Shiromani Gurdwara Parbandhak Committee's

Guru Nanak Khalsa College of

Arts, Science and Commerce (Autonomous)

Matunga, Mumbai – 400 019, Maharashtra, India

Syllabus for M.Sc. Semester I and II

Program: Master of Science

Course: Computer Science

(M.Sc. Computer Science –Part-1 Syllabus as per NEP-2020)

(With Effect from 2023-24)

Sem.-1 (Total credits-20)

Sr. No.	Paper Title	Papers	Theory credit (hours)	Practical credit (hours)
1	Algorithm for Optimization Software Defined Networking	Paper I	4 Credits	2 Credits
2	Applied Signal and Image Processing Advanced Database Techniques	Paper II	4 Credits	2 Credits
3	Principles of Compiler Design	Paper III	2 Credits	-
4	Data science technologies - I Data warehousing and Data mining	Elective Paper	3 Credits	1 Credits
5	Research Methodology		4 Credits	-

Sem.-2 (Total credits-20)

Sr. No.	Paper Title	Papers	Theory credit (hours)	Practical credit (hours)
1	Applied Machine and Deep Learning Natural Language Processing	Paper I	4 Credits	2 Credits
2	Web Mining Embedded and IoT Technology	Paper II	4 Credits	2 Credits
3	Web Data Analytics	Paper III	2 Credits	-
4	Data science technologies – II Data Analysis and Data mining	Elective Paper	3 Credits	1 Credits
5	On job training		4 Credits	-

Subject	**Credits**	**Internal Evaluation (Marks)**	**Semester End Examination (Marks)**	**Total**
Mandatory Paper – I	04	25	75	100
Mandatory Paper – II	04	25	75	100
Mandatory Paper – III	02			
Practical – I	02	-	75	75
Practical – II	02	-	75	75
Elective – Theory	03	25	75	100
Elective – Practical	01	-	50	50
Research Methodology (RM)	02 + 02 = 04	50	50	100
			Grand Total	600

Detailed Syllabus for Semester - I

Course Code	Course Title	Credits
GNKPSCS1501	**Algorithm for Optimization**	**02**

Course Outcome: -

- You will be able to effectively implement optimization techniques to theexisting algorithm to improve its performance.
- You will be able to work in the areas of Machine Learning and DataSciences Algorithms

Course Specific Outcome: -

- Optimization with a focus on practical algorithms for the design ofengineering systems
- Exposure to multivariable calculus, linear algebra, and probability concepts. Learn a wide variety of optimization topics, introducing the underlyingmathematical problem formulations and the algorithms for solving them.

Unit-1: Basic Optimization Problem, Conditions for Local Minima, Contour Plots, Fibonacci Search, Golden Section Search, Quadratic Fit Search, Gradient Descent, RMSProp, Adadelta, Second-Order Methods, Newton's Method, Secant Method, Quasi-Newton Methods	15L
Unit-2: Surrogate Models, Fitting Surrogate Models, Linear Models, BasisFunctions, Uncertainty, Set-Based Uncertainty,Probabilistic Uncertainty. Uncertainty Propagation, Sampling Methods, Taylor Approximation, Polynomial Chaos, Dynamic Programming, Ant Colony Optimization. Expression Optimization, Grammars, Genetic Programming	15L

TEXT BOOK:

. Algorithms for Optimization Mykel J. Kochenderfer, Tim A. Wheeler, TheMIT Press 2019.

REFERENCE BOOKS:

1. Think Julia: How to Think Like a Computer Scientist by Allen B. Downey and Ben Lauwens 1st Edition 2019 O'reilly.
2. Decision Making Under Uncertainty: Theory and Application by Mykel J. Kochenderfer MIT Lincoln Laboratory Series 2015.
3. Introduction to Algorithms, By Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein 3Ed. (International Edition) (MIT Press) 2009

Course Code	Course Title	Credits
GNKPSCS1P501	**Practical Course on Algorithm for Optimization**	**01**

Note: All the Practical's should be implemented using
Julia Link: Julia:https://julialang.org/

1	Implement Contour Plots.
2	Implement Fibonacci and Golden section search.
3	Implement Quadratic Fit Search.
4	Implement Gradient descent.

5	Implement quasi-Newton methods to find the local maxima.
6	Implement the Adagrad method with application, RMSprop andAdadelta.
7	Implement radial basis functions using surrogate modelling.
8	Apply Random Forest in surrogate Model.

Course Code	Course Title	Credits
GNKPSCS1501	**Software Defined Networking**	**02**

Course Outcome: -

- To make the students capable of understanding computer network basics.
- To obtain the knowledge of Software defined networks with understanding of data plane, control plane and application plane.
- To apply network virtualization for industry standard solutions.
- To improve skills in implementing network virtualization and SoftwareDefined Network (SDN).

Course Specific Outcome: -

- Learners will be able to understand basic concepts of Software DefinedNetworking and network virtualization.
- Learners will be able to explore OpenFlow specifications to build Softwaredefined networks.
- Learners will be able to analyse and implement theories and practicalrelated to Network management and Virtualization.

Learners will be able to apply knowledge of Software Defined Networkingas per industry standards.

