



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
TRIGUNA SEN SCHOOL OF TECHNOLOGY
ASSAM UNIVERSITY: SILCHAR**

PhD COURSE WORK STRUCTURE & SYLLABUS

PhD Course Work structure: Total credits 16

Course No	Course Code	Course Title	Scheme of Studies per week			Credits
			L	T	P	
1	Paper 501	Research and Publication Ethics	2	0	0	2
2	Paper 502	Research Methodology	4	0	0	4
3	Paper 503(A)	Data structures and Algorithms	2	0	0	4
	Paper 503(B)	Electives	2	0	0	
4	Paper 504	Term Paper	0	0	12	6

Detailed Syllabus

Paper 501: Research and Publication Ethics

Course structure

- The course comprises of six modules listed in table below. Each module has 4-5 units.

Modules	Unit title	Teaching hours
Theory		
RPE 01	Philosophy and Ethics	4
RPE 02	Scientific Conduct	4
RPE 03	Publication Ethics	7
Practice		
RPE 04	Open Access Publishing	4
RPE 05	Publication Misconduct	4
RPE 06	Databases and Research Metrics	7
	Total	30

Syllabus in detail

THEORY

- RPE 01: PHILOSOPHY AND ETHICS (3 hrs.)**
 - Introduction to philosophy: definition, nature and scope, concept, branches
 - Ethics: definition, moral philosophy, nature of moral judgements and reactions
- RPE 02: SCIENTIFIC CONDUCT (5hrs.)**
 - Ethics with respect to science and research
 - Intellectual honesty and research integrity
 - Scientific misconducts: Falsification, Fabrication, and Plagiarism (FFP)
 - Redundant publications: duplicate and overlapping publications, salami slicing
 - Selective reporting and misrepresentation of data
- RPE 03: PUBLICATION ETHICS (7 hrs.)**
 - Publication ethics: definition, introduction and importance
 - Best practices / standards setting initiatives and guidelines: COPE, WAME, etc.
 - Conflicts of interest
 - Publication misconduct: definition, concept, problems that lead to unethical behavior and vice versa, types
 - Violation of publication ethics, authorship and contributorship
 - Identification of publication misconduct, complaints and appeals
 - Predatory publishers and journals

PRACTICE

- RPE 04: OPEN ACCESS PUBLISHING(4 hrs.)**

1. Open access publications and initiatives
2. SHERPA/RoMEO online resource to check publisher copyright & self-archiving policies
3. Software tool to identify predatory publications developed by SPPU
4. Journal finder / journal suggestion tools viz. JANE, Elsevier Journal Finder, Springer Journal Suggester, etc.

• **RPE 05: PUBLICATION MISCONDUCT (4hrs.)**

A. Group Discussions (2 hrs.)

1. Subject specific ethical issues, FFP, authorship
2. Conflicts of interest
3. Complaints and appeals: examples and fraud from India and abroad

B. Software tools (2 hrs.)

Use of plagiarism software like Turnitin, Urkund and other open source software tools

• **RPE 06: DATABASES AND RESEARCH METRICS (7hrs.)**

A. Databases (4 hrs.)

1. Indexing databases
2. Citation databases: Web of Science, Scopus, etc.

B. Research Metrics (3 hrs.)

1. Impact Factor of journal as per Journal Citation Report, SNIP, SJR, IPP, Cite Score
2. Metrics: h-index, g index, i10 index, altmetrics

References

- Bird, A. (2006). *Philosophy of Science*. Routledge.
- MacIntyre, Alasdair (1967) *A Short History of Ethics*. London.
- P. Chaddah, (2018) *Ethics in Competitive Research: Do not get scooped; do not get plagiarized*, ISBN:978-9387480865
- National Academy of Sciences, National Academy of Engineering and Institute of Medicine. (2009). *On Being a Scientist: A Guide to Responsible Conduct in Research: Third Edition*. National Academies Press.
- Resnik, D. B. (2011). What is ethics in research & why is it important. *National Institute of Environmental Health Sciences*, 1–10. Retrieved from <https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm>
- Beall, J. (2012). Predatory publishers are corrupting open access. *Nature*, 489(7415), 179–179. <https://doi.org/10.1038/489179a>
- Indian National Science Academy (INSA), *Ethics in Science Education, Research and Governance*(2019), ISBN:978-81-939482-1-7. http://www.insaindia.res.in/pdf/Ethics_Book.pdf

Paper 502: Research Methodology

[L+T+P: 4+0+0 = 4 Credits]

Unit 1: Introduction to Research Methodology (10 hrs)

Objectives of research, Understanding research and its goals. Critical thinking. Techniques for generating research topics. Topic selection and justification.

