

# M.Phil COMPUTER SCIENCE

## Course Structure and Syllabus

(For the candidates admitted from the academic year 2020 - 2021 onwards)

# CHOICE BASED CREDIT SYSTEM (CBCS)



**THANTHAI HANS ROEVER COLLEGE (AUTONOMOUS)**



(Approved by NAAC, Affiliated to Bharathidasan University)

**ELAMBALUR, PERAMBALUR – 621 220**



**VISION:**

- To be a Centre of excellence in education and research in the frontier areas of Computer Science.

**MISSION:**

- To facilitate quality transformative education in Computer Science
- To promote quality research and innovation in technology for meeting global challenges
- To transform students to competent professionals to cater to the needs of the society.

**Programme Outcomes (POs):**

Upon completion of the programme, the scholar will be able to

- ✓ Adapt self-learning through reviews of previously acquired knowledge.
- ✓ Implement research by focusing on newer thrust areas of knowledge.
- ✓ Engage in quality and efficient designing, implementing and evaluating of the gathered information.
- ✓ Demonstrate technical and analytical competence with local and global perspective.
- ✓ Have professional integrity with knowledge of appropriate code of ethics and standards displaying social responsibilities.

**Program Specific Outcomes (PSOs):**

Students will be able to

- Have extensive knowledge in computer science concepts, ability to have an in-depth understanding of high level advanced techniques and extensive knowledge of the computer literature are applicable to the selected research areas.
- Enhance the social, cultural and ethical knowledge of the recent computer technologies and knowledge of research applied in real time problems solving used to develop the society in current and next level emerging technology world.
- Apply and analyze research-based computer knowledge and research methods to provide effective conclusions and decisions making in the critical environments.
- Apply the trending concepts of computer science in research work increase the moral, ethical and social values of the scholar.
- Encourage the theoretical knowledge of teaching learning skills inside the classroom, personal skill development and employment provisions in the relevant field of computer science.



**M.Phil. Computer Science (FT) Programme**

(For the candidates admitted from the academic year 2020-2021 onwards)



Semester	Course	Course Code	Title of the Course	Ins. Hours/ Weeks	Credit	Exam Hours	CIA (Max)	ESE (Max)	Total (Max)
I	Core Course – I	20MPCS1CC1	Research Methodology	3*	4	3	25	75	100
	Core Course – II	20MPCS1CC2	Advanced Concepts in Computer Science	3*	4	3	25	75	100
	Core Course – III	20MPCS1CC3	Teaching and Learning Skills (Common Paper)	3*	4	3	25	75	100
	Core Course – IV	20MPCS1CC4	Paper on Topic of Research(The syllabus will be prepared by the Guide and the examination will be conducted by the COE)	3*	4	3	25	75	100
	<b>* One Hour Library hour for each Course</b>								
			Total	<b>12</b>	<b>16</b>	<b>12</b>	<b>100</b>	<b>300</b>	<b>400</b>
II	Project Work	20MPCS2DW	Dissertation and Viva-Voce (Viva Voce 50 marks, Dissertation 150 marks)	-	08	-	-	-	200
			Total	-	<b>08</b>	-	-	-	<b>200</b>
			Grand Total	<b>12</b>	<b>24</b>	<b>12</b>	<b>100</b>	<b>300</b>	<b>600</b>

**SEMESTER – I**  
**CORE COURSE 1 – RESEARCH METHODOLOGY**

**Course Code: 20MPCS1CC1**

**Hours: 3**

**Credit: 4**

**Total Marks: 100**

**External Marks: 75**

**Internal Marks: 25**

**Course Outcomes (Cos)**

- ✓ To develop understanding of the basic framework of research process.
- ✓ To develop an understanding of various research designs and techniques.
- ✓ To identify various sources of information for literature review and data collection.
- ✓ To develop an understanding of the ethical dimensions of conducting applied research.
- ✓ Appreciate the components of scholarly writing and evaluate its quality.

**UNIT I: UNIT I:**

Overview of Research Methodology: Introduction– Meaning of Research – Objectives of Research – Motivation in Research –Types of Research – Research Approaches – Significance of Research –Research process- Data Collection and Data Preparation: Introduction - Collection of Primary Data - Collection of Secondary Data - Selection of Appropriate Method for Data Collection - Data Preparation Process.

