Indian Institute of Information Technology Bhagalpur

Computer Science and Engineering (CSE)

B.Tech. Course Curricula and Syllabus

Semester-VI

Curricula:

Course Code	Course name	L	T	P	C
CS304	Compiler Design	3	0	0	3
CS305	Computer Networks	3	0	0	3
CS306	Computer Graphics	3	0	2	4
CS307	**Machine Learning**	**3**	**0**	**0**	**3**
ME306	**Environmental Sciences & Green Technology**	**2**	**0**	**0**	**2**
	Elective-I	**3**	**1**	**0**	**4**
CS312	Compiler Design LAB	0	0	3	2
CS313	Computer Networks LAB	0	0	3	2
CS314	**Machine Learning LAB**	**0**	**0**	**3**	**2**

Syllabus:

Course Code	**Course Name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS304	Compiler Design	3	0	0	3	3rd	6th
Course Objective: This course allows one to understand the process involved in a compiler, create an overall view of various types of translators, linkers, loaders, and phases of a compiler. This course further ensures that students get to know about syntax analysis, various types of parsers especially the top down approach. Creating awareness among students about various types of bottom up parsers is also a part of the course. Apart from this, various other topics such as intermediate code generation, type checking, the role of symbol table and its organization, code generation, machine independent code optimization and instruction scheduling included in the course. It allows one to better understand language translation.							
Topic							Hour
Module I	Overview of Different Phases of a Compiler: Overview of The Translation Process of a Source Program, A Simple Compiler, Types of Compiler, Analysis of The Source Program, The Phases of a Compiler, Cousins of The Compiler, The Grouping of Phases, Front-End and Back-End of Compiler, Pass Structure, Compiler Construction Tools. A Simple One-Pass Compiler: Overview on Syntax definition, Syntax Directed Translation, Parsing, Symbol Tables.						6
Module II	Lexical Analysis: The Role of a Lexical Analyser, Input Buffering, Specification of Tokens, Recognition of Tokens, The Lexical-Analyser Generator E.G., Lex. Review On Finite Automata, Design of a Lexical						7

	Analyser Generator, And Optimization of DFA-Based Pattern Matches.	
Module III	Syntax Analysis and Syntax-Directed Translation: The Role of a Parser, Context Free Grammars, Top Down and Bottom Up Parsing Techniques, Construction of Efficient Parsers. Syntax-Directed Definitions, Construction of Syntax Trees, Bottom-Up Evaluation of S-Attributed Definitions, L-Attributed Definitions, And Translation Schemes.	8
Module IV	Semantic Analysis and Intermediate Code Generation: Declaration Processing, Type Checking, Symbol Tables, Error Recovery, Run-Time Environments, and Ad-Hoc and Systematic Methods. Variants of Syntax Trees, Different Intermediate Forms, Types and Declarations, Control Flow, Back-patching, Switch-Statements, and Intermediate Code for Procedures.	6
Module V	Code Optimization and Code Generation: Global Data Flow Analysis, A Few Selected Optimizations like Command Sub Expression Removal, Loop Invariant Code Motion, and Strength Reduction, Issues In The Design Of A Code Generator, The Target Machine, Run-Time Storage Management, Basic Blocks and Flow Graphs, Next-Use Information, A Simple Code Generator, Register Allocation and Assignment, The DAG Representation of Basic Blocks, Peephole Optimization, Generating Code from DAGs, Dynamic Programming Code-Generation Algorithm, Code-Generator Generators.	8
	Total	35
Text	1. Compilers: Principles, Techniques, and Tools ; A V Aho, L S Monica, R Sethi, J D Ullman ; 2nd Edition, Prentice Hall; 2014. 2. Engineering a Compiler ; K D Cooper, L Torczon; Morgan Kaufmann Publishers; 2004.	
Reference	1. Writing Compilers and Interpreters: A Software Engineering Approach ; Ronald Mak; 3rd Edition, Kindle Edition; 2. Compiler Design in C; Allen I Holub, ; , Pearson Education; 2015. 3. Principles of Compiler Design ; V Raghavan, ; Mc-Graw Hill; 2010	

