GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Master of Engineering

Level: PG

Branch: Computer Engineering (Artificial Intelligence And Data Science)

Course / Subject Code : ME01095061

Subject Name: Data Visualization & Pattern Recognition

WEF Academic Year :	2024-25
Semester :	1st Semester
Category of the Course :	PEC

Prerequisite :	Foundational Knowledge in Programming
Rationale :	This subject provides a comprehensive foundation in visualization and pattern recognition, equipping students with essential skills for analyzing and interpreting data in a variety of contexts. It prepares them for both practical challenges and ongoing developments in the field.

Course Scheme :

Teaching Scheme			Total Credits	Assessment Pattern and Marks				Total Marks
				Theory		Practical		
L	T	PR	C	ESE (E)	PA(M)	ESE (V)	PA (I)	
3	0	2	4	70	30	30	20	150

Course Content :

Sr. No.	Course Content	No. of Hours	% of Weightage
1	**Introduction** How visualization affects data interpretation, Role of visualization in data science, Two flavors of data visualization: exploratory and communicative	3	8
2	**Visualization Design Principles** Data and task abstraction, Best practices for encoding, Marks and channels, Effectiveness and expressiveness, How to critique visualizations, Design problems and consequences, How not to cause misinterpretation	6	18
3	**Exploratory Visualization Techniques**	6	24

	Handling high-dimensional data, Comparison techniques, small multiples, handling uncertainty, Depicting time		
4	**Interactive Visualization** Why interactivity is needed, handling multiple views, Brushing and Linking, D3 instructions and assignments	3	8
5	**Pattern Recognition system** Applications, Machine Perception, classification of pattern recognition systems. Design of Pattern recognition system, Pattern recognition Life Cycle. Statistical Pattern Recognition: Review of probability theory, Gaussian distribution, Bayes decision theory and Classifiers, Optimal solutions for minimum error and minimum risk criteria, Normal density and discriminant functions, Decision surfaces, Parameter estimation methods.	6	18
6	**Maximum-Likelihood and Bayesian Parameter Estimation:** Maximum-Likelihood Estimation, Bayesian Estimation, Bayesian Parameter Estimation: General Case & General Theory, Sufficient Statistics, Problems of Dimensionality, Component Analysis and Discriminants, Expectation-Maximization (EM)	6	14
7	**Multilayer Neural Networks :** Feedforward Operation and Classification, Backpropagation Algorithm, Error Surfaces, Backpropagation as Feature Mapping, Backpropagation, Bayes Theory and Probability, Practical Techniques for Improving Backpropagation, Second-Order Methods, Additional Networks and Training Methods, Regularization, Complexity Adjustment and Pruning	6	10
	Total		**100**

Reference Book :

1. Visualization Analysis and Design by Tamara Munzner.
2. Interactive Data Visualization for the Web: An Introduction to Designing with D3 by Scott Murray
3. The Visual Display of Quantitative Information by Edward Tufte.
4. 2nd Edition. Richard O. Duda, Peter E. Hart, David G. Stork, Pattern Classification, 2nd Edition John Wiley & Sons, Latest Edition.
5. C. Bishop, Pattern Recognition and Machine Learning, Springer Latest Edition.
6. Trevor Hastie, Robert Tibshirani, Jerome H. Friedman, The Elements of Statistical

Course Outcome :

After Completion of the Course, Student will able to :

No	Course Outcomes	RBT Level*
01	Understand how visualization affects data interpretation and the role of visualization in data science, including the differences between exploratory and communicative visualization.	UN
02	Apply visualization design principles such as data and task abstraction, encoding best practices, and the use of marks and channels to create effective visualizations.	AP
03	Evaluate and critique visualizations for effectiveness and expressiveness, identifying design problems and ensuring accuracy to prevent misinterpretation.	AP
04	Implement exploratory visualization techniques to manage high-dimensional data, use comparison techniques, small multiples, and handle uncertainty and time depiction.	CR
05	Apply interactive visualization techniques using D3, including managing multiple views, brushing, and linking to create dynamic and interactive visualizations.	CR
O6	Apply pattern recognition techniques, including statistical methods and machine perception, to design and implement classification systems, and address challenges in pattern recognition life cycles.	CR

*RM: Remember, UN: Understand, AP: Apply, AN: Analyse, EL: Evaluate, CR: Create

Suggested Course Practical List :

1. Create basic visualizations using sample datasets to understand the impact of visualization on data interpretation.
2. Design visualizations for given datasets focusing on encoding best practices, marks, and channels. Critique each other's designs for effectiveness and expressiveness.
3. Use techniques such as dimensionality reduction (e.g., PCA) and create visualizations using scatter plots, heatmaps, and small multiples.
4. Implement brushing and linking features to explore and interact with datasets. Create multiple linked views that update dynamically.
5. Perform EDA on a complex dataset using comparison techniques, handling uncertainty, and depicting temporal data.

6. Create a pattern recognition system using statistical methods (e.g., Gaussian distribution, Bayes decision theory) and evaluate its performance on a dataset.
7. Perform parameter estimation using maximum-likelihood methods on a dataset. Compare results with Bayesian estimation techniques.
8. Implement Bayesian estimation for a dataset, addressing problems of dimensionality and using expectation-maximization (EM) algorithms.
9. Develop a neural network using backpropagation, apply regularization techniques, and assess the model's performance on classification tasks.
10. Experiment with hyperparameter tuning for a neural network, including adjusting learning rates, batch sizes, and the number of hidden layers. Assess how different settings impact the model's accuracy and training time.

* * * * * * *