**UNIT II:**

Test of Hypothesis: Introduction- What is Hypothesis? -Basic Concepts Concerning Testing of Hypotheses- Tests of Hypotheses - Hypothesis Testing for: Mean, Proportion, Variance - Chi-Square Test: Test of Independence of Attributes - Categorized Data- Test of Goodness of Fit-Nonparametric Tests: Introduction- One Sample Test - Two Sample Tests-K Samples Tests-Advanced Multivariate Analysis: Introduction– Factor Analysis-Discriminant Analysis-Cluster Analysis.

**UNIT III:**

Algorithmic Research: Introduction-Algorithmic Research Problems- Types of Algorithmic Research Problems-Types of Solution Procedure/Algorithm-Steps of Development of Algorithm-Steps of Algorithmic Research-Design of Experiments and Comparison of Algorithms-Meta Heuristics for Combinatorial Problems.

**UNIT IV:**

Thesis Writing: Literature Survey – Writing Reviews and Journal Articles – Publication of Papers –Planning a Thesis – General Format – Page and Chapter Format – Footnotes – Tables and Figures –References and Appendices-Report Writing and Presentation: Introduction - Significance of Report Writing - Different Steps in Writing Report - Layout

of the Research Report - Types of Reports-Guidelines for Reviewing Draft-Oral Presentation.

**UNIT V:**

Research Tools: Introduction – SPSS – MATLAB – LaTeX – NS/2 –Rational Suite - Eclipse IDE.

**Reference Books:**

1. C. R. Kothari, Gaurav Garg, Research Methodology Methods and Techniques, NewAge International publishers, Third Edition, 2014.
2. Research Methodology – R.Panneerselvam, Prentice-Hall of India Private Limited, New Delhi-110 001, 2005.
3. Richard Johnson, Dean Wichern, Applied Multivariate Statistical Analysis. Prentice Hall (ISBN 0-1312-1973-1).
4. C Chatfield and A J Collins, Introduction to Multivariate Analysis. Chapman and Hall (SF 2 CHA).

**Relationship Matrix for Course Outcomes, Programme Outcomes and Programme Specific Outcomes:**

Semester	Code	Title of the Course					Hours	Credits			
I	20MPCS1CC1	RESEARCH METHODOLOGY					3	4			
Course Outcomes (Cos)	Programme Outcomes (POs)					Programme Specific Outcomes (PSOs)					
	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4	PSO5	
CO1	✓	✓	✓	✓			✓	✓			
CO2	✓	✓					✓			✓	
CO3	✓	✓	✓	✓		✓	✓	✓	✓	✓	
CO4	✓	✓	✓	✓		✓	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Number of matches (✓) = 38, Relationship: High											

<b>Mapping</b>	1-29%	30-59%	60-69%	70-89%	90-100%
<b>Matches</b>	1-14	15-29	30-34	35-44	45-50
<b>Relationship</b>	Very Poor	Poor	Moderate	High	Very High

**Prepared by:**

**Checked by:**

**HOD:**

**SEMESTER – I**  
**CORE COURSE I1 –ADVANCED CONCEPTS IN COMPUTER SCIENCE**

**Course Code: 20MPCS1CC2**

**Hours: 3**

**Credit: 4**

**Total Marks: 100**

**External Marks: 75**

**Internal Marks: 25**

**Course Outcomes(Cos)**

- ✓ Demonstrate a basic understanding of computer hardware and software.
- ✓ Demonstrate problem-solving skills.
- ✓ Apply logical skills to programming in a variety of languages and utilize web technologies.
- ✓ Present conclusions effectively, orally, and in writing.
- ✓ Demonstrate basic understanding of network principles.

**UNIT I:**

DATA MINING AND DATA ANALYTICS: Introduction - Data Mining & Data Preprocessing – Mining Frequent Patterns - Classification – Cluster Analysis- Data Mining Trends and Research Frontiers.Self-Learning Practice: WEKA / Rapid Miner / DB Miner.