Course Code	**Course Name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS305	Computer Networks	3	0	0	6	3rd	6th
Course Objective: The objective of this course is to get familiar with layered communication architectures (OSI and TCP/IP); To understand the concepts of data link, network, transportation and application layer protocols. Introduce the student with Socket interface; Network design and programming, which includes TCP/IP and many application layer protocols.							
Topic							Hour
Module I	Evolution of computer networks: Computer networks basics.						5
Module II	Data link layer: Framing, HDLC, PPP, sliding window protocols, medium access control, Token Ring, Wireless LAN; Virtual circuit switching: Frame relay, ATM; Network Layer.						7

Module III	Network Layer: Internet addressing, IP, ARP, ICMP, CIDR, routing algorithms (RIP, OSPF, BGP).	8
Module IV	Transport Layer: UDP, TCP, flow control, congestion control; Introduction to quality of service.	7
Module V	Application Layer: DNS, Web, email, authentication, encryption.	8
	Total	35
Text	1.L. L. Peterson and B. S. Davie, Computer Networks: A Systems Approach, 5th Ed., Elsevier India, 2011. 2.A. S. Tanenbaum, Computer Networks, 5th Ed., Pearson India, 2013.	
Reference	1. Kurose James F and Ross Keith W, Computer Networking: A Top-Down Approach, 6th Ed., Pearson, 2017. 2. D. E. Comer, Internetworking with TCP/IP Vol. 1, 6th Ed., Pearson, 2015. 3. S. Keshav, An Engineering Approach to Computer Networking, 1st Ed., Pearson India, 2002. 4. B. Forouzan, Data Communications and Networking, 5th Ed., Tata Mcgraw Hill, 2013.	

Course Code	**Course Name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS306	Computer Graphics	3	0	2	4	3rd	6th
Course Objective: The primary objective of this course is to provide students with the basic principles of 3-dimensional computer graphics such as transformational geometry, rendering of complex models, shading algorithms, Curves & surfaces etc. After successfully completing this course, students will demonstrate their ability to use modern 3D computer graphics techniques, models, and algorithms to solve graphics problems.							
Topic							Hour
Module I	Introduction to Graphics Systems: Visualization, GUI, Video Display Devices, Raster and Random Scan Systems, Vector Devices, Cathode Ray Tube Terminals, Input Devices, Display of Colors- Look Up Tables, Display of Gray Shades, Half Toning; Display and Drawing of Graphics Primitives: Point, Line, Polygon, Circle, Curves and Text. Coordinate Conventions: Scan Conversion- Line, Circle, and Ellipse. World Coordinates, Device Coordinates, Normalized Device Coordinates, View-Port and Window, Zooming and Panning by Changing Coordinate Reference Frames.						6
Module II	Computations and Filling on Polygons: Filling- Rectangle, Polygon, Ellipse, and Arc. Point Inclusion Problem, Polygon Filling, Polygon Intersection, Clipping, Polygonization of a Point Set, Convex Hull Computation, Triangulation of Polygons.						4
Module III	2D Geometric Transformations: Basic Transformations, 2D and 3D, Matrix Representations and Homogeneous Coordinates, Composite Transformations, Transformations between Coordinate Systems, Transformation Functions. 2D Viewing: Viewing Coordinate						8

