	Deep Learning		Semester	7
	Course Code	**BCS714A**	CIE Marks	50
	Teaching Hours/Week (L: T:P: S)	3:0:0:0	SEE Marks	50
	Total Hours of Pedagogy	40	Total Marks	100
	Credits	03	Exam Hours	03
	Examination type (SEE)	Theory		

Course objectives:

- Understand the basic concepts of deep learning.
- Know the basic working model of Convolutional Neural Networks and RNN in decision making.
- Illustrate the strength and weaknesses of many popular deep learning approaches.
- Introduce major deep learning algorithms, the problem settings, and their applications to solve real world problems

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation/Demonstration to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem/Practical Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills, and practical skill such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Use animations/videos to help the students to understand the concepts.
7. Demonstrate the concepts using PYTHON and its libraries wherever possible

Module-1

Introducing Deep Learning: **Biological and Machine Vision:** Biological Vision, Machine Vision: The Neocognitron, LeNet-5, The Traditional Machine Learning Approach, ImageNet and the ILSVRC, AlexNet, TensorFlow Playground. **Human and Machine Language:** Deep Learning for Natural Language Processing: Deep Learning Networks Learn Representations Automatically, Natural Language Processing, A Brief History of Deep Learning for NLP, Computational Representations of Language: One-Hot Representations of Words, Word Vectors, Word-Vector Arithmetic, word2viz, Localist Versus Distributed Representations, Elements of Natural Human Language.

Text book 2 : Chapter 1, 2

Module-2

Regularization for Deep Learning: Parameter Norm Penalties, Norm Penalties as Constrained Optimization, Regularization and Under-Constrained Problems, Dataset Augmentation, Noise Robustness, Semi- Supervised Learning, Multi-Task Learning, Early Stopping, Parameter Tying and Parameter Sharing, Sparse Representations, **Optimization for Training Deep Models:** How Learning Differs from Pure Optimization, Basic Algorithms. Parameter Initialization Strategies, Algorithms with Adaptive Learning Rates.

Text book 1 : Chapter 7 (7.1 to 7.10), Chapter 8 (8.1,8.3,8.4,8.5)

Module-3

	Convolution neural networks: The Convolution Operation, Motivation, Pooling, Convolution and Pooling as an Infinitely Strong Prior, Variants of the Basic Convolution Function, Structured Outputs, Data Types, Efficient Convolution Algorithms, Convolutional Networks and the History of Deep Learning. **Text book 1 : Chapter 9 (9.1 to 9.8, 9.11)**
	Module-4
	Sequence Modelling: Recurrent and Recursive Nets: Unfolding Computational Graphs, Recurrent Neural Networks, Bidirectional RNNs, Encoder-Decoder Sequence-to-Sequence Architectures, Deep Recurrent Networks, Recursive Neural Networks. Long short-term memory. **Text book 1 : Chapter 10 (10.1 to 10.6, 10.10)**
	Module-5
	Interactive Applications of Deep Learning: Natural Language Processing: Preprocessing Natural Language Data: Tokenization, Converting All Characters to Lowercase, Removing Stop Words and Punctuation, Stemming, Handling *n*-grams, Preprocessing the Full Corpus, Creating Word Embeddings with word2vec: The Essential Theory Behind word2vec, Evaluating Word Vectors, Running word2vec, Plotting Word Vectors, The Area under the ROC Curve: The Confusion Matrix, Calculating the ROC AUC Metric, Natural Language Classification with Familiar Networks: Loading the IMDb Film Reviews, Examining the IMDb Data, Standardizing the Length of the Reviews, Dense Network, Convolutional Networks, Networks Designed for Sequential Data: Recurrent Neural Networks, Long Short-Term Memory Units, Bidirectional LSTMs, Stacked Recurrent Models, Seq2seq and Attention, Transfer Learning in NLP. **Text book 2 : Chapter-8**

Course outcomes (Course Skill Set):

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

1. Interpret the concepts of neural networks learning processes.
2. Illustrate deep learning methods using regularization and Optimization process
3. Design deep learning models using convolutional operations.
4. Analyze sequential data to build recurrent and recursive models.
5. Demonstrate the different interactive applications of deep learning.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

- For the Assignment component of the CIE, there are 25 marks and for the Internal Assessment Test component, there are 25 marks.
- The first test will be administered after 40-50% of the syllabus has been covered, and the second test will be administered after 85-90% of the syllabus has been covered
- Any two assignment methods mentioned in the 22OB2.4, if an assignment is project-based then only one assignment for the course shall be planned. The teacher should not conduct two assignments at the end of the semester if two assignments are planned.
- For the course, CIE marks will be based on a scaled-down sum of two tests and other methods of assessment.

