INSTITUTE OF AERONAUTICAL ENGINEERING
(Autonomous)
Dundigal - 500 043, Hyderabad, Telangana

COURSE CONTENT

ENSEMBLE LEARNING								
VII Semester: CSE (AI&ML)								
Course Code	**Category**	**Hours / Week**			**Credits**	**Maximum Marks**		
		L	**T**	**P**	**C**	**CIA**	**SEE**	**Total**
ACAD23	**Elective**	3	0	0	3	40	60	100
Contact Classes: 48	**Tutorial Classes: Nil**	**Practical Classes: Nil**				**Total Classes: 48**		
Prerequisites: Python Programming / Machine Learning Algorithms / Deep Learning								

I. COURSE OVERVIEW:

Ensemble learning focuses on advanced machine leaning techniques, ensemble methods, boosting and bagging approaches. The course emphasizes combination methods, diversity and explores ensemble pruning and clustering ensembles. This course covers a solid grasp of ensemble learning principles, enabling students to enhance model accuracy and make informed decisions across domains in engineering applications. The applications include character recognition and automobile fuel efficiency.

II. COURSE OBJECTIVES:

The students will try to learn:

I. Foundational understanding of ensemble methods and machine learning concepts and algorithms for supervised learning tasks.

II. Different paradigms of boosting, beginning approaches, combination methods including combination by learning method, diversity measures and relevant methods.

III. Strategies of ensemble pruning, categorizations of ordering, clustering methods and diverse clustering, ensemble categories include graph and transformation-based methods.

III. COURSE OUTCOMES:

At the end of the course, student should be able to:

CO1 Outline popular supervised learning algorithms and neural networks to solve real world classification and regression problems.

CO2 Extend boosting and bagging ensemble methods, AdaBoost algorithms, multiclass extension, random tree ensemble for density estimation and anomaly detection.

CO3 Make use of combination methods averaging, voting, stacking, infinite ensemble methods to improve accuracy and generalization.

CO4 Interpret algebraic methods, behavior knowledge space methods, decision template methods, dynamic classifier selection, and mixture of experts to solve specific challenges in ensemble learning.

CO5 Compare ensemble diversity, pairwise diverse, non-pairwise diverse, visualization methods, strategically implement diversity generalization techniques to improve the robustness and effectiveness of ensemble methods.

CO6 Optimize ensemble models through strategic pruning and clustering technique, pruning methods, graph and transformation-based methods to enhance efficiency and performance.

IV. COURSE CONTENT:

MODULE – I: FOUNDATIONS OF MACHINE LEARNING (09)

Introduction: Basic concepts, popular learning algorithms: linear discriminant analysis, decision trees, neural networks, Navie bayes classifier, k-Nearest Neighbor, SVM and kernel methods, evaluation and comparison, ensemble methods, applications of ensemble methods.

MODULE – II: ENSEMBLE METHODS (09)

Boosting: boosting procedure, the AdaBoost algorithm, examples, theoretical issues, multiclass extension, and noise tolerance. Bagging: two ensemble paradigms, the bagging algorithm, examples, theoretical issues, Random Tree ensembles Random Forest, spectrum of randomizations, Random Tree ensembles for density estimation and anomaly detection.

MODULE – III: COMBINATION METHODS (10)

Combination Methods: Benefits of Combination, averaging simple and weighted. Voting: majority voting, plurality voting, weighted voting, soft voting, theoretical issues. Combining by learning: stacking, infinite ensemble. Functions of several variables: Partial differentiation, Jacobian, functional dependence, maxima and minima of functions of two variables and three variables, method of Lagrange multipliers.

Other combination methods: algebraic methods, behavior knowledge space methods, decision template method. Relevant methods: error-correcting output codes, dynamic classifier selection, mixture of experts.

MODULE – IV: DIVERSITY METRICS (10)

Diversity: Ensemble diversity, Error decomposition, Diversity measures: pairwise, non-pairwise, summary and visualization, limitation of diversity measures.

Information theoretic diversity: information theory and ensemble, interaction information diversity, multi-information diversity, estimation method. Diversity generation.

MODULE – V: CLUSTERING ENSEMBLE METHODS (10)

Ensemble Pruning: what is ensemble pruning, many could be better than all, categorization of pruning methods: ordering based, clustering based, optimization-based pruning methods. Clustering Ensembles: Clustering: methods, evaluation, need of clustering ensembles. Categorization of clustering ensemble methods, similarity based, graph-based, relabeling based, transformation-based methods.

V. TEXTBOOKS:

1. Zhou, Zhi-Hua. Ensemble Methods: Foundations and Algorithms. United Kingdom, Taylor & Francis, 2012.

VI. REFERENCE BOOKS:

1. Natarajan, Vijayalakshmi, and Sarkar, Dipayan. Ensemble Machine Learning Cookbook: Over 35 Practical Recipes to Explore Ensemble Machine Learning Techniques Using Python. India, Packt Publishing, 2019.
2. Rokach, Lior. Ensemble Learning: Pattern Classification Using Ensemble Methods (Second Edition). Singapore, World Scientific Publishing Company, 2019.

VII. WEB REFERENCES

1. https://www.analyticsvidhya.com/blog/2022/05/know-about-ensemble-methods-in-machine-learning/
2. https://machinelearningmastery.com/tour-of-ensemble-learning-algorithms/
3. https://towardsdatascience.com/

4. http://scikit-learn.org/stable/modules/ensemble.html

VIII. MATERIALS ONLINE

1. Course template
2. Tutorial question bank
3. Tech talk topics
4. Open-ended experiments
5. Definitions and terminology
6. Assignments
7. Model question paper – I
8. Model question paper – II
9. Lecture notes
10. PowerPoint presentation
11. E-Learning Readiness Videos (ELRV)