	Reference Frame, Window-To-Viewport Transformations, Clipping Operations- Line, Circle, Ellipse, Polygon, and Exterior.	
Module IV	GUI and Interactive Methods: Input of Graphical Data, Input Functions, Interactive Picture-Construction Techniques, Virtual-Reality Environments. Structure and Hierarchical Modelling: Structure Concepts, Editing Structures, Basic Modelling Concepts, Hierarchical Modelling with Structures. 3D Viewing: 3D Object Representations, Curves and Surfaces- Curved Lines, Polygon Meshes, Parametric Cubic Curves and Bicubic Surfaces, Hermite, Bezier, and B-Splines Curves and Surfaces. Quadric Surfaces. Projections, Specification and Implementation of 3D View.	8
Module V	Solid Modelling, Hidden Line and Surface Removal and shading: Boolean Set Operations, Spatial Partitioning Methods. Z-Buffer, List-Priority, Scan Line Algorithms, Algorithms for Binary Space Partitioning Trees and Octrees, and Ray Tracing. Illumination Model, Polygon Shading (Interpolated, Goursud, and Phong), Texture Mapping, Shadow Determination (Scan Line and Z-Buffer Algorithms), Transparency, Global Illumination Model.	8
	Total	
Text	1. D. Hearn and M. P. Baker, "Computer Graphics with OpenGL", 4th Ed., Pearson Education, 2013. 2. E. Angel, "Interactive Computer Graphics: A Top-Down Approach Using OpenGL", 5th Ed., Pearson Education, 2009. 3. J. D. Foley, A. Van Dam, S. K. Feiner and J. F. Hughes, "Computer Graphics: Principles and Practice in C", 2nd Ed., Addison-Wesley, 1995.	
Reference	1. P. Shirley and S. Marschner, "Computer Graphics", India Edition, Cengage Learning, 2009. 2. F. S. Hill, Computer Graphics Using Opengl, 3rd Ed., Pearson Eduaction, 2009. 3. Zhigang Xiang and Roy Plastock, "Computer Graphics", Schaum's Outlines, TMH, 2006. 4. John A. Vince, "Mathematics for Computer Graphics", 2/e, Springer, 2005.	

Course Code	**Course Name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS307	Machine Learning	3	0	0	3	3rd	6th
Course Objective: Machine learning is the science of getting computers to act without being explicitly programmed. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. This course will help the students to learn the necessary details to create next generation applications.							
Topic							Hour
Module I	Introduction: History of machine learning, Basic concepts						3
Module II	Supervised learning: Supervised learning setup, LMS, Logistic regression, Perceptron, Exponential family, Generative learning						10

	algorithms, Gaussian discriminant analysis, Naive Bayes, Support vector machines, Model selection and feature selection, Ensemble methods: Bagging, boosting.	
Module III	Learning theory: Bias/variance trade-off, Union and Chernoff/Hoeffding bounds, VC dimension, Worst case (online) learning.	7
Module IV	Unsupervised learning: Clustering K-means, EM. Mixture of Gaussians, Factor analysis, PCA (Principal components analysis), ICA (Independent components analysis).	8
Module V	Reinforcement learning and control: MDPs. Bellman equations, Value iteration and policy iteration, Linear quadratic regulation (LQR), Q-learning. Value function approximation, Policy search.	7
	Total	35
Text	1. Ethem Alpaydin, Introduction to Machine Learning, Second Edition, PHI, 2010. 2. Marsland, Stephen. Machine learning: an algorithmic perspective. Chapman and Hall/CRC, 2011.	
Reference	1. Murphy, Kevin P. "Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series)." (2018), MIT Press. 2. Brownlee, Jason. Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models and Work Projects End-To-End. Jason Brownlee, 2016.	

Course Code	Course name	L	T	P	C	Year	Semester
ME306	Environmental Sciences & Green Technology	2	0	0	2	3rd	6th

Course objective: To bring in the importance and the underlying principles of green and sustainable technology.