**UNIT II:**

CLOUD COMPUTING: Introduction – Companies Involved in Cloud Computing – Why Cloud Computing – Benefits of Cloud Computing – Cloud Computing Drawbacks – Major Components – Cloud Computing Issues – Cloud Services- Virtualization.Self-Learning Practice -VMware/ Cloud Foundry/ Open Stack. Globus Toolkit / Eucalyptus / Open Nebula.

**UNIT III:**

PARALLEL PROCESSING: Introduction – Parallel processing Architecture -Program Issue-Data Dependency Analysis- Shared Memory management- Thread based implementation- Using Parallelism Effectively-Self Learning Practice: CEPBA

**UNIT IV:**

EVOLUTIONARY COMPUTING: Origin- Evolutionary Algorithm-Representation, Mutation& Recombination-Fitness, Selection & Population Manage- Popular Evolutionary Algorithm Variants-Parameter Control- Constraint Handling-Evolutionary Robotics Self Learning Practice: MATLAB.

**UNIT V:**

CRYPTOGRAPHY: Basic Principles- Historical Cryptosystems – Theoretical versus Practical Security – Symmetric Encryption – Public-Key Encryption – Digital Signature Schemes – Key Management – Public Key Management-Self Learning Practice: GnuPG / KF Sensor/ Snort/ Net Stumbler- For Term Paper Writing: \* Self Learning Practice – One Problem is given to the scholar they have to solve it in any one of the tool.

**Reference Books:**

1. Jiawei Han, Micheline Kamber, Jian Pei, "Data Mining: Concepts and Techniques", Third Edition, Elsevier, 2011
2. Ivanka Menken, "A Complete Guide to Cloud Computing", Art of Service, 2008
3. M.Sasikumar, Dinesh Shikhare and P.RaviPrakesh "Introduction To Parallel Processing", 2014.
4. A.E.Eibe and J.E.Smith, "Introduction to Evolutionary Computing" 2015.
5. Keith M. Martin, "Everyday Cryptography: Fundamental Principles and Applications", Oxford University Press, 2012.

**Relationship Matrix for Course Outcomes, Programme Outcomes and Programme Specific Outcomes:**

Semester	Code	Title of the Course					Hours	Credits				
I	20MPCS1CC2	ADVANCED CONCEPTS IN COMPUTER SCIENCE					3	4				
Course Outcomes (Cos)	Programme Outcomes (POs)					Programme Specific Outcomes (PSOs)						
	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4	PSO5		
CO1	✓		✓			✓				✓		
CO2	✓	✓	✓	✓		✓	✓		✓	✓		
CO3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
CO4	✓	✓	✓	✓	✓		✓	✓	✓	✓		
CO5	✓			✓	✓			✓	✓	✓		
Number of matches (✓) = 37, Relationship: High												

<b>Mapping</b>	1-29%	30-59%	60-69%	70-89%	90-100%
<b>Matches</b>	1-14	15-29	30-34	35-44	45-50
<b>Relationship</b>	Very Poor	Poor	Moderate	High	Very High

Prepared by:

Checked by:

HOD:

**SEMESTER – I**  
**CORE COURSE III – TEACHING AND LEARNING SKILLS**

**Course Code: 20MPCS1CC3**

**Hours: 3**

**Credit: 4**

**Total Marks: 100**

**External Marks: 75**

**Internal Marks: 25**

**Course Outcomes (Cos)**

- ✓ Identify and describe the political, religious, economic, and social uses of art in Italy during the Renaissance
- ✓ Identify a range of works of art and artists
- ✓ Analyze the role of art and of the artist in Italy at this time and analyze the art of the period according to objective methods
- ✓ Link different materials and types of art to the attitudes and values of the period
- ✓ Evaluate and defend their response to a range of art historical issues

**Unit I :** Computer Application Skills: Information and Communication Technology (ICT): Definition, Meaning, Features, Trends – Integration of ICT in teaching and learning – ICT applications: Using word processors, Spread sheets, Power point slides in the classroom – ICT for Research: On-line journals, e-books, Courseware, Tutorials, Technical reports, Theses and Dissertations - ICT for Professional Development : Concept of professional development; institutional efforts for competency building; individual learning for professional development using professional networks, OERs, technology for action research, etc.