Unit 1: Introduction to Computer Networking and Software Defined Networking Basic Concepts and Definitions: LAN, MAN, WAN, AD-Hoc, Wireless Network. Elements of Modern Networking, Requirements and Technology,SDN: Background and Motivation, SDN Data Plane and OpenFlow, SDN Control Plane, SDN Application Plane	15L
Unit 2: Network Functions Virtualization and Design and implementation of Network Concepts and Architecture, NFV Functionality, Network Virtualization Quality of Service, Modern networking architecture: Cloud Computing, Fog computing, The Internet of Things: Components Understand and implement Layer 2/3 switching techniques Implementation of OSPF V2 and V3, Implementation BGP Implementation of Traffic Filtering by usingStandard and Extended Access Control List	15L

TEXT BOOK:

1. Behrouz A Forouzan ―TCPIP Protocol Suite‖ Fourth Edition 2010
2. William Stallings, ―Foundationsof Modern Networking‖, Pearson Ltd.,2016.
3. Software Defined Networks: A Comprehensive Approach by Paul Goransson and Chuck Black, Morgan Kaufmann Publications, 2014
4. SDN - Software Defined Networks by Thomas D. Nadeau & Ken Gray,O'Reilly, 2013

REFERENCE BOOKS:

1. Network Programmability and Automation-Jason Edelman, Matt OswaltFirst Edition 2018.

CourseCode	Course Title	Credits
GNKPSCS1P501	**Practical Course on Software Defined Networking**	**01**

Note: All the Practical's should be implemented using GNS3/EVE-NG/CISCO VIRL

Link: GNS3 :https://www.gns3.com/software/download
EVE-NG: https://www.eve-ng.net/index.php/download/CISCO
VIRL:https://learningnetwork.cisco.com/s/question/0D53i00000Kswpr/virl- 15-download

1	Implement Standard Access Control list.
2	Implement VTP
3	1) Implement Inter-VLAN Routing 2) Creating a VLAN on two Switch. (Inter-VLAN)
4	Observe STP Topology Changes and Implement RSTP
5	OSPF Implementation 1) Implement Single Area OSPFv2 2) Implement Multi Area OSPFv2
6	Implement BGP communities: 1) Implement EBGP 2) Implement IBGP
7	Implement Span Technology (Switch Port Analyzer)
8	Implement IPsec Site-to-site VPNs connection using cisco packet tracer.

Course Code	Course Title	Credits
GNKPSCS2501	**Applied Signal and Image Processing**	**02**

Course Outcome: -

- Introduce the concepts of signal processing terms and relate them toimage processing
- Learn about basic image processing techniques (e.g., noise removal andimage enhancement).
- Develop skills to design and implement algorithms for advanced imageanalysis
- Apply image processing to design solutions to real-life problems

Course Specific Outcome: -

- Understanding the terminologies of signal and digital image processing
- Ability to apply various images, intensity transformations, and spatialfiltering.
- Knowledge of Perform frequency domain operations on images.
- Ability to apply image segmentation and extract image features.Apply image processing algorithms in practical applications.

UNIT 1: Fundamentals of Digital Signals Processing Periodic signals, Spectral decomposition, Signals, Reading and writing Waves, Spectrums, Wave objects, Signal objects Noise: Uncorrelated noise, Integrated spectrum, Brownian noise, Pink Noise, Gaussian noise; Autocorrelation: Correlation, Serial correlation, Autocorrelation, Autocorrelation of periodic signals, Correlation as a dot product Frequency domain Operations: Representing Image as Signals, Sampling and Fourier Transforms, Discrete Fourier Transform, Convolution and Frequency Domain Filtering,	15L
Smoothing using low- pass filters, Sharpening using high-pass filters. Fast Fourier Transforms.	
UNIT 2: Image Processing fundamentals and Pixel- Transformation Definition, Application of Image Processing, Image Processing Pipeline, Tools and Libraries for Image Processing, Image types and files formats. Intensity Transformations- Log Transform, Power-law Transform, Contrast Stretching, Thresholding Histogram Processing- Histogram Equalization and Histogram Matching; Linear and Non-linear smoothing of Images, Sharpening of images Image Derivative: Derivatives and gradients, Laplacian, the effect of noise on gradient computation	15L

TEXT BOOK:

1. Digital Image Processing by Rafael Gonzalez & Richard Woods, Pearson; 4th edition, 2018
2. Think DSP: Digital Signal Processing in Python by Allen Downey, O'Reilly Media; 1st edition (August 16, 2016)

REFERENCE BOOKS:

1. Understanding Digital Image Processing, Vipin Tyagi, CRC Press, 2018
2. Digital Signal and Image Processing by Tamal Bose, John Wiley 2010
3. Hands-On Image Processing with Python by Sandipan Dey,Packt Publishing, 2018
4. Fundamentals of Digital Images Processing by A K Jain, Pearson, 2010

CourseCode	Course Title	Credits
GNKPSCS2P501	**Practical Course on Applied Signal and Image Processing**	**01**

Note: All the Practical's should be implemented using Python
Link:https://www.python.org/downloads/

1	Write program to demonstrate the following aspects of signal processing on suitable data 1. Upsampling and downsampling on Image/speech signal 2. Fast Fourier Transform to compute DFT
2	Write program to perform the following on signal 1. Create a triangle signal and plot a 3-period segment. 2. For a given signal, plot the segment and compute the correlation between them.

3	Write program to demonstrate the following aspects of signal on sound/image data 1. Convolution operation 2. Template Matching
4	Write program to implement point/pixel intensity transformations such as 1. Log and Power-law transformations 2. Contrast adjustments 3. Histogram equalization 4. Thresholding, and halftoning operations
5	Write a program to apply various enhancements on images using image derivatives by implementing Gradient and Laplacian operations.
6	Write a program to implement linear and nonlinear noise smoothing on suitable image or sound signal.
7	Write a program to apply various image enhancement using image derivatives by implementing smoothing, sharpening, and unsharp masking filters for generating suitable images for specific application requirements.
8	Write a program to Apply edge detection techniques such as Sobel and Canny to extract meaningful information from the given image samples

Course Code	Course Title	Credits
GNKPSCS2501	**Advanced Database Techniques**	**02**

Course Outcome: -

- To cover advanced topics of databases to become more proficient.
- To provide students with theoretical knowledge and practical skills in advanced topics in database systems, big data and modern data-intensivesystems.
- To Expand Students, view and introduce advanced topics and Business Intelligence.

