analysis Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25		
	DJS23DCPE612	Analysis of AI Algorithms	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100		
	DJS23DLPE612	Analysis of AI Algorithms Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25		
	DJS23DCPE613	Medical Imaging Informatics and Interoperability	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100		
	DJS23DLPE613	Medical Imaging Informatics and Interoperability Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25		
	DJS23DCPE614	Ethical Hacking and Digital Forensics	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100		
	DJS23DLPE614	Ethical Hacking and Digital Forensics Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25		
5@	DJS23DCPE616	Computer Vision	4	--	--	4	2	60	--	--	--	60	15	15	10	40	--	40	100	5	**5**
	DJS23DLPE616	Computer Vision Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25		
	DJS23DCPE617	Robotics and AI	4	--	--	4	2	60	--	--	--	60	15	15	10	40	--	40	100		
	DJS23DLPE617	Robotics and AI Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25		
	DJS23DCPE618	Applied Game Theory	4	--	--	4	2	60	--	--	--	60	15	15	10	40	--	40	100		
	DJS23DLPE618	Applied Game Theory Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25		
	DJS23DCPE619	Information Security	4	--	--	4	2	60	--	--	--	60	15	15	10	40	--	40	100		
	DJS23DLPE619	Information Security Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25		
6	DJS23DLMD601	Applied Data Science Engineering Laboratory	--	2	--	1	--	--	--	--	--	--	--	--	--	--	25	25	25	1	**1**
7	DJS23ICHSX10	Environmental Studies	--	--	1	1	--	--	--	--	--	--	--	--	--	--	25	25	25	1	**1**
8	DJS23IPSCX04	Innovative Product Development IV	--	2	--	1	2	--	--	--	25	25	--	--	--	--	25	25	50	1	**1**
		Total	**13**	**14**	**1**	**21**	**12**	**240**	**25**	**0**	**25**	**290**	**60**	**60**	**40**	**160**	**200**	**360**	**650**	**21**	**21**

Any one Programme Elective Course (PEC) from 4@ and 5@

Continuous Assessment (A):

Course	Assessment Tools	Marks	Time (hrs.)
Theory	a. Term Test 1 (based on 40 % syllabus)	15	45
	b. Term Test 2 (on next 40 % syllabus)	15	45
	c. Assignment / course project / group discussion / presentation / quiz/ any other.	10	--
	Total marks (a + b + c)	40	--
Audit course	Performance in the assignments / quiz / power point presentation / poster presentation / group project / any other tool.	--	As applicable
Laboratory	Performance in the laboratory and documentation.	25	
Tutorial	Performance in each tutorial & / assignment.	25	
Laboratory &Tutorial	Performance in the laboratory and tutorial.	50	

The final certification and acceptance of term work will be subject to satisfactory performance upon fulfilling minimum passing criteria in the term work / completion of audit course.

Semester End Assessment (B):

Course	Assessment Tools	Marks	Time (hrs.)
Theory / * Computer based	Written paper based on the entire syllabus.	60	2
	* Computer based assessment in the college premises.		
Oral	Questions based on the entire syllabus.	25	As applicable
Practical	Performance of the practical assigned during the examination and the output / results obtained.	25	2
Oral & Practical	Project based courses - Performance of the practical assigned during the examination and the output / results obtained. Based on the practical performed during the examination and on the entire syllabus.	As per the scheme	2

Prepared by Checked by Head of the Department Principal

Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Computer Science and Engineering (Data Science)

(Semester VI)

Third Year B.Tech. (DJS23)

Programme Core Courses

Program: B.Tech in Computer Science and Engineering (Data Science) Semester: VI

Course: Machine Learning – III (Reinforcement Learning) (DJS23DCPC601)

Machine Learning – III (Reinforcement Learning) Laboratory (DJS23DLPC601)

Pre-requisite: Machine Learning-I, Machine Learning-II(Deep Learning) and Artificial Intelligence.

Objectives: To make students learn to build programs that act in a stochastic environment, based on experience using various Reinforcement Learning methods.

Outcomes*:* On completion of the course, learner will be able to:

1. Analyze basic and advanced Reinforcement Learning techniques.
2. Identify suitable learning tasks to which Reinforcement learning and Deep Reinforcement Learning techniques can be applied.
3. Apply appropriate Reinforcement Learning method to solve a given problem.

Machine Learning – III (Reinforcement Learning) (DJS23DCPC601)		
Unit	**Description**	**Duration**
1	**Introduction:** Reinforcement Learning (RL), Elements of Reinforcement Learning, Reinforcement Learning Vs Supervised Learning, Approaches of solving Reinforcement Learning: Value based, policy based, model based, Exploration - Exploitation dilemma, Evolutionary methods, Immediate Reinforcement Learning.	**04**
2	**Immediate Reinforcement Learning:** **Bandit Problems:** Bandit problems, Value-action based methods (sample average), Greedy method, €-greedy method, Incremental Implementation, Non-stationary problem, Optimistic Initial values, UCB algorithm, Thompson Sampling. **Policy Gradient Approaches**: Linear reward Penalty Algorithm, Parameterized policy representation(Θ), Evaluation of policy(η(Θ)), REINFORCE algorithm.	**06**
3	**Full Reinforcement Learning:** Difference between Immediate and Full Reinforcement Learning, Agents and Environment, Goals, Rewards, Returns, Policy in Full Reinforcement Learning, Episodic and Continuing Tasks. **Markov Decision Process (MDP)** Markov Property, Finite Markov Decision Process, Value functions, Bellman's equations, optimal value functions, Definition of MDP in Reinforcement Learning, Solution of the Recycling Robot problem	**08**
4	**Dynamic Programing:** Policy evaluation, policy improvement, policy iteration, value iteration, bootstrap, full back up. **Monte Carlo Method** Advantages of Monte Carlo over Dynamic Programing, Monte Carlo Control, on-policy, off- policy, Issues/Assumptions in Monte Carlo Methods, Solution of BlackJack using Monte Carlo Method.	**08**

5	**Temporal Difference Learning:** What is Temporal Difference learning, Advantages of Temporal Difference methods over Monte Carlo and Dynamic Programming methods, TD (0), On-policy vs off-policy, SARSA, Q learning. **Eligibility traces** N-step Temporal Difference methods, On-line vs Off-line updation, TD (λ): forward view, backward view, Traces: Accumulating trace Vs Replacing trace, Equivalence of forward and backward view, SARSA (λ)	**08**
6	**Deep Reinforcement Learning:** **Function Approximation** Drawbacks of tabular implementation, Function Approximation, Gradient Descent Methods, Policy gradient with function approximation. **Deep Reinforcement Learning** Intro of Deep Learning in Reinforcement Learning, Deep Learning training workflow, Categories of Deep Learning, Deep Q-Network, Ways of improving Deep Q-Network, REINFORCE in Full Reinforcement Learning, Actor-Critic Algorithm, A2C, A3C, DDPG. Proximal Policy Optimization (PPO).	**08**
	Total	**42**

Machine Learning – III (Reinforcement Learning) Laboratory (DJS23DLPC601)	
Sr. No	**Suggested experiments**
1.	Bandit Problem: • Implement Greedy and Epsilon greedy methods. • Comparison between Greedy and Epsilon Greedy Policy • UCB: Upper Confidence Bound
2.	Policy Gradient (Convergence) • Implement REINFORCE algorithm on a CartPole/ Lunar Lander.
3.	Dynamic Programming and Monte Carlo Methods • Implementation of GridWorld using Dynamic Programming • Jack's Car Rental using Dynamic Programming • Gamblers Problem using Dynamic Programming • BlackJack using Monte Carlo • Racetrack Problem
4.	Temporal Difference • Implement Frozen Lake using SARSA • Implement Grid world using Q learning
5.	Deep Reinforcement Learning • Compare the performance of Reinforcement Learning and Deep Reinforcement Learning on a Cart pole problem. • Implementation of Deep Q-Network algorithm • Actor Critic: Find the optimal policy using the Actor Critic method. • Analyze the effects of PPO's clipping parameter and learning rate on policy stability and convergence.

Minimum 10 experiments from the above-suggested list or any other experiment based on syllabus will be included, which would help the learner to apply the concept learnt.

* The Term Work will be calculated based on Laboratory Performance (15marks) and Quizzes/ Assignments (10marks).

Books Recommended:

Textbooks:

1. Richard S. Sutton and Andrew G. Barto, "Reinforcement Learning: An Introduction", MIT Press, 2nd Edition, 2022.
2. Laura Graesser Wah Loon Keng, "Foundations of Deep Reinforcement Learning," Pearson Education, 1st Edition, 2020.

Reference Books:

1. Phil Winder, "Reinforcement Learning Industrial Applications of Intelligent Agents", O'Reilly, 1st Edition, 2020.
2. Csaba Szepesvari, "Algorithms for Reinforcement Learning," Morgan & Claypool Publishers, 1st Edition, 2019.
3. Enes Bilgin, "Mastering Reinforcement Learning with Python", Packt publication, 1st Edition, 2020.
4. Brandon Brown, Alexander Zai, "Deep Reinforcement Learning in Action", Manning
5. Publications, 1st Edition, 2020.
6. Micheal Lanham, "Hands-On Reinforcement Learning for Games," Packt Publishing, 1st Edition, 2020
7. Abhishek Nandy, Manisha Biswas, "Reinforcement Learning: With Open AI, TensorFlow and Keras using Python," Apress, 1st Edition, 2018.

Weblinks:

1. NPTEL Course in Reinforcement Learning
   Reinforcement Learning - Course
2. Reinforcement Learning Course (Stanford University):
   https://www.youtube.com/watch?v=FgzM3zpZ55o
3. AI Games with Deep Reinforcement Learning: https://towardsdatascience.com/how-to-teach-an-ai- to-play-games-deep-reinforcement-learning-28f9b920440a
4. Deep Reinforcement Learning: https://www.v7labs.com/blog/deep-reinforcement-learning-guide

Program: B.Tech in Computer Science and Engineering (Data Science) Semester: VI

Course: Natural Language Text Processing (DJS23DCPC602)

Natural Language Text Processing Laboratory (DJS23DLPC602)

Pre-requisite: Machine Learning-II(Deep Learning), Foundations of Data Analysis, Statistics for Data Science

Objectives: To introduce basics of language computation fundamental through morphological computation, syntax, semantic and discourse analysis. Apply these concepts to develop Computational Models for Real World Applications.

Outcomes: **On completion of the course, learner will be able to:**

1. Apply appropriate pre-processing techniques on linguistic data.
2. Analyze different Machine Learning and deep learning algorithms to develop applications based on natural language processing.
3. Evaluate Natural Language Processing Applications.