**Unit II :** Communications Skills Communication: Definitions – Elements of Communication: Sender, Message, Channel, Receiver, Feedback and Noise – Types of Communication: Spoken and Written; Non-verbal communication – Intrapersonal, interpersonal, Group and Mass communication – Barriers to communication: Mechanical, Physical, Linguistic & Cultural – Skills of communication: Listening, Speaking, Reading and Writing – Methods of developing fluency in oral and written communication – Style, Diction and Vocabulary – Classroom communication and dynamics.

**Unit III :** Pedagogy Instructional Technology: Definition, Objectives and Types – Difference between Teaching and Instruction – Lecture Technique: Steps, Planning of a Lecture, Delivery of a Lecture – Narration in tune with the nature of different disciplines – Lecture with power point presentation - Versatility of Lecture technique – Demonstration: Characteristics, Principles, planning Implementation and Evaluation – Teaching-learning Techniques: Team Teaching, Group discussion, Seminar, Workshop, Symposium and Panel Discussion.

**Unit IV :** E- Learning, Technology Integration and Academic Resources in India Concept and types of e-learning (synchronous and asynchronous instructional delivery and means),



m-learning (mobile apps); blended learning; flipped learning; E-learning tools (like LMS; software's for word processing, making presentations, online editing, etc.); subject specific tools for e-learning; awareness of e-learning standards- Concept of technology integration in teaching- learning processes; frameworks guiding technology integration (like TPACK; SAMR); Technology Integration Matrix- Academic Resources in India: MOOC, NMEICT; NPTEL; e-pathshala; SWAYAM, SWAYAM Prabha, National academic depository, National Digital Library; e-Sodh Sindhu; virtual labs; eYantra, Talk to a teacher, MOODLE, mobile apps, etc.

**Unit V :** Skills of Teaching and Technology based assessment Teaching skills: Definition, Meaning and Nature- Types of Teaching Skills: Skill of Set Induction, Skill of Stimulus Variation, Skill of Explaining, Skill of Probing Questions, Skill of Black Board Writing and Skill of Closure – Integration of Teaching Skills – Evaluation of Teaching Skills- Technology for Assessment: Concept of assessment and paradigm shift in assessment; role of technology in assessment 'for' learning; tools for self & peer assessment (recording devices; e-rubrics, etc.); online assessment (open source software's; e-portfolio; quiz makers; e- rubrics; survey tools); technology for assessment of collaborative learning like blogs, discussion forums; learning analytics.

### **References :**

1. Bela Rani Sharma (2007), Curriculum Reforms and Teaching Methods, Sarup and sons, New Delhi
2. Brandon Hall , E-learning, A research note by Namahn, found in: [www.namahn.com/resources/ .../note-e-learning.pdf](http://www.namahn.com/resources/.../note-e-learning.pdf), Retrieved on 05/08/2011
3. Don Skinner (2005), Teacher Training, Edinburgh University Press Ltd., Edinburgh 4. Information and Communication Technology in Education: A Curriculum for schools and programmed of Teacher Development, Jonathan Anderson and Tom Van Weart, UNESCO, 2002.
5. Jereb, E., & Šmitek, B. (2006). Applying multimedia instruction in elearning. Innovations in Education & Teaching International, 43(1), 15-27.
6. Kumar, K.L. (2008) Educational Technology, New Age International Publishers, New Delhi.
7. Learning Management system : [https://en.wikipedia.org/wiki/Learning\\_management\\_system](https://en.wikipedia.org/wiki/Learning_management_system) , Retrieved on 05/01/2016
8. Mangal, S.K (2002) Essential of Teaching – Learning and Information Technology, Tandon Publications, Ludhiana.
9. Michael, D and William (2000), Integrating Technology into Teaching and Learning: Concepts and Applications, Prentice Hall, New York.
10. Pandey, S.K (2005) Teaching communication, Commonwealth Publishers, New Delhi.
11. Ram Babu, A abd Dandapani, S (2006), Microteaching (Vol.1 & 2), Neelkamal Publications, Hyderabad.