Topic	Contents	No. of Lectures
Module-I	Introduction to Environmental Pollution: Environmental Awareness, Concept of an ecosystem, structure and function of an ecosystem, energy and nutrient flow, biogeochemical cycles, sources, pathways and fate of environmental pollutants.	**05**
Module-II	Air pollution- Introduction, Segments of environment, Layers of atmosphere and their significance; Mechanism, Causative factors, Consequences and Preventive measures – Ozone depletion, Greenhouse effect and Global warming; Earth's radiation budget, Classification of air pollutants, Indoor air pollution, Smog-photochemical and sulphurous, Acid rain, Air Quality Standards, Human health effects-Bhopal gas tragedy.	**05**
Module-III	Water Resource; Water Pollution : Definition, Classification , Sources of Contamination, Pollutants & their Detrimental Effects; Water Quality: Portability limit – WHO and PHED Specification; Water Quality Monitoring, Municipal Water Treatment: Slow and Rapid Sand Filter, Disinfection – Methods, Advantages & Disadvantages, Sterilization	**05**

Module-IV	Soil and Noise pollution: Lithosphere and Soil profile, Soil contamination, sources of soil contamination, Important environmental properties of soil contaminants, Ecological & Health effects, Exposure & Risk Assessment; Noise pollution: Brief introduction to noise pollution, source, measurement and prevention of noise pollution	**05**
Module-V	Radioactive Pollution & Solid Waste Management: Radioactive pollutant: units of radiation and instruments for their measurements, types of radioactive pollutants and risk factor associated with these radiations Radioactive waste and their disposal, accidental leakage of radiation from nuclear reactors (discuss Chernobyl and Fukushima) Solid waste management different types of solid waste, composting, biological methods of detoxification of hazardous waste Onsite handling and composting, integrated solid waste management,	**05**
	Total	**42**
Text	1. Miller, T. G. Jr., *Environmental Science*, Wadsworth Publishing House, USA. 2. Masters, G.M, *Introduction to Environmental Engineering*.	

Elective-I

List of choices

Course code	Course Name	Area of Specialization
CS351	Introduction to Cryptography	[Network & Security]
CS352	Data Compression and Protection	[Network & Security]
CS353	Advanced Computer Architecture	[Hardware & Systems]
CS354	Distributed Operating System	[Hardware & Systems]
CS355	Formal Methods & Verification	[Theory & Machine intelligence]
CS356	Introduction to Data Mining	[Theory & Machine intelligence]

Syllabus for Elective-I Courses

Course Code	Course name	L	T	P	C	Year	Semester
CS351	**Introduction to Cryptography**	3	1	0	4	3rd	6th

Course Objective: The goal of this course is to provide students with the core principles of modern cryptography, including the modern, computational approach to security that overcomes the limitations of perfect secrecy. The goal of this course to give an excellent introduction to the theoretical background of cryptography.

Topic	Contents	No. of Lectures
Module 1	Mathematical Background for Cryptography: Solving Modular Linear Equations, the Chinese Remainder Theorem, Modular Exponentiation, and Discrete Logarithm Problem GCD Computation: Euclid's Algorithm, Extended Euclid's	**7**

	Algorithm Key Exchange: Diffie Hellman, ElGamal, Massey-Omura, Computation of Generators of Primes Public Key	
Module 2	Cryptosystem: RSA, Different Attacks & Remedies Primality Testing: Pseudoprimality Testing, Quadratic Residues, Randomized Primality Test & Deterministic Polynomial Time Algorithm Factorization: Quadratic-Sieve Factoring Algorithm.	**10**
Module 3	Method Elliptic Curve Cryptosystem: Theory of Elliptic Curves, Elliptic Curve Encryption & Decryption Algorithms, Security of Elliptic Curves Cryptography, Elliptic Curve Factorization	**10**
Module 4	Cryptographic Hash Functions: MD5 Message Digest Algorithm, Secure Hash Algorithm (SHA-1), Security of Hash Functions & Birthday Attack Digital Signatures: Authentication Protocols, Digital Signature Standards (DSS).	**10**
Module 5	System Security, Firewalls and Intrusion Detection Systems, Side Channel Analysis of Cryptographic Implementations,	**5**
	Total	**42**
Text	1. Neal Koblitz, A Course in Number Theory and Cryptography, Springer-Verlag, New York, May 2001. 2. Johnson Jr., Peter D., Greg A. Harris, D.C. Hankerson, Introduction to Information Theory and Data Compression, Chapman and Hall/CRC, 2003. 3. Oded Goldrich, Foundations of Cryptography-Basics, vol-1, Cambridge Univ. Press, 2005. 4. W. Trappe, L. C. Washington, Introduction to Cryptography With Coding Theory, Pearson, 2007.	
Reference	1. Oded Goldrich, Foundations of Cryptography-Applications, vol-2, Cambridge Univ. Press, 2005. 2. Titu Andreescu, Dorin Andrica, Number Theory: Structures, Examples, and Problems, Birkhäuser, 2009. 3. Cryptography, Network Security and Cyber Laws – Bernard Menezes, Cengage Learning, 2010. 4. Behrouz A Forouzan, Debdeep Mukhopadhyay, Cryptography and Network Security- Mc-GrawHill, 2015. 5. William Stallings, Cryptography and Network Security, Pearson Education, 2013.	