Natural Language Text Processing (DJS23DCPC602)		
Unit	**Description**	**Duration**
1	**Introduction:** Generic Natural Language Processing (NLP) system, levels of NLP, Knowledge in language processing, Ambiguity in Natural language, stages in NLP, challenges of NLP, Applications of NLP Machine Translation, Sentiment Analysis etc. **Text Processing:** Word Tokenization and Segmentation, Lemmatization, Bag of words, N-gram language model, N-gram for spelling correction. Edit distance – Dynamic Programming Approach, Weighted Edit Distance, Finding Dictionary Entries with Small Edit Distances, Noisy Channel Model, Non-word errors Real-word errors. Evaluation of Language Models, Basic Smoothing, Advanced Smoothing Models. Advanced: Perplexity's Relation to Entropy.	**5**
2	**Computational Semantics and Semantic Parsing:** **Vector Semantics:** Words and Vectors, Term Frequency-Inverse Document Frequency (TFIDF), Word2vec, Continuous Bag of Words, ELMO, GloVe Vector Visualizing Embedding's, Semantic properties of embedding's, Bias and Embedding's Evaluating Vector Models, Cosine for measuring similarity, Pointwise Mutual Information (PMI), PPMI vector models. **Lexical Semantics:** Word Senses -Relations Between Senses, WordNet: A Database of Lexical Relations, Word Sense Disambiguation Alternate WSD algorithms and Tasks.	**9**
3	**Text Classification:** Text classification definition and datasets, Generative text classifiers (Naïve Bayes) Discriminative text classifiers (Support Vector Machine), Bag-of-words Generative Classifier, BOW Discriminative Model, Multi-class Classification: Softmax, Gradient Descent, Statistical significance testing, Dataset understanding and creation.	**5**

4	**Recurrent Neural Networks:** Recurrent Neural Network, RNNs as Language, RNNs for Sequence Classification, Stacked Recurrent Neural network, Bidirectional RNNs, Managing Context in RNNs: Long Short-Term Memory (LSTMs) and gated Recurrent Unit (GRUs).	**7**
5	**Computational Morphology and Syntax Analysis:** **Computational Morphology:** Morphological Processes, Morphological Analysis- Inflectional morphology & Derivational morphology, Regular expression, Finite State Automata, Finite State Transducer, Morphological parsing with FST, Lexicon free FST Porter stemmer, Two - level Morphology. **Syntax Analysis:** Introduction to POS Tagging, Probabilistic Tagging, Markov Models, Hidden Markov Models (HMM) for POS Tagging, Conditional Random Fields (CRF), Named Entities and Named Entity Tagging, Context-Free Grammars-Derivation, Constituency Parsing, Dependency Parsing.	**12**
6	**Discourse Coherence:** Coherence Relation, Discourse Structure Parsing, Centring and Entity-Based Coherence, Global Coherence.	**4**
	Total	**42**

Natural Language Text Processing Laboratory (DJS23DLPC602)	
Sr. No	**Suggested experiments**
1	Perform Pre-processing steps in Natural Language Processing (Tokenization, Stop Word detection, Stemming and Lemmatization.
2	Implement Parts of Speech tagging using HMM
3	Implement word-embedding and TF-IDF vectors in Natural language Processing
4	Implement language model using Ngram language model
5	Generate recursive set of sentences using Context Free Grammar. Identify the word senses using "synset" in NLTK
6	Implement Spelling Check, Spelling Correction and Auto complete using Language models or CFG.
7	Implement a Spam classifier in Natural Language Processing
8	Implement Fake News Classifier Using LSTM-Deep Learning in NLP
9	Implement a Sentiment Analysis in Natural Language Processing
10	Implement NLP application on Regional Language
11	Implement Question Answering in NLP
12	Implement Chatbot in NLP
13	Implement Information Retrieval for extracting Text from Webpages and Images
14	Mini Project

Minimum 10 experiments from the above-suggested list or any other experiment based on syllabus will be included, which would help the learner to apply the concept learnt.

* The Term Work will be calculated based on Laboratory Performance (15m) and Quizzes (10m).

Books Recommended:

Textbooks:

1. Jurafsky and Martin, "Speech and Language Processing", Prentice Hall, 3rd Edition, 2020.
2. Uday Karnath, "Deep Learning for NLP and Speech Recognition", 1'1 Edition, 2019.

Reference Books:

1. Jelinek, F., "Statistical Methods for Speech Recognition", The MIT Press, 2022.
2. Yuli Vasiliev "Natural Language Processing with Python and spaCy - A Practical Introduction", No Starch Press, 2022.
3. Sowmya Vajjala, Bodhisattwa Majumder, Anuj Gupta, Harshit Surana, "Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems", O'Reilly, 1st Edition, 2020.

Web Links

1. Virtual Lab: https://nlp-iith.vlabs.ac.in/
2. Virtual Lab: http://vlabs.iitb.ac.in/vlabs/dev/vlab_bootcamp/bootcamp/The_Big_Bang_Nerds/index.html
3. Nptel Course: https://nptel.ac.in/courses/106105158

Program: B.Tech in Computer Science and Engineering (Data Science) **Semester: VI**

Course: Advanced Statistics Laboratory (DJS23DLPC603)

Pre-requisite: Statistics for Data Science, Python and Machine Learning.

Objectives: This course introduces the theoretical and computational foundations of statistical inference, covering parameter estimation, hypothesis testing, and model diagnostics. It integrates classical, Bayesian, and nonparametric approaches with resampling and model assessment techniques to enable robust, data-driven decision-making.

Outcomes: On completion of the course, learner will be able to:

1. Derive and evaluate estimators under classical and Bayesian paradigms.
2. Conduct and interpret hypothesis tests for linear and logistic regression models.
3. Apply nonparametric testing methods when parametric assumptions fail.
4. Critically assess model performance and generalization.
5. Integrate statistical reasoning with computational implementation.

Advanced Statistics Laboratory (DJS23DLPC603)		
Unit	**Description**	**Duration**
1	**Estimation Theory Fundamentals:** • Estimation of Parameters using Maximum Likelihood Estimation (MLE) and Method of Moments • Estimate μ for a Normal (μ, σ^2) using MLE and compute Cramér–Rao Lower Bound (CRLB). • Derive MLE for Bernoulli(p), Normal Distributions and analyze estimator properties.	**06**
2	**Linear Regression Inference & Hypothesis Testing:** • Fit and Evaluate a Simple Linear Regression Model using t-test, F-test, and Confidence Intervals for Coefficients. • Compare Nested and Non-Nested Regression Models using Likelihood Ratio, AIC, and BIC Criteria.	**04**
3	**Regression Diagnostics and Model Assumptions:** • Perform Regression Diagnostics: Multicollinearity (VIF) and Normality Tests on Residuals. • Detect and Correct Heteroscedasticity using Breusch–Pagan Test, White Test, and Log/Weighted Least Squares Transformations.	**04**
4	**Logistic Regression and GLMs:** • Estimate and Interpret Logistic Regression Coefficients, Odds Ratios, and Model Fit using Hosmer–Lemeshow Test and Pseudo-R^2. • Extend Logistic Regression to Generalized Linear Models (GLMs) using Different Link Functions and Evaluate Model Deviance.	**06**
5	**Multinomial Logistic and Generalized Models:** • Model Count Data using Poisson Regression under the Generalized Linear Models (GLMs) Framework. • Fit and Interpret Multinomial and Ordinal Logistic Regression Models.	**06**

	• Compare Logit and Probit Link Functions within the Exponential Family of GLMs.	
6	**Nonparametric Hypothesis Testing**: • Compare Two Independent Samples using Mann–Whitney U and Wilcoxon Rank-Sum Tests. • Perform Goodness-of-Fit Testing using Kolmogorov–Smirnov and Chi-Square Tests. • Construct Bootstrap-Based Confidence Intervals and Conduct Nonparametric Inference.	**06**
7	**Bayesian Inference:** • Perform Bayesian Estimation for Normal and Binomial Models using Conjugate Priors and Credible Intervals.	**04**
8	**Markov Chain Monte Carlo (MCMC) Techniques:** • Implement MCMC Methods (Metropolis–Hastings and Gibbs Sampling) for Posterior Inference using PyMC/Stan. • Analyze Convergence Diagnostics and Visualize Posterior Distributions using Trace and Density Plots.	**04**
9	**Advanced Bayesian Methods** • Perform Bayesian Regression • Model Comparison using Posterior Predictive Checks and Information Criteria • Model Comparison (e.g., WAIC, LOO).	**02**
10	**Bayesian Hypothesis Testing:** • Bayes Factors (BF_{10} and BF_{01}) • Posterior Odds vs. Prior Odds • Evidence Interpretation Scale (Jeffreys scale) • Bayesian One-Sample & Two-Sample Tests (Mean difference) • Bayesian A/B Testing (Beta-Binomial model) • Hypothesis Testing using Credible Intervals	**02**
11	**Uncertainty Quantification and Resampling:** • Integrate Frequentist, Bayesian, and Resampling Approaches using Cross-Validation, Bootstrapping, and Permutation Methods. • Quantify Uncertainty through Confidence vs. Credible Intervals and Posterior Predictive Checks.	**06**
12	**High-Dimensional Inference** • Perform High-Dimensional Inference with Multiple Testing Corrections • Bonferroni • Benjamini–Hochberg FDR.	**02**
13	**Causal Inference, and Regularization Techniques:** • Apply Causal Inference and Regularization Methods: • Propensity Scores • Instrumental Variables • Ridge, Lasso, and Elastic Net.	**04**
	Total	**56**

* The Term Work will be calculated based on Laboratory Performance (15marks), Computer based Assessment (25 Marks) and Quizzes/ Assignments (10marks).

Books Recommended:

Text Books

1. G. Casella & R. L. Berger, Statistical Inference, 2nd Edition, Cengage Learning (2021).

2. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, Bayesian Data Analysis, CRC Press, 4rd Edition, 2025.
3. B. Efron and R. J. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall/CRC, Reprint Edition, 2003.
4. Agresti, Categorical Data Analysis, Wiley, 3rd Edition, 2013.

Reference Books

1. G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning with Applications in R, Springer, 2nd Edition, 2021.
2. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
3. R. McElreath, Statistical Rethinking: A Bayesian Course with Examples in R and Stan, CRC Press, 2nd Edition, 2020.
4. C. Davison and D. V. Hinkley, Bootstrap Methods and Their Application, Cambridge University Press, Reprint Edition, 2006.

Web Links:

1. L. Wasserman, All of Statistics: A Concise Course in Statistical Inference, Springer, 2004.
   All of Statistics: A Concise Course in Statistical Inference | SpringerLink
2. Bayesian Data Analysis in Python Course with Datacamp.
   Data Analytics A-Z with Python | Udemy

Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Computer Science and Engineering (Data Science)

(Semester VI)

Third Year B.Tech. (DJS23)

Programme Elective Courses
SET I

Program: B.Tech. in Computer Science and Engineering (Data Science) **Semester: VI**

Course: Time Series Analysis (DJS23DCPE611)

Time Series Analysis Laboratory (DJS23DLPE611)

Pre-requisite: Probability, Statistics and Linear Models.

