Aided PGP Credit Structure for 2023-24 (Across All courses) … COMPUTER SCIENCE (Under NEP 2020)

Level	Sem	Major		RM	OJT, FP	Cum Cr/Sem	Degree/Cum Cr
		Mandatory	Elective				
6 (2023-24)	Sem 1	14	4	4	0	22	
		Advanced Database System (4+2)	Cyber and Information Security(3+1) OR Image Processing(3+1)	Research Methodology			
		Programming Paradigms(4)					
		Analysis of Algorithm(3+1)					
	Sem 2	14	4	0	4	22	
		Business Intelligence(4+2)	Data visualization (4) OR Enterprise Application Integration(3+1)		Internship		
		Social Network Analysis(3+1)					
		Machine Learning- I (3+1)					
	Cum Cr	28	8	4	4	44	PG Diploma after TY

Course: CS401	Advanced Database Systems (Credits : 4 Lectures/Week: 4)	Lecture
Expected Course Outcomes **After successful completion of this course, students would be able to:** 1. Develop knowledge and understanding of the underlying principles of Distributed Database . 2. Management System there Architecture and design strategies in detail with XML. 3. Evaluate the proper functions of transaction management using ACID properties, Deadlock 4. management, database reliability and parallel query Evaluation . 5. Explain the concept of NoSQL Basics & Accessing Data from MongoDB. 6. Describe the Gaining Proficiency With NoSQL,MongoDB Query ,HBase,RDBMS and ACID.		
Unit I	**Enhanced Database Models** Object–Oriented Databases: Need of Object-oriented databases, Complex Data Types, Structured Types and Inheritance, ObjectIdentity and Reference, ODL and OQL, Implementing O-R Features, Persistent Programming Languages, Object-Oriented versus ObjectRelational, Example of Object oriented and object relational database implementation, comparison of RDBMS, OODBMS, ORDBMS XML Databases: Structured Semi structure and unstructured data, XML hierarchical tree data model, Documents DTD and XML schema, XML Documents & Database, XML query and transformation, Storage of XML data, Xpath. XQuery, Join and Nesting Queries, XML database applications. Spatial Databases: Types of spatial data, Geographical Information Systems (GIS), Conceptual Data Models for spatial databases, Logical data models for spatial databases: Raster and vector model. Physical data models for spatial databases: Clustering methods (space filling curves), Storage methods (R-tree). Query processing. Temporal Databases: Time ontology, structure, and granularity, Temporal data models, Temporal relational algebra.	15L
Unit II	**Cooperative Transaction Model** :Parallel and Distributed Databases: Architecture of parallel databases, Parallel query evaluation, Parallelizing individual operations, Sorting Joins Distributed Databases: Concepts, Data fragmentation, Replication and allocation techniques for distributed database design, Query processing, Concurrency control and recovery in distributed databases, Architecture and Design: Centralised versus non centralized Databases, Homogeneous and Heterogeneous DDBMS, Functions and Architecture, Distributed database design, query processing in DDBMS, Distributed concurrency management, deadlock management, Distributed Commit Protocols: 2 PC and 3 PC, Concepts of replication servers. Mobile Database: Overview, Features, Advantages and Disadvantages, Mobile databases in Android System	15L
Unit III	**Learning the NoSQL Basics** Introduction to NoSQL: Characteristics of NoSQL, NoSQL Storage types, Advantages and Drawbacks, NoSQL Products Interfacing and interacting with NoSQL: Storing Data In and Accessing Data from MongoDB, Redis, HBase and Apache Cassandra, Language Bindings for NoSQL Data Stores Understanding the storage architecture: Working with Column Oriented Databases, HBase Distributed	15L

	Storage Architecture, Document Store Internals, Understanding Key/Value Stores in Memcached and Redis, Eventually Consistent Non-relational Databases Performing CRUD operations: Creating Records, Accessing Data, Updating and Deleting Data	
Unit IV	**Gaining Proficiency With NoSQL** Querying NoSQL Stores: Similarities Between SQL and MongoDB Query Features, Accessing Data from Column-Oriented Databases Like HBase, Querying Redis Data Stores Indexing And Ordering Data Sets: Essential Concepts Behind a Database Index, Indexing and Ordering in MongoDB, ouchDB and Apache Cassandra Managing Transactions And Data Integrity: RDBMS and ACID, Distributed ACID Systems, Upholding CAP, Consistency Implementations Using NoSQL in The Cloud: Google App Engine Data Store, Amazon SimpleDB	15L

Text book:
- Distributed Database; Principles & Systems By Publications, Stefano Ceri and Giuseppo Pelagatti,, McGraw-Hill International Editions (1984)
- Database Management Systems, 3rd edition, Raghu Ramakrishnan and Johannes Gehrke, McGraw-Hill (2002).
- Fundamentals of Database Systems, 6thEdition, Elmasri and Navathe, Addison. Wesley (2003).
- Unifying temporal data models via a conceptual model, C.S. Jensen, M.D. Soo, and R.T. Snodgrass: Information Systems, vol. 19, no. 7, pp. 513-547, 1994.
- Spatial Databases: A Tour by Shashi Shekhar and Sanjay Chawla, Prentice Hall, 2003 (ISBN 013-017480-7)
- Principles of Multimedia Database Systems, Subramanian V. S. Elsevier Publishers, 2013.

References:
- Principles of Distributed Database Systems; 2nd Editied By M. Tamer Ozsu and Patrick Valduriez, Person Education Asia.
- Database System Concepts, 5th edition, Avi Silberschatz , Henry F. Korth , S. Sudarshan: McGraw-Hill (2010)
- Database Systems: Concepts, Design and Applications, 2nd edition, Shio Kumar Singh, Pearson Publishing, (2011).
- Multi-dimensional aggregation for temporal data. M. Böhlen, J. Gamper, and C.S. Jensen. In Proc. of EDBT-2006, pp. 257-275, (2006).
- Moving objects databases (chapter 1 and 2), R.H. Güting and M. Schneider: Morgan Kaufmann Publishers, Inc., (2005)
- Advanced Database Systems, (chapter 5, 6, and 7), Zaniolo et al.: Morgan Kaufmann Publishers, Inc., (1997).