Course Specific Outcome: -

- To form professional competencies related to design and implementation of non-relational databases, including object-oriented, parallel and Distributed.
- Learners will be able to explore XML, and Mobile databases.
- Learners will be able to deal with methods used for dealing with spatial and Temporal Databases.Learner will have a solid grasp on business intelligence tools and XML.

Unit-1 Object–Oriented Databases: Need of Object-oriented databases, Complex Data Types, Structured Types and Inheritance, Object- Identity and Reference, ODL and OQL, Implementing O-R Features, Persistent Programming Languages, Object-Oriented versus Object- Relational, 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. Temporal Databases: Time ontology, structure, and granularity,Temporal data models, Temporal relational algebra.	15L
Unit-2 Introduction to NoSQL: Characteristics of NoSQL, NoSQL Storage types, Advantages and Drawbacks, NoSQL Products Performing CRUD operations: Creating Records, Accessing Data,Updating and Deleting Data from MongoDB, Redis,HBase and Apache Cassandra	15L

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,	

TEXT BOOK:

1. Database Management Systems by Raghu Ramakrishnan and JohannesGehrke, McGraw Hill, 3rd Edition, 2014
2. Professional NoSQL By Shashank Tiwari, Wrox-John Wiley & Sons, Inc,2011
3. Getting Started with NoSQL, Gaurav Vaish, Packt Publishing Ltd, 2013

REFERENCE BOOKS:

1. Advanced Database Management System by byRini Chakrabarti and Shilbhadra Dasgupta, Dreamtech Press, 2017
2. SQL & NoSQL Databases, Andreas Meier · Michael Kaufmann, SpringerVieweg, 2019
3. Parallel and Distributed Systems by Arun Kulkarni, Nupur Prasad Giri,Wiley, Second edition, 2017
4. Practical Hadoop Migration: How to Integrate Your RDBMS with the Hadoop Ecosystem and Re-Architect Relational Applications to NoSQL ByBhushan Lakhe, Apress; 1st edition, 2016.

Course Code	Course Title	Credits
GNKPSCS2P501	**Practical Course on Advanced Database Techniques**	**01**

Note: All the Practical's should be implemented using NoSQL
Link: https://www.oracle.com/database/technologies/nosql-database-server-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 anddelete operations on these tables. Issue queries on it.
3	Demonstrate distributed databases environment by dividing given global database 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	Demonstrate the indexing and ordering operations in 1. MongoDB 2. CouchDB 3. Apache Cassandra

Course Code	Course Title	Credits
	Principles of Compiler Design	**02**
Course Outcome: -Understand the theoretical foundations and concepts underlying the design and implementation of compilers. **Course Specific Outcome**: - Acquire knowledge about the different phases of the compilation process, lexical analyzers and parsers, code generation and Implement optimization techniques.		
Unit-1: Front end of Compiler: Introduction to Compiler Design: Role and importance of compilers, Phases of compilation process Lexical Analysis: Role of lexical analyzer, Regular expressions and finite automata, Lexical analyzer generators (e.g., Lex) Syntax Analysis: Role of parser, Context-free grammars, Top-down parsing (LL parsing) Bottom-up parsing (LR parsing)		15L
Unit-2: Back end of Compiler Code Optimization: Data flow analysis, Common subexpression elimination, Constant folding and propagation, Loop optimization techniques Code Generation: Code generation techniques Runtime Environments: Activation records and stack management. Heap memory management, Call and return mechanisms, Exception handling Lexical and Syntax Error Handling: Error recovery strategies Error reporting and Handling The DAG representation of basic blocks, Dominators, Reducible flow graphs, Depth-first search, Loop-invariant computations, Induction variable elimination, Some other loop optimizations.		15L
TEXT BOOK: Compilers: Principles, Techniques, and Tools" by Alfred V. Aho, Monica S. Lam, Ravi Sethi, and Jeffrey D. Ullman 2nd Edition, Pearson Publication, 2006 ISBN-13: 978- 0321486813		
REFERENCE BOOKS: Modern Compiler Implementation in C" by Andrew W. Appel, 3rd Edition, Cambridge University Press, 2020, ISBN-13: 978-1108426631 Principles of Compiler Design" by D. M. Dhamdhere, 2nd Edition Publisher: McGraw-Hill Education, 2017, ISBN-13: 978-9339204608		

Course Code	Course Title	Credits

GNKPSCS3A501	**Elective Paper -Data science technologies - I**	**03**

Course Outcome: -

- students should:Have a solid understanding of big data concepts and the Hadoop ecosystem.
- Be proficient in using Hadoop commands and managing Hadoop clusters.
- Be able to develop Map-Reduce programs and run them in different modes.
- Understand data processing with HIVE, SQOOP, and PIG, and be able to use them effectively.
- Possess knowledge of NoSQL databases and their suitability for various data scenarios.
- Be able to design and implement data workflows using Oozie.
- Gain practical skills through hands-on illustrations and exercises.
- Be well-prepared to work with big data technologies and contribute to data processing projects in real-world scenarios.

Course-Specific Outcome:

Upon successful completion of this course, students will have acquired the knowledge and skills needed to:

- Effectively work with Hadoop and its ecosystem components, including HIVE, SQOOP, PIG, and Oozie.
- Develop Map-Reduce programs for data processing.
- Manage and manipulate large datasets using Hadoop technologies.
- Integrate data from various sources into Hadoop clusters.
- Choose appropriate database solutions (SQL and NoSQL) based on data requirements.
- Design and implement data workflows for efficient data processing.