Objectives: Learn basic analysis of time series data; concepts in time series regression, auto-regressive and model averaging models, learn basic concepts of spectral analysis and space-time models.

Outcomes: On completion of the course, learner will be able

1. Interpret a correlogram and a sample spectrum
2. Apply appropriate model for a time series dataset.
3. Compute forecasts for a variety of linear and non-linear methods and models.

Time Series Analysis (DJS23DCPE611)		
Unit	**Description**	**Duration**
1	**Introduction** Types of forecasting methods, Types of Time Series, simple descriptive techniques, trends in time series (Parametric trends, differencing, nonparametric methods, noise), seasonality. **Stationary Time Series** Formal definition of a time series, the sample mean and its standard error, Stationary processes: types of stationarity, statistical inference of time series.	**06**
2	**Linear Time Series** Motivation, Linear time series and moving average models, The AR model, simulating from an autoregressive process, The ARMA model, The ARIMA model, Unit roots, integrated and non-invertible processes, Box – Jenkins Model Selection, Seasonality, The SARIMA model.	**06**
3	**Prediction** Using prediction in estimating, forecasting for autoregressive processes, forecasting for AR, forecasting for general time series using infinite past, One-step ahead predictors based on the finite past: Levinson -Durbin algorithm; Forecasting for ARMA processes, State space model, The Kalman filter. A**utomated Forecasting Systems:** Auto-ARIMA, Auto Prophet, Auto ML for time series, Integration of Auto ML and Hashing for Time Series Forecasting.	**08**
4	**Models with Trend** Removing trend, Unit Root and Regression Residuals, The Monte Carlo Method, Dickey-Fuller tests. **Multi equation Time Series Models** Intervention Analysis, ADLs and Transfer Functions, Introduction to VAR, Vector Error Correction Model (VECM), Structural VAR (SVAR), Time-Varying Parameter VAR, Bayesian VAR (BVAR)	**10**

5	**Multivariate Time Series** Background: Sequences and Functions, Convolution, Spectral Representations and mean squared errors; Multivariate time series regression: Conditional independence, Partial correlation and coherency between time series.	**06**
6	**Non Linear Time series** The ARCH model: Feature of an ARCH, Existence of a strictly stationary solution, The GARCH model: Existence of stationary solution of a GARCH(1,1) and Bilinear models: Bilinear auto regression model(BAR), Advanced Nonlinear Models: Stochastic Volatility (SV) models, Nonlinear Autoregressive Models (NAR, NARX)	**06**
	Total	**42**

Time Series Analysis Laboratory (DJS23DLPE611)	
Exp.	**Suggested Experiments**
1.	Time Series Characteristics: Time Series Data, Cross-Section Data, Panel data/ Longitudinal data
2.	Trends: (1) Detecting trends using Hodrick -Prescott Filter. (2) Detrending a Time Series
3.	Seasonality: (1) Multiple Box Plots (2) Autocorrelation Plot (3) Deseasoning of Time-Series Data (4) Seasonal Decomposition (5) Detecting Cyclic Variations
4.	Data Wrangling and Preparation for Time Series Data
5.	Smoothing Methods: Simple exponential, Double exponential and Triple exponential.
6.	Making Data Stationary: Plots, Summary Statistics, Statistics Unit Root Tests, Augmented Dickey – Fuller Test.
7.	Automated Forecasting Systems using Auto-ARIMA.
8	Prophet, AutoML
9	Multivariate Time Series Analysis using VAR,VECM,SVAR modeling
10	Evaluation and Model Selection for Time Series Forecasting (Compare ARIMA vs XGBoost vs other candidate models using metrics like RMSE, MAE, and MAPE)

*The Term Work will be calculated based on Laboratory Performance (15m) and Quizzes/ Assignments (10m).

Books Recommended:

Text Books:

1. Walter Enders," Applied Econometric Time Series," Fourth Edition, Wiley, 2014.
2. B. V. Vishwas and Ashish Patel, "Hands-on Time Series Analysis with Python," First Edition, Apress, 2020

Reference Books:

1. Chris Chatfield, "Time- Series Forecasting," First Edition, Chapman & Hall/CRC, 2001.
2. Douglas C. Montgomery, Cheryl L. Jennings and Nurat Kulahci, "Introduction to Time Series Analysis and Forecasting," Second Edition, Wiley, 2015.
3. Aileen Nielsen, "Practical Time Series Analysis," O'Reilly, 2019.
4. James D Hamilton, "Time Series Analysis," Princeton University Press, 1994.
5. Robert H. Shumway and David S. Stoffer, "Time Series Analysis and Its Applications," Springer, 2000.

Web Links:

1. A course on Time Series Analysis. Course Dashboard | Time Series Analysis
2. A comprehensive guide to Time Series Analysis. Time Series Analysis: Definition, Components and Examples
3. The Complete Guide to Time Series Analysis and Forecasting. https://towardsdatascience.com/the-complete- guide-to-time-series-analysis-and-forecasting-70d476bfe775

Prepared by Checked by Head of the Department Principal

Program: B.Tech in Computer Science and Engineering (Data Science) **Semester: VI**

Course: Analysis of AI Algorithms (DJS23DCPE612)

Analysis of AI Algorithms Laboratory (DJS23DLPE612)

Pre-requisite: Data Structures, Design and Analysis of Algorithms, Machine Learning

Objectives: The course aims to develop analytical skills to understand algorithmic foundations, complexity, and performance trade-offs in designing efficient intelligent systems. It also equips students to apply, analyze, and evaluate data structures, optimization, probabilistic, randomized, and approximation algorithms for solving computationally complex AI problems effectively.

Outcomes: On completion of the course, learner will be able to:

1. Analyze algorithmic and computational complexity foundations, including time–space trade-offs and machine-learning model complexities during training and testing phases.
2. Apply various searching, indexing, and data-structuring techniques such as hash-based search, trees, graphs, and spatial indices to develop efficient retrieval and reasoning systems in AI.
3. Evaluate advanced optimization, probabilistic, randomized, and approximation algorithms to address complex AI problems considering efficiency, convergence, and scalability aspects.

Analysis of AI Algorithms (DJS23DCPE612)		
Unit	**Description**	**Duration**
1	**Algorithmic and Complexity Foundations:** Review of algorithm analysis: Time and space complexity Complexity classes: P, NP, NP-hard, NP-complete Trade-offs in AI algorithm design (accuracy, scalability, efficiency) **Complexity Analysis of Machine Learning Algorithms:** Training Time Complexity and Testing Time Complexity Train/Test Complexity of Linear Regression Train/Test Complexity of Random Forest Train/Test Complexity of Naïve Bayes Classifier Train/Test Complexity of SVM	**04**
2	**Searching and Indexing for AI:** Linear, Binary, Hash-based Searching, Heuristic Search, Search problems: Sliding tile puzzles, the Rubik's Cube, Sokoban, Inverted index and TF-IDF Vector space models and similarity search. High-dimensional indexing: KD-Trees, Ball Trees, Approximate Nearest Neighbor (ANN) search and LSH, Modern AI retrieval systems - IVF, FNSW/HNSW (FAISS, Annoy, ScaNN).	**06**
3	**Graphs and Trees in AI:** Flow networks and Ford–Fulkerson algorithm, Bipartite matching and Hungarian algorithm, Minimum spanning trees and disjoint sets. Tries, Tango tree, R Tree, Splay Tree, 2-3 Tree, Max spanning Tree, Binomial tree, and Binomial Heap, Operations on binomial heap.	**08**

4	**Optimization Algorithms:** Kernel Trick, Advanced First-Order Optimization Methods: Momentum and Gradient Methods (MGD, NAG, AdaGrad, RMSProp) Quasi Newton's Method, Hessian Approximation. (BFGS / L-BFGS) Swarm-based algorithms: PSO, ACO, GA, Differential Evolution Exploration vs. exploitation, convergence, and complexity trade-offs	**12**
5	**Probabilistic & Randomized Algorithms:** Probabilistic data structures: Bloom Filters, Count-Min Sketch, LogLog and HyperLogLog, random projections, and the Johnson–Lindenstrauss Lemma Randomized Algorithms: Monte Carlo and Las Vegas algorithm, Randomized gradient methods, Randomized matrix algorithms and sketching, Markov Chain Monte Carlo (MCMC), Gibbs sampling.	**06**
6	**Approximation Algorithms:** Approximation ratios and performance bounds, Low-rank matrix decomposition (SVD, PCA, NMF), Semi-definite programming and convex relaxations Applications in recommendation systems and clustering	**06**
	Total	**42**

Analysis of AI Algorithms Laboratory (DJS23DLPE612)	
Exp.	**Suggested Experiments**
1	Analyze and demonstrate how kernel functions map data to higher dimensions for better separability.
2	Analyze the implementation of Gradient Descent, Momentum, and Nesterov Accelerated Gradient (NAG) and compare their convergence speeds.
3	Analyze the Ford–Fulkerson algorithm by implementing it for a given flow network and evaluating how it determines the maximum flow.
4	Analyze Quasi-Newton optimization by implementing it and comparing its convergence behavior against first-order methods.
5	Analyze swarm-based optimization algorithms through implementation and evaluation of their search dynamics.
6	Analyze the structure of a 2–3 tree by implementing insertion, split, and deletion operations and validating balancing invariants under random inputs.
7	Analyze an R-Tree by designing and implementing it for multidimensional spatial indexing and evaluating its performance on range and nearest-neighbor queries.
8	Analyze Binomial Trees of various orders by constructing them and verifying their structural properties.
9	Analyze a randomized rounding algorithm by implementing it and evaluating its approximation quality.
10	Analyze differences between Monte Carlo and Las Vegas algorithms by implementing them on a sample problem and comparing probabilistic accuracy and runtime behavior.
11	Analyze the MAX-CUT problem by applying SDP relaxation and randomized rounding to evaluate approximation performance.
12	Analyze low-rank approximation techniques such as SVD, Tucker, or Tensor Train decomposition by implementing them and examining reconstruction error.

*The Term Work will be calculated based on Laboratory Performance (15marks) and Assignments/Quizzes (10marks).

Books Recommended:

Textbooks:

1. Sanjoy Dasgupta, Christos Papadimitriou, Umesh Vazirani, 'Algorithms', Tata McGraw-Hill, 1st Edition, 2023.
2. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, "Introduction to Algorithms", 4th Edition, The MIT Press, 2022.
3. Probabilistic Graphical Models: Principles and Techniques — Daphne Koller & Nir Friedman, 2018.