Course Code	Course name	L	T	P	C	Year	Semester
CS352	**Data Compression and Protection**	3	1	0	4	3rd	6th

Course Objective: The goal of this course is to provide students with basic data compression techniques and technology that are ever-evolving with new applications in image, speech, text, audio and video. This course will also enhance the students' ability to understand the vital aspects of data protection and the full dimensions of data protection which leads to poor data protection management, costly resource allocation issues, and exposure to unnecessary risks.

Topic	Contents	No. of Lectures
Module 1	Lossless Compression, Huffman Coding, Arithmetic Coding, Dictionary Techniques	**6**
Module 2	Context Based Compression, Lossless Image Compression, Lossy Coding, Scalar Quantization, Vector Quantization	**8**
Module 3	Differential Encoding, Transforms, Subbands, and Wavelets, Transform Coding, Subband Coding, Wavelet-Based Compression	**8**
Module 4	Audio Coding, Analysis/Synthesis and Analysis by Synthesis Schemes, Video Compression	**8**
Module 5	Data Protection—Where the Problems Lie, Setting the Right Objectives. Information Lifecycle Management Changes the Data Protection Technology Mix. The Critical Role of Data Retention, Where Data Protection Technologies Fit, Special Requirements for Compliance, Governance, and Data Security, eDiscovery and the Electronic Discovery Reference Model.	**10**
	Total	**40**

Text	1. Khalid Sayood, Introduction to Data Compression, Morgan Kaufmann, 2018. 2. Mark Nelson, Jean-Loup Gailly, The Data Compression Book, Wiley, 1995. 3. Peter D. Johnson Jr., Greg A. Harris, D.C. Hankerson, Introduction to Information Theory and Data Compression, Chapman and Hall/CRC, 2003. 4. David G. Hill, Data Protection: Governance, Risk Management, and Compliance, CRC Press, 2009.
Reference	1. Roy Hoffman, Data Compression in Digital Systems, Springer, 1997. 2. David Salomon, Giovanni Motta, D. Bryant, Handbook of Data Compression, Springer, 2010. 3. Gilbert Held, Thomas R. Marshall, Data and Image Compression: Tools and Techniques, Wiley, 1996. 4. Salomon David, Data Compression: The Complete Reference, Springer, 2014. 5. Preston De Guise, Data Protection, Routledge Publisher, 2017.

Course Code	Course name	L	T	P	C	Year	Semester

CS353	Advanced Computer Architecture	3	1	0	4	3rd	6th

Course Objective:

1. An understanding of the fundamental computer architectural issues and the inherent limitations of the traditional approaches.
2. Familiarity with the principles and the terminologies involved in computer architecture, organization and design.
3. Introduction to methods of specification, description, measurement and evaluation of processors and systems.
4. An appreciation of the historical developments in computer architecture and an acquaintance with many of the current innovative designs, providing a basis for understanding the new computer architectures that are on the horizon.