Internal Assessment Test question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours).**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scored shall be proportionally reduced to 50 marks

Suggested Learning Resources:

Books

1. Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press, 2016. https://www.deeplearningbook.org/lecture_slides.html
2. John Krohn, Grant Beyleveld, Aglae Bassens, Deep Learning Illustrated, A Visual, Interactive Guide to Artificial Intelligence, Pearson, 2022.

Web links and Video Lectures (e-Resources):

https://www.youtube.com/watch?v=VyWAvY2CF9c
https://www.youtube.com/watch?v=7sB052Pz0sQ
https://www.youtube.com/watch?v=Mubj_fqiAv8
https://www.coursera.org/learn/neural-networks-deep-learning

- **https://onlinecourses.nptel.ac.in/noc20_cs62/preview**

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

- Programming Assignments, such as implementation of CNN and Recurrent neural network models - 10 Marks
- Group assignment (Group of two) on recent developments in Deep learning – Refer IEEE/ACM/Elsevier etc publications - 15 Marks

	NATURAL LANGUAGE PROCESSING		Semester	7
	Course Code	**BCS714B**	CIE Marks	50
	Teaching Hours/Week (L: T:P: S)	3:0:0:0	SEE Marks	50
	Total Hours of Pedagogy	40	Total Marks	100
	Credits	03	Exam Hours	03
	Examination type (SEE)	Theory		

Course objectives:

- Learn the importance of natural language modelling
- Understand the applications of natural language processing
- Study spelling, error detection and correction methods and parsing techniques in NLP
- Illustrate the information retrieval models in natural language processing

Teaching-Learning Process (General Instructions)

These are sample strategies; which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) does not mean only the traditional lecture method, but different types of teaching methods may be adopted to achieve the outcomes.
2. Utilize video/animation films to illustrate the functioning of various concepts.
3. Promote collaborative learning (Group Learning) in the class.
4. Pose at least three HOT (Higher Order Thinking) questions in the class to stimulate critical thinking.
5. Incorporate Problem-Based Learning (PBL) to foster students' analytical skills and develop their ability to evaluate, generalize, and analyze information rather than merely recalling it.
6. Introduce topics through multiple representations.
7. Demonstrate various ways to solve the same problem and encourage students to devise their own creative solutions.
8. Discuss the real-world applications of every concept to enhance students' comprehension.
9. Use any of these methods: Chalk and board, Active Learning, Case Studies.

Module-1

Introduction: What is Natural Language Processing? Origins of NLP, Language and Knowledge, The Challenges of NLP, Language and Grammar, Processing Indian Languages, NLP Applications.

Language Modeling: Statistical Language Model - N-gram model (unigram, bigram), Paninion Framework, Karaka theory.

Textbook 1: Ch. 1, Ch. 2.

Module-2

Word Level Analysis: Regular Expressions, Finite-State Automata, Morphological Parsing, Spelling Error Detection and Correction, Words and Word Classes, Part-of Speech Tagging.

Syntactic Analysis: Context-Free Grammar, Constituency, Top-down and Bottom-up Parsing, CYK Parsing.

Textbook 1: Ch. 3, Ch. 4.

	Module-3
	Naive Bayes, Text Classification and Sentiment: Naive Bayes Classifiers, Training the Naive Bayes Classifier, Worked Example, Optimizing for Sentiment Analysis, Naive Bayes for Other Text Classification Tasks, Naive Bayes as a Language Model. **Textbook 2:** Ch. 4.
	Module-4
	Information Retrieval: Design Features of Information Retrieval Systems, Information Retrieval Models - Classical, Non-classical, Alternative Models of Information Retrieval - Custer model, Fuzzy model, LSTM model**,** Major Issues in Information Retrieval. **Lexical Resources**: WordNet, FrameNet, Stemmers, Parts-of-Speech Tagger, Research Corpora. **Textbook 1:** Ch. 9, Ch. 12.
	Module-5
	Machine Translation: Language Divergences and Typology, Machine Translation using Encoder-Decoder, Details of the Encoder-Decoder Model, Translating in Low-Resource Situations, MT Evaluation, Bias and Ethical Issues. **Textbook 2:** Ch. 13.

Course outcome (Course Skill Set)

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

1. Apply the fundamental concept of NLP, grammar-based language model and statistical-based language model.
2. Model morphological analysis using Finite State Transducers and parsing using context-free grammar and different parsing approaches.
3. Develop the Naïve Bayes classifier and sentiment analysis for Natural language problems and text classifications.
4. Apply the concepts of information retrieval, lexical semantics, lexical dictionaries such as WordNet, lexical computational semantics, distributional word similarity.
5. Identify the Machine Translation applications of NLP using Encode and Decoder.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

- For the Assignment component of the CIE, there are 25 marks and for the Internal Assessment Test component, there are 25 marks.
- The first test will be administered after 40-50% of the syllabus has been covered, and the second test will be administered after 85-90% of the syllabus has been covered
- Any two assignment methods mentioned in the 22OB2.4, if an assignment is project-based then only one assignment for the course shall be planned. The teacher should not conduct two assignments at the end of the semester if two assignments are planned.
- For the course, CIE marks will be based on a scaled-down sum of two tests and other methods of assessment.