Reference Books:

1. Approximation Algorithms — Vijay V. Vazirani, 2021 (New edition).
2. Search Engines: Information Retrieval in Practice — Bruce Croft, Donald Metzler & Trevor Strohman, 3rd Edition, 2019.
3. Randomized Algorithms - Rajeev Motwani & Prabhakar Raghavan, 2020

Web Link

1. https://www.coursera.org/learn/advanced-algorithms-and-complexity
2. https://onlinecourses.nptel.ac.in/noc23_cs64/preview

Program: B.Tech in Computer Science and Engineering (Data Science) Semester: VI

Course: Medical Imaging Informatics and Interoperability (DJS23DCPE613)
Medical Imaging Informatics and Interoperability Laboratory (DJS23DLPE613)

Pre-requisite: Machine Learning, Cloud Computing and Security

Objectives: To equip students with the skills to understand medical imaging and digital pathology workflows, apply interoperability standards like HL7 v2.x, FHIR, and DICOM, and evaluate enterprise imaging systems including PACS, VNA, RIS, LIS, and Mirth Connect. Students will also learn to manage imaging data using DICOMweb and cloud platforms, and apply data science and AI methods for imaging analytics, digital pathology, and workflow integration.

Outcomes: On completion of the course, learner will be able to:

1. Describe the structure, components, and workflows of medical imaging and digital pathology systems.
2. Apply healthcare interoperability standards such as HL7 v2.x, FHIR, and DICOM for accurate and secure data exchange.
3. Design imaging data pipelines using Mirth Connect, PACS, VNA, and cloud-based DICOMweb services.
4. Analyze medical imaging and digital pathology datasets using data science and AI techniques, including federated learning workflows.

Medical Imaging Informatics and Interoperability (DJS23DCPE613)		
Unit	**Description**	**Duration**
1	**Modalities in Medical Imaging** Overview of medical imaging modalities (Radiology, Cardiology, Pathology): X-ray, XA, mammography, CT, MRI, PET, Ultrasound, Echocardiogram, Whole slide imaging (WSI).	**06**
2	**Healthcare Interoperability and Data Exchange Standards** HL7 v2.x messaging: ADT, ORM, ORU, SIU, MDM, DFT, BAR, RDS, RDE, ACK **Medical Coding & Terminology Standards**: LOINC, SNOMED CT, ICD, CPT, Healthcare API Standards, API Workflow in Healthcare, Challenges in Healthcare API Adoption.	**10**
3	**DICOM and Imaging Data Standards** DICOM architecture, DICOM file format, SOP classes, DIMSE operations (C-ECHO, C-STORE, C-FIND, C-MOVE). DICOM association, metadata, IODs. PACS workflow using DICOM DICOM for Digital Pathology (VL Whole Slide Microscopy Image IOD).	**08**
4	**DICOMweb & Cloud Imaging** JSON and REST-based imaging workflows, DICOM web services (WADO-RS, STOW-RS, QIDO-RS). AWS Health Imaging, S3, Sage Maker, IAM – DICOM storage and AI integration. GCP Healthcare API, Vertex AI, Big Query – DICOM web and FHIR interoperability. Azure Health Data Services, Blob Storage, Power BI – secure imaging workflows.	**08**

5	**Enterprise Imaging Systems** Imaging workflow and personas: Radiology, cardiology and pathology workflows, PACS, RIS, LIS, VNA, HIS, viewer, WSI scanner PACS and VNA architectures. IHE profiles – SWF, PIR, XDS-I.	**04**
6	**AI in Medical Imaging:** Imaging datasets (DICOM, NIFTI, Pathology data), federated learning, MonAI models and integrate in imaging workflow, Data anonymization	**06**
	Total	**42**

Medical Imaging Informatics and Interoperability Laboratory (DJS23DLPE613)	
Sr. No	**Suggested experiments**
1	Deploy a cloud-hosted PACS using Orthanc on AWS and verify DICOM C-STORE and C-FIND operations.
2	Configure DICOM Web endpoints on Azure and perform QIDO-RS, WADO-RS, and STOW-RS operations using REST tools.
3	Launch OHIF viewer on GCP and integrate with Orthanc to visualize CT/MRI studies and annotations.
4	Use Google Cloud Healthcare API to create a DICOM datastore, upload studies, run QIDO queries, and export metadata to Big Query.
5	Simulate a Radiology workflow using Mirth Connect by sending HL7 ADT and ORM messages to Orthanc and generating a DICOM worklist.
6	Convert Whole Slide Image (WSI) files to DICOM-VL format, upload to Orthanc, and view multi-resolution tiles in OHIF pathology mode.
7	Implement Federated Learning for medical imaging using Flower/NVFlare and train a shared CNN model without sharing raw data.
8	Train a MONAI-based UNet on MRI tumor segmentation data, export the trained model, and deploy an inference API using FastAPI.
9	Deploy MONAI Label and perform AI-assisted interactive annotation on CT/MRI studies using 3D Slicer or OHIF.
10	Mini Project

* The Term Work will be calculated based on Laboratory Performance (15m) and Quizzes/ Assignments (10m).

Books Recommended:

Text Books:

1. Tim Benson & Grahame Grieve, Principles of Health Interoperability: HL7, FHIR, and SNOMED, Springer, 2021.
2. Darren Treanor & Keith J. Dreyer, Digital Pathology: Current Practices and Future Directions, Springer, 2021.
3. Barton F. Branstetter, Practical Imaging Informatics: Foundations and Applications for PACS Professionals, Springer, 2021.

4. H.K. Huang, PACS and Imaging Informatics: Basic Principles and Applications, Wiley-Blackwell, 2010.
5. Oleg S. Pianykh, Digital Imaging and Communications in Medicine (DICOM): A Practical Introduction and Survival Guide, Springer, 2012.

Reference Books:

1. HL7 International, HL7 v2.x and FHIR Specifications, Available at: www.hl7.org, 2024.
2. IHE International, IHE Technical Frameworks (Radiology, IT Infrastructure, Pathology), 2024
3. NEMA, DICOM Standard (PS3.1–PS3.20) and Supplement 145: Whole Slide Imaging, National Electrical Manufacturers Association, 2023.
4. Daniel Rueckert et al., Deep Learning in Medical Image Analysis, Academic Press, 2020.
5. Maier, A., Steidl, S., Christlein, V., & Hornegger, J. (Eds.). (2018). Medical Imaging Systems: An Introductory Guide. Springer.
6. George C. Kagadis & Steve G. Langer, Informatics in Medical Imaging, CRC Press, 2011.

Web Links:

1. Health Informatics Specialization: Coursera, offered by Johns Hopkins University
2. Health Informatics on FHIR (edX/Georgia Tech)
3. HL7 FHIR Official Site: https://hl7.org/fhir
4. DICOM Standard: https://dicomstandard.org/
5. IHE Technical Frameworks: https://profiles.ihe.net/
6. NextGen Mirth Connect Documentation: https://docs.nextgen.com/display/mirthconnect
7. NIH Digital Pathology Resources: https://www.pathologyatlas.nih.gov/

Prepared by Checked by Head of the Department Principal

Program: Third Year B.Tech. in Computer Science and Engineering (Data Science) **Semester: VI**

Course: Ethical Hacking and Digital Forensics (DJS23DCPE614)

Ethical Hacking and Digital Forensics Laboratory (DJS23DLPE614)

Prerequisite: Information Security.

Objectives:

- To understand ethical hacking concepts, hacker classifications, and hacking methodologies.
- To understand and use basic tools and methods to find information about computer systems through footprinting, scanning, and enumeration.
- To introduce the phases and tools used in penetration testing and system hacking.
- To explain the fundamentals and significance of digital forensics in various domains.
- To develop the ability to collect, preserve, and analyze digital evidence using proper techniques and tools while considering legal and anti-forensic challenges.
- To familiarize students with modern forensic tools and techniques used in email and mobile device investigations.

Outcomes: On completion of the course, learner will be able to:

1. Explain ethical hacking concepts, hacker types, legal aspects, and real-world applications.
2. Use footprinting, scanning, and enumeration tools to gather system and network information.
3. Perform penetration testing and demonstrate basic system hacking techniques.
4. Identify digital evidence and apply file system and disk forensics tools.
5. Apply evidence collection, hashing, and anti-forensics detection methods.
6. Analyze network traffic, email artifacts, and mobile device data using forensic tools.

Ethical Hacking and Digital Forensics (DJS23DCPE614)		
Unit	**Description**	**Duration**
1	**Introduction to Ethical Hacking** Introduction to Ethical Hacking, Classification of Hackers (White Hat, Black Hat, Grey Hat), Phases of Ethical Hacking, Cybersecurity vs Ethical Hacking, Cyber Laws and Ethical Responsibilities, Introduction to Artificial Intelligence in Cybersecurity and Hacking, Real-World Case Studies of Ethical Hacking	**6**

2	**Footprinting, Scanning, and Enumeration** Footprinting Techniques, DNS Interrogation, Email Harvesting, Social Engineering, Footprinting Tools (Maltego, Recon-ng), Scanning Methodology, Port Scanning Types and Tools (Nmap, Netcat), Enumeration Techniques, Enumeration Tools (SNMP, SMB, LDAP), Banner Grabbing, Use of AI for Automated Reconnaissance and Threat Detection	**6**
3	**Penetration Testing and System Hacking** **Penetration Testing:** Fundamentals of Penetration Testing, Types of Penetration Testing (Black Box, White Box, Grey Box), Phases of Penetration Testing, Penetration Testing Tools (Metasploit, Burp Suite, Nikto,etc) **System Hacking:** Password Attacks (Brute-force, Dictionary, Rainbow Tables), Privilege Escalation Techniques, Malware and Rootkits, Executing Applications and Hiding Files, System Hacking Tools, Introduction to Dark Web and TOR Network, Dark Web-based Threat Intelligence and Anonymity Tools	**8**
4	**Introduction to Digital Forensics** Definition and Scope of Digital Forensics, Types and Characteristics of Digital Evidence, Phases of a Digital Investigation, File System Forensics (FAT, NTFS), Disk Imaging and Cloning (Bit-by-Bit Copy), Data Recovery Concepts, Deleted File and Slack Space Analysis, Metadata Extraction and Timestamp Interpretation, Disk Forensics Tools (FTK Imager, Autopsy, EnCase), Role of AI and Machine Learning in Digital Forensics	**7**
5	**Evidence Collection and Data Analysis** Evidence Collection Techniques (Live vs Dead), Chain of Custody and Legal Considerations, Volatile and Non-Volatile Evidence Acquisition, Remote Evidence Acquisition, Hashing Algorithms (MD5, SHA1, SHA256) for Verification, Write Blockers and Imaging Devices, Anti-Forensics Techniques (Data Hiding, Steganography, File Obfuscation), Detection and Countering Anti-Forensics, Use of AI Tools for Pattern Detection and Data Anomaly Analysis, Data Carving and Signature-Based Recovery.	**8**
6	**Network, Email, and Mobile Forensics** Network Forensics (Packet Capture, Flow Analysis), Live Traffic Monitoring Tools (Wireshark, TCPDump), Log File and Firewall Analysis, Intrusion Detection Logs, Honeynet and Sandbox Environments, Email Forensics (Header and Server Log Analysis, MIME Format), Mobile Device Forensics (Android and iOS), Acquisition Techniques (Logical, Physical, Cloud), SIM and App Data Extraction, Tools (Cellebrite, XRY, MOBILedit), Open Source Intelligence (OSINT) and Threat Attribution Techniques	**7**
	Total	**42**