Links:
1) http://aries.ektf.hu/~hz/pdf-tamop/pdf-xx/Radvanyi-hdbms-eng2.pdf
2) https://www.studyyaar.com/index.php/learning-program/7-advanced-database-management-system

Sr. No.	Practical Course on CS401(Credits : 2 Lectures/Week: 4)
	Note: All the Practical's should be implemented using NoSQL Link: https://www.oracle.com/database/technologies/nosql-databaseserver-downloads.html
1.	Create different types that include attributes and methods. Define tables for these types by adding a sufficient number of tuples. Demonstrate insert, update and delete operations on these tables. Execute queries on them.
2.	Create an XML database and demonstrate insert, update and delete operations on these tables. Issue queries on it.
3.	Demonstrate distributed databases environment by dividing given global conceCreate a table that stores spatial data and issue queries on it. ptual schema, into vertical and Horizontal fragments and place them on different nodes. Execute queries on these fragments.
4.	Create a table that stores spatial data and issues queries on it.
5.	Create a temporal database and issue queries on it.
6.	Demonstrate the Accessing and Storing and performing CRUD operations in 1. MongoDB 2. Redis
7.	Demonstrate the Accessing and Storing and performing CRUD operations in 1. HBase 2. Apache Cassandra
8.	Demonstrating MapReduce in MongoDB to count the number of female (F) and male (M) respondents in the database.
9.	Demonstrate the indexing and ordering operations in 1. MongoDB 2. CouchDB 3. Apache Cassandra
10.	Demonstrate the use of data management and operations using NoSQL in the Cloud.

CS402	**Programming Paradigms** **(Credits : 4 Lectures/Week: 4)**	Lecture
	Course Outcomes: After completion of the course, students will be able to 1)Describe characteristics and design principles of imperative programming language paradigms 2) Demonstrate different forms of declaration, typing, binding, visibility, scoping, and lifetime management for various programming language constructs 3)Describe characteristics and design principles of declarative programming language paradigms 4)Choose a language or paradigm suitable for solving a particular problem	
I	**Introduction to Programming Languages:** Role of programming languages, towards high level programming languages, problems of Scale, programming paradigms, Language implementation, Expression notations, Abstract syntax trees, Lexical syntax, Context free grammar, variants of grammars, The need of structured programming, syntax directed control flow, design considerations, handling special cases in loops, programming with invariants, control flow .	15L
II	**Data Representation and procedure activation** :The role of Types, Basic Types, assignment and local state, the environment model of evaluation, Arrays, Records, Union and variant records, sets, pointers, two string tables, types and error checking, modeling with mutable data. Introduction to procedures, Parameter passing methods, Scope rules for names, nested scope in source text, activation records, Lexical scope	15L
III	**Functional Programming Paradigm:** Elements of functional programming, a little language of expressions, types: values and operations, function declarations, approaches to expression evaluation, lexical scope, type checking. Exploring a list, function ML/Scheme/Lisp: Implicit types, Data types, Exception handling in ML/Scheme/Lisp. Functional programming features in Python. **Logic Programming** Computing with relations, Introduction to Prolog, Data Structures in Prolog, Programming Techniques, Control in Prolog, Cuts.	15L
IV	**Object Oriented Programming and Recent Advances in Programming**: Constructs of program structuring, information hiding, program design with modules, modules and defined types, Object oriented programming in Python. Recent developments in the world of programming. Overview of emerging programming languages – Elm, Rust, Kotlin, Go, Ruby, Scala, Swift etc.	15L

Text Books:
1. Sethi R., "Programming Languages concepts & constructs", 2nd Edition, Pearson Education, ISBN 81 - 7808 - 104 - 0
2. Harold Abelson, Gerald Jay Sussman, Julie Sussman, "Structure and Interpretation of Computer Programs", 2nd Edition, ISBN 0-07-000484-6 (McGraw-Hill hardcover)
Reference Books:
1. Roosta S., "Foundations of Programming Languages", Thomson Brookes/Cole, ISBN 981 -243-141-1
2. Sebesta R., "Concepts Of Programming Languages", Pearson Education, (10th Edition)(2014)
3. Allen Tucker, Robert Noonan, "Programming Languages: Principles and Paradigms", Tata McGraw Hill, (2nd edition),(2007)
4. Carlo Ghezzi, Mehdi Jazayeri, "Programming Language Concepts",3rd Ed, Wiley Publication ISBN : 978-81-265-1861-6.
Online/Web/OtherReferences:
1. https://nptel.ac.in/courses/106/102/106102067/

Course: CS403	**Analysis of Algorithms** (Credits : 3 Lectures/Week: 3)	
Expected Course Outcomes : **After successful completion of this course, students will be able to:** 1. Apply the algorithms and design techniques to solve problems. 2. Compare two or more algorithms in terms of time and space complexity on growth functions. 3. Evaluate Algorithms using Dynamic Programming approach, Greedy strategy and Minimum spanning trees (MST) method. 4. Apply the concept of Lower Bound on RSA public-key cryptosystem NP-Completeness and Approximation algorithms.		
Unit I:	**Design strategies** The Role of Algorithms in Computing: Algorithms as a technology. Getting Started: Insertion sort, Analyzing algorithms, Designing algorithms. Growth of Functions: Asymptotic notation, Standard notations and common functions. Divide-and-Conquer: The maximum-subarray problem, Strassen's algorithm for matrix multiplication, The substitution method for solving recurrences. Probabilistic Analysis and Randomized Algorithms: The hiring problem, Indicator random variables, Randomized algorithms.	
Unit II:	**Advanced Design and Analysis Techniques** Dynamic Programming: Rod cutting, Elements of dynamic programming, longest common subsequence. Greedy Algorithms: An activity-selection problem, Elements of the greedy strategy, Huffman codes.	