12. Singh, V.K and Sudarshan K.N. (1996), Computer Education, Discovery Publishing Company, New York.

13. Sharma, R.A., (2006) Fundamentals of Educational Technology, Surya Publications, Meerut

14. Vanaja, M and Rajasekar, S (2006), Computer Education, Neelkamal Publications, Hyderabad.

**Relationship Matrix for Course Outcomes, Programme Outcomes and Programme Specific Outcomes:**

Semester	Code	Title of the Course					Hours	Credits				
I	20MPCS1CC3	TEACHING AND LEARNING SKILLS					3	4				
Course Outcomes (Cos)	Programme Outcomes (POs)					Programme Specific Outcomes (PSOs)						
	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4	PSO5		
CO1	✓	✓					✓	✓				
CO2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
CO3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
CO4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
CO5	✓	✓	✓				✓		✓	✓		
Number of matches (✓) = 40, Relationship: High												

<b>Mapping</b>	1-29%	30-59%	60-69%	70-89%	90-100%
<b>Matches</b>	1-14	15-29	30-34	35-44	45-50
<b>Relationship</b>	Very Poor	Poor	Moderate	High	Very High

**Prepared by:**

**Checked by:**

**HOD:**

**SEMESTER – I**  
**CORE COURSE IV – GUIDE PAPER**

**Course Code: 20MPCS1CC4**

**Hours: 3**

**Credit: 4**

**Total Marks : 100**

**External Marks: 75**

**Internal Marks: 25**

**Name of the Supervisor : Dr.S.Sivakumar**

**Title of the Course : Data Mining and Knowledge Discovery**

**Course Outcomes(Cos):**

- ✓ Understand what Is Data Mining, what kinds of data can be mined, what kinds of patterns can be mined, and what kinds of applications are targeted.
- ✓ Explain major Issues in data mining.
- ✓ Apply machine learning, pattern recognition, statistics, visualization, algorithm, database technology and high-performance computing in data mining applications.
- ✓ Identify what kinds of technologies are used for different application.
- ✓ Manipulate data preprocessing, data Warehouse and OLAP technology, data cube technology; mining frequent patterns and association, classification, clustering, and outlier detection.

**UNIT-1**

Data mining: Introduction: Motivation, Important, What kind of data?, functionalities, Patterns, classifications -Data warehouse and OLAP technology for data mining: Introduction-multidimensional data model, Architecture, implementation- Data preprocessing: Why?, Data cleaning, Integration & Transformation, Reduction.

**UNIT-2**

Primitives for data mining: what defines a data mining task?- Concepts description: Characterization and Comparison : what is concept description?, data generalization and summarization-based characterization, efficient implementation, mining class comparisons.

**UNIT-3**

Classification & Prediction:-Classification Algorithm: C4.5, Apriori, SVM, AdaBoost, kNN, Naive Bayes, CART - Cluster Analysis: what is CA?, Type of data in CA, Hierarchical methods, Density-based, Grid-based, Model based-Clustering Algorithm: K-Means, EM(expectation maximization), Pagerank.

**UNIT-4**

Mining complex types of data: Generalization & multidimensional analysis, Spatial Databases, Time-series & Temporal, Mining text databases, Multimedia databases, Mining in WWW.

## UNIT-5

Data Mining Applications and trends in data mining: data mining applications, other themes on data mining, social impacts of data mining, trends& Research issues in data mining-Data Mining for Business Applications: Domain Driven KDD Methodology.

### Reference Books:

1. Jiawei Han, Micheline Kamber "Data Mining: Concepts and Techniques"
2. "The Top Ten Algorithms in Data Mining" Edited by XindongWu, Vipin Kumar
3. LongbingCao, Philip.Yu,ChengqiZhang, Huaifeng Zhang "Data Mining for Business Applications"

### Relationship Matrix for Course Outcomes, Programme Outcomes and Programme Specific Outcomes:

Semester	Code	Title of the Course					Hours	Credits			
V	20MPCS1CC4	Data Mining and Knowledge Discovery					3	4			
Course Outcomes (Cos)	Programme Outcomes (POs)					Programme Specific Outcomes (PSOs)					
	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4	PSO5	
CO1	✓	✓			✓		✓	✓			
CO2	✓	✓	✓	✓	✓	✓	✓				
CO3	✓	✓	✓	✓	✓	✓	✓	✓	✓		
CO4	✓		✓	✓		✓	✓	✓	✓	✓	
CO5		✓		✓	✓	✓		✓	✓	✓	
Number of matches (✓) = 36, Relationship: High											