Ethical Hacking and Digital Forensics Laboratory (DJS23DLPE614)	
Exp.	**Suggested experiments**
1	**Footprinting and Reconnaissance:** Objective: Gather information about a target using passive and active footprinting techniques. Tools: Recon-ng, theHarvester, SpiderFoot (for AI-powered OSINT), ThreatFox, VirusTotal
2	**DNS Interrogation and Email Harvesting** Objective: Perform WHOIS lookups, DNS zone transfers, and identify email addresses. Tools: nslookup, dig, whois, theHarvester
3	**Scanning and Enumeration** Objective: Identify open ports, services, and perform OS detection. Tools: Nmap, Netcat
4	**Enumeration** Objective: Extract user and system info via SMB and SNMP protocols. Tools: enum4linux, snmpwalk
5	**Web Application Scanning** Objective: Scan a web server for vulnerabilities. Tools: Nikto, OWASP ZAP, Burp Suite Community Edition
6	**Exploitation Using Metasploit Framework** Objective: Exploit vulnerabilities in a virtual test machine. Tools: Metasploit Framework, DVWA, Metasploitable VM
7	**Password Cracking** Objective: Perform brute-force and dictionary attacks on password hashes and login services. Tools: John the Ripper, Hydra, Hashcat
8	**Disk Imaging and Basic Forensic Analysis** Objective: Create and analyze a disk image. Tools: FTK Imager, Autopsy, dd
9	**File System and Deleted File Analysis** Objective: Recover deleted files and analyze file system metadata. Tools: Autopsy, Sleuth Kit (fls, icat), Scalpel
10	**Hashing and Data Integrity Verification** Objective: Generate and verify file hashes to maintain evidence integrity. Tools: md5sum, sha256sum, HashCalc
11	**Network Traffic Capture and Protocol Analysis** Objective: Capture and analyze live network traffic for suspicious activity. Tools: Wireshark, TCPDump
12	Email Header Analysis and Evidence Extraction Objective: Trace the source of an email and extract digital evidence. Tools: Autopsy (email plugin), ExifTool

13	**Live Memory Acquisition and Analysis (Windows/Linux)** Objective: Acquire and examine volatile memory for evidence. Tools: WinPmem (Windows), LiME (Linux), Volatility
14	**Steganography and Anti-Forensics Detection** Objective: Detect hidden data in images and analyze steganographic files. Tools: OpenStego, Steghide, binwalk
15	**Dark Web Exploration & TOR-based Threat Discovery** Objective:To explore TOR and Dark Web Securely Tools: TOR browser, OnionScan, Ahmia
16	**Mobile Device Forensics and App Data Extraction** Objective: TOR browser, OnionScan, Ahmia Tools: MOBILedit, ADB, Cellebrite (demo)

Any other practical covering the syllabus topics and subtopics can be conducted.

*The Term Work will be calculated based on Laboratory Performance (15m) and Assignments (10m).

Books Recommended:

Text books:

1. EC-Council "Ethical Hacking and Countermeasures Attack Phases", Cengage Learning.2nd Edition, 2017
2. Rafay Boloch, "Ethical Hacking and Penetration Testing Guide", CRC Press, 2014.
3. John R. Vacca, "Computer Forensics", Computer Crime Investigation Firewall Media, New Delhi. 2012
4. Nelson, Phillips, Steuart, "guide to Computer Forensics and Investigations", CENGAGE Learning, 6th Edition, 2020.
5. E. Casey, "Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet", 3rd ed. Burlington, MA, USA: Academic Press, 2011.
6. S. Davidoff and J. Ham, "Network Forensics: Tracking Hackers through Cyberspace". Upper Saddle River, NJ, USA: Prentice Hall, 2012.

Reference Books:

1. Kevin Smith, "Hacking How to Hack - The ultimate Hacking Guide", Hacking Intelligence.2018
2. Kevin Beaver, "Ethical Hacking for Dummies", Sixth Edition, Wiley, 2018.
3. Keith J. Jones, Richard Bejtiich, Curtis W. Rose, "Real Digital Forensics", Addison- Wesley Pearson Education 2006
4. Tony Sammes and Brian Jenkinson, "Forensic Compiling", A Tractitioneris Guide, Springer International edition.
5. Christopher L.T. Brown, "Computer Evidence Collection & Presentation", Firewall Media.
6. Jesus Mena, "Homeland Security, Techniques & Technologies", Firewall Media.
7. J. T. Luttgens, M. Pepe, and K. Mandia, "Incident Response and Computer Forensics", 3rd ed.

New York, NY, USA: McGraw-Hill Education, 2014.

Web Links:

1. **Ethical Hacking – IIT Kharagpur (NPTEL)**
   nptel.ac.in/courses/106/105/106105217/
2. **Digital Forensics – UTAustin / Prof. Matt L. (Free Online Course)**
   https://digital-forensics.utexas.edu/
3. **Introduction to Cybersecurity Tools & Cyber Attacks – IBM (Coursera Free Audit)**
   https://www.coursera.org/learn/intro-cybersecurity-tools-cyber-attacks
4. **Cybersecurity Fundamentals – University of Washington (edX Free Audit)**
   https://www.edx.org/learn/cybersecurity/university-of-washington-cybersecurity-fundamentals
5. **Digital Forensic Techniques – OpenLearn (The Open University)**
   https://www.open.edu/openlearn/science-maths-technology/digital-forensics/content-section-0
6. **Network Forensics – University of California, Davis (Coursera Free Audit)**
   https://www.coursera.org/learn/network-security
7. **Open-Source Intelligence (OSINT) Training – EUROPOL / ENLETS**
   https://osintframework.com/

Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Computer Science and Engineering (Data Science)

(Semester VI)

Third Year B.Tech. (DJS23)

Programme Elective Courses
SET II

Program: B.Tech. in Computer Science and Engineering (Data Science) **Semester: VI**

Course: Computer Vision (DJS23DCPE616)

Computer Vision Laboratory (DJS23DLPE616)

Pre-requisite: Machine Learning -I, Machine Learning- II(Deep Learning)

Objectives: To equip students with advanced skills in image, video, and computer vision processing, emphasizing feature extraction, shape and motion analysis, object detection, and action recognition through both classical and deep learning approaches for effective visual understanding and prediction.

Outcomes: On completion of the course, learner will be able

1. Apply fundamental image and video processing techniques for visual data analysis.
2. Analyze object detection and recognition algorithms for visual understanding.
3. Illustrate generative and segmentation-based visual learning models for scene understanding.

Computer Vision (DJS23DCPE616)		
Unit	**Description**	**Duration**
1	**Foundations of Computer Vision and Image Preprocessing** Basics of image processing. Image acquisition. Preprocessing. Enhancement. Segmentation. Representation. Image transforms: Fourier Transform, Discrete Cosine Transform, Wavelet Transform. **Morphological Image Processing:** Dilation, Erosion, Opening and Closing, Hit-or-Miss Transformation. Morphological Algorithms: Boundary Extraction, Region Filling, Extraction of Connected Components, Thinning, Thickening, Skeletons, Pruning, Morphological Reconstruction. Image Preprocessing Challenges: Clutter, deformation, intra-class variation, Gaussian blur, noise removal, illumination variance, normalization techniques.	**09**
2	**Shape Analysis and Segmentation Techniques in Vision** Contour-Based Methods: Chain Codes, Geometric Border Representation, Fourier Transform of Boundaries, Boundary Description using Segment Sequences, B-Spline Representation, Shape Invariants. Region-Based Methods: Scalar Region Descriptors, Moments, Convex Hull, Graph Representation Using Region Skeletons, Region Decomposition, Region Neighborhood Graphs. **Thresholding:** Foundation, Role of illumination, Basic Global thresholding, Otsu's method Region Based segmentation: Region Growing, Region Splitting and merging, Relationships between pixels, Hough transform.	**10**
3	**Object Detection:** Two Stage/Proposal: Convolutional Neural Networks for Detection: R-CNN, Fast R-CNN, Faster R-CNN, RFCN and Mask RCN; Architecture and Issues in each algorithm. Backprop-to-image/Deconvolution Methods. **One Stage/Proposal Free**: YOLO, SSD, evaluation metrics (IoU, AP), Non-max suppression YOLO Loss function, Variants of YOLO. **Face Recognition and Verification:** Zero-shot, One-shot, Few-shot Learning;	**10**

	Siamese Networks, Triplet Loss, Contrastive Loss, Ranking Loss; Attention Models in Vision.	
4	**Generative Models:** Types of generative models: Implicit and Explicit density; Generative Adversarial Network; Vanilla GAN, Mode Collapse in GAN (Strategies to address Mode Collapse and Convergence Issues); Conditional GAN, DC GAN, Wasserstein GAN (WGAN), CycleGAN, StyleGAN; GAN objective functions, JSD Divergence, EM Distance, Least Squares, Evaluation Metrics: Inception Score (IS), Fréchet Inception Distance (FID)	**09**
5	**Object Segmentation:** Semantic segmentation, Scene Parsing, semantic flow, Bilinear Interpolation, Symmetry in Segmentation, Featured image pyramid, pixel-wise softmax, PSPNet, FPN, UNet, clustering method for segmentation, Distance metrics(Euclidean, Cosine, Hamming, Manhattan, Minkowski, Chebyshev, Jaccard, Haversine), Linkage Types (Single, Average, Complete, Centroid).	**10**
6	**Motion Analysis and Optical Flow:** Basics of motion estimation and optical flow using classical and deep learning methods. 3D Vision Components: Stereo Matching, disparity estimation, depth computation. **Action Recognition and Object Tracking:** Introduction to recognizing actions and tracking moving objects in videos. Video Processing and Spatio-Temporal Features Overview of video signals, motion models, and basic spatio-temporal analysis.	**08**
	Total	**56**

Computer Vision Laboratory (DJS23DLPE616)	
Exp.	**Suggested Experiments**
1	To perform morphological operations on Image.
2	To perform image enhancement in frequency domain.
3	To perform segmentation using region growing , merging and splitting.
4	To detect cancer cells using medical image processing techniques through CNN-based object detection.
5	To identify vehicles from road traffic CCTV video footage using the YOLO object detection algorithm.
6	To convert black-and-white images into colored images using GAN.
7	To detect deepfakes in digital media using GAN.
8	To perform image segmentation using advanced deep learning models such as PSPNet, FPN, and U-Net.
9	To analyze body postures through motion analysis using spatio-temporal feature extraction techniques.
10	Mini Project

*The Term Work will be calculated based on Laboratory Performance (15marks) and Quizzes/Assignments (10marks).

