

Advanced Data Mining

Course Code	20CS4702A	Year	IV	Semester	I
Course Category	PEC	Branch	CSE	Course Type	Theory
Credits	3	L-T-P	3-0-0	Prerequisites	Machine Learning
Continuous Evaluation :	30	Semester End Evaluation:	70	Total Marks:	100

Course Outcomes

Upon successful completion of the course, the student will be able to

CO1	Understand the basic techniques utilized in data mining.	L2
CO2	Apply suitable technique to extract patterns/information from various application areas of mining.	L3
CO3	Develop the knowledge for application of data mining and social impacts of data mining.	L3
CO4	Analyze the given scenario and use the appropriate technique based on application areas for mining.	L4

Contribution of Course Outcomes towards achievement of Program Outcomes & Strength of correlations

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	√													
CO2	√					√	√							
CO3						√	√							√
CO4		√				√	√		√	√				

Syllabus		Mapped CO
Unit No.	Contents	
I	<p>Mining Frequent Patterns: Basic Concepts, Efficient and scalable frequent item set mining methods: Apriori algorithm, Generating Association Rules from Frequent Item sets, Improving the efficiency of Apriori , A pattern –Growth Approach for Mining Frequent Item sets.</p> <p>Advanced Pattern Mining: Mining Multilevel Associations, Mining Multidimensional Associations, Mining Quantitative Association Rules, and Applications of Pattern Mining.</p>	CO1,CO2,CO4
II	<p>Classification: Advanced Methods Bayesian Belief Networks, Classification Using Frequent Patterns, Other Classification methods: Genetic Algorithms, Rough set Approach, Fuzzy set Approaches, Additional topics Regarding Classification: Multiclass Classification. Semi supervised classification, Active Learning, Transfer Learning</p>	CO1,CO2,CO4
III	<p>Cluster Analysis: Additional Issues and Algorithms Prototype Clustering, Density Based Clustering, Graph-Based Clustering, Scalable Clustering Algorithm , Which Clustering Algorithm</p>	CO1,CO2,CO4
IV	<p>Social Network Analysis: Social Network Analysis - Centrality, Prestige. Co-Citation and Bibliographic Coupling - Co-Citation, Bibliographic Coupling. PageRank - PageRank Algorithm. HITS - HITS Algorithm, Finding Other Eigenvectors, Relationships with Co-Citation and Coupling, Strengths and Weaknesses of HITS.</p>	CO1,CO3,CO4
V	<p>Data Mining Trends and Research Frontiers Mining Complex Data Types: Sequence Mining, Spatial Mining, Web mining, Stream Mining Data Mining Applications, Data Mining and Society: Ubiquitous and Invisible Data Mining, Privacy , security and Social impacts of data Mining</p>	CO1,CO3,CO4

Learning Resources

Text Books

1. **Data Mining: Concepts and Techniques**, Jiawei Han and Micheline Kamber, Morgan Kaufmann Publishers, Third edition ,2013
2. Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, Second edition, Pearson, 2019.
3. Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, Bing Liu, Springer Publications.

References

1. Principles of Data Mining , Hand, D., Mannila, H. and Smyth, P., MIT Press: Massachusetts, Third edition, 2013, Pearson.
2. Data Warehousing, Data Mining & OLAP, Alex Berson and Stephen J. Smith, 2008, Tata McGraw Hill Edition.

e- Resources & other digital material

1. https://onlinecourses.nptel.ac.in/noc21_cs06