Unit-1 Big data and Hadoop: Hadoop architecture, Single node & Multi-node Hadoop, Hadoop commands, Hadoop daemon, Task instance, Hadoop ecosystem and its installation, Illustrations.	15L
Unit-2 Framework, Developing Map-Reduce program, Life cycle method, Serialization, Running Map-Reduce in local and pseudo- distributed mode, Illustrations.	15L

Unit-3 HIVE: Data types and commands, Illustrations. SQOOP: Importing data, exporting data, Running, Illustrations PIG: Schema, Commands, Illustrations NoSQL database: Features, Types, NoSQL vs. SQL, Advantages and Disadvantages Oozie: What is Oozie? Workflow, Oozie workflow application, Features.	15L
TEXT BOOK: 1. Hadoop The Definitive Guide : Tom White , 4th Edition, 2017 2. Data-intensive Text Processing with Map Reduce : Jimmy Lin and Chris Dyer, Morgan & Claypool Publishers, 2010	
REFERENCE BOOKS: 1. Hadoop in Action : Chuck Lam, 2010	

CourseCode	Course Title	Credits
GNKPSCS3AP501	**Practical Course on Elective Paper -Data science technologies - I**	**01**
1	Multi-node Hadoop eco-system configuration with HDFS	
2	Improving Map-Reduce performance using combiners	
3	Creating map reduce jobs	
4	Creating Map-Reduce Programs in local and pseudo-distributed mode	
5	Working of HADOOP with HIVE	
6	Working of HADOOP with SQOOP	
7	Working of HADOOP with PIGs	
8	Working of HADOOP with NoSQL	

Course Code	Course Title	Credits
GNKPSCS3B501	**Elective Paper - Data warehousing and Data mining**	**03**

Course Outcome:

Upon completion of the Data Warehousing and Data Mining course, students should:

- Have a solid understanding of data warehousing concepts, architecture, and implementation.
- Be able to design, implement, and manage data warehouses effectively.
- Possess the knowledge and skills to preprocess and prepare data for data mining.
- Be proficient in using data mining algorithms and tools to discover patterns and insights in data.
- Be able to evaluate and interpret the results of data mining models.
- Understand the ethical and legal considerations associated with data mining.
- Be prepared to apply data warehousing and data mining techniques to real-world problems and projects.
- Be aware of the latest trends and developments in the field of data warehousing and data mining.
- Collaborate in teams to solve complex data-related challenges using data warehousing and data mining techniques.

Upon successful completion of the Data Warehousing and Data Mining course, students will:

Course-Specific Outcome:

Be capable of designing, implementing, and managing data warehousing solutions.
Possess the skills to preprocess and integrate data from diverse sources into a data warehouse.
Have the ability to apply data mining algorithms to discover patterns and insights in data.
Be skilled in evaluating and interpreting data mining results in practical contexts.
Be proficient in using data mining tools and libraries for real-world applications.

Unit-1 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, Discretization: Equal Width Binning; Equal Depth Binning, Normalization, Smoothing.	15L
Unit-2 Definition of Data warehouse, Logical architecture of Data Warehouse, Data Warehouse model- Enterprise warehouse; Data Marts, Populating business Data Warehousing: data integration and extract, transform, load (ETL). OLTP and OLAP systems, Data cubes, Data cube operations, data cube schemas.	15L

Unit-3 Data mining definitions and process: business and data understanding. Association Analysis: Definition of association rule, General issues: Support; Confidence; Lift; Conviction Definition of association rule, General issues: Support; Confidence; Lift; Conviction. Frequent Item sets: APriori Algorithm; Issues with APriori Algorithm, Data structures:	
Hash tree and FP tree.	

TEXT BOOK:

1. Building the Data Warehouse, Inmon: Wiley (1993).
2. Data Mining: Introductory and Advanced Topics, Dunham, Margaret H, Prentice Hall (2006)
3. Data Mining: Practical Machine Learning Tools and Techniques, Second Edition, Witten, Ian and Eibe Frank, Morgan Kaufmann (2011)

REFERENCE BOOKS:

1. Data Modeling Techniques for Data Warehousing by IBM; International Technical Support Organization, Chuck Ballard, Dirk Herreman, Don Schau, Rhonda Bell, Eunsaeng Kim, Ann Valencic :http://www.redbooks.ibm.com
2. Data Mining: Concepts and Techniques, The Morgan Kaufmann Series in Data Management Systems, Han J. and Kamber M. Morgan Kaufmann Publishers, (2000).
3. Data Mining with Microsoft SQL Server 2008, MacLennan Jamie, Tang ZhaoHui and Crivat Bogdan, Wiley India Edition (2009).

Course Code	Course Title	Credits
GNKPSCS3BP501	**Practical Course on Elective Paper - Data warehousing and Data mining**	**01**

Note:

1	Create tables using different applications.
2	Develop an application to design a warehouse by importing various tables from external sources
3	Develop an application to design a warehouse by importing various tables from external sources
4	Develop an application to create dimension tables in a cube and form star schema.
5	Create association rules by considering suitable parameters
6	Create association rules by considering suitable parameters
7	Develop an application to pre process data imported from external sources.
8	Create association rules by considering suitable parameters

Course Code	Course Title	Credits
GNKPSCS4501	**Research Methodology**	**04**

Course Outcome:

Upon successful completion of the Research Methodology course, students will:

- Possess a strong foundation in research methodology, including research design and data collection techniques.
- Understand the ethical principles and guidelines governing research.
- Be able to identify research problems, develop research questions, and formulate hypotheses.
- Have knowledge of both quantitative and qualitative research approaches.
- Be proficient in data analysis techniques, including statistical analysis and qualitative data analysis.
- Be capable of writing comprehensive research proposals and research reports**.**

Course-Specific Outcome:

- Students will have: Developed a heightened awareness of research principles and practices.
- Acquired practical skills in various research methods and techniques.
- Demonstrated ethical research conduct and a commitment to research integrity.
- Gained the ability to formulate research questions and design research projects effectively.
- Mastered data collection and analysis techniques, both quantitative and qualitative.
- Acquired proficiency in research proposal and report writing, with proper citation and referencing.