Unit III:	**Elementary Graph Algorithms:** Representations of graphs, Breadth-first search, Depth-first search. **Minimum Spanning Trees:** Growing a minimum spanning tree, Algorithms of Kruskal and Prim. Single-Source Shortest Paths: The Bellman-Ford algorithm, Single-source shortest paths in directed acyclic graphs, Dijkstra's algorithm.	15 L

Text book:
- Introduction to Algorithms, Third Edition, Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, PHI Learning Pvt. Ltd-New Delhi (2009).
- Researching Information Systems and Computing, Brinoy J Oates, Sage Publications India Pvt Ltd (2006).

References:
- Algorithms, Sanjoy Dasgupta , Christos H. Papadimitriou, Umesh Vazirani, McGraw-Hill Higher Education (2006)
- Grokking Algorithms: An illustrated guide for programmers and other curious people, MEAP, Aditya Bhargava, http://www.manning.com/bhargava
- Research Methodology, Methods and Techniques, Kothari, C.R.,1985, third edition, New Age International (2014) .
- Basic of Qualitative Research (3rd Edition), Juliet Corbin & Anselm Strauss:, Sage Publications (2008).

Links:
1) https://www.tutorialspoint.com/design_and_analysis_of_algorithms/analysis_of_algorithms.htm
2) https://www.geeksforgeeks.org/analysis-of-algorithms-set-1-asymptotic-analysis/
3) https://mcdtu.files.wordpress.com/2017/03/introduction-to-algorithms-3rd-edition-sep-2010.pdf

Sr. No.	List of Practical Experiments on CS403(Credits : 1 Lectures/Week: 2)
1.	Write a program to implement insertion sort and find the running time of the algorithm.
2.	Write a program to implement a merge sort algorithm. Compare the time and memory complexity.
3.	Given an array of numbers of length l. Write a program to generate a random permutation of the array using (i) permute-by-sorting() and(ii) permute-by-cyclic().
4.	Write a program to implement Longest Common Subsequence (LCS) algorithm
5.	Write a program to implement Huffman's code algorithm
6.	Write a program to implement Kruskal's algorithm.
7.	Write a program to implement Dijkstrass's algorithm

8.	Write a program to implement Euclid's algorithm to implement gcd of two non negative integers a and b. Extend the algorithm to find x and y such that gcd(a,b) = ax+by. Compare the running time and recursive calls made in each case.
9.	Write a program to verify (i) Euclid's theorem (ii) Fermat's theorem.
10.	Write a program to implement greedy set cover

Course: CS406	**Research Methodology** (Credits : 3 Lectures/Week: 3)	
Unit I	**Introduction:** Role of Business Research, Information Systems and Knowledge Management, Theory Building, Organization ethics and Issues	**12L**
Unit II	**Beginning Stages of Research Process:**Problem definition, Qualitative research tools, Secondary data research	**12 L**
Unit III	**Research Methods and Data Collection:** Survey research, communicating with respondents, Observation methods, Experimental research	**12 L**
Unit IV	**Measurement Concepts, Sampling and Field work:** Levels of Scale measurement, attitude measurement, questionnaire design, sampling designs and procedures, determination of sample size	**12 L**
Unit-V	**Data Analysis and Presentation:** Editing and Coding, Basic Data Analysis, Univariate Statistical Analysis and Bivariate Statistical analysis and differences between two variables. Multivariate Statistical Analysis.	**12 L**

Text book:
1. Business Research Methods, William G.Zikmund, B.J Babin, J.C. Carr,Atanu Adhikari, M.Griffin, 8th Edition. 2016.
2. Business Analytics, Albright Winston, 5th Edition,2015
3.Research Methods for Business Students Fifth Edition, Mark Saunders, 2011.
4.Multivariate Data Analysis, Hair, Pearson, 7th Edition, 2014.

References:
Links:
1)http://www.library.auckland.ac.nz/subject-guides/med/pdfs/Hindex%20and%20impact%20factors.pdf
2)www.openintro.org/stat/down/OpenIntroStatFirst.pdf

Sr No		Practical of CS406 (Credits : 1 Lectures/Week: 2)
1	A	Write a program for obtaining descriptive statistics of data.
	B	Import data from different data sources (from Excel,csv, mysql, sql server, oracle to R/Python/Excel)
2	A	Design a survey form for a given case study, collect the primary data and analyze it
	B	Perform suitable analysis of given secondary data.
3	A	Perform testing of hypothesis using one sample t-test.
	B	Perform testing of hypothesis using two sample t-test.
	C	Perform testing of hypothesis using paired t-test.
4	A	Perform testing of hypothesis using chi-squared goodness-of-fit test.
	B	Perform testing of hypothesis using chi-squared Test Of Independence
5		Perform testing of hypothesis using Z-test.
6	A	Perform testing of hypothesis using one-wayANOVA.
	B	Perform testing of hypothesis using two-wayANOVA.
	C	Perform testing of hypothesis using multivariateANOVA (MANOVA).
7	A	Perform the Random sampling for the given data and analyse it.
	B	Perform the Stratified sampling for the given data and analyse it.
8		Compute different types of correlation.
9	A	Perform linear regression for prediction.
	B	Perform polynomial regression for prediction.
10	A	Perform multiple linear regression.
	B	Perform Logistic regression.