<b>Mapping</b>	1-29%	30-59%	60-69%	70-89%	90-100%
<b>Matches</b>	1-14	15-29	30-34	35-44	45-50
<b>Relationship</b>	Very Poor	Poor	Moderate	High	Very High

Prepared by:

Checked by:

HOD:

**SEMESTER – I**  
**CORE COURSE IV – GUIDE PAPER**

**Course Code: 20MPCS1CC4**

**Hours: 3**

**Credit: 4**

**Total Marks : 100**

**External Marks: 75**

**Internal Marks: 25**

**Name of the Supervisor : Dr.V.Poongodi**

**Title of the Course : Cloud Computing and Software Services**

**Course Outcomes(Cos):**

- ✓ Explain the core concepts of the cloud computing paradigm: how and why this paradigm shift came about, the characteristics, advantages and challenges brought about by the various models and services in cloud computing.
- ✓ Apply fundamental concepts in cloud infrastructures to understand the tradeoffs in power, efficiency and cost, and then study how to leverage and manage single and multiple datacenters to build and deploy cloud applications that are resilient, elastic and cost-efficient.
- ✓ Discuss system, network and storage virtualization and outline their role in enabling the cloud computing system model.
- ✓ Illustrate the fundamental concepts of cloud storage and demonstrate their use in storage systems such as Amazon S3 and HDFS.
- ✓ Analyze various cloud programming models and apply them to solve problems on the cloud.

**Unit – I: Cloud Computing Fundamentals**

Introduction- Layers of Cloud Computing- Types of Cloud Computing- Cloud Computing Versus Cloud Services- Enabling Technologies- Cloud Computing Features- Cloud Computing Platforms- Cloud Computing Challenges-Cloud Types –Cloud Services and Cloud Roles- Infrastructure as a Service- Platform as a Service- Software as a Service- The Amazon Family

**Unit – II: Science Gateways**

Science Gateways—Background and Motivation-Clouds and Software Services-Science Clouds, Public and Private –Eucalyptus - Open-Source IaaS - Engineering Challenge-Eucalyptus Architecture-Cloud Computing for Science - Nimbus Goals and Architecture - Science Clouds Applications - Nimbus Helps Meet STAR Production Demands - Building a Cloud Computing Ecosystem with CernVM - CloudBLAST: Creating a Distributed Cloud Platform -Architecture of an SaaS Science Gateway -Dynamic Provisioning of Large-Scale Scientific Datasets - Science Gateways for Data - Cloud Computing and Data

### **Unit-III: Enterprise KnowledgeClouds**

Introduction - Enterprise Knowledge Management: Architecture and Technologies - Enterprise Knowledge Cloud- Real Cases and Applications of Cloud Computing -Cloud Computing: IT as a Service-Cloud Computing Security-Cloud Computing Model Application Methodology-Cloud Computing in Development/Test-Cloud-Based High Performance Computing Clusters-Use Cases of Cloud Computing

### **Unit-IV:Large-Scale DataProcessing**

Introduction -MapReduce - Programming Model -Implementation Sketch- Failure Handling Optimizations - GridBatch - DFS Extension -GridBatch Operator- MapReduce Implementation on a Cloud OS -What Is a Cloud OS? - Advantages Offered by a Cloud OS- Challenges Posed by Cloud OS -Advantages of Cloud MapReduce - Cloud MapReduce Architecture and Implementation - Architecture

### **Unit-V: Clustering**

Introduction to Clustering Techniques-Points, Spaces, and Distances -Clustering Strategies-The Curse of Dimensionality- Hierarchical Clustering -Hierarchical Clustering in a Euclidean Space-Efficiency of Hierarchical Clustering-Alternative Rules for Controlling Hierarchical Clustering - K-means Algorithms- K-Means Basics - The Algorithm of Bradley, Fayyad, and Reina - The CURE Algorithm- Initialization in CURE - Completion of the CURE Algorithm - Clustering in Non-Euclidean Spaces- Clustering for Streams and Parallelism

### **References**

1. Borko Furht , Armando Escalante , "Hand Book of Cloud Computing " – Springer
2. Syed A. Ahson , Mohammad Ilyas , "Cloud Computing and Software Services"- CRC Press.
3. Anand Rajaraman and Jeffrey D. Ullman, "Mining of MassiveDatasets".