Unit-1 — 15L

1.1 : Research Fundamentals and Terminology

Meaning and Objective of research, features of a good research study, types of Research (qualitative and quantitative research)

Study designs and variations: basic, applied, historical, exploratory, experimental, ex-post-facto, case study, diagnostic research, crossover design, case control design, cohort study design, multifactorial design.

1.2 : Literature Survey Methods

Journal and abbreviation, current tites and review, monographs, textbooks, introduction to abstract, Beilstein, subject and author index

Digital: Web sources, E-journals, Journal access, TOC alerts, Hot articles, Citation Index, Impact factor, H-index, E- consortium, UGC infonet, E-books, Internet discussion groups and communities, Blogs, preprint servers, Search engines, Sciurus, Google Scholar, Wiki-databases, Science Direct, SciFinder, Scopus.

Unit-2 **2.1 : Research writing** Scientific writing- Reporting practical and project work, writing literature surveys and reviews, organizing a poster display, giving an oral presentation. Writing Scientific Papers: Justification for scientific contributions, bibliography, description of methods, conclusions, the need for illustration, style, publications of scientific work. Project Proposal and research funding agencies, Research grants, scholarships and funding (CSIR, DBT, DST, DST- INSPIRE Fellowship, ICMR, INSA, BRNS, MoEFCC, UGC- RFSMS, Fulbright Fellowships for Indian students, Lady Tata Memorial Trust, EPA, Bill and Melinda Gates Foundation, Wellcome Trust, Erasmus Mundus) **2.2 : Publication Ethics and Bibliography** Publication ethics : definition, introduction Best practices/ Standards settings initiative and guidelines COPE, WAME Conflict of interest Publication Misconduct: definition, concept, problems that lead to unethical behavior Violation of publication ethics, authorship and contributorship, Identification of publication misconduct, Predatory publisher and journals	15L

Use of reference management software (MS Word / Zotero / Mendeley)	
Unit 3 **Research Design** Choosing appropriate research methods, Experimental, observational, case study, survey, etc., Sampling techniques and sample size determination, Validity and reliability in research **Data Collection Methods** Surveys and questionnaires, Interviews: structured, semi-structured, unstructured, Observations and participant observations, Ethnographic studies and fieldwork, Data Collection Tools and Technologies **Online survey platforms** Data collection software, Sensor-based data collection, Wearable devices and Internet of Things (IoT) for data collection, Data Analysis Techniques **Quantitative data analysis using statistical tools (e.g., SPSS, R)** Qualitative data analysis: coding, thematic analysis, content analysis, Mixed-methods data analysis	15
Unit 4 **Research Presentation and Communication** Writing research papers and technical reports, Creating effective research presentations, Peer review process and responding to feedback, Presenting research findings at conferences, Research Presentation and Communication **Writing research papers and technical reports** Creating effective research presentations, Peer review process and responding to feedback Presenting research findings at conferences	15

Detailed Syllabus for Semester – II

Course Code	Course Title	Credits
GNKPSCS1502	**Applied Machine and Deep Learning**	**02**

Course Outcome: -

- Developing projects in machine learning for industrial applications.
- Understanding and implementing algorithms and techniques of Machine Learning useful in the field of Data Science, Image Processing, NLP, etc.

Course Specific Outcome:

- Understand core concepts of ML through implementations in python.
- Working with diverse toolkits and packages useful for developing projectsin ML

Implement and understand deep learning and ANNs useful for industrytoday.

Unit-1 **The Fundamentals of Machine Learning** What is Machine Learning? Why use Machine Learning? Types of Machine Learning, Supervised Learning, Unsupervised Learning & Reinforcement Learning. Challenges of Machine Learning, Testing and Validation. A First Application: Classification, MNIST Dataset, Performance Measures, Confusion Matrix, Precision and Recall, Precision/Recall Tradeoff, The ROC Curve, Multiclass Classification. **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.	15L
Unit-2 **Support Vector Machines** Linear SVM Classification, Soft Margin Classification, Nonlinear SVMClassification, 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. **Fundamentals of Deep Learning** What is Deep Learning? Need Deep Learning? Introduction to Artificial Neural Network (ANN), Core components of neural networks, Multi-Layer Perceptron (MLP), Activation functions, Sigmoid, Rectified Linear Unit (ReLU), Introduction to Tensors and Operations, TensorFlow framework	15L

TEXT BOOK:

1. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems byAurélienGéron, Second Edition, O'reilly 2019
2. Deep Learning with Python by François Chollet Published by Manning 2018

Reinforcement Learning: An Introduction by Richard S. Sutton and Andrew G. Barto, Second Edition 2014

REFERENCE BOOKS:

1. Introduction to Machine with Python - A Guide for Data Scientists byAndreas C. Müller & Sarah Guido O'reilly 2016
2. Artificial Neural Networks with TensorFlow 2 ANN Architecture MachineLearning Projects Poornachandra Sarang by Apress 2021

CourseCode	Course Title	Credits
GNKPSCS1P502	**Practical Course on Applied Machine and Deep Learning**	**01**

Note: All the Practical's should be implemented using Python andTensorFlow.
Link:Python :https://www.python.org/downloads/
TensorFlow :https://www.tensorflow.org/install

1	a.Implement Linear Regression (Diabetes Dataset) b.Implement Logistic Regression (Iris Dataset)
2	a.Implements Multinomial Logistic Regression (Iris Dataset) b.Implement SVM classifier (Iris Dataset)
3	Train and fine-tune a Decision Tree for the Moons Dataset
4	Train an SVM regressor on the California Housing Dataset
5	Implement MLP for classification of handwritten digits (MNISTDataset)
6	Classification of images of clothing using Tensorflow (FashionMNIST dataset)
7	Implement Regression to predict fuel efficiency using Tensorflow(Auto MPG dataset)
8	Implement Batch GradientSoftmax Regression

Course Code	Course Title	Credits
GNKPSCS1502	**Natural Language Processing**	**02**

Course Outcome: -

- Understanding the importance and concepts of Natural Language Processing (NLP)
- Applying algorithms available for the processing of linguistic informationand computational properties of natural languages.
- Knowledge on various morphological, syntactic, and semantic NLPtasks.
- Introducing various NLP software libraries and data sets publicly available.