Course: CS404	Cyber and Information Security (Network Security) (Credits : 2 Lectures/Week: 2)	Lecture

Expected Course Outcomes :
After successful completion of this course, students will be able to:
   1. Identify when attacks are happening inside networks
   2. Collect evidence of network intrusions
   3.   Test networks and systems for vulnerabilities
   4. Prepare for defend against network attacks		

Unit I	**Computer Security** Principles of Security, Different Attacks: malicious and non-malicious program, Types of Computer Criminals. Protected objects and methods of protection. Memory address protection: Fence, Relocation, Base/Bound Registers, Tagged Architecture, Segmentation, Paging, Directory, access control list. Database Security: Security requirements, Integrity, Confidentiality, Availability, Reliability of Database, Sensitive data, Multilevel database, Proposals for multilevel security.	**15l**
Unit II	**Network Security** Different types of network layer attacks. Firewall (ACL, Packet Filtering, DMZ, Alerts and Audit Trials) IDS,IPS and its types (Signature based, Anomaly based, Policy based, Honeypot based). Web Server Security: SSL/TLS Basic Protocol-computing the keys- client authentication-PKI as deployed by SSL Attacks fixed in v3-Exportability-Encoding-Secure Electronic Transaction (SET), Kerberos	**15l**
Unit III	**Mobile Security:** Mobile system architectures, Overview of mobile cellular systems, GSM and UMTS Security & Attacks, Vulnerabilities in Cellular Services, Cellular Jamming Attacks & Mitigation, Security in Cellular VoIP Services, Mobile application security. Securing Wireless Networks: Overview of Wireless Networks, Scanning and Enumerating 802.11 Networks, Attacking 802.11 Networks, Bluetooth Scanning and Reconnaissance, Bluetooth Eavesdropping, Attacking & Exploiting Bluetooth, Zigbee Security & Attacks	**15l**

Text book:
  ● Security in Computing 4th edition, Charles P. Pfleeger, Charles P. Pfleeger, Shari Lawrence Pfleeger, Prentice Hall; 4th edition (2006)
  ● Mobile and Wireless Security and Privacy, Kia Makki, Peter Reiher, Springer, (2007).		

- Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance (Theory and practice), Tim Mather, Subra Kumaraswamy, Shahed Latif., O'Reilly Media; 1 edition (2009).

References:
- Cloud Security: A Comprehensive Guide to Secure Cloud Computing, Ronald L. Krutz, Russell Dean Vines, Wiley (2010)
- Network Security, Charlie Kaufman, Radia Perlam, Mike Speciner, Prentice Hall, 2nd Edition (2002)
- Cryptography and Network Security 3rd edition, Atul Kahate, Tata McGraw Hill Education Private Limited (2013)
- Network Security, Charlie Kaufman, Radia Perlam, Mike Speciner, Prentice Hall, 2nd Edition (2002)
- Cryptography and Network Security: Principles and practice 6th edition, William Stallings, Pearson Education (2013).

Links:
1) https://elearning.dei.unipd.it/pluginfile.php/17665/mod_resource/content/1/slides.pdf
2) http://www.vssut.ac.in/lecture_notes/lecture1428550736.pdf

Sr. No.	Practical of CS405(Credits : 2 Lectures/Week: 4)
1.	Write a program to store username and password in an encrypted form in a database to implement integrity lock.
2.	Write SQL query to retrieve sensitive information from less sensitive queries
3.	Write SQL query to create a view to implement the concept of views and commutative filter in distributed databases.
4.	Write a program to implement SSL.
5.	Write a program to send an encrypted email.
6.	Write a program to digitally sign MIME to create an 'opaque' signature.
7.	Write a program to generate DSA SSH key.
8.	Write a program to implement multilevel security.
9.	Write a program to Demonstrates how to encrypt and decrypt the content of an XML node using 128-bit CBC AES encryption

Course: CS405	Image Processing and Pattern Recognition (Credits : 3 Lectures/Week: 3)	Lecture
Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Define the fundamental concepts of a digital image processing system. 2. Analyze images in the frequency domain using various transforms. 3. Evaluate the techniques for image enhancement and image restoration. 4. Develop image patterns using statistical patterns		
Unit I	**Introduction to digital image processing:** Digital image representation, Digital image processing: Problems and applications, Elements of visual perception, Sampling and quantization, relationships between pixels. **Two-dimensional System:**Fourier transform and Fast Fourier Transform, Other image transforms and their properties: Cosine transform, Sine transform, Hadamard transform, Haar transform.	**15 L**
Unit II	**Image Restoration:** Image Restoration-Constrained and unconstrained restoration Wiener filter , motion blur remover, geometric and radiometric correction Image data compression-Huffman and other codes transform compression, predictive compression two tone Image compression, block coding, run length coding, and contour codin+g.	**15 L**
Unit III	**Segmentation Techniques:** Segmentation Techniques-thresh holding approaches, region growing, relaxation, line and edge detection approaches, edge linking, supervised and unsupervised classification techniques, remotely sensed image analysis and applications, Shape Analysis – Gestalt principles, shape number, moment Fourier and other shape descriptors, Skeleton detection, Hough trans-form, topological and texture analysis, shape matching.	**15L**
Unit IV	**Pattern Recognition:** Basics of pattern recognition, Design principles of pattern recognition system, Learning and adaptation, Pattern recognition approaches, Mathematical foundations – Linear algebra, Probability Theory, Expectation, mean and covariance, Normal distribution, multivariate normal densities, Chi squared test. **Statistical Pattern Recognition:** Bayesian Decision Theory, Classifiers- k-NN classifier, Normal density and discriminant functions, Parameter estimation methods: Maximum-Likelihood estimation, Bayesian Parameter estimation, Dimension reduction methods – Principal Component Analysis (PCA), Fisher Linear discriminant analysis, Expectation-maximization (EM), Hidden Markov Models (HMM),Gaussian mixture models.	**15L**