Books Recommended:

Text Books:

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning*, MIT Press, Reprint Edition with updates, 2022.
2. Shafqat Alauddin, Mrutyunjaya S. Yalawar, S. Bharathidasan, T. Thiyagarajan, Image Processing Techniques and its Applications in Computer Vision and Artificial Intelligence, 2024.

Reference Books:

1. Richard Szeliski, Computer Vision: Algorithms and Applications, 2nd Edition, Springer, 2022.
2. Shafqat Alauddin, Mrutyunjaya S. Yalawar, S. Bharathidasan, T. Thiyagarajan, Image Processing Techniques and its Applications in Computer Vision and Artificial Intelligence, 2024.

3. Kevin P. Murphy, Probabilistic Machine Learning*:* An Introduction, MIT Press, 2022.
4. Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing*, 4th Edition, Pearson Education, 2021.

Web Links:

1. Virtual Lab on Vision and deep learning Lab https://www.ee.iitb.ac.in/~viplab/
2. Virtual Lab on Computer Vision Laboratory https://www.iitk.ac.in/ee/computer-vision-lab
3. Course on Modern Computer Vision https://www.youtube.com/playlist?list=PLzWRmD0Vi2KVsrCqA4VnztE4t71KnTnP5
4. Coursera course on Advanced Computer Vision with TensorFlow https://www.coursera.org/learn/advanced-computer-vision-with-tensorflow
5. Udemy course on Deep Learning and Computer Vision A-Z™: OpenCV, SSD & GANs | Udemy
6. Vision Lab: Computer Vision http://cse.iitm.ac.in/lab_details.php?arg=NQ
7. Funded Projects on Computer Vision at NAVER LABS Europe https://europe.naverlabs.com/research/computer-vision/

Prepared by Checked by Head of the Department Principal

Program: B.Tech in Computer Science and Engineering (Data Science) Semester: VI

Course: Robotics and AI (DJS23DCPE617)
Robotics and AI Laboratory (DJS23DLPE617)

Pre-requisite: Linear algebra and Probability theory

Objectives: To introduce fundamental concepts, kinematics, perception, planning, and control in robotics with an emphasis on data acquisition, machine learning, and autonomous navigation.

Course Outcomes: On completion of the course, learner will be able to:

1. Illustrate the fundamental principles, components, and kinematics of robotic systems.
2. Apply sensor data acquisition, preprocessing, and fusion techniques using Python and ROS.
3. Analyze computer vision, path planning, and SLAM algorithms for perception and navigation.
4. Develop control and decision-making strategies for autonomous robot operation.

Robotics and AI (DJS23DCPE617)		
Unit	**Description**	**Duration**
1	**Introduction to Robotics** Evolution of Robotics, Types of robots – industrial, mobile, humanoid, and autonomous, Components of robotic system (sensing, actuation, control, AI), Types of Sensors – IMU, LiDAR, Camera, GPS, Ultrasonic, Proximity Sensors (Infrared), Vision Sensors, Accelerometers, Gyroscopes, and Encoders (Linear or Rotary), Types of Actuators – Electric (DC, Stepper, Servo Motors), Hydraulic, Pneumatic, and Emerging Smart Actuators (Piezoelectric, Shape Memory Alloy, Electroactive Polymer).	**08**
2	**Robot Kinematics and Motion planning** Coordinate frames and transformations – homogeneous transformation matrices, Forward and inverse kinematics; Denavit–Hartenberg (D–H) parameters, Differential kinematics and Jacobians – conceptual and numerical examples, Configuration space and robot motion representation, Motion planning – introduction, path and trajectory concepts,Types of trajectory planning – Point-to-Point (PTP) and Continuous Path (CP).	**12**
3	**Data Acquisition and Preprocessing** Data collection and preprocessing from multiple sensors through frameworks and middleware (e.g., ROS) for acquisition, synchronization, and refinement of sensor data., Understanding sensor data formats and structures (CSV, JSON, image/video files and ROS bag files). Noise handling using Gaussian, Median, and Kalman filtering methods. Aligning multi-sensor data streams (camera, LiDAR, IMU) using temporal synchronization, spatial calibration, and sensor fusion methods. Handling missing or corrupted data using interpolation, statistical imputation, and smoothing methods, Sensor calibration using intrinsic and extrinsic methods for camera–LiDAR and IMU sensors.	**08**
4	**Robot Perception** Introduction to robot perception and environment understanding, Computer vision in robotics – image acquisition, feature extraction, and object recognition, Application of machine learning and deep learning for perception: Classification using CNN and SVM, Object detection using YOLO and SSD, Image segmentation using U-Net and SegNet, building 3D representations from LiDAR using Voxel	**08**

	Grid Mapping or from stereo vision using SGBM (Semi-Global Block Matching).	
5	**Path Planning** Path planning algorithms –Rapidly-exploring Random Tree(RRT), and PRM, Simultaneous Localization and Mapping (SLAM),Components – Localization, mapping, sensor data processing, Types of SLAM – Visual SLAM, LiDAR-based SLAM, RGB-D SLAM, Data association and map building, Learning-based SLAM and navigation using neural implicit mapping and policy learning, Integration of SLAM with path planning and control for autonomous navigation, Evaluation metrics – accuracy, drift, real-time performance, and scalability, Reinforcement Learning for Navigation using Q-Learning, Sim-to-Real Transfer – Domain Randomization, Imitation Learning, Challenges, and Evaluation.	**12**
6	**Control and Decision Making** Control architectures – Open Loop, Closed Loop, and Hierarchical, PID and Feedback Control, Model Predictive Control (MPC) for trajectory tracking, Behavior-Based Robotics – Subsumption Architecture and Reactive Control, Reinforcement Learning for Robot Control – DDPG, PPO, A3C, TRPO, TD3, Human-Robot Interaction – Shared Autonomy, Intent Recognition, and Safety Considerations.	**08**
	Total	**56**

Robotics and AI Laboratory (DJS23DLPE617)	
Sr. No	**Suggested Experiments**
1	Install and explore ROS or Webots; visualize a simple robot model.
2	Simulate a robotic arm and perform forward and inverse kinematics using Python or MATLAB.
3	Acquire and preprocess sensor data (camera or LiDAR dataset) using Python.
4	Implement visual feature detection using SIFT or SURF, or deep object recognition using YOLO or MobileNet.
5	Simulate SLAM using ROS or Gazebo datasets
6	Implement EKF-based localization or 2D SLAM using Python
7	Implement a RRT-based path planner in Webots or Gazebo
8	Implement PID control for a mobile robot simulation using Python or Webots.
9	Implement a simple reinforcement learning control policy for navigation or balancing tasks using Python or Webots.
10	Integrate perception, localization, and planning on a simulated robot using TurtleBot or drone simulation.

* The Term Work will be calculated based on Laboratory Performance (15marks) and Assignments/ Quizzes (10marks).

Books Recommended

Text Books:

1. Roland Siegwart, Illah R. Nourbakhsh, and Davide Scaramuzza**,** "Introduction to Autonomous Mobile Robots", 3rd Edition, MIT Press, 2022.
2. Peter Corke**,** "Robotics, Vision and Control: Fundamental Algorithms in Python", 3rd Edition, Springer, 2023.
3. Jonathan Cacace**,** "Ultimate Robotics Programming with ROS 2 and Python", 1st Edition, 2024.

Reference Books:

1. Mohamed M. Atia**,** "Sensor Fusion Approaches for Positioning, Navigation, and Mapping: How Autonomous Vehicles and Robots Navigate in the Real World with MATLAB Examples", 1st Edition, Springer, 2025.
2. Christoph Bartneck, Tony Belpaeme, Friederike Eyssel, Takayuki Kanda, Merel Keijsers & Selma Šabanović**,** "Human-Robot Interaction – An Introduction", 2nd Edition, Cambridge University Press, 2024.

3. Larry T. Ross, Stephen W. Fardo & Michael F. Walach**,** "Industrial Robotics Fundamentals", 4th Edition, Jones & Bartlett Learning, 2023.

Web Links:

1. https://swayam.gov.in/nc_details/NPTEL (Introduction to Robotics by IIT Madras and Robotics by IIT Kharagpur)
2. https://www.udemy.com/course/robotics-course/
3. https://www.coursera.org/courses?query=robotics

Prepared by Checked by Head of the Department Principal

Program: B.Tech in Computer Science and Engineering (Data Science) **Semester: VI**

Course: Applied Game Theory (DJS23DCPE618)

Applied Game Theory Laboratory (DJS23DLPE618)

Pre-requisite: Linear Algebra, Calculus, Probability, Statistics and Basic algorithm design and analysis.

Objectives:

- Introduce fundamental principles of strategic, zero-sum, non-zero-sum, and repeated games with computational simulation.
- Develop the ability to model, analyse, and solve multi-agent interactions using algorithmic and Python-based approaches.
- Familiarize students with evolutionary, cooperative, and Bayesian game-theoretic frameworks, including reinforcement learning and human-in-the-loop feedback.
- Apply game-theoretic reasoning to real-world applications such as auctions, market design, bargaining, and resource allocation.

Outcomes: On completion of the course, learner will be able to:

1. Analyse and simulate strategic, zero-sum, and non-zero-sum games to identify Nash equilibria and optimal strategies.
2. Apply algorithmic techniques and reinforcement learning to compute equilibria and model multi-agent interactions.
3. Evaluate evolutionary, cooperative, and Bayesian game scenarios for fairness, stability, and decision-making efficiency.
4. Design, implement, and assess real-world games and applications, including auctions, repeated games, market design, and RL-based strategy optimization.