Internal Assessment Test question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours).**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scored shall be proportionally reduced to 50 marks

Suggested Learning Resources:

Text Books:

1. Tanveer Siddiqui, U.S. Tiwary, "Natural Language Processing and Information Retrieval", Oxford University Press.
2. Daniel Jurafsky, James H. Martin, "Speech and Language Processing, An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition", Pearson Education, 2023.

Reference Books:

1. Akshay Kulkarni, Adarsha Shivananda, "Natural Language Processing Recipes - Unlocking Text Data with Machine Learning and Deep Learning using Python", Apress, 2019.
2. T V Geetha, "Understanding Natural Language Processing – Machine Learning and Deep Learning Perspectives", Pearson, 2024.
3. Gerald J. Kowalski and Mark.T. Maybury, "Information Storage and Retrieval systems", Kluwer Academic Publishers.

Web links and Video Lectures (e-Resources):

- https://www.youtube.com/watch?v=M7SWr5xObkA
- https://youtu.be/02QWRAhGc7g
- https://www.youtube.com/watch?v=CMrHM8a3hqw
- https://onlinecourses.nptel.ac.in/noc23_cs45/preview
- https://archive.nptel.ac.in/courses/106/106/106106211/

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

Text Classification Game (15 Marks)

- Objective: Learn supervised learning and text classification.
- Activity: Provide students with a set of documents (e.g., movie reviews) labeled as positive or negative. Divide them into groups and have them create a simple classification model using keywords or phrases. They can then test their model on new reviews.

Grammar Check and Correction (10 Marks)

- Objective: Learn about language structure and NLP tools.
- Activity: Provide sentences with grammatical errors. Students can use grammar checking tools (like Grammarly or LanguageTool) to identify errors and suggest corrections, discussing why each suggestion is made.

Social Network Analysis		Semester	7
Course Code	**BAD714D**	CIE Marks	50
Teaching Hours/Week (L: T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03
Examination type (SEE)	Theory		

Course objectives:

- To introduce the fundamentals of Social Network Analysis and its significance in understanding societal connections and behaviors.
- To analyze various models of network growth and understand the properties of real-world networks.
- To explore link analysis algorithms and their applications in understanding relationships within a network.
- To study community detection methods and their relevance in identifying meaningful clusters within networks.
- To understand link prediction techniques and their application in forecasting future connections within a network.

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) needs not to be only traditional lecture method, can make use of digital tools to visually demonstrate key ideas that could be adopted to attain the outcomes.
2. Use think-pair-share strategies where students collaborate in pairs or groups to discuss concepts and solve small problems before sharing their understanding with the class.
3. Use real-world examples such as social media platforms or professional networks (e.g., LinkedIn) to demonstrate the concepts of Social Network Analysis.
4. Conduct practical sessions using software like Python with network libraries (e.g., NetworkX) to model and visualize network growth.
5. Use step-by-step explanations to demonstrate algorithms like PageRank and SimRank, followed by coding sessions for implementation.
6. Use network visualization tools (e.g., Gephi, Cytoscape) to help students identify and analyze communities in networks.
7. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information.
8. Demonstrate ways to solve the same problem and encourage the students to come up with their own creative solutions.

Module-1

Networks and Society - What is Social Network Analysis, why do We Study Social Networks, Applications of Social Network Analysis, Preliminaries, Three Levels of Social Network Analysis.

Network Measures - Network Basics, Node Centrality, Assortativity, Transitivity and Reciprocity, Similarity, Degeneracy.

T1 – Chapter 1 (1.1. – 1.5), Chapter 2 (2.1 – 2.6)

Module-2

Network Growth Models - Properties of Real-World Networks, Random Network Model, Ring Lattice Network Model, Watts–Strogatz Model, Preferential Attachment Model, Price's Model, Local-world Network Growth Model, Network Model with Accelerating Growth, Aging in Preferential Attachment.