Text book:
- R. C. Gonzalez and P. Wintz, "Digital Image Processing", Second Edition, Addison-Wesley Publishing, 1987.
- K. Castlemann. "Digital Image Processing", Prentice Hall of India Ltd., 1996.
- A. K. Jain, "Fundamentals of Digital Image Processing", Prentice Hall of India Pvt. Ltd., 1995.
- Digital Image Processing – Ganzalez and Wood, Addison Wesley, 1993.
- Pattern Classification – R.O. Duda, P.E. Hart and D.G. Stork, Second Edition John Wiley, 2006

References:
- Digital Picture Processing – Rosenfeld and Kak, vol.I & vol.II, Academic,1982
- Computer Vision – Ballard and Brown, Prentice Hall, 1982
- An Introduction to Digital Image Processing – Wayne Niblack, Prentice Hall, 1986
- Pattern Recognition and Machine Learning – C. M. Bishop, Springer, 2009.
- Pattern Recognition – S. Theodoridis and K. Koutroumbas, 4th Edition, Academic Press,2009

Sr. No.	Practical Course on CS405 (Credits : 1 Lectures/Week: 2)
1.	To study the Image Processing concept. (image acquisition, image storage, image processing and display)
2.	To obtain histogram equalization image from original image histogram.
3.	To Implement smoothing or averaging filters in spatial domain.
4.	Write a program for opening and closing of the image.
5.	Write a program to fill the region of interest for the image.
6.	Write a program to perform an edge detection algorithm.
7.	Write a program of sharpening images using gradient masks.
8.	Write Program for morphological operation: erosion and dilation.
9.	Write a Program for DCT/IDCT to express the finite sequence.
10.	Implement the k-NN classifier for an unknown image and for a general K value.

Course: CS451	**Business Intelligence** (Credits :6 Lectures/Week:4)	**Lectures**
	Expected Learning Outcomes: After successful completion of this course, students will be able to: 1. Interpret different Data mining algorithms and Techniques. 2. Critically evaluate and apply Mining algorithms to real-world case studies. 3. Compare MapReduce and traditional Software paradigms. 4. Assess the similarity of documents 5. Apply different link analysis techniques for specific use cases.	
Unit I	**Introduction to Business Intelligence** Operational and Decision Support System, Data-Information-Knowledge, Decision making-Action cycle. Basic definitions- Business Intelligence; Data warehousing, Business Intelligence architecture, Use and benefits of Business Intelligence. **Knowledge Discovery in Databases**: KDD process model. **Data Pre-processing**: Cleaning: Missing Values; Noisy Values; Inconsistent values; redundant values. Outliers, Integration, transformation, reduction, transformation.	**15 L**
Unit II	**Statistical Modelling:** What is data Statistical Modelling?, Machine Learning, Computational Approaches to Modelling, Summarization, Feature Extraction Statistical Limits on Data Mining. **Clustering:** Introduction to Clustering Techniques, Hierarchical Clustering, K-means Algorithms, The CURE Algorithm, Clustering in Hierarchical Non-Euclidean Spaces. **Recommendation Systems:** A Model for Recommendation Systems, Content-Based Recommendations,Collaborative Filtering, Dimensionality Reduction.	**15 L**
Unit III	**Finding Similar Items:** Applications of homo Near-Neighbor Search, Shingling of Documents, Similarity-Preserving Summaries of Sets, **Locality-Sensitive Hashing** for Documents, Distance Measures, The Theory of Locality-Sensitive Functions, LSH Families for Other Distance Measures, Applications of Locality-Sensitive Hashing, Methods for High Degrees of Similarity. **Mining Data Streams:** The Stream Data Model Sampling Data in a Stream Filtering Streams Counting Distinct Elements in a Stream, Estimating Moments, Counting Ones in a Window, Decaying Windows	**15 L**
Unit IV	**Link Analysis:** PageRank, Efficient Computation of PageRank, Topic-Sensitive PageRank, Link Spam, Hubs and Authorities. **Frequent Itemset**: A-Priori Algorithm Handling Larger Datasets in Main Memory, Limited-Pass Algorithms, Counting Frequent Items in a Stream.	**15 L**

	MapReduce and the New Software Stack: Distributed File Systems, MapReduce, Algorithms Using MapReduce, Extensions to MapReduce, The Communication Cost Model, Complexity Theory for MapReduce.	

Text book:
Text book:
1) Business Intelligence (2nd Edition), Efraim Turban, Ramesh Sharda, Dursun Delen, David King, Pearson (2013)
2) Mining of Massive Datasets by Jure Leskovec Stanford Univ., Anand Rajaraman Milliway Labs, Jeffrey D. Ullman Stanford Univ.

Reference book:
- Business Intelligence RoadMap, Larissa T. Moss, Shaku Atr, Addison-Wesley
- Data mining concepts and techniques by Jiawei Han, Micheline Kamber, Jian Pei 3rd Edition.

Links:
1) http://infolab.stanford.edu/~ullman/mmds/book.pdf
2) http://freecomputerbooks.com/Mining-of-Massive-Datasets.html

Sr. No.	Practical of CS451
1.	Develop an application to perform clustering using various techniques.
2.	Develop a content-based recommendations system.
3.	Demonstrate an application of Near Neighbor search.
4.	Write a program for measuring similarity among documents and detecting passages which have been reused.
5.	Demonstrate an application of Locality sensitive hashing technique for large datasets.
6.	Write a program to explain links to establish higher-order relationships among entities in Link Analysis.
7.	Demonstrate page ranking with an appropriate application.
8.	Develop an application to implement the apriori algorithm
9.	Write a map-reduce program to count the number of occurrences of each alphabetic character in

		the given dataset. The count for each letter should be case-insensitive (i.e., include both upper-case and lower-case versions of the letter; Ignore non-alphabetic characters).
10.		Write a map-reduce program to count the number of occurrences of each word in the given dataset. (A word is defined as any string of alphabetic characters appearing between non-alphabetic characters like nature's is two words. The count should be case-insensitive. If a word occurs multiple times in a line, all should be counted)

Course: CS452	Social Network Analysis (Credits : 3 Lectures/Week: 3)	No. of Lectures

Expected Learning Outcomes:
After successful completion of this course, students will be able to:
   Represent social network relationships through graph theory
   Analyze the social network relationship and ties between different egos networks using a page ranking algorithm.
   Compare different Similarity and dissimilarity distance measuring approaches
   Analyze two mode Bi-partite data structures.		