Course Specific Outcome: -

- The ability to describe the concepts of morphology, syntax, semantics, discourse & pragmatics of natural language
- Discover various linguistic and statistical features relevant to the basic NLP task, namely, spelling correction, morphological analysis, parts-of- speech tagging, parsing, and semantic analysis
- Assess and Evaluate NLP based systems
- Ability to choose appropriate solutions for solving typical NLP subproblems

(tokenizing, tagging, parsing)

Analyse NLP problems to decompose them inadequate independent components and develop real-life applications

Unit-1 Introduction to NLP: Introduction and applications, NLP phases, Difficulty of NLP including ambiguity Language Modelling: N-gram and Neural Language Models Language Modelling with N-gram, Simple N-gram models, smoothing (basic techniques) Computational morphology & Parts-of-speech Tagging:basic concepts; Tagset; Lemmatization, Early approaches: Rule-based andTBL; POS tagging using HMM; Introduction to POS Tagging using Neural Model.	15L
Unit-2 Parsing Basic concepts: top-down and bottom-up parsing, treebank; Syntactic parsing: CKY parsing; Statistical Parsing basics: Probabilistic Context-Free Grammar (PCFG);	15L
Probabilistic CKY Parsing of PCFGs. Introduction to WordNet Intelligent Work Processors: Machine Translation; User Interfaces; man-machine Interfaces: Natural language Querying Tutoring and Authoring Systems. Speech Recognition Commercial use of NLP: NLP in customer Service, Sentiment Analysis, Emotion Mining, Handling Frauds and SMS, Bots, LSTM & BERT models, Conversations	

TEXT BOOK:

1. Speech and Language Processing‖, Jurafsky Dan and Martin James H.,3rd Edition, Pearson, 2018.
2. Natural Language Processing with Python‖, Steven Bird, Ewan Klein, andEdward Loper, 2nd Edition, O'Reilly, 2016.

REFERENCE BOOKS:

1. Practical NaturalLanguage Processing with Python‖, Mathangi Sri, Apress, 2021
2. "Handbook of Computational Linguistics and Natural Language Processing‖, Martin Whitehead, Clanrye International, 2020
3. Handbook of Natural Language Processing‖, Nitin Indurkhya, and Fred J. Damerau, Pearson; 2nd edition, 2008

CourseCode	Course Title	Credits
GNKPSCS1P502	**Practical Course on Natural Language Processing**	**01**

Note: - The following set of practicals can be performed using any PythonLibraries for NLP such as NLTK, spaCy, genism:
Link:-https://www.python.org/downloads/

1	Write a program to implement sentence segmentation and wordtokenization
2	Write a program to Implement stemming and lemmatization
3	Write a program to Implement a tri-gram model
4	Write a program to Implement PoS tagging using HMM & NeuralModel
5	Write a program to Implement syntactic parsing of a given text
6	Write a program to Implement dependency parsing of a given text
7	Write a program to Implement Named Entity Recognition (NER)
8	Write a program to Implement Text Summarization for the givensample text

Course Code	Course Title	Credits
GNKPSCS2502	**Web Mining**	**02**

Course Outcome: -

- To understand the difference between Web Mining and Data mining.
- To Understand the Basics and Needs of Web Mining.
- To Understand Web-based Data.
- To Understand Opinion Mining and Sentiment classification.

Course Specific Outcome:

- Develop deep understanding of mining techniques exclusively for the Internet
- Understand and develop analytics for social media data.
- Design and implementation of various web analytical tool to understand complex unstructured data on the Internet for aiding individuals and Businesses to grow their business

Unit-1 Introduction to Web Mining: Web Mining, Data Mining, Basic Concepts, Difference. Basic Concepts of Information Retrieval, Information Retrieval Models, Relevance feedback, Evaluation measures Text and Web Page Preprocessing, Web Search, Web Spamming, Web Information Retrieval, Feature based Opinion Mining and Summarization, Web Usage Mining	15L
Unit-2 Social Network Analysis, PageRank, HITS, Community Discovery, Basic Crawler Algorithm, Implementation Issues, Universal Crawlers, Focused Crawlers, Topical Crawlers, Crawler Ethics and Conflicts, Data modelling and webpage usage mining, Discovery and analysisof web usage patterns.	15L

TEXT BOOK:

1. Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data byBing Liu (Springer Publications) 2017 publication

REFERENCE BOOKS:

1. Data Mining: Concepts and Techniques, Second Edition Jiawei Han,Micheline Kamber (Elsevier Publications),2017
2. Web Mining: Applications and Techniques by Anthony Scime,2010
3. Mining the Web: Discovering Knowledge from Hypertext Data by SoumenChakrabarti 2010

CourseCode	Course Title	Credits

GNKPSCS2P502	**Practical Course on Web Mining**	**01**
Note: - The following set of practical's should be implemented in Scrape,python: Link:-Python : https://www.python.org/downloads/		
1	Scrape an online E-Commerce Site for Data. Extract product data from Amazon - be it any product and put these details in the MySQL database. One can use pipeline. Like 1 pipeline to process the scraped data and other to put data in the database and since Amazon has some restrictions on scraping of data, ask them to work onsmall set of requests otherwise proxies and all would haveto be used	
2	Page Rank for link analysis using python Create a small set of pages namely page1, page2, page3 andpage4 apply random walk on the same	
3	Develop a basic crawler for the web search for user definedkeywords.	
4	Demonstrate Text Mining and Webpage Pre-processing usingmeta information from the web pages (Local/Online).	
5	Perform Spam Classifier.	
6	Write a code to remove stop words from given sentence.	
7	Write a code to perform Stemming on given sentence.	
8	Write a code to perform Chunking on given sentence.	