T1 – Chapter 3 (3.1 – 3.9)

Module-3

Link Analysis - Applications of Link Analysis, Signed Networks, Strong and Weak Ties, Link Analysis Algorithms, PageRank, Personalised PageRank, DivRank, SimRank, PathSIM.
T1 – Chapter 4 (4.1 – 4.8)

Module-4

Community Structure in Networks - Applications of Community Detection, Types of Communities, Community Detection Methods, Disjoint Community Detection, Overlapping Community Detection, Local Community Detection, Community Detection vs Community Search, Evaluation of Community Detection Methods.
T1 – Chapter 5 (5.1 – 5.8)

Module-5

Link Prediction - Applications of Link Prediction, Temporal Changes in a Network, Problem Definition Evaluating Link Prediction Methods, Heuristic Models, Probabilistic Models, Supervised Random Walk, Information-theoretic Model, Latest Trends in Link Prediction.
T1 – Chapter 6 (6.1 – 6.9)

Course outcome (Course Skill Set)

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

1. Illustrate the core concepts of Social Network Analysis and its levels of study.
2. Demonstrate the different network growth models for real-world networks
3. Apply algorithms of PageRank and SimRank to analyze and interpret link relationships.
4. Apply community detection methods and evaluating their effectiveness in real-world scenarios.
5. Analyze heuristic, probabilistic, and supervised models to predict network link formations and changes.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

- For the Assignment component of the CIE, there are 25 marks and for the Internal Assessment Test component, there are 25 marks.
- The first test will be administered after 40-50% of the syllabus has been covered, and the second test will be administered after 85-90% of the syllabus has been covered
- Any two assignment methods mentioned in the 22OB2.4, if an assignment is project-based then only one assignment for the course shall be planned. The teacher should not conduct two assignments at the end of the semester if two assignments are planned.
- For the course, CIE marks will be based on a scaled-down sum of two tests and other methods of assessment.

Internal Assessment Test question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours).**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scored shall be proportionally reduced to 50 marks

Suggested Learning Resources:

Books

1. Tanmoy Chakraborty, "Social Network Analysis", Wiley India Pvt. Ltd., 2021

Reference Books

1. Albert-Laszlo Barabasi, "Network Science", Cambridge University Press, 2016
2. Stanley Wasserman, Katherine Faust, "Social Network Analysis: Methods and Applications", Cambridge University Press, 1994

Web links and Video Lectures (e-Resources):

- https://onlinecourses.nptel.ac.in/noc22_cs117/preview
- https://social-network-analysis.in/
- https://www.coursera.org/learn/social-network-analysis

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

Activity 1: Network Visualization and Analysis (10 Marks)

Understand network basics, measures, and visualization techniques using a real-world dataset.

Instructions:

1. Choose a small real-world dataset (e.g., social media connections, collaboration networks, or communication networks).
2. Use a network analysis tool such as Gephi, NetworkX, or Cytoscape to visualize the dataset.
3. Analyze the following:
   - Node centrality measures (degree, closeness, and betweenness).
   - Network transitivity and reciprocity.
   - Similarity or assortativity in the network.
4. Submit a report that includes the network visualization and a summary of key findings.

Assessment Criteria:

- Clarity of visualization (3 marks)
- Accuracy in calculating and interpreting network measures (5 marks)
- Quality of the report (2 marks)

Activity 2: Community Detection and Link Prediction Project (15 Marks)

Apply community detection techniques and predict future connections within a network.

Instructions:

1. Select a medium-sized dataset (e.g., email communications, citation networks, or transport networks).
2. Perform the following tasks:
   - Identify and visualize communities using two different community detection methods (e.g., Disjoint and Overlapping Community Detection).
   - Evaluate the detected communities using appropriate evaluation metrics (e.g., modularity).
   - Use a link prediction algorithm (e.g., supervised random walk or probabilistic models) to forecast future connections within the network.
3. Prepare a detailed report with visuals and findings.

Assessment Criteria:

- Accuracy and comparison of community detection methods (7 marks)
- Implementation and results of link prediction (5 marks)
- Overall presentation and report quality (3 marks)

BIG DATA ANALYTICS		Semester	7
Course Code	**BCS714D**	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	3
Examination nature (SEE)	Theory		

Course objectives:

1. To implement MapReduce programs for processing big data.
2. To realize storage and processing of big data using MongoDB, Pig, Hive and Spark.
3. To analyze big data using machine learning techniques.

Teaching-Learning Process (General Instructions)

These are sample Strategies; that teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) needs not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.
6. Use any of these methods: Chalk and board, Active Learning, Case Studies.

MODULE-1

Classification of data, Characteristics, Evolution and definition of Big data, What is Big data, Why Big data, Traditional Business Intelligence Vs Big Data,Typical data warehouse and Hadoop environment.

Big Data Analytics: What is Big data Analytics, Classification of Analytics, Importance of Big Data Analytics, Technologies used in Big data Environments, Few Top Analytical Tools , NoSQL, Hadoop.