Applied Game Theory (DJS23DCPE618)		
Unit	**Description**	**Duration**
1	**Introduction to Game Theory:** Strategic games, players, strategies, payoffs; dominance, minimax, saddle points; pure & mixed strategy Nash equilibria; real-world examples: auctions, pricing, simple multi-agent interactions	**08**
2	**Zero-Sum and Non-Zero-Sum Games** Zero-sum games, saddle points, matrix games; mixed strategies; non-zero-sum games, iterated elimination of dominated strategies; Lemke–Howson algorithm.	**08**
3	**Evolutionary and Cooperative Game Theory** Evolutionarily Stable Strategies (ESS), replicator dynamics, fictitious play; cooperative games: transferable utility, core, Shapley value, nucleolus; correlated equilibria; Multi-Agent Reinforcement Learning (MARL) and reward shaping with human feedback.	**10**
4	**Bayesian and Algorithmic Game Theory** Bayesian games, Bayes–Nash equilibrium, auctions, bilateral trading; complexity of equilibrium computation; mechanism design basics; Vickrey auction, incentive-compatible resource allocation.	**10**

5	**Repeated and Extensive Form Games** Repeated games: Nash Folk Theorem, subgame perfect equilibrium, one-shot deviation principle; extensive form games: game trees, backward induction, sequential equilibria	**10**
6	**Game Design, Simulation, and Real-World Applications** Game design principles: payoff engineering, fairness, multiplayer dynamics; applications: oligopoly models, voting games, matching markets, resource allocation, utility theory; RL-based strategy optimization and human-in-the-loop feedback	**10**
	Total	**56**

Applied Game Theory Laboratory (DJS23DLPE618)	
Exp.	**Suggested experiments**
1	Strategic-Form Games and Payoff Matrix Construction: Construct and simulate matrix games; compute best responses and pure Nash equilibria.
2	Dominance Analysis and Mixed Strategy Nash Equilibria: Analyse dominance relations, compute mixed strategies, and visualize payoffs.
3	Zero-Sum Game Simulation and Minimax / Saddle Points: Implement zero-sum games and compute saddle points using Python.
4	Non-Zero-Sum Game Simulation and Iterated Elimination of Dominated Strategies: Solve non-zero-sum games computationally and identify NE.
5	Lemke–Howson Algorithm Implementation: Compute Nash equilibria of bimatrix games algorithmically.
6	Evolutionarily Stable Strategies (ESS) and Replicator Dynamics: Simulate population dynamics and ESS in multi-agent systems.
7	Fictitious Play and Multi-Agent Reinforcement Learning (MARL): Implement learning strategies for repeated interactions and observe convergence.
8	Cooperative Game Theory: Core, Shapley Value, Nucleolus Computation: Compute fair allocations in transferable utility games.
9	Bayesian Games and Bayes–Nash Equilibrium: Implement games with incomplete information and compute equilibria.
10	Auction Simulation: First-Price, Second-Price, Sealed-Bid, and Online Auctions; Analyse strategy and outcomes.
11	Repeated and Extensive Form Games: Simulate repeated strategies, subgame-perfect equilibrium, and backward induction in game trees.
12	Matching Market and Resource Allocation Applications: Implement market design, matching, and allocation scenarios computationally.
13	Custom Game Design with RL and Human-in-the-Loop Feedback: Design a multi-agent game, implement RL agents, incorporate human feedback, and analyse strategic outcomes.

*****The Term Work will be calculated based on Laboratory Performance (15m) and Assignments (10m).

Books Recommended:

Text Books:

1. Game Theory: An Introduction, 3rd Edition by E.N. Barron, Wiley, 2024.
2. Richard S. Sutton and Andrew G. Barto, "Reinforcement Learning: An Introduction", MIT Press, 2nd Edition, 2022.
3. Binmore, K., Game Theory: A Very Short Introduction, OUP, 2010

Reference Books:

1. Shoham Y., Leyton-Brown K., Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations, Cambridge University Press, 2024
2. Thomas Ferguson, Game Theory, World Scientific, 2018

Weblinks:

1. NPTEL Course: https://nptel.ac.in/courses/106105237
2. IIT Bombay: Useful Lecture Notes on Game Theory | IEOR @ IIT Bombay

Program: B.Tech in Computer Science and Engineering (Data Science) **Semester: VI**

Course: Information Security (DJS23DCPE619)

Information Security Laboratory (DJS23DLPE619)

Pre-requisite:
Computer Communication and Networks

Objectives:
- To understand the fundamental principles of cryptography, network security, and secure communication mechanisms in modern computer systems.
- To apply symmetric and asymmetric encryption techniques to ensure confidentiality, authentication, and data integrity.
- To analyse and evaluate security threats, vulnerabilities, and countermeasures, including the application of machine learning techniques for intelligent threat detection.

Outcomes: On completion of the course, the learner will be able to:
1. Explain the fundamental concepts of cyber-attacks, defence strategies, guiding principles of modern security practices, and apply number theory concepts such as modular arithmetic, Euclid's algorithm, and classical ciphers to illustrate basic cryptographic operations.
2. Apply appropriate encryption, hashing, and authentication protocols to design and implement secure systems and data communication models.
3. Analyse different network threats and attacks such as DDoS, spoofing, and phishing, and evaluate suitable defence mechanisms, including IDS, IPS, and firewalls.
4. Evaluate the role of machine learning models in cybersecurity, examining their effectiveness in detecting and mitigating cyber threats in real-world scenarios.

Information Security (DJS23DCPE619)		
Unit	**Description**	**Duration**
1	**Introduction:** Cyber Attacks, Defense, Strategies and Techniques, Guiding Principles of Modern Security Practices. OSI security model. **Number Theory:** Modulo Arithmetic, Euclid's Algorithm, Fermat's and Euler's Theorem, Chinese Remainder Theorem, Cipher Properties, Substitution Ciphers – Mono-alphabetic Ciphers, Polyalphabetic Ciphers, Transposition Ciphers.	**10**
2	**Symmetric Cryptography:** Block Cipher, Fiestel Structure, Block Cipher Modes of Operation, S-DES, Double DES, Triple DES, AES Algorithm. **Asymmetric Cryptography:** Private Key and Public Key Cryptography, The RSA algorithm, Key Management, Diffie-Hellman Key Exchange, Key Exchange Algorithm.	**10**
3	**Integrity and Authentication:** **Cryptographic Hash Functions:** Message Authentication, Secure Hash Algorithm (SHA-512), **Message authentication codes:** Authentication requirements, HMAC, CMAC, Digital signatures, Elgamal Digital Signature Scheme. **Key Management and Distribution:** Symmetric Key Distribution Using Symmetric & Asymmetric Encryption, Distribution of Public Keys, Kerberos, X.509 Authentication Service, Public – Key Infrastructure	**10**

4	**Network Security:** Network attacks, DoS and DDoS attack, Sniffing, Session hijacking, Spoofing, Phishing, Cross-site Scripting (XSS), IPSec Protocol, SSL Handshake Protocol, Firewalls, IDS Prevention and Detection.	**12**
5	**Basics of Machine Learning in Cyber Security:** Cyber Threat Landscape, The Cyber Attackers Economy, why use machine learning in cybersecurity? Real-World Uses of Machine Learning in Security, Spam Fighting: An Iterative Approach, Limitations of Machine Learning in Security.	**05**
6	**Machine Learning in Cybersecurity:** Machine Learning: Problems and Approaches, Classification and Clustering, an ML approach for Security, Time Series Analysis and Ensemble Modeling: Analysis of time series in cyber security, Prediction of DDoS attack, Ensemble learning methods and voting ensemble methods to detect cyber attacks	**09**
	Total	56

Information Security Laboratory (DJS23DLPE619)	
Exp.	**Suggested experiments**
1	Implement Playfair Cipher with key entered by user.
2	Implement polyalphabetic Cipher
3	Implement Simple and Advanced Columner Transposition technique
4	Implement Simplified DES
5	Implement Simple RSA Algorithm with small numbers.
6	Implement Diffie-Hellman Key Exchange
7	Implement DoS and DDoS attack using Hping.
8	Implement phishing attack using HTTrack Website Cloning.
9	Implement static code analysis using Flawfinder Python Distribution.
10	Implement packet sniffing using Wireshark and TCP Dump.
11	Implement cross site request forgery in a controlled virtual environment using DVWA Web Server.
12	Implement firewalls using IP tables.
13	Implement Network Intrusion Detection System (NIDS).
14	Implement Host based Intrusion Detection System (HIDS).
15	Implementing and Evaluating an Email Spam Classifier Using Naive Bayes or Logistic Regression
16	Detection of DDoS Attacks in Network Traffic using Random Forest and Ensemble Methods

Minimum eight experiments from the above suggested list or any other experiment or mini project based on syllabus will be included, which would help the learner to apply the concept learnt.

*****The Term Work will be calculated based on Laboratory Performance (15m) and Assignments (10m).

Books Recommended:

Text Books:

1. William Stallings, "Cryptography and Network Security Principles and Practices", Pearson/PHI, 8th Edition, 2023.
2. Behrouz A. Forouzan, Debdeep Mukhopadhyay, "Cryptography and Network Security", McGraw Hill, 3rd edition 2017
3. Clarence Chio and David Freeman (2018). "Machine Learning and Security: Protecting Systems with Data and Algorithms", O'REILLY Publications.
4. Soma Halder, Sinan Ozdemir"Hands-On Machine Learning for Cybersecurity", 2018 Packt Publishing

Reference Books

1. Atul Kahate, "Cryptography and Network Security", McGraw Hill, 3rd Edition, 2013.
2. Bernard Menezes, Network Security and Cryptography, Cengage Learning: 2nd Edition, 2011.
3. Wade Trappe, Lawrence C Washington, "Introduction to Cryptography with coding theory", Pearson: 2nd Edition, 2006.
4. W. Mao, "Modern Cryptography – Theory and Practice", Pearson Education: 1st Edition, 2003.
5. Charles P. Pfleeger, Shari Lawrence Pfleeger – Security in computing – Prentice Hall of India, 2015.

Web Links:

1. Damn Vulnerable Web Application (DVWA): https://dvwa.co.uk/
2. Open Web Application Security Project: https://owasp.org
3. Web penetration testing: https://pentesterlab.com
4. Penetration Testing: https://kali.org

Prepared by Checked by Head of the Department Principal

Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Computer Science and Engineering (Data Science)

(Semester VI)

Third Year B.Tech. (DJS23)

Multidisciplinary Minor

Program: B.Tech in Computer Science and Engineering (Data Science) **Semester: VI**

Course: Applied Data Science Engineering Laboratory (DJS23DLMD601)

Prerequisite: Foundations of Data Analysis, Database Systems, Python Laboratory.

Objectives: To introduce students to the fundamentals of big data processing, analytics, and performance monitoring tools.

Outcomes: On completion of the course, learner will be able to:

1. Analyze large-scale datasets on real-time messaging workflows using Hadoop, HDFS, MapReduce, and AMPS.
2. Develop data processing pipelines using ML models, and evaluate model drift using Apache Spark.
3. Evaluate system metrics, KPIs, and workflows using Prometheus, Grafana, and Apache Airflow.