Unit I	**Introduction to social network analysis (SNA)** Introduction to networks and relations- analyzing relationships to understand people and groups, binary and valued relationships, symmetric and asymmetric relationships, multimode relationships, Using graph theory for social networks analysis- adjacency matrices, edge-lists, adjacency lists, graph traversals and distances, depth-first traversal, breadth-first traversal paths and walks, Dijkstra's algorithm, graph distance and graph diameter, social networks vs. link analysis, ego-centric and socio-centric density.	**15 L**
Unit II	**Networks, Centrality and centralization in SNA** Understanding networks- density, reachability, connectivity, reciprocity, group-external and group-internal ties in networks, ego networks, extracting and visualizing ego networks, structural holes, Centrality- degree of centrality, closeness and betweenness centrality, local and global centrality, centralization and graph centers, notion of importance within network, Google pagerank algorithm, Analyzing network structure bottom-up approaches using cliques, N-cliques, N-clans, K-plexes, K-cores, F-groups and top-down approaches using components, blocks and cut-points, lambda sets and bridges, and factions.	**15 L**
Unit III	**Measures of similarity and structural equivalence in SNA** Approaches to network positions and social roles- defining equivalence or similarity, structural equivalence, automorphic equivalence, finding equivalence	**15 L**

	sets, brute force and Tabu search, regular equivalence, equivalence of distances: Maxsim, regular equivalence, Measuring similarity/dissimilarity- valued relations, Pearson correlations covariance and cross-products, Understanding clustering-agglomerative and divisive clusters, Euclidean, Manhattan, and squared distances, binary relations, matches: exact, Jaccard, Hamming, **Two-mode networks for SNA** Understanding mode networks- Bi-partite data structures, visualizing two-mode data, quantitative analysis using two-mode Singular value decomposition (SVD) analysis,	

Text book:
- Introduction to Social Network Methods: Robert A. Hanneman, Mark Riddle, University of California, 2005 [Published in digital form and available at http://faculty.ucr.edu/~hanneman/nettext/index.html].
- Social Network Analysis for Startups- Finding connections on the social web: MaksimTsvetovat, Alexander Kouznetsov, O'Reilly Media, 2011.
- Social Network Analysis- 3rd edition, John Scott, SAGE Publications, 2012.

Reference book:
- Exploratory Social Network Analysis with Pajek, Second edition: Wouter de Nooy, Andrej Mrvar, Vladimir Batagelj, Cambridge University Press, 2011.
- Analyzing Social Networks, Stephen P Borgatti, Martin G. Everett, Jeffrey C. Johnson, SAGE Publications, 2013.
- Statistical Analysis of Network Data with R: Eric D. Kolaczyk, GáborCsárdi, Springer, 2014.
- Network Analysis: Methodological Foundations, (Editors) UlrikBrandes, Thomas Erlebach. Springer, 2005.
- Models and Methods in Social Network Analysis: (Editors) Peter J. Carrington, John Scott,Stanley Wasserman, Cambridge University Press, 2005.

Links:
3) http://www.faculty.ucr.edu/~hanneman/nettext/C11_Cliques.html
4) https://www.safaribooksonline.com/library/view/social-network-analysis/9781449311377/ch04.html

Sr. No.	Practical of CS452
11.	Write a program to compute the following for a given a network: (i) number of edges, (ii) number of nodes; (iii) degree of node; (iv) node with lowest degree; (v)the adjacency list; (vi) matrix of the graph.
12.	Perform following tasks: (i) View data collection forms and/or import onemode/ two-mode datasets; (ii) Basic Networks matrices transformations

13.	Compute the following node level measures: (i) Density; (ii) Degree; (iii) Reciprocity; (iv) Transitivity; (v) Centralization; (vi) Clustering.
14.	For a given network find the following: (i) Length of the shortest path from a given node to another node; (ii) the density of the graph; (iii) Draw egocentric network of node G with chosen configuration parameters.
15.	Write a program to distinguish between a network as a matrix, a network as an edge list, and the network as a sociogram (or "network graph") using 3 distinct networks representatives of each.
16.	Write a program to exhibit structural equivalence, automatic equivalence, and regular equivalence from a network.
17.	Create sociograms for the persons-by-persons network and the committee-by committee network for a given relevant problem. Create one-mode network and two-node network for the same.
18.	Perform SVD analysis of a network.
19.	Identify ties within the network using two-mode core periphery analysis.
20.	Find "factions" in the network using two-mode faction analysis.

Course: CS453	Machine Learning (Credits :3 Lectures/Week:3)	Lectures
Expected Learning Outcomes: After successful completion of this course, students will be able to: 1. Develop a solid understanding of the fundamentals of machine learning, including its types and applications. 2. Apply various machine learning techniques such as linear regression, logistic regression, and decision trees to train models and make predictions. 3. Gain proficiency in using support vector machines (SVM) for classification and regression tasks. 4. Acquire practical skills in implementing machine learning algorithms using the TensorFlow framework and analyzing performance measures for model evaluation.		
Unit I	The Fundamentals of Machine Learning Understanding Machine Learning,Need and Relevance of Machine Learning, Types of Machine Learning, Supervised Learning, Unsupervised Learning & Reinforcement Learning. Challenges of Machine Learning, Testing and Validation. Classification, MNIST Dataset, Performance Measures, Confusion Matrix,	15 L

	Precision and Recall, Precision/Recall Tradeoff, The ROC Curve, Multiclass Classification, Error Analysis.	
Unit-II	Training Models Linear Regression, Gradient Descent, Batch Gradient Descent, Stochastic Gradient Descent, Mini-batch Gradient Descent, Polynomial Regression, Learning Curves, The Bias/Variance Tradeoff, Ridge Regression, Lasso Regression, Early Stopping, Logistic Regression, Decision Boundaries, Softmax Regression, Cross Entropy.	15 L
Unit-III	Support Vector Machines Linear SVM Classification, Soft Margin Classification, Nonlinear SVM Classification, Polynomial Kernel, Gaussian RBF Kernel, SVM Regression, Decision Trees, Training and Visualizing a Decision Tree, Making Predictions, The CART Training Algorithm, Gini Impurity vs Entropy, Regularization Hyperparameters.	15 L

Text Books:
- Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems by AurélienGéron, Second Edition, O'reilly, 2019
- Reinforcement Learning: An Introduction by Richard S. Sutton and Andrew G. Barto, Second Edition ,2014

Reference book:
- Introduction to Machine with Python - A Guide for Data Scientists by Andreas C. Müller & Sarah Guido O'reilly 2016

Course Code	Course Title	Credits
CS453	Machine Learning Practical	01
Note: • All the Practical's should be implemented using Python and TensorFlow. Link:Python :https://www.python.org/downloads/ TensorFlow :https://www.tensorflow.org/install		
1	Data Preprocessing – I	
2	Data Preprocessing – II	
3	Regression Analysis- Linear regression	
4	Regression Analysis- Multiple regression	
5	Regression Analysis- Logistic Regression	

6	Classification Techniques- Decision tree	
7	Classification Techniques- SVM	
8	Classification Techniques- Naïve Bayes, KNN	
9	Clustering- K- Means clustering	
10	Market Basket Analysis	

Course: CS454	Data Visualization (Credits :3 Lectures/Week:3)	Lectures
Expected Learning Outcomes: After successful completion of this course, students will be able to: 1. Understand various data formats, sources and storage mechanisms. 2. Handle missing data and manage data wrangling and manipulation 3. Create data visualization and report making using various software tools 4. Demonstrate the visualizations and make interpretations 5. Create a data story using various software tools.		
Unit I	Concept of Data Science Pipeline, different stages within a pipeline,Introduction to data Visualization,Categories of Data Visualization,Advantages of Data Visualization,Disadvantages of data visualization ,Importances of Data Visualization,Top Data Visualization Tools,Top Data Visualization Libraries. Data Visualization libraries(matplotlib):Environment Setup for Matplotlib, Introduction to Matplotlib,Using Matplotlib with Jupyter Notebook,Pyplot in Matplotlib Multiple Plots**,** Line Graph,Bar Chart,Histogram, Scatter Plot, Pie Chart, 3D Plots, Working with Images, Customizing Plots	15 L
Unit-II	Data Visualization Intro to data visualization - Introduction to Visualization and Dashboarding Software - Visualization Tools - Getting started with Tableau Desktop – Connecting to the dataset - Creating charts – Creating common visualizations (bar charts, line charts etc.) - Filtering and sorting data - Adding Titles, Labels, and descriptions - Publish your work to Tableau Cloud - Interactivity with text and visual tooltips - Interactivity with actions (filter, highlight, URL) – Assembling dashboards from multiple charts	15 L

Unit-III	Story Telling Introduction to Power BI - Understanding Desktop - Understanding Power BI Report Designer - Report Canvas, Report Pages: Creation, Renames - Report Visuals, Fields and UI Options - Experimenting Visual Interactions, Advantages - Reports with Multiple Pages and Advantages - Pages with Multiple Visualizations - PUBLISH Options and Report Verification in Cloud - Adding Report Titles. Report Format Options - Introduction to data storytelling - Creating a data story	15 L

Reference book:
- Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython by McKinney, W., 2nd edition. O‟Reilly Media, 2017
- Doing Data Science: Straight Talk from the Frontline by O‟Neil, C., & Schutt, R, O‟Reilly Media, 2013
- The Big Book of Dashboards by Steve Wexler, Jeffrey Shaffer, Andy Cotgreave, John Wiley & Sons, 2017
- Practical Tableau by Ryan Sleeper, O‟Reilly Media, 2018
- Power BI. Book-1, Business Intelligence Clinic: Create and Learn by Roger F Silva, 2018
- Introducing Microsoft Power BI by Alberto Ferrari and Marco Russo, Microsoft Press, Washington, 2016

Course Code	Course Title	Credits
CS454	**Data Visualization Practical**	01
Note: ● The following set of practicals could be implemented in Python/ R/ Power BI/ Tableau or any other suitable software		
1	**Multiple Plots:**Create multiple subplots in Matplotlib in Python, Add Title to Subplots in Matplotlib, Set a Single Main Title for All the Subplots in Matplotlib, Turn Off the Axes for Subplots in Matplotlib, Create Different Subplot Sizes in Matplotlib, set the spacing between subplots in Matplotlib in Python	
2	**Working with Legends:** Matplotlib.pyplot.legend() in Python, Matplotlib.axes.Axes.legend() in Python, Change the legend position in Matplotlib, Change Legend Font Size in Matplotlib , Change the vertical spacing between legend entries in Matplotlib, Use multiple columns in a Matplotlib legend , Create a Single Legend for All Subplots in Matplotlib, manually add a legend with a color box on a Matplotlib figure, Place Legend Outside of the Plot in Matplotlib, Remove the Legend, move the legend border in Matplotlib ,multiple columns in a Matplotlib legend	

3	**Line Chart:** Line chart in Matplotlib - Python, Line plot styles in Matplotlib, Plot Multiple lines in Matplotlib, Change the line opacity in Matplotlib , Increase the thickness of a line with Matplotlib t, Fill Between Multiple Lines in Matplotlib	
4	**Bar Plot:** Bar Plot in Matplotlib, Draw a horizontal bar chart with Matplotlib, Create a stacked bar plot in Matplotlib ,Stacked Percentage Bar Plot In MatPlotLib ,Plotting back-to-back bar charts Matplotlib ,display the value of each bar in a bar chart using Matplotlib , Annotate Bars in Barplot with Matplotlib in Python	
5	**Histogram:** Plotting Histogram in Python using Matplotlib, Create a cumulative histogram in Matplotlib ,plot two histograms together in Matplotlib ,Overlapping Histograms with Matplotlib in Python , Bin Size in Matplotlib Histogram	
6	**Scatter Plot:** Matplotlib.pyplot.scatter() in Python, add a legend to a scatter plot in Matplotlib , Connect Scatterplot Points With Line in Matplotlib , create a Scatter Plot with several colors in Matplotlib, increase the size of scatter points in Matplotlib .	
7	**Pie Chart :** Plot a pie chart in Python using Matplotlib , Set border for wedges in Matplotlib pie chart ,Radially displaced pie chart wedge in Matplotlib	
6	Create common charts with title, labels and descriptions using Tableau.	
7	Perform sorting and filtering using tableau, create visualizations and publish it on Tableau Cloud.	
8	Perform data visualization using Power BI.	
9	Create reports using Power BI.	
10	Create a data story in Tableau or power BI.	

Course: CS455	**Enterprise Application Integration** **(Credits :3 Lectures/Week:3)**	**Lectures**
Expected Learning Outcomes: After successful completion of this course, students will be able to: 1. Define your specific integration problem in a useful form that enables a real solution . 2. Develop your own EAI architecture and ensure interoperability of legacy technology applications. 3. Choose the best among architecture, object architecture, transaction architecture . 4. Work with the technologies .		

UNIT I	**What Is EAI?,** Traditional Systems, Microcomputer Systems, Distributed Systems, Packaged Applications, Making the Business Case for EAI, The Virtual System, e-Business, Types of EAI, Middleware and EAI **Data-Level EAI:** Data-Level EAI by Example, Database-to-Database EAI, Federated Database EAI, Consider the Data Source, Relational Data,Object-Oriented, Multidimensional, Other Data Storage Models,Hierarchical, ISAM and VSAM, CODASYL,Adabas, Working with Data-Level EAI	15L
UNIT II	**Application Interface-Level EAI:** Application Interfaces, What's an API?, Interface by Example, Approaching Application Interfaces, The Interface Tradeoff, Packaged Applications, Packaged Application Technology Architecture, Packaged Application APIs, Types of Services, Types of Interfaces, Other Interfaces, Vertical Market Application Interfaces, SWIFT, FIX, HL7, Custom Applications, Rolling Your Own API, Application Wrapping, Using Application, Interfaces **Method-Level EAI:** Method-Level Example, What's Process?Scenarios,Rules, Logic, Data, Objects, Process?Scenarios, Rules, Logic, Data, Objects, Method Warehousing, Leveraging, frameworks for EAI, The Value of Frameworks, Framework Functionality, framework Types, Service Frameworks, Procedural Frameworks, Component Frameworks, Framework Categories, Application Service Frameworks, Domain Frameworks, Support Frameworks, Enabling Technology, Application or Transaction Servers, Message Brokers, Distributed Objects, Sharing Methods to Bind Your Enterprise	15L
UNIT III	User Interface-Level EAI: Leveraging User Interface-Level EAI, Going to the User Interface, Understanding the Application, Creating the Screen Catalog, Mapping Screens, Finding the Information, Static Extraction, Dynamic Extraction, Error Technology, Screen Access Tricks, HLLAPI, ASCII or ANSI, OLE Automation, Screens as Objects, Screens-as-Objects, Enabling cessing, Approaches, Screens-as-Data, The EAI Process: 12 Step process model. Step 1: Understanding the Enterprise and Problem Domain Making Step 2: Sense of the Data Step 3: Making Sense of the Processes Step 4: Identifying Application Interfaces Step 5: Identifying the Business Events Step 6: Identifying the Schema and Content Transformation Scenarios Step 7: Mapping Information Movement Step 8: Applying Technology Step 9: Testing, Testing, Testing Step 10: Considering Performance Step 11: Defining the Value Step 12: Creating Maintenance Procedures	15L

UNIT IV	Java Middleware and EAI: Categories of Java Middleware Standards, Database-Oriented, Interprocess, Message-Oriented, Messaging Models, JMS and Application Development, Application-Hosting, Distributed Objects XML and EAI: the Rise Structures, DTDs,XML XML,What's Metadata,XML XML?,Data and Middleware, Persistent XML,RDF and EAI, XSL and EAI, XML and EAI	15L

Text book:
1.Enterprise Application Integration by David S. Linthicum Addison-Wesley Information Technology Series
2.Next-Generation Application Integration: From Simple Information to Web Services 1st Edition by Mary O'Brien (Author), David Linthicum (Author), John Fuller (Series Editor)

Course Code	Course Title	Credits
CS455	**Enterprise Application Integration**	01
1	Demonstrate usage of middleware in e-business application.	
2	Develop a database connector to integrate data from various sources into a single data model.	
3	Demonstrate the working of distributed objects in an application.	
4	Demonstrate Message broker as a Middleware for EAI	
5	Bank teller application to demonstrate API.	
6	Develop API for a dummy travel website on appropriate EAI framework.	
7	Demonstrate working of User interface level EAI	
8	Demonstrate data exchange in EAI app using XML.	