TB1: Ch 1: 1.1, Ch2: 2.1-2.5,2.7,2.9-2.11, Ch3: 3.2,3.5,3.8,3.12, Ch4: 4.1,4.2

MODULE-2

Introduction to Hadoop: Introducing hadoop, Why hadoop, Why not RDBMS, RDBMS Vs Hadoop, History of Hadoop, Hadoop overview, Use case of Hadoop, HDFS (Hadoop Distributed File System),Processing data with Hadoop, Managing resources and applications with Hadoop YARN(Yet Another Resource Negotiator).

Introduction to Map Reduce Programming: Introduction, Mapper, Reducer, Combiner, Partitioner, Searching, Sorting, Compression.

TB1: Ch 5: 5.1-,5.8, 5.10-5.12, Ch 8: 8.1 - 8.8

MODULE-3

Introduction to MongoDB: What is MongoDB, Why MongoDB, Terms used in RDBMS and MongoDB, Data Types in MongoDB, MongoDB Query Language.

TB1: Ch 6: 6.1-6.5

MODULE-4

Introduction to Hive: What is Hive, Hive Architecture, Hive data types, Hive file formats, Hive Query Language (HQL), RC File implementation, User Defined Function (UDF).

Introduction to Pig: What is Pig, Anatomy of Pig, Pig on Hadoop, Pig Philosophy, Use case for Pig, Pig Latin Overview, Data types in Pig, Running Pig, Execution Modes of Pig, HDFS Commands, Relational Operators, Eval Function, Complex Data Types, Piggy Bank, User Defined Function, Pig Vs Hive.

TB1: Ch 9: 9.1-9.6,9.8, Ch 10: 10.1 - 10.15, 10.22

MODULE-5

Spark and Big Data Analytics: Spark, Introduction to Data Analysis with Spark.

Text, Web Content and Link Analytics: Introduction, Text Mining, Web Mining, Web Content and Web Usage Analytics, Page Rank, Structure of Web and Analyzing a Web Graph.
TB2: Ch5: 5.2,5.3, Ch 9: 9.1-9.4

Course outcomes (Course Skill Set):
At the end of the course, the student will be able to:

- Illustrate Big Data concepts, tools and applications.
- Develop programs using HADOOP framework.
- Use Hadoop Cluster to deploy Map Reduce jobs, PIG,HIVE and Spark programs.
- Analyze the given data set to identify deep insights.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

For the Assignment component of the CIE, there are 25 marks and for the Internal Assessment Test component, there are 25 marks.

The first test will be administered after 40-50% of the syllabus has been covered, and the second test will be administered after 85-90% of the syllabus has been covered

Any two assignment methods mentioned in the 22OB2.4, if an assignment is project-based then only one assignment for the course shall be planned. The teacher should not conduct two assignments at the end of the semester if two assignments are planned.

For the course, CIE marks will be based on a scaled-down sum of two tests and other methods of assessment.

Internal Assessment Test question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

- Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (duration 03 hours).
- The question paper will have ten questions. Each question is set for 20 marks.
- There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), should have a mix of topics under that module.
- The students have to answer 5 full questions, selecting one full question from each module.
- Marks scored shall be proportionally reduced to 50 marks.

Suggested Learning Resources:

Books:

1. Seema Acharya and Subhashini Chellappan "Big data and Analytics", Wiley India Publishers, 2nd Edition, 2019.
2. Rajkamal and Preeti Saxena, "Big Data Analytics, Introduction to Hadoop, Spark and Machine Learning", McGraw Hill Publication, 2019.

Reference Books:

1. Adam Shook and Donald Mine, "MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems" - O'Reilly 2012

2. Tom White, "Hadoop: The Definitive Guide" 4th Edition, O'reilly Media, 2015.
3. Thomas Erl, Wajid Khattak, and Paul Buhler, Big Data Fundamentals: Concepts, Drivers & Techniques, Pearson India Education Service Pvt. Ltd., 1st Edition, 2016
4. John D. Kelleher, Brian Mac Namee, Aoife D'Arcy -Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, MIT Press 2020, 2nd Edition

Web links and Video Lectures (e-Resources):

- https://www.youtube.com/watch?v=bAyrObl7TYE&list=PLEiEAq2VkUUJqp1k-g5W1mo37urJQOdCZ
- https://www.youtube.com/watch?v=VmO0QgPCbZY&list=PLEiEAq2VkUUJqp1kg5W1mo37urJQOdCZ&index=4
- https://www.youtube.com/watch?v=GG-VRm6XnNk https://www.youtube.com/watch?v=JglO2Nv_92A

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

1. Implement MongoDB based application to store big data for data processing and analyzing the results [15 marks]
2. Install Hadoop and Implement the following file management such as Adding files and directories, Retrieving files, Deleting files and directories and execute Map- Reduce based programs.[10]

Introduction to DBMS		Semester	7
Course Code	**BCS755A**	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	3
Examination nature (SEE)	Theory		

Course objectives:

- To Provide a strong foundation in database concepts, technology, and practice.
- To Practice SQL programming through a variety of database problems.
- To Understand the relational database design principles.
- To Demonstrate the use of concurrency in database.
- To Design and build database applications for real world problems.

Teaching-Learning Process (General Instructions)

These are sample Strategies; that teachers can use to accelerate the attainment of the various course outcomes.

- Lecturer method (L) needs not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
- Use of Video/Animation to explain functioning of various concepts.
- Encourage collaborative (Group Learning) Learning in the class.
- Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
- Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.
- Use any of these methods: Chalk and board, Active Learning, Case Studies.

MODULE-1

Introduction to Databases: Introduction, Characteristics of database approach, Advantages of using the DBMS approach, History of database applications.

Overview of Database Languages and Architectures: Data Models, Schemas, and Instances. Three schema architecture and data independence, database languages, and interfaces, The Database System environment.

Textbook 1:Ch 1.1 to 1.8, 2.1 to 2.6

MODULE-2

Conceptual Data Modeling using Entities and Relationships: Entity types, Entity sets and structural constraints, Weak entity types, ER diagrams, Specialization and Generalization.

Mapping Conceptual Design into a Logical Design: Relational Database Design using ER-to-Relational mapping

Textbook 1: Ch 3.1 to 3.10, 9.1 & 9.2

MODULE-3

Relational Model: Relational Model Concepts, Relational Model Constraints and relational database schemas, Update operations, transactions, and dealing with constraint violations.

Relational Algebra: Unary and Binary relational operations, additional relational operations (aggregate, grouping, etc.) Examples of Queries in relational algebra.

Textbook 1: Ch 5.1 to 5.3, Ch 8.1 to 8.5

MODULE-4

SQL: SQL data definition and data types, Schema change statements in SQL, specifying constraints in SQL, retrieval queries in SQL, INSERT, DELETE, and UPDATE statements in SQL, Additional features of SQL

Normalization: Database Design Theory – Introduction to Normalization using Functional and Multivalued Dependencies: Informal design guidelines for relation schema, Functional Dependencies, Normal Forms based on Primary Keys, Second and Third Normal Forms, Boyce-Codd Normal Form, Multivalued Dependency and Fourth Normal Form, Join Dependencies and Fifth Normal Form.

Textbook 1: Ch 6.1 to 6.5,14.1 to 14.7

MODULE-5

SQL: Advanced Queries: More complex SQL retrieval queries, Specifying constraints as assertions and action triggers, Views in SQL.

Concurrency Control in Databases: Two-phase locking techniques for Concurrency control, Concurrency control based on Timestamp ordering, Multiversion Concurrency control techniques, Validation Concurrency control techniques, Granularity of Data items and Multiple Granularity Locking.

Textbook 1: Ch 7.1 to 7.3, 21.1 to 21.5

Course outcomes (Course Skill Set):
At the end of the course, the student will be able to:

- Demonstrate the basic elements of a database management system.
- Design entity relationship and convert entity relationship diagrams into RDBMS.
- Use Structured Query Language (SQL) for database manipulation.
- Apply normalization to increase the efficiency of database design.
- Illustrate the concepts of concurrency control techniques.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

For the Assignment component of the CIE, there are 25 marks and for the Internal Assessment Test component, there are 25 marks.

The first test will be administered after 40-50% of the syllabus has been covered, and the second test will be administered after 85-90% of the syllabus has been covered

Any two assignment methods mentioned in the 22OB2.4, if an assignment is project-based then only one assignment for the course shall be planned. The teacher should not conduct two assignments at the end of the semester if two assignments are planned.

For the course, CIE marks will be based on a scaled-down sum of two tests and other methods of assessment.

Internal Assessment Test question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

- Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (duration 03 hours).
- The question paper will have ten questions. Each question is set for 20 marks.
- There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), should have a mix of topics under that module.
- The students have to answer 5 full questions, selecting one full question from each module.
- Marks scored shall be proportionally reduced to 50 marks.

Suggested Learning Resources:

Text Books:

1. Fundamentals of Database Systems, Ramez Elmasri and Shamkant B. Navathe, 7th Edition, 2017, Pearson.

Reference Books:

1. Database management systems, Ramakrishnan, and Gehrke, 3rd Edition, 2014, McGraw Hill

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

Course Project (25 marks)

- For any problem selected
    - Develop the application having at least five tables & domain areas shall include health care, agriculture & so on.

	Introduction to Algorithms		Semester	7
	Course Code	BCS755B	CIE Marks	50
	Teaching Hours/Week (L: T:P: S)	3:0:0:0	SEE Marks	50
	Total Hours of Pedagogy	40	Total Marks	100
	Credits	03	Exam Hours	03
	Examination type (SEE)	Theory		

Course objectives:

- To learn the methods for analyzing algorithms and evaluating their performance.
- To demonstrate the efficiency of algorithms using asymptotic notations.
- To solve problems using various algorithm design methods, including brute force, greedy, divide and conquer, decrease and conquer, transform and conquer, dynamic programming, backtracking, and branch and bound.
- To learn the concepts of P and NP complexity classes.

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) does not mean only the traditional lecture method, but different types of teaching methods may be adopted to achieve the outcomes.
2. Utilize video/animation films to illustrate the functioning of various concepts.
3. Promote collaborative learning (Group Learning) in the class.
4. Pose at least three HOT (Higher Order Thinking) questions in the class to stimulate critical thinking.
5. Incorporate Problem-Based Learning (PBL) to foster students' analytical skills and develop their ability to evaluate, generalize, and analyze information rather than merely recalling it.
6. Introduce topics through multiple representations.
7. Demonstrate various ways to solve the same problem and encourage students to devise their own creative solutions.
8. Discuss the real-world applications of every concept to enhance students' comprehension.

Module-1

INTRODUCTION: What is an Algorithm?, Fundamentals of Algorithmic Problem Solving, Important problem Types, Fundamental Data Structures, Analysis Framework, Asymptotic Notations and Basic Efficiency Classes, ,Analysis Framework, Asymptotic Notations and Basic Efficiency Classes,

Chapter 1 (Sections 1.1 to 1.4), Chapter 2 (2.1, 2.2)

Module-2

FUNDAMENTALS OF THE ANALYSIS OF ALGORITHM EFFICIENCY: Mathematical Analysis of Non-recursive Algorithms, Mathematical Analysis of Recursive Algorithms.

BRUTE FORCE APPROACHES: Selection Sort and Bubble Sort, Sequential Search and Brute Force String Matching.

Chapter 2(Sections 2.3,2.4), Chapter 3(Section 3.1,3.2)

	Module-3
	Exhaustive Search (Travelling Salesman problem and Knapsack Problem). **Depth First search and Breadth First search.** **DECREASE-AND-CONQUER:** Insertion Sort, Topological Sorting. **DIVIDE AND CONQUER:** Merge Sort, Binary Tree Traversals. **Chapter 3(3.4,3.5), Chapter 4 (Sections 4.1,4.2), Chapter 5 (Section 5.1,5.3)**
	Module-4
	TRANSFORM-AND-CONQUER: Balanced Search Trees (AVL Trees), Heaps and Heapsort. **SPACE-TIME TRADEOFFS:** Sorting by Counting: Comparison counting sort, Input Enhancement in String Matching: Horspool's Algorithm, Hashing. **Chapter 6 (Sections 6.3,6.4), Chapter 7 (Sections 7.1,7.2, 7.3)**
	Module-5
	DYNAMIC PROGRAMMING: Three basic examples, The Knapsack Problem and Memory Functions. **THE GREEDY METHOD:** Kruskal's Algorithm, Dijkstra's Algorithm, Huffman Trees and Codes. **Chapter 8 (Sections 8.1,8.2), Chapter 9 (Sections 9.2,9.3,9.4)**

Course outcome (Course Skill Set)

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

1. Explain the algorithm design steps and computational problem types.
2. Apply the asymptotic notational method to analyze the performance of the algorithms in terms of time complexity.
3. Demonstrate divide & conquer approaches and decrease & conquer approaches to solve computational problems.
4. Make use of the transform & conquer design approach to solve the given real-world or complex computational problems.
5. Apply greedy and dynamic programming methods to solve graph & string-based computational problems.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

- For the Assignment component of the CIE, there are 25 marks and for the Internal Assessment Test component, there are 25 marks.
- The first test will be administered after 40-50% of the syllabus has been covered, and the second test will be administered after 85-90% of the syllabus has been covered
- Any two assignment methods mentioned in the 22OB2.4, if an assignment is project-based then only one assignment for the course shall be planned. The teacher should not conduct two assignments at the end of the semester if two assignments are planned.
- For the course, CIE marks will be based on a scaled-down sum of two tests and other methods of assessment.

Internal Assessment Test question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours).**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scored shall be proportionally reduced to 50 marks

Suggested Learning Resources:

Textbooks

1. Introduction to the Design and Analysis of Algorithms, By Anany Levitin, 3rd Edition (Indian), 2017, Pearson.

Reference books

1. Computer Algorithms/C++, Ellis Horowitz, SatrajSahni and Rajasekaran, 2nd Edition, 2014, Universities Press.
2. Introduction to Algorithms, Thomas H. Cormen, Charles E. Leiserson, Ronal L. Rivest, Clifford Stein, 3rd Edition, PHI.
3. Design and Analysis of Algorithms, S. Sridhar, Oxford (Higher Education)

Web links and Video Lectures (e-Resources):

- Design and Analysis of Algorithms: https://nptel.ac.in/courses/106/101/106101060/

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

1. Problem Solving - Competitive programming (Hacker Rank/ Hacker Earth / Leetcode) – 10 Marks
2. Problem solving (Numerical examples) related to different algorithms – 15 Marks

	SOFTWARE ENGINEERING		Semester	7
	Course Code	**BCS755C**	CIE Marks	50
	Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
	Total Hours of Pedagogy	50	Total Marks	100
	Credits	04	Exam Hours	3
	Examination type (SEE)	Theory		

Course objectives:

To understand foundational principles and the evolving nature of software engineering.
- To learn various software process models and their practical applications.
- To acquire skills in gathering, modeling, and validating software requirements.
- To apply Agile methodologies and understand core software engineering practices.
- To build a foundation for software design, testing, and quality assurance.

Teaching-Learning Process

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) needs not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding
9. Use any of these methods: Chalk and board, Active Learning, Case Studies

Module-1

Software and Software Engineering: The nature of Software, The unique nature of WebApps, Software Engineering, The software Process, Software Engineering Practice, Software Myths.

Process Models: A generic process model, Process assessment and improvement, Prescriptive process models: Waterfall model, Incremental process models, Evolutionary process models, Concurrent models, Specialized process models. Unified Process , Personal and Team process models

Textbook 1: Chapter 1: 1.1 to 1.6, Chapter 2: 2.1 to 2.5

Module-2

	Understanding Requirements: Requirements Engineering, Establishing the ground work, Eliciting Requirements, Developing use cases, Building the requirements model, Negotiating Requirements, Validating Requirements. **Requirements Modeling Scenarios, Information and Analysis classes**: Requirement Analysis, Scenario based modeling, UML models that supplement the Use Case, Data modeling Concepts, Class-Based Modeling. Requirement Modeling Strategies : Flow oriented Modeling , Behavioral Modeling. **Textbook 1: Chapter 5: 5.1 to 5.7, Chapter 6: 6.1 to 6.5, Chapter 7: 7.1 to 7.3**
	Module-3
	Agile Development: What is Agility?, Agility and the cost of change. What is an agile Process?, Extreme Programming (XP), Other Agile Process Models, A tool set for Agile process . **Principles that guide practice**: Software Engineering Knowledge, Core principles, Principles that guide each framework activity. **Textbook 1: Chapter 3: 3.1 to 3.6, Chapter 4: 4.1 to 4.3**
	Module-4
	Software Design: Design within the context of software engineering, Design process and quality, Design concepts: abstraction, modularity, architecture, patterns. **Architectural Design**: Architectural styles and patterns, reference architectures, component-level design, designing class-based components, conducting component-level design, design for reuse. **Textbook 1:Chapter 8: 8.1–8.6, Chapter 9: 9.1–9.5**
	Module-5
	Software Testing: Introduction to software testing, Strategic approach, Test strategies for conventional and object-oriented software, Validation testing, System testing, White-box and Black-box testing, Basis Path Testing, Control structure testing. **Software Quality:** Concepts of quality, Software quality assurance, Reviews, Software reliability and metrics. **Textbook 1: Chapter 14: Sections 14.1 to 14.5,Chapter 15: Sections 15.1 to 15.5, Chapter 19: Sections 19.1 to 19.5**

Course outcome

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

1. Explain the software nature, engineering practices, myths, and software process models.
2. Apply requirements engineering, elicitation, modeling, and validation in software development.
3. Demonstrate agile principles, practices, and tools for software development agility.
4. Apply design concepts, process, and architecture for quality software development.
5. Explain software testing strategies and quality assurance for reliable software.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

- For the Assignment component of the CIE, there are 25 marks and for the Internal Assessment Test component, there are 25 marks.
- The first test will be administered after 40-50% of the syllabus has been covered, and the second test will be administered after 85-90% of the syllabus has been covered
- Any two assignment methods mentioned in the 22OB2.4, if an assignment is project-based then only one assignment for the course shall be planned. The teacher should not conduct two assignments at the end of the semester if two assignments are planned.
- For the course, CIE marks will be based on a scaled-down sum of two tests and other methods of assessment.

Internal Assessment Test question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours).**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scored shall be proportionally reduced to 50 marks

Suggested Learning Resources:

Textbook

Roger S. Pressman: Software Engineering – A Practitioner's Approach, 7th Edition, Tata McGraw Hill, 2010.

Web links and Video Lectures (e-Resources):

https://www.geeksforgeeks.org/software-engineering/software-engineering/

Activity-Based Learning (Suggested Activities in Class)/Practical-Based learning

- Course project (Group of two students): Simulation that covers all the phases of SDLC - 25 marks