**Relationship Matrix for Course Outcomes, Programme Outcomes and Programme Specific Outcomes:**

Semester	Code	Title of the Course					Hours	Credits			
V	20MPCS1CC4	Cloud Computing and Software Services					3	4			
Course Outcomes (Cos)	Programme Outcomes (POs)					Programme Specific Outcomes (PSOs)					
	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4	PSO5	
CO1	✓	✓	✓	✓		✓		✓			
CO2	✓	✓				✓	✓	✓	✓		
CO3	✓	✓	✓	✓		✓	✓	✓	✓		
CO4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Number of matches (✓) = 40, Relationship: High											

<b>Mapping</b>	1-29%	30-59%	60-69%	70-89%	90-100%
<b>Matches</b>	1-14	15-29	30-34	35-44	45-50
<b>Relationship</b>	Very Poor	Poor	Moderate	High	Very High

**Prepared by:**

**Checked by:**

**HOD:**

**SEMESTER – I**  
**CORE COURSE IV – GUIDE PAPER**

**Course Code: 20MPCS1CC4**

**Hours: 3**

**Credit: 4**

**Total Marks : 100**

**External Marks: 75**

**Internal Marks: 25**

**Name of the Supervisor : Dr.S.Yuvarani**

**Title of the Course : Data Mining**

**Course Outcomes(Cos)**

- ✓ Analyze data mining problems and reason about the most appropriate methods.
- ✓ To apply to a given dataset and knowledge extraction need.
- ✓ Implement basic pre-processing, association mining, classification and clustering algorithms.
- ✓ Apply and reflect on advanced pre-processing, association mining, classification and clustering algorithms.
- ✓ Work efficiently in groups and evaluate the algorithms on real-world problems.

**Unit-I:**

Introduction: Basic Data Mining Tasks, Data Mining Issues, Data Mining Metrics, Data Mining from a Database Perspective.

**Unit-II:**

Data Mining Techniques: A Statistical Perspective on Data Mining, Similarity Measures, Decision Trees, Neural Networks, Genetic Algorithms.

**Unit-III:**

Classification: Statistical-Based Algorithms, Distance-Based Algorithms, Decision Tree-Based Algorithms, Neural Network-Based Algorithms, Rule-Based Algorithms, Combining Techniques.

**Unit-IV:**

Clustering: Similarity and Distance Measures, Hierarchical Algorithms, Partitional Algorithms, Clustering Large Databases, Clustering with Categorical Attributes.

**Unit-V:**

Association Rules: Basic Algorithms, Parallel and Distributed Algorithms, Incremental Rules, Advanced Association Rule Techniques, Measuring the Quality of Rules. Advanced Techniques : Web Mining, Spatial Mining, Temporal Mining.

**Text Books**

1. J. Han and M. Kamber. Data Mining: Concepts and Techniques, 2nd Ed. Morgan Kaufman. 2006.



## References

1. M. H. Dunham. Data Mining: Introductory and Advanced Topics. Pearson Education. 2001.
2. I. H. Witten and E. Frank. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann. 2000.
3. D. Hand, H. Mannila and P. Smyth. Principles of Data Mining. Prentice-Hall. 2001.

Semester	Code	Title of the Course					Hours	Credits			
V	20MPCS1CC4	Data Mining					3	4			
Course Outcomes (Cos)	Programme Outcomes (POs)					Programme Specific Outcomes (PSOs)					
	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4	PSO5	
CO1		✓			✓		✓		✓		
CO2	✓	✓	✓			✓	✓			✓	
CO3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
CO4	✓	✓	✓	✓	✓	✓		✓	✓	✓	
CO5	✓		✓	✓	✓	✓		✓	✓	✓	
Number of matches (✓) = 37, Relationship: High											

<b>Mapping</b>	1-29%	30-59%	60-69%	70-89%	90-100%
<b>Matches</b>	1-14	15-29	30-34	35-44	45-50
<b>Relationship</b>	Very Poor	Poor	Moderate	High	Very High

Prepared by:

Checked by:

HOD: