Rayat Shikshan Sanstha's

Karmaveer Bhaurao Patil College Vashi, Navi Mumbai

Autonomous College

[University of Mumbai]

Syllabus for Approval

Sr. No.	Heading	Particulars
1	Title of Course	M.Sc. II Computer Science
2	Eligibility for Admission	M.Sc. I
3	Passing Marks	40%
4	Ordinances/Regulations (if any)	
5	No. of Years/Semesters	One year/Two semester
6	Level	P.G.
7	Pattern	Semester
8	Status	Revised
9	To be implemented from Academic year	2022-23

AC - 7.6
Item No-

RayatShikshanSanstha's
KARMAVEER BHAURAO PATIL COLLEGE, VASHI.
NAVI MUMBAI
(AUTONOMOUS COLLEGE)
Sector-15- A, Vashi, Navi Mumbai - 400 703

Syllabus for M.Sc. II Computer Science

Program: M.Sc. Computer Science

Course: M.Sc. II Computer Science

(Choice Based Credit, Grading and Semester System with effect from the academic year 2022-23)

Rayat Shikshan Sanstha's
KARMAVEER BHAURAO PATIL COLLEGE, VASHI.
NAVI MUMBAI (Autonomous)
Department of Computer Science
M. Sc. Computer Science

Program Outcomes (POs)

Learners are able to–

PO-1	**Disciplinary Knowledge**	Acquire the comprehensive and in-depth knowledge of various subjects in sciences such as Physics, Chemistry, Mathematics, Microbiology, Bio-analytical Science, Computer Science, Data Science, Information Technology and disciplinary skills and ability to apply these skills in the field of science, technology and its allied branches
PO-2	**Communication and Presentation Skills**	Develop various communication skills including presentation to express ideas evidently to achieve common goals of the organization.
PO-3	**Creativity and Critical Judgment**	Facilitate solutions to current issues based on investigations, evaluation and justification using evidence based approach.
PO-4	**Analytical Reasoning and Problem Solving**	Build critical and analytical attitude in handling the problems and situations.
PO-5	**Sense of Inquiry**	Curiously raise relevant questions based on highly developed ideas, scientific theories and its applications including research.
PO-6	**Use of Digital Technologies**	Use various digital technologies to explore information/data for business, scientific research and related purposes.
PO-7	**Research Skills**	Construct, collect, investigate, evaluate and interpret information/data relevant to science and technology to adapt, evolve and shape the future.
PO-8	**Application of Knowledge**	Develop a scientific outlook to create consciousness against the social myths and blind faith.
PO-9	**Moral and Ethical Reasoning**	Imbibe ethical, moral and social values to develop virtues such as justice, generosity and charity as beneficial to individuals and society at large.
PO-10	**Leadership and Teamwork**	Work cooperatively and lead proactively to achieve the goals of the organization by implementing the plans and projects in various field-based situations related to science, technology and society at large.
PO-11	**Environment and Sustainability**	Create social awareness about the environment and develop sustainability for betterment of the future.

PO-12	**Lifelong Learning**	Realize that pursuit of knowledge is a lifelong activity and in combination with determined efforts, positive attitude and other qualities to lead a successful life.

Preamble

As mentioned in the syllabus of semester I and II, the intended philosophy of the new syllabus is to meet following guidelines:

- Give a strong foundation on core Computer Science subjects.
- Expose students to emerging trends in a gradual and incremental way.
- Prepare the student community for the demands of the ICT industry.
- Offer specialization in a chosen area.
- Create research temper among students in the whole process.

This syllabus for the semester - III and semester – IV has tried to continue the steps initiated in the semester- I and semester –II to meet the goals set. This proposes two core compulsory subjects in semester III. The student has to continue with the tracks they have taken in semester II as elective subjects. The syllabus also includes project proposals as part of the practical course in elective subjects.

The semester – IV will have one compulsory subject. Students can choose one subject as specialization out of the two electives he or she has been pursuing since the semester – II. That means, there will be four specializations in the semester IV as mentioned below:

- Cloud Computing
- Cyber and Information Security
- Business Intelligence and Big Data Analytics
- Machine Learning
- Skill Based Courses

The syllabus also offers an internship and project implementation in the semester – IV, each of which has weights equivalent to a full course. By introducing different electives as tracks in

semester –II, espousing more of that tracks in the semester –III and offer students the added advantage of high level competency in the advanced and emerging areas of computer science. This will definitely equip the student with industry readiness as an internship in an IT or IT-related organization gives a practical exposure to what is learned and what is practiced. The strong foundation given in the core courses in different semesters will give enough confidence to the learner to face and adapt to the changing trends and requirements of industry and academia.

As one can easily notice, the syllabus offers lots of emphasis on student driven learning and learning through experience. Research is embedded in the course structure. By introducing Researching Computing in semester – I, Case study in semester – II, Project Proposal in semester – III and Project Implementation in semester – IV (which together has a weightage equivalent to almost two theory courses), the syllabus prepares a strong army of budding computer science researchers. The syllabus designed by the firm believes that by focusing on student driven research on cutting edge and emerging trends with lots of practical experience will make the learning more interesting and stimulating. It is hoped that the student community and teacher colleagues will appreciate the thrust, direction and treatment given in the syllabus.

Thanks to one and all who have directly or indirectly helped in this venture.

Scheme of examination for Each Semester:

Continuous Internal Evaluation: 40 Marks (Common Test-20 Marks & 20 Marks for- Mini Projects, Presentation, Online Course, Case Study, Assignment, Analysis In Statistics, Report Writing, Interviews, Paper Review, Surprise Test, Research Paper, Data Analysis).

Semester End Examination: 60 Marks will be as follows-

I.	**Theory:** The Semester End Examination for theory course work will be conducted as per the following scheme.	
	Each theory paper shall be of two and half hour duration.	
	All questions are compulsory and will have internal options.	
	Q – I	From Unit – I (having internal options.) 12 M
	Q – II	From Unit – II (having internal options.)12 M
	Q – III	From Unit – III (having internal options.)12 M
	Q – IV	From Unit – IV (having internal options.)12 M
	Q – V	Questions from all the FOUR Units with equal weightage of marks allotted to each Unit. 12 M
II.	Practical	The Semester End Examination for practical course will be conducted as per the following scheme.

Sr. No.	**Particulars of Semester End Practical Examination**	**Marks**
1	Laboratory Work	40
2	Journal	05
3	Viva	05
	TOTAL	**50**

M.Sc. Computer Science Syllabus
Choice Based Credit, Grading and Semester System
Academic year 2022-2023
SEMESTER - III

CODE	COURSE TYPE	SUBJECT	SCHEME OF INSTRUCTION (PERIOD PER WEEK)		SCHEME OF EXAMINATION (MAX MARKS)			NO. OF CREDITS
			TH	LAB	CIA	SEE	TOTAL	
PGCS301	CORE	Business Intelligence and Big Data Analytics –II	4		40	60	100	4
PGCSP301	Core Subject Practical	PGCS301		4			50	2
PGCS302	CORE	Distributed Computing	4		40	60	100	4
PGCSP302	Core Subject Practical	PGCS302		4			50	2
PGCS303	CORE	Natural Language Processing	4		40	60	100	4
PGCSP303	Core Subject Practical	PGCS303		4			50	2
PGCS304	Skill Enhancement Course	Data Visualization using Tableau	2				50	2
PGCSP304	Skill Enhancement Subject Practical	PGCS304		4			50	2
PGCS305A	Discipline Specific Elective - I	Cyber and Information Security- II (Cyber Forensics)	4		40	60	100	4
PGCSP305A	Discipline Specific Elective - I Subject Practical **OR**	PGCS305A		4			50	2
PGCS305B	Discipline Specific Elective - I	Machine Learning –II (Advanced Machine Learning)	4		40	60	100	4
PGCSP305B	Discipline Specific Elective - I Subject Practical	PGCS305B		4			50	2
						TOTAL	750	28

SEMESTER - IV

CODE	COURSE TYPE	SUBJECT	SCHEME OF INSTRUCTION (PERIOD PER WEEK)		SCHEME OF EXAMINATION (MAX MARKS)			NO. OF CREDITS
			TH	LAB	CIA	SEE	TOTAL	
PGCS401	CORE	Business Intelligence and Big Data Analytics –III	4	-	40	60	100	4
PGCSP401	Core Subject Practical	PGCS401	-	4			50	2
PGCS402	CORE	Internship	-	-				
PGCSP402		Internship	-	8			150	6
PGCS403	CORE	Project	-	-				
PGCSP403		Project	-	8			150	6
PGCS404	Skill Enhancement Course	MOOCs or GoLang	2	-			50	2
PGCSP404		MOOCs or GoLang	-	4			50	2
PGCS405A	Discipline Specific Elective - I	Cyber and Information Security- II (Cryptography and Crypt Analysis)	4	-	40	60	100	4
PGCSP405A	Discipline Specific Elective - I Subject Practical **OR**	PGCS405A	-	4			50	2
PGCS405B	Discipline Specific Elective - I	Machine Learning –III (Computational Intelligence)	4	-	40	60	100	4
PGCSP405B	Discipline Specific Elective - I Subject Practical	PGCS405B	-	4			50	2
						TOTAL	750	28

Note: TH-Theory, CIA- Continuous Internal Assessment, SEE-Semester End Examination.

Detailed syllabus of Semester – III

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** III	
Subject : Business Intelligence and Big Data Analytics –II			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS301 Business Intelligence and Big Data Analytics –II

Course Outcomes: After successful completion of this course, students will be able to:

CO1:Describe the Big Data, Statistical concepts, Data Analysis,neural network and fuzzy logic.
CO2: Illustrate the various algorithms using mapreduce.
CO3: Explain shingling of documents using various applications such as jaccard's similarity and methods of high degree of similarity,locality sensitive hashing.
CO4:Summarize the stream concepts, decaying windows, Real time analytics platform(RTAP)

ICT Tools Used: Videos, PPT,

Students Centric Methods: Problem Solving and Participative
(Experimental, Participative, Problem Solving)

Links: SWAYAM / MOOCS:
1) https://www.coursera.org/courses?query=big%20data%20analytics

The CO-PO Mapping Matrix

CO\PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO1	-	-	-	2	1	-	1	-	-	-	-	-
CO2	-	-	1	1	-	2	-	-	-	-	-	-
CO3	-	-	-	1	1	-	2	-	-	-	-	-
CO4	-	1	1	1	-	1	-	-	-	-	-	-

Course: PGCS301	Business Intelligence and Big Data Analytics –II **(Credits : 4 Lectures/Week: 4)**	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Describe the Big Data, Statistical concepts, Data Analysis,neural network and fuzzy logic. 2. Illustrate the various algorithms using mapreduce. 3. Explain shingling of documents using various applications such as jaccard's similarity and methods of high degree of similarity,locality sensitive hashing. 4. Summarize the stream concepts, decaying windows, Real time analytics platform(RTAP)	
Unit I:	**Introduction To Big Data** Big data: Introduction to Big data Platform, Traits of big data, Challenges of conventional systems, Web data, Analytic processes and tools, Analysis vs Reporting, Modern data analytic tools, Statistical concepts: Sampling distributions, Re-sampling, Statistical Inference, Prediction error. Data Analysis: Regression modeling, Analysis of time Series: Linear systems analysis, Nonlinear dynamics, Rule induction, 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.	**15 L**
Unit II:	**MAP REDUCE** Introduction to Map Reduce: The map tasks, Grouping by key, The reduce tasks, Combiners, Details of MapReduce Execution, Coping with node failures. Algorithms Using MapReduce: Matrix-Vector Multiplication, Computing Selections and Projections, Union, Intersection, and Difference, Natural Join. Extensions to MapReduce: Workflow Systems, Recursive extensions to MapReduce, Common map reduce algorithms.	**15 L**
Unit III:	**SHINGLING OF DOCUMENTS** Finding Similar Items, Applications of Near-Neighbor Search, Jaccard similarity of sets, Similarity of documents, Collaborative filtering as a similar-sets problem, Documents, k-Shingles, Choosing the Shingle Size, Hashing Shingles, Shingles built from Words. Similarity-Preserving Summaries of Sets, Locality-Sensitive hashing for documents. The Theory of Locality-Sensitive functions. Methods for high degrees of similarity.	**15 L**
Unit IV:	**MINING DATA STREAMS** Introduction to streams concepts – Stream data model and architecture, Stream computing, Sampling data in a stream, Filtering streams, Counting distinct elements in a stream, Estimating moments, Counting oneness in a Window, Decaying window, Real time analytics Platform(RTAP).	**15 L**

Text book:

- Mining of Massive Datasets, AnandRajaraman and Jeffrey David Ullman, Cambridge University Press, 2012.
- Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, Michael Minelli, Wiley, 2013

References:

- Big Data for Dummies, J. Hurwitz, et al., Wiley, 2013
- Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data, Paul C. Zikopoulos, Chris Eaton, Dirk deRoos, Thomas Deutsch, George Lapis, McGraw-Hill, 2012.
- Big data: The next frontier for innovation, competition, and productivity, James Manyika ,Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Byers, McKinsey Global Institute May 2011.
- Big Data Glossary, Pete Warden, O'Reilly, 2011.
- Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph, David Loshin, Morgan Kaufmann Publishers, 2013

Links:

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

Sr. No.	List of Practical Experiments on PGCSP301
1	Generate regression model and interpret the result for a given data set.
2	Generate forecasting model and interpret the result for a given data set.
3	Write a map-reduce program to count the number of occurrences of each alphabetic character in the given dataset. The count for each letter should be case-insensitive (i.e., include both upper-case and lower-case versions of the letter; Ignore non-alphabetic characters).
4	Write a map-reduce program to count the number of occurrences of each word in the given dataset. (A word is defined as any string of alphabetic characters appearing between non-alphabetic characters like nature's is two words. The count should be case-insensitive. If a word occurs multiple times in a line, all should be counted)
5	Write a map-reduce program to determine the average ratings of movies. The input consists of a series of lines, each containing a movie number, user number, rating and a timestamp.
6	Write a map-reduce program: (i) to find matrix-vector multiplication; (ii) to compute selections and projections; (iii) to find union, intersection, difference, natural Join for a given dataset.
7	Write a program to construct different types of k-shingles for given document.
8	Write a program for measuring similarity among documents and detecting passages which have been reused.
9	Write a program to compute the n- moment for a given stream where n is given.
10	Write a program to demonstrate the Alon-Matias-Szegedy Algorithm for second moments.
Note: The experiments may be done using software/tools like Hadoop / WEKA / R /Java etc.	

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** III	
Subject : Distributed Cloud Computing			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS302 Distributed Cloud Computing

Course Outcomes: After successful completion of this course, students will be able to:

CO1: Describe distributed system and cloud models
CO2: Elaborate distributed computation model on large datasets using parallel
CO3: Illustrate distributed programming approaches over cloud platforms.
CO4:Analyze the synchronization process

ICT Tools Used: Videos, PPT, Open Source Databases (Scopus/ Google Scholar)

Students Centric Methods: Videos, PPT, Open Source Cloud service provider

Links: SWAYAM / MOOCS:
1)https://class.coursera.org/pgm/lecture/preview

The CO-PO Mapping Matrix

CO\PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	1	-	-	-	1	-	-	-	-	-	-	-
CO2	1	-	-	-	-	-	-	-	-	-	-	-
CO3	-	-	-	1	2	-	-	-	-	-	-	-
CO4	-	-	-	-	-	1	2	-	-	-	-	-

Course: PGCS302	**Distributed System and Cloud Computing** **(Credits : 4 Lectures/Week: 4)**	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1) Explain distributed system and cloud models 2) Explain distributed computation model on large datasets using parallel and distributed programming approaches over cloud platforms. 3) Analyze the synchronization process	
Unit I:	**Fundamental:**Introduction, Distributed Computing Models, Software 4 Concepts, Issues in designing Distributed System, Client – Server Model **Communication:**Message Passing , Introduction to Message Passing, Advantages and features of Message Passing, Message Format, Message Buffering, Multi Data gram Messaging , Group Communication Remote Procedure Call (RPC): Basic RPC Operations, Parameter Passing, Extended RPC Models Remote Object Invocation: Distributed Objects, Binding a Client to an Object, Static Vs Dynamic RMI, Parameter Passing, Java RMI Message Oriented Communication: Persistence and synchronicity in communication, Message Oriented Transient and Persistent Communications	**15 L**
Unit II:	**Processes:**Threads, Code Migration: Approaches to Code Migration, 4 Migration and Local Resources, Migration in Heterogeneous Systems **Synchronization:**Clock Synchronization, Physical and Logical Clocks, Global State, Election Algorithms, Mutual Exclusion, Distributed Transactions, Deadlocks	**15 L**
Unit III:	**Distributed Technologies and Frameworks :** Overview of EJB S/W Architecture, view of EJB Conversation, Building and Deploying EJB, Roles in EJB, Types of Enterprise Beans, Lifecycle of Beans , Developing Applications using EJB Framework. Introduction to CORBA, CORBA Components and 4 architecture, Method Invocation, Static and Dynamic Invocation in CORBA, CORBA IDL, Developing Application using CORBA.Comparison of RMI, CORBA, EJB.	**15 L**
Unit IV:	**Introduction** – Essentials – Benefits – Why cloud – Business and IT perspective – cloud and virtualization – cloud service requirements – dynamic cloud infrastructure – cloud computing characteristics – cloud adoption – cloud rudiments. Cloud deployment models: introduction – cloud characteristics – measured service accounting – cloud deployment models – security in a public cloud – public versus private clouds – cloud infrastructure self-service.	**15 L**

Text book:

- Sunita Mahajan, Seema Shah, " Distributed Computing", Oxford, second edition.
- Andrew S. Tanenbaum & Maarten van Steen " Distributed Systems : Principles and paradigms" Prentice Hall of India Private Limited
- G. Sudha Sadasivam, Radha Shankarmani, "Middleware and Enterprise Integration Technologies " , Wiley Precise Textbook .

References:

- Pradeep K. Sinha "Distributed Operating Systems", Prentice Hall of India Private Limited
- 2. Thomas Erl "Service Oriented Architecture : Concepts, Technology and Design" Prentice Hall
- 3. G. Coulouris, J. Dollimore and T. Kindberg "Distributed Systems :

Links:
1) https://archive.mu.ac.in/myweb_test.

Sr. No.	List of Practical Experiments on PGCSP302
1	Client Server based program using RPC
2	Client Server based program using RMI
3	Implementation of Clock Synchronization (logical/physical)
4	Implementation of Election algorithm.
5	Implementation of Mutual Exclusion algorithms
6	Program multithreaded client/server processes..
7	Write a distributed application using EJB
8	Write a program using CORBA to demonstrate object brokering.
9	Designing of sample cloud services.
10	Creating an Amazon Account to store images.

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** III	
Subject : Natural Language Processing			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS303 Natural Language Processing

Course Outcomes: After successful completion of this course, students will be able to:

CO1: Describe the concepts of morphology, syntax, semantics, discourse & pragmatics of natural language

CO2: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

CO3:Introduce various NLP software libraries and data sets publicly available

CO4:Analyze NLP problems to decompose them inadequate independent components and develop real-life applications

ICT Tools Used: Videos, PPT, Open Source Cloud service provider

Students Centric Methods: Problem Solving and Participative
(Experimental, Participative, Problem Solving)

Links: SWAYAM / MOOCS:
1) https://www.coursera.org/specializations/natural-language-processing

The CO-PO Mapping Matrix

CO\PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO1	1	-	-	-	-	1	2	-	-	-	-	-
CO2	-	-	-	1	-	1	1	-	-	-	-	-
CO3	-	-	1	2	-	-	-	-	-	-	-	-
CO4	-	-	-	1	-	-	1	-	-	-	-	1

Course: PGCS303	Natural Language Processing **(Credits : 4 Lectures/Week: 4)**	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Describe the concepts of morphology, syntax, semantics, discourse & pragmatics of natural language 2. 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 3. Introduce various NLP software libraries and data sets publicly available 4. Analyze NLP problems to decompose them inadequate independent components and develop real-life applications	
Unit I:	**INTRODUCTION** Introduction to NLP, Regular Expressions, Words, Corpora, Text Normalization, Minimum Edit distance, N gram Language Models, Evaluating Language Models	**15 L**
Unit II:	**SYNTACTIC ANALYSIS** English Word Classes, The Penn Treebank Part-of-Speech Tagset, Part-of-Speech Tagging, HMM Part- of-Speech Tagging, Maximum Entropy Markov Models, Grammar Rules for English, Treebanks,Grammar Equivalence and Normal form, Lexicalized Grammar.	**15 L**
Unit III:	**SEMANTIC ANALYSIS** Representation of Sentence Meaning: Computational Desiderata for Representations, Model-Theoretic Semantics, First-Order Logic, Event and State Representations, Description Logics,Semantic roles, Semantic role labeling.	**15 L**
Unit IV:	**SEQUENCE PARSING WITH RECURRENT NETWORKS** Simple Recurrent Networks, Applications of RNNs, Deep Networks: Stacked and Bidirectional RNNs, Managing Context in RNNs: LSTMs and GRUs, Words, Characters and Byte-Pairs. **Case Study** Sentiment Classification, Dialog Systems and Chatbots	**15 L**

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, and Edward Loper, 2nd Edition, O'Reilly, 2016

References:

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
4. Foundations of Statistical Natural Language Processing‖, Manning, Christopher and Heinrich, Schutze, MIT Press, 1997

Links:

- https://www.cs.vassar.edu/~cs366/docs/Manning_Schuetze_StatisticalNLP.pdf

- https://www.nltk.org/book/

Sr. No.	List of Practical Experiments on PGDSP303
1	a. Convert the text into tokens b. Find the word frequency c. Demonstrate a bigram language model d. Demonstrate a trigram language model e. Generate regular expression for a given text
2	a. Perform Lemmatization b. Perform Stemming c. Identify parts-of Speech using Penn Treebank tag set. d. Implement HMM for POS tagging e. Build a Chunker
3	a. Find the synonym of a word using WordNet b. Find the antonym of a word c. Implement semantic role labeling to identify named entities d. Resolve the ambiguity e. Translate the text using First-order logic
4	a. Implement RNN for sequence labeling b. Implement POS tagging using LSTM c. Implement Named Entity Recognizer d. Word sense disambiguation by LSTM/GRU
5	a. Develop a Movie review system b. Create a chatbot for HITS.

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** III	
Subject : **Data Visualization in Tableau**			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS303 Data Visualization using Tableau

Course Outcomes: After successful completion of this course, students will be able to:

CO1: Understand types of data and data visualization methods
CO2: Understand the need of data visualization.
CO3: Create and design visualizations and dashboards
CO4: Evaluate the performance of visualization technique
CO5: Apply data visualization using open source tool Tableau

ICT Tools Used: Videos, PPT, Open Source Cloud service provider

Students Centric Methods: Problem Solving and Participative
(Experimental, Participative, Problem Solving)

Links: SWAYAM / MOOCS:
1)https://www.mygreatlearning.com/academy/learn-for-free/courses/data-visualization-using-tableau

The CO-PO Mapping Matrix

CO\PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	1	-	-	-	1	-	-	-	-	-	-	-
CO2	1	-	-	-	-	-	-	-	-	-	-	-
CO3	-	-	-	1	2	-	-	-	-	-	-	-
CO4	-	-	-	-	-	1	2	-	-	-	-	-
CO5	-	-	-	-	-	2	-	-	-	-	-	-

Course: PGDS304	Data Visualization in Tableau (Credits : 2 Lectures/Week: 2)	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Understand types of data and data visualization methods 2. Understand the need of data visualization. 3. Create and design visualizations and dashboards 4. Evaluate the performance of visualization technique 5. Apply data visualization using open source tool Tableau	

Unit I:	**Introduction to Data Visualization:** Need for data visualization. Visualization an aspect of business analytics, importance of data visualization. Types of Data, Stages of Data visualization, Fitts Law, Human visual perception and cognition Comparison between tableau and Power BI **Installation of Tableau:** Using the Workspace Control Effectively,Tableau Desktop Workspace,Navigation, Data Terminology, Data Types & Default Properties,different aggregation types, File Types **Data Connection:** Data Connection with Text File,Connection with Microsoft Excel, Extracting data,data joining, data blending, sorting and replacing data source.	**15 L**
Unit II:	**TableauCalculation:**Operations,functions,string,Numeric,date, table calculation,LOD Expressions **Filter data:**Basic filter,filter operation, Extract filter,Quick Filters,Context filter,conditional filters,data source filters, Sort data, Build Groups, Hierarchy and sets **Tableau Charts and Graphs:** Bar chart,Line Graphs with Date & Without Date,Pie Chart,Tree maps, Scatter Plots, Individual Axes, Blended Axes, Dual Axes & Combination chart, Edit axis,Bins/Histograms,Parts of Views, Sorting, Trend lines, Reference Lines, Forecasting,View data & Actions (across sheets), latitude and longitude, Default location/Edit locations, Symbol Map & Filled Map	**15 L**
Unit III:	**Building Interactive Dashboards**: (Building & Customizing) Combining multiple visualizations into a dashboard (overview), Making your worksheet interactive by using actions Filter URL, Highlight, Options in Formatting your Visualization, Working with Labels and Annotations, Effective Use of Titles and Captions	**15 L**

Text Books:

- Learning Tableau 10 – Second Edition, by Joshua Milligan
- Practical Tableau by Ryan Sleeper
- Communicating Data with Tableau by Ben Jones
- Mastering Tableau by David Baldwin

References:

- Big data black book, Dream tech publication
- Handbook for visualizing : a handbook for data driven design by Andy krik

Links:

Sr. No.	List of Practical Experiments on PGDSP304
1	Installation of Tableau and working with Tableau Desktop Workspace
2	Data Connection with various File
3	Import and manage data (join, relationship, replace)
4	Perform various Tableau Calculation

5	Adding filters and quick filters to dashboards
6	Implement Bar charts, Pie chart, Line chart, Multiple chart and distribution
7	Implement Highlight tables, Scatterplot, Trendline
8	Implement Heatmap, Geographic mapping, Impressive bar chart, bullet graph
9	Implement Gantt chart, data calendar, circle view, general operation
10	Building Interactive Dashboards

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** III	
Subject : Cyber and Information Security- II (Cyber Forensics)			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS305A Cyber and Information Security- II (Cyber Forensics)

Course Outcomes: After successful completion of this course, students will be able to:

CO1: Define the objectives of computer forensics in law enforcement, evidence, case studies, investigation, computer forensic techniques, Vulnerabilities and Computer forensics Technologies.

CO2:Apply the various techniques of data recovery, data hiding, evidence collection rules, Verification & Authentication.

CO3:Analyze Network based evidence, Principles of internetworking, protocols, various acquisition methods, NIDS & NIPS Systems.

CO4: Determine the concepts of mobile forensics using identification and data interception using web proxies and evidence analysis .

ICT Tools Used: Videos, PPT, Open Source Databases (Scopus/ Google Scholar)

Students Centric Methods: Problem Solving and Participative
(Experimental, Participative, Problem Solving)

Links: SWAYAM / MOOCS:

1) https://www.coursera.org/specializations/computer-network-security#about
2) https://www.coursera.org/courses?query=network%20security

The CO-PO Mapping Matrix												
CO\PO	**PO 1**	**PO 2**	**PO 3**	**PO 4**	**PO 5**	**PO 6**	**PO 7**	**PO 8**	**PO 9**	**PO 10**	**PO 11**	**PO 12**
CO1	1	-	-	2	-	-	-	1	-	-	1	-
CO2	3	-	-	1	-	2	-	1	-	-	-	-
CO3	-	-	-	2	2	-	2	-	1	-	-	-
CO4	-	-	1	-	2	-	-	-	3	-	-	-

Course: PGCS305A	Cyber and Information Security- II (Cyber Forensics) **(Credits : 4 Lectures/Week: 4)**	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Define the objectives of computer forensics in law enforcement, evidence, case studies, investigation, computer forensic techniques, Vulnerabilities and Computer forensics Technologies. 2. Apply the various techniques of data recovery, data hiding, evidence collection rules, Verification & Authentication. 3. Analyze Network based evidence, Principles of internetworking, protocols, various acquisition methods, NIDS & NIPS Systems. 4. Determine the concepts of mobile forensics using identification and data interception using web proxies and evidence analysis .	
Unit I:	**Computer Forensic Fundamentals:** Introduction to Computer Forensics and objective, the Computer Forensics Specialist, Use of Computer Forensic in Law Enforcement, Users of Computer Forensic Evidence, Case Studies, Information Security Investigations. Types of Computer Forensics Technology: Types of Military Computer Forensic Technology, Types of Law Enforcement Computer Forensic Technology, Types of Business Computer Forensic Technology, Specialized Forensics Techniques, Hidden Data, Spyware and Adware, Encryption Methods and Vulnerabilities, Protecting Data from Being Compromised, Internet Tracing Methods, Security and Wireless Technologies. Types of Computer Forensics Systems: Study different Security System: Internet, Intrusion Detection, Firewall, Storage Area, Network Disaster Recovery, Public Key Infrastructure, Wireless Network, Satellite Encryption, Instant Messaging (IM), Net Privacy, Identity Management, Biometric, Identity Theft.	**15 L**
Unit II:	**Data Recovery:** Data Recovery and Backup, Role of Data Recovery, Hiding and Recovering Hidden Data. Evidence Collection: Need to Collect the Evidence, Types of Evidence, The Rules of Evidence, Collection Steps. Computer Image Verification and Authentication: Special Needs of Evidence Authentication. Identification of Data: Timekeeping, Forensic Identification and Analysis of Technical Surveillance Devices, Reconstructing Past Events: How to Become a	**15 L**

	Digital Detective, Useable File Formats, Unusable File Formats, Converting Files.	
Unit III:	**Network Forensics:** Sources of Network Based Evidence, Principles of Internetworking, Internet Protocol Suite. Evidence Acquisition: Physical Interception, Traffic Acquisition Software, Active Acquisition. Traffic Analysis: Protocol Analysis, Packet Analysis, Flow Analysis, Higher-Layer Traffic analysis. Statistical Flow Analysis: Sensors, Flow Record Export Protocols, Collection and Aggregation, Analysis. Wireless: the IEEE Layer 2 Protocol Series, Wireless Access Point, Wireless Traffic Capture and Analysis, Common Attacks, Locating Wireless Devices. Network Intrusion Detection and Analysis: NIDS/NIPS Functionality, Modes of Detection, Types of NIDS/NIPS, NIDS/NIPS Evidence Acquisition.	**15 L**
Unit IV:	**Network Devices and Mobile Phone Forensics:** Sources of Logs, Network Architecture, Collecting and Analyzing Evidence, switches, routers, firewalls, interfaces Web Proxies: Need to Investigate Web Proxies, Functionality, Evidence, Squid, Web Proxy Analysis, Encrypted Web Traffic. Mobile Phone Forensics: Crime and Mobile Phones, Voice, SMS and Identification of Data Interception in GSM, Mobile Phone Tricks, SMS Security, Mobile Forensic.	**15 L**

Text book:

- Computer Forensics Computer Crime Scene Investigation, John R. Vacca, Second Edition, 2005.
- Network Forensics, Sherri Davidoff, Jonathan HAM, Prentice Hall, 2012.
- Mobile Phone Security and Forensic: A Practical Approach, Second Edition, Iosif I. Androulidkis, Springer, 2012.

References:

- Digital forensics: Digital evidence in criminal investigation", Angus M.Marshall, John – Wiley and Sons, 2008.
- Computer Forensics with FTK, Fernando Carbone, PACKT Publishing, 2014.
- Practical Mobile Forensics, SatishBommisetty, RohitTamma, Heather Mahalik, PACKT Publishing, 2014.

Links:

1) http://sitlib.sethu.ac.in/e-books/CSE%20and%20IT%20books/Computer%20Forenscis%20John%20R.%20Vacca.pdf

2) https://books.google.co.in/books?id=hqjWCwAAQBAJ&printsec=frontcover&dq=Mobile+Phone+Security+and+Forensic:+A+Practical+Approach,+Second+Edition+pdf&hl=en&sa=X&ved=0ahUKEwiP3ZKCr4XdAhWafH0KHReYC0YQ6AEIKDAA#v=onepage&q&f=false

Sr. No.	List of Practical Experiments on PGCSP305A
1	Write a program to take backup of mysql database
2	Write a program to restore mysql database
3	Use DriveImage XML to image a hard drive
4	Write a program to create a log file
5	Write a program to find a file in a directory
6	Write a program to find a word in a file

7	Create forensic images of digital devices from volatile data such as memory using Imager for: (i) Computer System; (ii) Server; (iii) Mobile Device
8	Access and extract relevant information from Windows Registry for investigation process using Registry View, perform data analysis and bookmark the findings with respect to: (i) Computer System; (ii) Computer Network; (iii) Mobile Device; (iv) Wireless Network
9	Generate a report based on the analysis done using Registry View for different case scenario of the following: (i) Computer System; (ii) Computer Network; (iii) Mobile Device; (iv) Wireless Network
10	Create a new investigation case using Forensic Tool: (i) Computer System; (ii) Computer Network; (iii) Mobile Device ;(iv) Wireless Network.

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** III	
Subject : Machine Learning –II (Advanced Machine Learning)			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS305B Machine Learning –II (Advanced Machine Learning)

Course Outcomes: After successful completion of this course, students will be able to:

CO1: Discuss detailed knowledge on Basics of Probability, Conditional Probability and various distributions

CO2: Explain the concepts of Monte Carlo simulation inference in Machine learning

CO3: Summarize the distribution of scenarios using confidence intervals

CO4: Describe the various terms related to neural networks, such as activation, backpropagation and feedforward.

ICT Tools Used: Videos, PPT, Weka, Python

Students Centric Methods: Problem Solving and Participative (Experimental, Participative, Problem Solving)

Links: SWAYAM / MOOCS:

1) https://class.coursera.org/pgm/lecture/preview

2)https://www.coursera.org/lecture/excel-vba-for-creative-problem-solving-part-3-projects/introduction-to-monte-carlo-simulation-iozbm

The CO-PO Mapping Matrix

CO\PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO1	1	-	3	1	-	-	1	-	-	2	-	-
CO2	-	-	-	3	-	-	2	-	--	-	-	-
CO3	-	-	3	2	-	-	1	-	-	-	-	-
CO4	1	-	-	-	-	-	-	2	-	-	-	-

Course: PGCS305 B	Machine Learning –II (Advanced Machine Learning) **(Credits : 4 Lectures/Week: 4)**	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Discuss detailed knowledge on Basics of Probability, Conditional Probability and various distributions 2. Explain the concepts of Monte Carlo simulation inference in Machine learning 3. Summarize the distribution of scenarios using confidence intervals 4. Describe the various terms related to neural networks, such as activation, backpropagation and feedforward.	
Unit I:	**Probability** A brief review of probability theory, Some common discrete distributions, Some common continuous distributions, Joint probability distributions, Transformations of random variables, Monte Carlo approximation, Information theory. Directed graphical models (Bayes nets): Introduction, Examples, Inference, Learning, Conditional independence properties of DGMs. Mixture models and EM algorithm: Latent variable models, Mixture models, Parameter estimation for mixture models, The EM algorithm.	**15 L**
Unit II:	**Kernels** Introduction, kernel function, Using Kernel inside GLMs, kernel trick, Support vector machines, Comparison of discriminative kernel methods. Markov and hidden Markov models: Markov models, Hidden Markov Models (HMM), Inference in HMMs, Learning for HMMs. Undirected graphical models (Markov random fields): Conditional independence properties of UGMs, Parameterization of MRFs, Examples of MRFs, Learning, Conditional random fields (CRFs), applications of CRFs.	**15 L**

Unit III:	**Monte Carlo inference** Introduction, Sampling from standard distributions, Rejection sampling, Importance sampling, Particle filtering, Applications: visual object tracking, time series forecasting, Rao-Blackwellised Particle Filtering (RBPF). Markov chain Monte Carlo (MCMC) inference: Gibbs sampling, Metropolis Hastings algorithm, Speed and accuracy of MCMC.	**15 L**
Unit IV:	**Graphical model structure learning** Structure learning for knowledge discovery, Learning tree structures, Learning DAG structure with latent variables, Learning causal DAGs, Learning undirected Gaussian graphical models, Learning undirected discrete graphical models. Deep learning: Deep generative models, Deep neural networks, Applications of deep networks.	**15 L**

Text book:

- Machine Learning: A Probabilistic Perspective: Kevin P Murphy, The MIT Press Cambridge (2012).

References:

- Introducing Monte Carlo Methods with R, Christian P. Robert, George Casella, Springer, 2010
- Introduction to Machine Learning (Third Edition): EthemAlpaydın, The MIT Press (2015).
- Pattern Recognition and Machine Learning: Christopher M. Bishop, Springer (2006)
- Bayesian Reasoning and Machine Learning: David Barber, Cambridge University Press (2012).
- Statistical And Machine Learning Approaches For Network Analysis, Edited By
- Matthias Dehmer, Subhash C. Basak: John Wiley & Sons, Inc (2012)
- Practical Graph Mining with R: Edited by Nagiza-F-Samatova et al, CRC Press (2014)

Links:

1) https://class.coursera.org/pgm/lecture/preview
2) https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning_%20A%20Probabilistic%20Perspective%20%5BMurphy%202012-08-24%5D.pdf

Sr. No.	List of Practical Experiments on PGCSP305B
1	Find probability density function or probability mass function, cumulative distribution function and joint distribution function to calculate probabilities and quantiles for standard statistical distributions.
2	Create a Directed Acyclic Graph (DAG) using (i) set of formulae (ii) set of vectors and (iii) set of matrices. Find parents and children of nodes. Read conditional independence from DAG. Add and remove edges from graph.
3	Create a Bayesian network for a given narrative. Set findings and ask queries [One may use narratives like ‘chest clinic narrative’ and package gRain for the purpose].
4	Implement EM algorithm
5	Use string kernel to find the similarity of two amino acid sequence where similarity is defined as the number of a substring in common.
6	Demonstrate SVM as a binary classifier.
7	Create a random graph and find its page rank.
8	Apply random walk technique to a multivariate time series.

9	Implement two stage Gibbs Sampler.
10	Implement Metropolis Hastings algorithm.

Detailed syllabus of semester – IV

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** IV	
Subject : Business Intelligence and Big Data Analytics –III			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS401 Business Intelligence and Big Data Analytics –III
(Intelligent Data Analysis

Course Outcomes: After successful completion of this course, students will be able to:

CO1: Describe the distance based algorithms, KNN methods, trees, document classification, regression models and trees.

CO2: Determine the concept of Eigenvalue, Eigen vectors and decompositions.

CO3: Analyze the Chi Squared Automatic Interaction Detector.

CO4: Interpret the Evaluation techniques.

ICT Tools Used: Videos, PPT, Open Source Databases (Scopus/ Google Scholar)

Students Centric Methods: Problem Solving and Participative (Experimental, Participative, Problem Solving)
Links: SWAYAM / MOOCS: 1) https://www.coursera.org/learn/business-intelligence-data-analytics 2) https://www.udemy.com/course/the-business-intelligence-analyst-course-2018/

The CO-PO Mapping Matrix

CO\PO	PO 1	PO2	PO3	PO4	PO 5	PO 6	PO7	PO8	PO9	PO10	PO 11	PO 12
CO1	1	-	3	1	-	-	1	-	-	2	-	-
CO2	-	-	-	3	-	-	2	-	--	-	-	-
CO3	-	-	3	2	-	-	1	-	-	-	-	-
CO4	1	-	-	-	-	-	-	2	-	-	-	-

Course: PGCS401	Business Intelligence and Big Data Analytics –III **(Credits : 4 Lectures/Week: 4)**	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Describe the distance based algorithms, KNN methods, trees, document classification, regression models and trees. 2. Determine the concept of Eigenvalue, Eigen vectors and decompositions. 3. Analyze the Chi Squared Automatic Interaction Detector. 4. Interpret the Evaluation techniques.	
Unit I:	**Clustering** Distance/Similarity, Partitioning Algorithm: K-Means; K-Medoids, Partitioning Algorithm for large data set: CLARA; CLARANS, Hierarchical Algorithms: Agglomerative (AGNES); Divisive (DIANA), Density based clustering: DBSCAN, Clustering in Non- Euclidean Spaces, Clustering for Streams and Parallelism.	**15 L**
Unit II:	**Classification** Challenges, Distance based Algorithm: K nearest Neighbors and kD-Trees, Rules and Trees based Classifiers, Information gain theory, Statistical based classifiers: Bayesian classification, Document classification, Bayesian Networks. Introduction to Support Vector Machines, Evaluation: Confusion Matrix, Costs, Lift Curves, ROC Curves, Regression/model trees: CHAID (Chi Squared Automatic Interaction Detector). CART (Classification And Regression Tree).	**15 L**

Unit III:	**Dimensionality Reduction** Introduction to Eigen values and Eigen vectors of Symmetric Matrices, Principal-Component Analysis, Singular-Value Decomposition, CUR Decomposition.	**15 L**
Unit IV:	**Link Analysis And Recommendation Systems** Link analysis: PageRank, Efficient Computation of PageRank, Topic-Sensitive PageRank, Link Spam. Recommendation Systems: A Model for Recommendation Systems, Content-Based Recommendations, Collaborative Filtering, Dimensionality Reduction.	**15 L**

Text book:

- Mining of Massive Datasets, AnandRajaraman and Jeffrey David Ullman, Cambridge University Press, 2012.
- Data Mining: Introductory and Advanced Topics, Margaret H. Dunham, Pearson, 2013.

References:

- Big Data for Dummies, J. Hurwitz, et al., Wiley, 2013.
- Networks, Crowds, and Markets: Reasoning about a Highly Connected World, David Easley and Jon Kleinberg, Cambridge University Press, 2010.
- Lecture Notes in Data Mining, Berry, Browne, World Scientific, 2009.
- Data Mining: Concepts and Techniques third edition, Han and Kamber, Morgan Kaufmann, 2011.
- Data Mining Practical Machine Learning Tools and Techniques, Ian H. Witten, Eibe Frank, The Morgan Kaufmann Series in Data Management Systems, 2005.
- Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL and Graph, David Loshin, Morgan Kaufmann Publishers, 2013.

Links:

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

Sr. No.	List of Practical Experiments on PGCSP401
1	Pre-process the given data set and hence apply clustering techniques like K Means, K-Medoids. Interpret the result.
2	Pre-process the given data set and hence apply partition clustering algorithms. Interpret the result
3	Pre-process the given data set and hence apply hierarchical algorithms and density based clustering techniques. Interpret the result.
4	Pre-process the given data set and hence classify the resultant data set using tree classification techniques. Interpret the result.
5	Pre-process the given data set and hence classify the resultant data set using Statistical based classifiers. Interpret the result.
6	Pre-process the given data set and hence classify the resultant data set using support vector machine. Interpret the result.
7	Write a program to explain different functions of Principal Components.
8	Write a program to explain CUR Decomposition technique.
9	Write a program to explain links to establish higher-order relationships among entities in Link Analysis.

10	Write a program to implement step-by-step a Collaborative Filtering Recommender System.
The experiments may be done using software/ tools like R/Weka/Java etc.	

GUIDELINES FOR INTERNSHIP IN SEMESTER – IV (PGCS402)

- Internship should be of 2 to 3 months with 8 to 12 weeks duration.
- A student is expected to find internship by himself or herself. However, the institution should assist their students in getting internship in good organizations.
- The home institution cannot be taken as the place of internship.
- A student is expected to devote at least 300 hours physically at the organization.
- Internship can be on any topic covered in the syllabus mentioned in the syllabus, not restricted to the specialization.
- Internship can be done, in one of the following, but not restricted to, types of organizations:
    - Software development firms
    - Hardware/ manufacturing firms
    - Any small scale industries, service providers like banks
    - Clinics/ NGOs/professional institutions like that of CA, Advocate etc
    - Civic Depts like Ward office/post office/police station/ panchayat.
    - Research Centers/ University Depts/ College as research Assistant for research projects or similar capacities.

GUIDELINES FOR MAKING INTERNSHIP REPORT IN SEMESTER –IV (PGCS402)

A student is expected to make a report based on the internship he or she has done in an organization. It should contain the following:

- **Certificate:** A certificate in the prescribed Performa (given in appendix 1) from the organization where the internship was done.
- **Evaluation form:** The form filled by the supervisor or to whom the intern was reporting, in the prescribed Performa (given in appendix 2).

- **Title:** A suitable title giving the idea about what work the student has performed during the internship.
- **Description of the organization:** A small description of 1 to 2 pages on the organization where the student has interned
- **Description about the activities done by the section where the intern has worked**: A description of 2 to 4 pages about the section or cell of the organization where the intern actually worked. This should give an idea about the type of activity a new employee is expected to do in that section of the organization.
- **Description of work allotted and actually done by the intern:** A detailed description of the work allotted and actual work performed by the intern during the internship period. Intern may give a weekly report of the work by him or her if needed. It shall be of around 7 to 10 pages.
- **Self assessment:** A self assessment by the intern on what he or she has learnt during the internship period. It shall contain both technical as well as inter personal skills learned in the process. It shall be of around 2 to 3 pages.

The internship report may be around 15 pages and this needs to be submitted to the external examiner at the time of University examination.

GUIDELINES FOR RESEARCH IMPLEMENTATION IN SEMESTER – IV (PGCS403)

- A student is expected to devote at least 3 to 4 months of efforts for the implementation.
- Student should submit a detailed project implementation report at the time of viva.

GUIDELINES FOR DOCUMENTATION OF PROJECT PROPOSAL IN SEMESTER –IV (PGCS403)

A Student should submit project implementation report with following details:

- **Title:** Title of the project (Same as the one proposed and evaluated at the semester II examination).
- **Implementation details:** A description of how the project has been implemented. It shall be of 2 to 4 pages.
- **Experimental set up and results:** A detailed explanation on how experiments were conducted, what software used and the results obtained. Details like screen shots, tables and graphs can come here. It shall be of 6 to 10 pages.
- **Analysis of the results:** A description on what the results means and how they have been arrived at. Different performing measures or statistical tools used etc may be part of this. It shall be of 4 to 6 pages.
- **Conclusion:** A conclusion of the project performed in terms of its outcome (May be half a page).
- **Future enhancement:** A small description on what enhancement can be done when more time and resources are available (May be half a page).

- **Program code:** The program code may be given as appendix. The report may be of around 20 pages (excluding program code), which needs to be signed by the teacher in charge and head of the Department. Student should submit the signed project implementation report along with evaluated copy of the project proposal documentation at the time of Project evaluation and viva.

Appendix 1

(Proforma for the certificate for internship in official letter head)

This is to certify that Mr/Ms_____ of _____College/Institution worked as an intern as part of her MSc course in Computer Science of University of Mumbai. The particulars of internship are given below:

Internship starting date: _____

Internship ending date:_____

Actual number of days worked:_____

Tentative number of hours worked:_____ Hours

Broad area of work:_____

A small description of work done by the intern during the period:

_____

_____

_____

_____

_____

_____

Signature:

Name:

Designation:

Contact number:

Email:

(seal of the organization)

Appendix 2

(Proforma for the Evaluation of the intern by the supervisor/to whom the intern was reporting in the organization)

Professional Evaluation of intern

Name of intern:_____

College/institution:_____

[Note: Give a score in the 1–5 scale by putting √ in the respective cells]

Sr. No.	Particular	Excellent	Very Good	Good	Moderate	Satisfactory
1	Attendance					
2	Punctuality					
3	Adaptability					
4	Ability to shoulder responsibility					
5	Ability to work in a team					
6	Written and oral communication skills					
7	Problem solving skills					
8	Ability to grasp new concepts					

9	Ability to complete task					
10	Quality of work done					

Comments:

_____

_____

Signature:

Name:

Designation:

Contact number:

Email:

(seal of the organization)

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** IV	
Subject : Cyber and Information Security- II (Cryptography and Crypt Analysis)			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS405A Cyber and Information Security- II (Cryptography and CryptAnalysis)

Course Outcomes: After successful completion of this course, students will be able to:

CO1: Describe elementary number theory, algorithms,quadratic residues,reciprocity and theorems.

CO2: Summarize the cryptography as ciphers, algorithm modes, cryptographic hash functions.

CO3: Apply the concepts of RSA algorithms, public key cryptosystems, Diffie-Hellman Key Agreement and Knapsack problem.

CO4: Memorize key agreement and key agreement scheme, public key infrastructures and models privacy in cryptosystems & Trust model.
ICT Tools Used: Videos, PPT, Open Source Databases (Scopus/ Google Scholar)
Students Centric Methods: Problem Solving and Participative (Experimental, Participative, Problem Solving)
Links: SWAYAM / MOOCS: 1) https://onlinecourses.swayam2.ac.in/nou19_cs08/preview 2)https://nptel.ac.in/courses/106/106/106106178/

The CO-PO Mapping Matrix

CO\PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO1	1	-	3	1	-	-	1	-	-	2	-	-
CO2	-	-	-	3	-	-	2	-	--	-	-	-
CO3	-	-	3	2	-	-	1	-	-	-	-	-
CO4	1	-	-	-	-	-	-	2	-	-	-	-

Course: PGCS405 A	Cyber and Information Security- II (Cryptography and Crypt Analysis) **(Credits : 4 Lectures/Week: 4)**	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Describe elementary number theory, algorithms,quadratic residues,reciprocity and theorems. 2. Summarize the cryptography as ciphers, algorithm modes, cryptographic hash functions. 3. Apply the concepts of RSA algorithms, public key cryptosystems, Diffie-Hellman Key Agreement and Knapsack problem. 4. Memorize key agreement and key agreement scheme, public key infrastructures and models privacy in cryptosystems & Trust model.	
Unit I:	**Introduction to Number Theory** Topics in Elementary Number Theory: O and notations, time estimates for doing arithmetic-divisibility and the Euclidean algorithm, Congruence: Definitions and properties, linear congruence, residue classes, Euler's phi function, Fermat's Little Theorem, Chinese Reminder Theorem, Applications to factoring, finite fields, quadratic residues and reciprocity: Quadratic residues, Legendre symbol, Jacobi Symbol. (proofs of the theorems are not expected to cover).	**15 L**
Unit II:	**Simple Cryptosystems** Shift Cipher, Substitution Cipher, Affine Cipher, Vigenère Cipher, Vermin Cipher, Hill Cipher, Permutation Cipher, Stream Cipher, Cryptanalysis of Affine Cipher, Substitution Cipher, Vigenère Cipher and Hill Cipher, Block Ciphers, Algorithm	**15 L**

	Modes, DES, Double DES, Triple DES, Meet-in-Middle Attack, AES, IDEA algorithm. Cryptographic Hash Functions: Hash Functions and Data Integrity, Security of Hash Functions, Secure Hash Algorithm, Message Authentication Code, Nested MACs, HMAC.	
Unit III:	**RSA Cryptosystem** The RSA Algorithm, Primarily Testing, Legendre and Jacobi Symbols, The Solovay- Strassen Algorithm, The Miller-Rabin Algorithm, Factoring Algorithm: The pollard p-1 Algorithm, Dixon's Random Squares Algorithm, Attacks on RSA, The Rabin Cryptosystem. Public Key Cryptosystems: The idea of public key Cryptography, The Diffie-Hellman Key Agreement, ElGamal Cryptosystem, The Pollard Rho Discrete Logarithm Algorithm, Elliptic Curves, Knapsack problem.	**15 L**
Unit IV:	**Key Distribution and Key Agreement Scheme** Diffie-Hellman Key distribution and Key agreement scheme, Key Distribution Patterns, Mitchell-Piper Key distribution pattern, Station-to-station protocol, MTI Key Agreement scheme. Public-Key Infrastructure: What is PKI?, Secure Socket Layer, Certificates, Certificate Life cycle, Trust Models: Strict Hierarchy Model, Networked PKIs, The web browser Model, Pretty Good Privacy.	**15 L**

Text book:

- Discrete Mathematics and Its Applications, Kenneth H. Rosen, 7th Edition, McGraw Hill, 2012.
- Cryptography Theory and Practice, 3rd Edition, Douglas R. Stinson, 2005.

References:

- Network Security and Cryptography, AtulKahate, McGraw Hill, 2003.
- Cryptography and Network Security: Principles and Practices, William Stalling, Fourth Edition, Prentice Hall, 2013.
- Introduction to Cryptography with coding theory, second edition, Wade Trappe, Lawrence C. Washington, Pearson, 2005.

Links:

1) http://www.icst.pku.edu.cn/F/course/Cryptography/CryptographyTheoryandpractice(3ed).pdf
2) http://www2.fiit.stuba.sk/~kvasnicka/Mathematics%20for%20Informatics/Rosen_Discrete_Mathematics_and_Its_Applications_7th_Edition.pdf

Sr. No.	List of Practical Experiments on PGCS405A
1	Write a program to implement following: ● Chinese Reminder Theorem ● Fermat's Little Theorem
2	Write a program to implement the (i) Affine Cipher (ii) Rail Fence Technique (iii) Simple Columnar Technique (iv) Vermin Cipher (v) Hill Cipher to perform encryption and decryption.
3	Write a program to implement the (i) RSA Algorithm to perform encryption and decryption.
4	Write a program to implement the (i) Miller-Rabin Algorithm (ii) pollard p-1 Algorithm to perform encryption and decryption.
5	Write a program to implement the ElGamal Cryptosystem to generate keys and perform encryption and decryption.

6	Write a program to implement the Diffie-Hellman Key Agreement algorithm to generate symmetric keys.
7	Write a program to implement the MD5 algorithm compute the message digest.
8	Write a program to implement different processes of DES algorithm like (i) Initial Permutation process of DES algorithm, (ii) Generate Keys for DES algorithm, (iii) S-Box substitution for DES algorithm.
9	Write a program to encrypt and decrypt text using IDEA algorithm.
10	Write a program to implement HMAC signatures.

Class: M.Sc Part II	**Branch**: Computer Science	**Semester:** IV	
Subject : Machine Learning –III (Computational Intelligence)			
Period per Week(Each 48 min)	Lecture	04	
	Practical	04	
Evaluation System		**Hours**	**Marks**
	Semester End Examination	2 hrs 30 min	60
	Continuous Internal Assessment	—	40
	Semester End Practical Examination	2	50
	Total	—	150

PGCS405 Machine Learning –III (Computational Intelligence)

Course Outcomes: After successful completion of this course, students will be able to:

CO1: Discuss Artificial neural networks and reinforcement in machine learning.
CO2: Explain the concepts of genetics algorithms and working in machine learning.
CO3: Analyze the Particle Swarm Optimization(PSO).
CO4: Develop fuzzy logic systems for various applications.

ICT Tools Used: Videos, PPT, Python, Weka

Students Centric Methods: Problem Solving and Participative
(Experimental, Participative, Problem Solving)

Links: SWAYAM / MOOCS:
1) https://www.edx.org/learn/artificial-intelligence
2) https://www.udemy.com/course/fuzzy-logic/

The CO-PO Mapping Matrix

CO\PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO1	1	-	3	1	-	-	1	-	-	2	-	-

CO2	-	-	-	3	-	-	2	-	--	-	-	-
CO3	-	-	3	2	-	-	1	-	-	-	-	-
CO4	1	-	-	-	-	-	-	2	-	-	-	-

Course: PGCS405 B	Machine Learning –III (Computational Intelligence) **(Credits : 4 Lectures/Week: 4)**	
	Expected Course Outcomes : After successful completion of this course, students will be able to: 1. Discuss Artificial neural networks and reinforcement in machine learning. 2. Explain the concepts of genetics algorithms and working in machine learning. 3. Analyze the Particle Swarm Optimization(PSO). 4. Develop fuzzy logic systems for various applications.	
Unit I:	**Artificial Neural Networks** The Artificial Neuron, Supervised Learning Neural Networks, Unsupervised Learning Neural Networks, Radial Basis Function Networks, Reinforcement Learning, Performance Issues.	**15 L**
Unit II:	**Evolutionary Computation** Introduction to Evolutionary Computation, Genetic Algorithms, Genetic Programming, Evolutionary Programming, Evolution Strategies, Differential Evolution, Cultural Algorithms, Co-evolution.	**15 L**
Unit III:	**Computational Swarm Intelligence** Particle Swarm Optimization(PSO) - Basic Particle Swarm Optimization, Social Network Structures, Basic Variations and parameters, Single-Solution PSO. Advanced Topics and applications. Ant Algorithms- Ant Colony Optimization Meta-Heuristic, Cemetery Organization and Brood Care, Division of Labor, Advanced Topics and applications.	**15 L**
Unit IV:	**Artificial Immune systems, Fuzzy Systems and Rough Sets** Natural Immune System, Artificial Immune Models, Fuzzy Sets, Fuzzy Logic and Reasoning, Fuzzy Controllers, Rough Sets.	**15 L**
Text book: • Computational Intelligence- An Introduction (Second Edition): Andries P. Engelbrecht, John Willey & Sons Publications (2007).		
References: • Computational Intelligence And Feature Selection: Rough And Fuzzy Approaches, Richard Jensen QiangShen, IEEE Press Series On Computational Intelligence, A John Wiley & Sons, Inc., Publication, 2008. • Computational Intelligence And Pattern Analysis In Biological Informatics, (Editors). UjjwalMaulik, SanghamitraBandyopadhyay, Jason T. L.Wang, John Wiley & Sons, Inc, 2010.		

- Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition 1st Edition, SandhyaSamarasinghe, Auerbach Publications, 2006.
- Introduction to Evolutionary Computing (Natural Computing Series) 2nd ed, A.E. Eiben , James E Smith, Springer; 2015.
- Swarm Intelligence, 1st Edition, Russell C. Eberhart, Yuhui Shi, James Kennedy, Morgan Kaufmann, 2001
- Artificial Immune System: Applications in Computer Security, Ying Tan, Wiley- IEEE Computer Society, 2016.
- Computational Intelligence and Feature Selection: Rough and Fuzzy Approaches 1st Edition, Richard Jensen, QiangShen, Wiley-IEEE Press, 2008

Links:

1) https://leseprobe.buch.de/images-adb/9b/b8/9bb89e22-4a13-48ad-b9c5-5d341e2bcda4.pdf
2) http://www.shahed.ac.ir/stabaii/Files/CompIntelligenceBook.pdf

Sr. No.	List of Practical Experiments on PGCS405B
1	Implement feed forward neural network for a given data.
2	Implement Self Organizing Map neural network.
3	Implement Radial Basis Function neural network with gradient descent.
4	Implement a basic genetic algorithm with selection, mutation and crossover as genetic operators.
5	Implement evolution strategy algorithm.
6	Implement general differential evolution algorithm.
7	Implement gbest and lbest of PSO.
8	Implement simple Ant colony optimization algorithm.
9	Implement basic artificial immune system algorithm.
10	Apply different defuzzification methods for centroid calculation of a given fuzzy rule base.
Note: The above practical experiments may use programming languages like C, Java, R etc.	