Topic	Contents	No. of Lectures
Module 1	Review of Memory Hierarchy: Set-Associative Cache, Cache Performance, Six Basic Cache Optimizations, Cache Coherence and the MESI Protocol, Virtual Memory	**6**
Module 2	Overview of Pipelined Architecture: Basic Pipelining, Pipelined data path and Control, Performance evaluation of pipelined architecture. Limitations of scalar pipelines, Data and Control Hazards., Pipeline Exceptions and Control, Dynamic Pipelines.	**8**
Module 3	Multicore, Multiprocessors, and Scalars: Classification of Computer Architectures, shared memory Multiprocessors, SISD, SIMD, MIMD, SPMD, and Vector Schemes, Superscalar architecture, superscalar techniques, performance evaluation of superscalar architectures, VLIW architecture, Multiple-Issue Processors, Symmetric Shared-Memory Architectures, Performance of Symmetric Shared-Memory Multiprocessors, Distributed Shared Memory and Directory-Based Coherence.	**10**
Module 4	Instruction-Level Parallelism: Concepts and Challenges, Basic Compiler Techniques for Exposing ILP, Loop Unrolling, Reducing Branch Costs with Prediction, Dynamic Scheduling, Overcoming Data Hazards with Dynamic Scheduling, Scoreboard, Tomasulo, Hardware-Based Speculation, Exploiting ILP Using Multiple Issue and Static Scheduling, Limits on Instruction-Level Parallelism. Thread-Level Parallelism: Multithreading, simultaneous multi-threaded architectures, instruction fetch policies in multi-threaded architectures, Performance and Efficiency in Advanced	**10**
Module 5	Parallel Processing: Multiple Processor Organizations, Symmetric Multiprocessors, Multithreading and Chip Multiprocessors, Clusters, Non-uniform Memory Access,	**8**

	Vector Computation, Multicore Computers- Hardware Performance Issues, Software Performance Issues.	
	Total	**42**
Text	1. J. L. Hennessy and D. A. Patterson, Computer Architecture: A Quantitative Approach, 5th Ed., Morgan Kaufmann, 2011. 2. W. Stallings, Computer Organization and Architecture: Designing for Performance, 8th Ed., Pearson Education India. 2010.	
Reference	1. Culler, David E. Singh, Jaswinder Pal. Gupta, Anoop, Parallel Computer Architecture: A Hardware/Software Approach, 1st Edition, Morgan Kaufmann, 2003. 2. Kai Hwang and Naresh Jotwani, Advanced Computer Architecture, 3rd ed., Tata Mcgraw Hill, 2010.	

Course Code	Course name	L	T	P	C	Year	Semester
CS354	Distributed Operating System	3	1	0	4	3rd	6th

Course Objective:

1. To learn the architectural differences and issues related to Advanced Operating System.
2. To get a comprehensive knowledge of the distributed systems and Real time operating system.
3. To get a thorough knowledge of database operating systems and cloud operating System

Topic	Contents	No. of Lectures
Module 1	Process synchronization – Overview, Synchronization mechanisms, process deadlocks	**5**
Module 2	Distributed Operating Systems – Architectures of Distributed Systems, Distributed Mutual Exclusion, Distributed Deadlock Detection, Agreement Protocols	**10**
Module 3	Distributed Research Management – Distributed File System, Distributed Shared Memory, Distributed Scheduling	**10**
Module 4	Distributed Fault Handling - Failure recovery and Fault tolerance	**8**
Module 5	Multiprocessor and Database Operating System – Multiprocessor System Architectures and Operating Systems, Introduction to Database Operating Systems and handling concurrency control	**8**
	Total	**41**

Text Books	1. Mukesh Singhal, Niranjan G.Shivaratri, "Advancedconcepts in operating systems: Distributed,Database and multiprocessor operating systems" .MC Graw Hill education, 2017 2. Pradeep K.Sinha, "Distributed Operating System-Concepts and design",PHI, 1998
Reference Books	1. Andrew S.Tanenbaum, " Modern Operating Systems ". Pearson Education, 2016. 2. Andrew S.Tanenbaum, "Distributed Operating System", PearsonEducation, 2002. 3. Rajib Mall, "Real-Time Systems: Theory and Practice", Pearson education, 2006