Course Code	**Course Title**	**Credits**
GNKPSCS2502	**Embedded and IoT Technology**	**02**

Course Outcome: -

- The course is designed to enable students, to understand and implementIoT in industry.
- Design and executive projects in IoT with Automatic Identification and Data Capture

Course Specific Outcome: -

- Understand basic components and functionalities of Embedded System including its hardware.
- Effectively achieve collaboration of various technologies in IoT and enablethe same using software programming like Python, Embedded C etc.Understand case studies in IoT and replicate the same for more detailedanalysis of the IoT development.

Unit 1: Embedded System Basics and Basics of IoT Introduction to Embedded Systems, Design of Embedded Systems, Memory Architecture, Input/Output. Basic electronics: Semiconductors, Transistors, BJT, Flip Flops, Resistors, Capacitors, CMOS, MOSFET, FPGA, Relays. Microcontrollers. Introduction IoT: Evolution of the IoT concept, basic characteristics of IoT, distinguish the IoT from other related technologies, IoT architectures. IoT Building Blocks -Hardware and Software: The basic IoT building blocks, smart thing components and capabilities, basics of IoT gateway, Cloud, and analytics	15L
Unit 2: Advanced IoT Technologies IoT Gateway: IoT architecture domains, IoT gateway architecture, IoT gateway functionalities, IoT gateway selection criteria, IoT gateway and edge computing, edge computing-based solution for specific IoT applications IoT Cloud and Fog Computing: Components of IoT Cloud architecture, usage of application domains of IoT Cloud platforms, layered architecture of Fog computing. IoT Security: Security constraints in IoT systems. IoT attacks, security threats at each layer of IoT architecture. Social IoT: Nature of social relationships among IoT Devices.	15L

TEXT BOOK:

1. Introduction to Embedded Systems – Cyber physical systems ApproachEdward Ashford Lee &SanjitArunkumarSeshia Second Edition — MIT Press — 2017
2. Enabling the Internet of Things Fundamentals, Design and Applications byMuhammad Azhar Iqbal, Sajjad Hussain, Huanlai Xing, Muhammad Ali Imran Wiley Pub.1st Edition 2021

REFERENCE BOOKS:

1. Introduction Embedded Systems by K.V. Shibu Second Edition McGrawHills–2017
2. Build your own IoT Platform Develop a Fully Flexible and Scalable Internet of Things Platform in 24 Hours by Anand Tamboli 2019 Apress

CourseCode	Course Title	Credits
GNKPSCS2P502	**Practical Course on Embedded and IoT Technology**	**01**

Note: - The following set of practical's should be implemented inCodeVisionAVR, Proteus8, Cisco Packet Tracer, Keli V5, PythonLink: -
Python:https://www.python.org/downloads/ CodeVisionAVR :https://www.codevision.be/ Proteus8:https://www.labcenter.com/downloads/
Cisco Packet Tracer:https://www.netacad.com/courses/packet-tracerKeli V5: https://www.keil.com/download/

1	Design and implement basics embedded circuits. 1. Automatic Alarm system- Alarm should get trigger bysenor 2. Timer based buzzer 3. Sensor based Counting device
2	Demonstrate communication between two embeddeddevices using UART port.
3	Built an IoT system to send ticket before entering thebus.
4	Develop an IoT application for Motion detection.
5	Develop an IoT application that will raise an alarmwhenever with going to rain outside based on the weather prediction data.
6	Develop an IoT application for monitoring water levelsin tanks and automatically start the motor to fill the tank if the level goes below the critical level.
7	Develop an IoT module to which measure the intensityof light and send the same to your PC/ Phone.
8	Develop a IoT application which will record the movement and orientation of your phone and give thedata back to the PC

Course Code	Course Title	Credits
	Web Data Analytics	**02**

Course Outcome: -

- Understand the concepts and techniques of web mining, including sequential pattern mining and rule generation.
- Learn about opinion mining and sentiment classification in web information retrieval.
- Explore social network analysis, link analysis, and the implementation of webpage crawlers.

Course Specific Outcome: -

- Gain knowledge of information retrieval models, text preprocessing, and web search techniques.
- Understand web usage mining, including the discovery and analysis of web usage patterns, and the use of recommender systems and query log mining.

Unit 1: Introduction to Web Mining Web Mining-Data Mining, Basic Concepts, Difference, Mining Sequential Patterns on Prefix Span, Generating Rules from Sequential Patterns. Basic Concepts of Information Retrieval, Information Retrieval Models, Relevance feedback, Evaluation measures Text and Web Page Preprocessing, Inverted Index and Its Compression, latent semantic indexing, Web Search, Web Spamming **Opinion Mining and Web Usage Mining:** Web Information Retrieval, Sentiment Classification, Feature based Opinion Mining and summarization, Comparative Sentence and Relation Mining, Opinion Search and Opinion Spam. Web Usage Mining.	15L
Unit 2:Social Network & Link Analysis Social Network-Link Analysis, Scrapy using python (without pipelining), Social Network Analysis, Co-Citation and Bibliographic Coupling, PageRank, HITS, Community Discovery **Webpage crawlers and usage mining:** Basic Crawler Algorithm, Implementation Issues, Universal Crawlers, Focused Crawlers, Topical Crawlers, Crawler Ethics and Conflicts, Data modelling and webpage usage mining., Discovery and analysis of web usage patterns, Recommender systems and collaborative filtering, query log mining	15L