Applied Data Science Engineering Laboratory (DJS23DLMD601)		
Sr. No	**Suggested List**	**Duration**
1	**Introduction to Big Data and Hadoop** • Big Data Concepts and Hadoop Architecture • Core Hadoop Components. • Hadoop Ecosystem • Data profiling and documentation (EDA)	**2**
2	**Hadoop Distributed File System** • Understand Hadoop Distributed File System (HDFS) architecture and commands. • Introduction to MapReduce • Write and run a basic MapReduce program (WordCount). • View MapReduce job output and logs.	**2**
3	**Messaging Services using AMPS** • Real-time messaging using AMPS. • Pub / Sub models • Using SOW Topics and Historical Replay in AMPS • Filtering, Message Queuing and Expiration/Acknowledgement in AMPS • Complete Real-Time Messaging Workflow with AMPS: From Publish to SOW	**4**
4	**Introduction to Apache Spark** • Apache Spark ecosystem • Setup and shell • Real-time and Batch processing of high volume of data. • Processing high volume records in-memory (SQL)	**4**

5	**Data analytics and Visualization** • Spark Dataframe • Build a classification pipeline using MLlib • Apply feature transformers/regression model • Data visualization using GraphX / Graph Frames.	**4**
6	**Model Drift** • Introduction to Model Drift • Types of Model Drift • Drift Detection Techniques: Statistical tests, PSI, and distribution analysis • Implementing Drift Detection using Apache Spark.	**2**
7	**Prometheus – Metric Instrumentation and Monitoring** • Learn to install and configure Prometheus • Instrument an application or system for metrics collection • Scrape metrics using Prometheus • Configure alerts using Alertmanager	**2**
8	**Grafana – Dashboard Visualization** • Connect Grafana to Prometheus data source • Create dashboards with time-series graphs, heatmaps, and charts • Configure alerts in Grafana dashboards • Visualize trends and analyze system/application performance	**2**
9	**Performance Indicator** • Gather structured data from Grafana dashboards • Identify all Key performance indicators (KPIs) • Build a KPI matrix linking technical, operational, and business KPIs • Present KPIs as insights to business stakeholders	**2**
10	**Apache Airflow** • Install and configure Apache Airflow • Create a Cron-based scheduled workflow (Airflow Cron Job) • Configure and schedule a recurring job using Cron expressions • Monitor DAG execution performance through Airflow's UI	**4**
	Total	**28**

*The term work will be calculated based on Laboratory performance (10marks) ,Computer Based Assessment (10marks) and Quiz(05marks)

Books Recommended:

Textbooks:

1. Big Data with Hadoop and Spark: Analyze Massive Datasets with Apache Hadoop, Spark, and NoSQL , Thompson Carter 2024.
2. Data Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control by Steven L. Brunton & J. Nathan Kutz. Cambridge University Press (2nd Edition 2022).

Reference Books:

1. Big Data Computing: Advances in Technologies, Methodologies, and Applications, Tanvir Habib Sardar & Bishwajeet Kumar Pandey, 2023
2. Mastering Prometheus – William Hegedus, 2024

3. Data Pipelines with Apache Airflow, Second Edition, Julian de Ruiter, Ismael Cabral, 2024

Web Links:

1. https://www.udemy.com/course/grafana-tutorial
2. https://www.pluralsight.com/courses/prometheus-grafana-building-dashboards-data
3. https://airflow.apache.org/

Program: B. Tech. in Computer Science and Engineering (Data Science) **Semester: VI**
Course: Environmental Studies (DJS23ICHSX10)

Pre-requisite: Interest in Environment and its impact on Human

Objectives: Familiarize students with environment related issues such as depleting resources, pollution, ecological problems and the renewable energy scenario; Give overview of Green Technology options.

Outcomes: On successful completion of this course, student should be able to:

1. Understand how human activities affect environment
2. Understand the various technology options that can make a difference

Environmental Studies (DJS23ICHSX10)		
Unit	**Description**	**Duration**
1	**Air Pollution** Sources of Air pollution. Definition of Air Quality Index and how it is measured.	**02**
2	**Water Pollution** Sources of water pollution. Ground water pollution and eutrophication.	**02**
3	**Noise Pollution** Noise pollution and sources. Decibel limits for hospital, library, silence zone.	**01**
4	**Biodiversity loss** Value of Biodiversity. Endangered species.	**01**
5	**Deforestation** Product and services provided by forests. Relationship between forests and climate change.	**02**
6	**Renewable Energy sources** Our energy needs and global energy crisis. Renewable energy sources.	**02**
7	**Climate change** Greenhouse gases and climate change.	**02**
8	**Green Technology** Data Center Energy Efficiency, Thin-Client and Energy Efficiency.	**02**
	Total	**14**

Environmental Studies (DJS23ICHSX10)	
Sr. No.	**Tutorial List**
1	Case study on Smog.
2	Qualitative and Quantitative methods for Air Pollution Monitoring
3	Presentation on Water Pollution (Industrial, Sewage) explaining any specific case.
4	General Techniques in Water Quality Monitoring
5	List effects of noise pollution on human health. Measure decibel level in college library, canteen, classroom
6	Case study on effect of Pollution on Biodiversity loss.
7	Radioactive and Hazardous Pollutants.
8	Debate for and against “To promote Economic growth, Deforestation is required.”
9	Presentation on different Renewable Energy Technologies.
10	Green Hydrogen: The Energy of Future

11	Report on major impact of Global warming on Environment giving real examples.
12	Report on advantages and examples of Green Building for Sustainable development,
13	Sustainable Software Design: Reducing energy/carbon footprint (e.g., green coding, optimized AI/ML, efficient cloud usage)
14	Reducing AI's carbon footprint (efficient models, serverless) and applying AI for environmental good**,** like optimizing energy grids, smart buildings, precision agriculture (water/fertilizer use), waste management, and climate modeling.

Tutorial: (Term work: 25 marks)

- Minimum of 10 Tutorials required.
- The distribution of marks for term work shall be as follows:
  i. Performance in Tutorial: 15 Marks
  ii. Write-up/ Report: 10 marks

The final certification and acceptance of term work will be subject to satisfactory term work performance and fulfilling minimum passing criteria in the term work.

Books Recommended:

Text books:

1. R. Rajagopalan, "Environmental Studies from Crisis to Cure", 2nd Edition, Oxford university press, 2011.
2. Erach Bharucha, "Textbook of Environmental Studies for Undergraduate Courses", University Grants Commission (UGC) & Bharati Vidyapeeth Institute of Environment Education and Research (BVIEER), 2005.
3. Narayanan, P., "Environmental Pollution: Principles, Analysis and Control", CBS Publishers & Distributors, 2009.
4. Mohammad Dastbaz, Colin Pattinson, Babak Akhgar, Morgan and Kaufman, "Green Information Technology: A Sustainable Approach", Elsevier, 2015.

Reference Books:

1. Paulina Golinska, Marek Fortsch, Jorge Marx-Gómez, "Information Technologies in Environmental Engineering: New Trends and Challenges", Springer, 2011

Web Links:

1. CITES: https://cites.org/eng
2. Convention on Biological Diversity: www.biodiv.org
3. Kalpvriksh: www.kalpvriksh.org
4. Water pollution: http://en.wikipedia.org/wiki/Water_polution
5. Ecosan: www.eco-solutions.org

Program: Computer Science and Engineering (Data Science) **Semester: VI**

Course: Innovative Product Development IV (DJS23IPSCX04)

Objectives:

1. To acquaint the students with the process of identifying the need (considering a societal requirement) and ensuring that a solution is found out to address the same by designing and developing an innovative product.
2. To familiarize the students with the process of designing and developing a product, while they work as part of a team.
3. To acquaint the students with the process of applying basic engineering fundamentals, so as to attempt at the design and development of a successful value-added product.
4. To inculcate the basic concepts of entrepreneurship and the process of self-learning and research required to conceptualize and create a successful product.

Outcomes: On completion of the course, student should be able to:

1. Identify the requirement for a product based on societal/research needs.
2. Apply knowledge and skills required to solve a societal need by conceptualising a product, especially while working in a team.
3. Use standard norms of engineering concepts/practices in the design and development of an innovative product.
4. Draw proper inferences through theoretical/ experimental/simulations and analyse the impact of the proposed method of design and development of the product.
5. Develop interpersonal skills, while working as a member of the team or as the leader.
6. Demonstrate capabilities of self-learning as part of the team, leading to life-long learning, which could eventually prepare themselves to be successful entrepreneurs.
7. Demonstrate product/project management principles during the design and development work and also excel in written (Technical paper preparation) as well as oral communication.

Guidelines for the proposed product design and development:

- Students shall convert the solution designed in semester 3 and 4 into a working model, using various components drawn from their domain as well as related interdisciplinary areas.
- The working model is to be validated with proper justification and the report is to be compiled in a standard format and submitted to the department. Efforts are to be made by the students to try and publish the extended technical paper, either in the institute journal, "Techno Focus: Journal for Budding Engineers" or at a suitable publication, approved by the department research committee/ Head of the department.
- Faculty supervisor may provide inputs to students during the entire span of the activity, spread over 2 semesters, wherein the main focus shall be on self-learning.
- A record in the form of an activity log-book is to be prepared by each team, wherein the team can record weekly progress of work. The guide/supervisor should verify the recorded notes/comments and approve the same on a weekly basis.
- The focus should be on self-learning, capability to design and innovate new products as well as on developing the ability to address societal problems. Advancement of entrepreneurial capabilities and quality development of the students through the year long course should ensure that the design and development of a product of appropriate level and quality is carried out, spread over two semesters, i.e. during the semesters V and VI.

Guidelines for Assessment of the work:

- The review/ progress monitoring committee shall be constituted by the Head of the Department. The progress of design and development of the product is to be evaluated on a continuous basis, holding a minimum of two reviews in each semester.
- In the continuous assessment, focus shall also be on each individual student's contribution to the team activity, their understanding and involvement as well as responses to the questions being raised at all points in time.
- Oral examination should be conducted by Internal and External examiners. Students have to give presentation and demonstration on their working model.
- The distribution of marks for term work shall be as follows:
  1. Marks awarded by the supervisor based on log-book: 10
  2. Marks awarded by review committee: 10
  3. Quality of the write-up : 05

  The overall work done by the team shall be assessed based on the following criteria:

1. Quality of survey/ need identification of the product.
2. Clarity of Problem definition (design and development) based on need.
3. Innovativeness in the proposed design.
4. Feasibility of the proposed design and selection of the best solution.
5. Cost effectiveness of the product.
6. Societal impact of the product.
7. Functioning of the working model as per stated requirements.
8. Effective use of standard engineering norms.
9. Contribution of each individual as a member or the team leader.
10. Clarity on the write-up and the technical paper prepared.
11. The semester reviews (V and VI) may be based on relevant points listed above, as applicable.

Guidelines for Assessment of Semester Reviews:

- The write-up should be prepared as per the guidelines given by the department.
- The design and the development of the product shall be assessed through a presentation and demonstration of the working model by the student team to a panel of Internal and External Examiners, preferably from industry or any research organisations having an experience of more than five years, approved by the Head of the Institution. The presence of the external examiner is desirable only for the 2nd presentation in semester VI. Students are compulsorily required to present the outline of the extended technical paper prepared by them during the final review in semester VI.

Prepared by Checked by Head of the Department Principal