Course Code	Course name	L	T	P	C	Year	Semester
CS355	Formal Methods and Verification	3	1	0	4	3rd	6th

Course Objective: The goal of this course is to provide students with an overall understanding of basic concepts in the Formal Methods and Verification. The course introduces students with the mathematical foundations of specification languages, theorem provers, and model checkers. It will also enhance the students' ability to understand a wonderful example of what a modern text on logic for computer science should be like

Topic	Contents	No. of Lectures
Module 1	Introduction to Propositional Logic, Predicate Logic.	**4**
Module 2	Introduction to Formal Methods, Formal Methods and Agent-Based Systems.	**6**
Module 3	Verification by Model Checking: Temporal Logic; Model Checking Systems, Tools, and Properties; Branching Time Logic; Model Checking Algorithms.	**12**
Module 4	Program Verification: Software Verification Framework, Proof Calculus for Partial Correctness and Total Correctness, Programming by Contract.	**8**
Module 5	Modal Logics and Agents: Models of Truth, Basic Modal Logic, Logic Engineering, Natural Deduction. Binary Decision Diagram: Representation of Boolean Functions, Algorithms for Reduced OBDDS, Symbolic Model Checking, A Relational mu-Calculus.	**10**
	Total	**40**

Text Books	1. Michael Huth and Mark Ryan, Logic in Computer Science: Modelling and Reasoning about Systems, Cambridge University Press, 2018. 2. Monin-Jean Francois, Understanding Formal Methods, Springer-Verlag, 2003.
Reference Books	1. Christopher A. Rouff, Michael Hinchey, James Rash, Walter Truszkowski, and Diana Gordon-Spears, Agent Technology from a Formal Perspective, Springer-Verlog, 2006. 2. Jean-Louis Boulanger, Formal Methods, Wiley, 2012. 3. Kenneth Rosen, Discrete Mathematics and Its Applications, Mc-Graw Hill, 2017.

	C Liu and D. Mohapatra, Elements of Discrete Mathematics: A Computer Oriented Approach, Mc-Graw Hill, 2017

Course Code	Course name	L	T	P	C	Year	Semester
CS356	Introduction to Data Mining	3	1	0	4	3rd	6th

Course Objective: The increasing volume of data in modern business and science calls for more complex and sophisticated tools. This course allows one to gain insight of the large data sets by discovering patterns through scientific procedures.

Topic	Contents	No. of Lectures
Module 1	Introduction to Data mining and knowledge discovery, evolution of decision support systems.	**5**
Module 2	Steps In Data Mining Process, architecture of a typical Data Mining systems, data pre-processing, data Integration, data transformation and data reduction.	**7**
Module 3	Basic concepts of pattern discovery, such as frequent pattern, closed pattern, max-pattern, and association rules; Identify efficient pattern mining methods, such as Apriori, FPgrowth etc.	**8**
Module 4	Basic concepts, methods, and applications of cluster analysis, including the concept of clustering, the requirements and challenges of cluster analysis, a multi-dimensional categorization of cluster analysis, and an overview of typical clustering methodologies.	**8**
Module 5	Applications of Data Mining, Social Impacts of Data Mining, Case Studies involving text mining, mining Spatial Databases and web mining.	**7**
	Total	**35**

Text Books	1. Jiawei Han, Micheline Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, 2011. 2. Dunham, Margaret H. *Data mining: Introductory and advanced topics*. Pearson Education India, 2006.
Reference Books	1. Pujari, Arun K. *Data mining techniques*. Universities press, 2001. 2. Squire, Megan. *Mastering Data Mining with Python–Find patterns hidden in your data*. Packt Publishing Ltd, 2016.