TEXT BOOK: 1. Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data by Bing Liu (Springer Publications) 2017 publication 2. Data Mining: Concepts and Techniques, Second Edition Jiawei Han, MichelineKamber (Elsevier Publications),2017	
REFERENCE BOOKS: 1. Web Mining: Applications and Techniques by Anthony Scime,2010 2. Mining the Web: Discovering Knowledge from Hypertext Data by SoumenChakrabarti 2010	

Course Code	**Course Title**	**Credits**
GNKPSCS3A502	**Elective Paper -Data science technologies – II**	**03**

Course Outcome:
Upon completing the "Data Science Technologies – II" course, students will:
Possess a strong foundation in clustering and classification techniques, enabling them to segment data and make predictive models.
Be proficient in Apache Spark, capable of processing large datasets, and developing Spark applications for various data processing tasks.
Master the Scala programming language, empowering them to work with data and implement complex data processing tasks.
Understand the features and installation process of Apache Mahout and be able to implement basic recommendation systems.

Unit-1 Clustering: exploring distance measure, data representation, clustering algorithms in Mahout, evaluating and improving clustering quality Classification: training a classifier, evaluating and tuning a classifier, deploying a Classifier.	15L
Unit-2 Spark: Features, Architecture, Components of Spark, Resilient Distributed Datasets – data structure of Spark, working with Key/Value pairs, Loading and Saving data, Core Programming - RDD Transformations, Actions. Executing a Spark Application. Scala: Features, Basic Syntax, Data types, Variables, Classes and objects, Access modifiers, operators, if construct, loop statement, functions, OOP concepts, Array, String, Exceptions, Collections, File Handling, Multithreading	15L

Unit-3 Mahout: Features, Installation, Recommendations: Introducing recommenders, representing recommender data, making recommendations.	15L
TEXT BOOK: 1. Learning Spark: Holden Karau, Andy Konwinski, Patrick Wendell, and MateiZaharia, OReilly 2. Programming in Scala: Martin Odersky, Lex Spoon, BillVenners,2008	
REFERENCE BOOKS: 1. Mahout in Action: Sean Owen, Robin Anil, Ted Dunning, EllenFriedman,2012	

CourseCode	Course Title	Credits
GNKPSCS3AP502	**Practical Course on Elective Paper -Data science technologies – II**	**01**
Note:		
1	Spark Installation	
2	Demonstrate RDD Transformation.	
3	Demonstrate RDD Action.	
4	Scala Program to demonstrate strings and arrays.	
5	Scala Program to demonstrate collections.	
6	Scala Program to demonstrate file handling.	
7	Implement clustering algorithms supported.	
8	Implement classification algorithms.	

Course Code	Course Title	Credits
GNKPSCS3B502	**Elective Paper - Data analysis and Data mining**	**03**

Course Outcome:

This elective paper aims to provide students with a comprehensive understanding of data analysis and data mining techniques. Students will learn how to process, analyze, and extract valuable insights from large datasets using various methods and tools. The course will cover both theoretical concepts and practical applications.

Course Specific outcome:

These course outcomes are designed to provide students with a well-rounded understanding of data analysis and data mining principles, along with practical experience in using various tools and techniques to extract meaningful insights from diverse datasets.

Unit-1 Big data: Introduction to Big data Platform, Traits of big data, Challenges of conventional systems, Web data, Analytic processes and tools Analytic processes and tools, Analysis vs Reporting, Modern data analytic tools. Statistical concepts: Sampling distributions, Resampling, Statistical Inference, Prediction error. Data Analysis: Regression modeling, Analysis of time Series: Linear systems analysis, Nonlinear dynamics, Rule induction.	15L
Unit-2 Neural networks: Learning and Generalization, Competitive Learning, Principal Component Analysis and Neural Networks, Fuzzy Logic: Extracting Fuzzy Models from Data, Fuzzy Decision Trees, Stochastic Search Methods. Collaborative filtering as a similar-sets problem, Documents, k-Shingles, Choosing the Shingle Size, Hashing Shingles, Shingles built from Words	15L

Unit-3 Similarity-Preserving Summaries of Sets, Locality- Sensitive hashing for documents. The Theory of Locality-Sensitive functions. Methods for high degrees of similarity Introduction to streams concepts – Stream data model and architecture, Stream computing Sampling data in a stream, Filtering streams, Counting distinct elements in a stream,	15L

Estimating moments. Counting oneness in a Window, Decaying window, Real time analytics Platform (RTAP).	
TEXT BOOK: 1. Mining of Massive Datasets, Anand Rajaraman and Jeffrey David Ullman, Cambridge University Press, 2012. 2. Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, Michael Minelli, Wiley, 2013	
REFERENCE BOOKS: 1. Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph, David Loshin, Morgan Kaufmann Publishers, 2013	

CourseCode	**Course Title**	**Credits**
GNKPSCS3BP502	**Practical Course on Elective Paper - Data warehousing and Data mining - II**	**01**
Note:		
1	Counting oneness in a Window, Decaying window, Real time analytics Platform (RTAP).	
2	Generate forecasting model and interpret the result for a given data set.	
3	Write a program to construct different types of k 8 shingles for given document.	
4	Write a program for measuring similarity among documents and detecting passages which have been reused.	
5	Write a program to compute the n 10 moment for a given stream where n is given	
6	Create a stream data model.	
7	Sampling streams.	
8	Estimating moments	

Course Code	**Course Title**	**Credits**
GNKPSCS4502	**On Job Training**	**04**
	A learner has to show an on job training of 60 hours with proper documentation. -Joining letter -Job training letter	