Methods of acquiring knowledge: Traditional, empirical, natural, interpretative, dialectical & scientific. Meaning of role of assumption, constructs, laws, theory, hypothesis, propositions, axioms and postulates.

Types of research: Basic research, applied research, and action research. Method of research: theoretical, descriptive, experimental & analytical.

Unit 2: Research Problem and Reporting (10 hrs)

Research Problem: selection of problem, process of designing a research proposal, steps to be followed in formulating a research proposal - theoretical and experimental processes, sources of research materials, review of literature - sources of information.

Preparation of Research Report/Thesis.

Unit 3: Quantitative Methods (10 hrs)

Probability distribution: Normal, Binomial, Poisson, Exponential, Weibull and Geometric Distributions. Estimation, Hypothesis testing & application; Correlation & regression analysis. Parametric and Non-parametric statistics. Confidence interval, Errors.

Quantitative Techniques: Levels of significance, Regression and Correlation coefficient. Statistical analysis and fitting of data; Chi-Square Test, Association of Attributes t-Test Anova, Standard deviation, Co-efficient of variations.

Open source software for quantitative and statistical analysis.

Unit 4: Computer Applications (10 hrs)

Concept of algorithms, Flow Charts, Data flow diagrams

Matlab programming: Basics of MATLAB programming, Array operations in MATLAB, Loops and execution control, Working with files: Scripts and Functions, Plotting and program output

Suggested Reading:

1. Ranjit Kumar, Research Methodology: A Step-by-Step Guide for Beginners, 3rd Edition, SAGE Publication (2010).
2. C.R. Kothari, Research Methodology- Methods and Techniques, New Age International, 2nd Edition. New Delhi (2013).
3. Y.K.Singh and R.B.Bajpai, Research Methodology-Techniques and Trends, APH. Publishing, 2nd Edn., New Delhi (2007).
4. Dan Gookin, MS Word (2003) for Dummies, John Wiley & Sons, (2003).
5. Dong Lowe, MS Power point (2003) for Dummies, John Wiley & Sons, (2003).
6. Wallace Wang, MS Office (2003) for Dummies, John Wiley & Sons, (2003).
7. Walpole R.A., Myers R.H., Myers S.L. and Ye, King, Probability & Statistics for Engineers and Scientists, 9th Edition, Pearson Prentice Hall, Pearson Education, Inc. (2011).
8. Anderson B.H., Dursaton, and Poole M., Thesis and assignment writing, John Wiley & Sons, Inc. (1979).
9. Bjorn Gustavii, How to write and illustrate scientific papers, Cambridge University Press; 3rd edition (2017).
10. Bordens K.S. and Abbott, B.b.: Research Design and Methods : A Process Approach, McGraw-Hill Education; 10 edition (2017).

Paper 503 (A)

Credit:2

Data structures and Algorithms

Complexity of Algorithms, Sorting techniques, Searching techniques, Hashing concepts, Search Trees – Binary search trees, AVL trees, Red-Black Trees, B & B+ Trees, KD tree, Splay Trees, Heaps and related algorithms,

Graph Search Algorithms-Generic Search-Breadth-First Search-Dijkstra's Shortest-Weighted-Path - Depth-First Search, Minimum Spanning Trees , Kruskal and Prim's algorithms. Single-Source Shortest Paths: The Bellman-Ford algorithm, Single source shortest paths in DAG's, Dijkstra's algorithm

Paper 503(B)

Credit:2

Elective papers

List of Elective courses:

- | | |
|-------------------------------------|---------------------------------|
| 01. Cryptography & Network Security | 06. Natural Language processing |
| 02. Wireless and Mobile Technology | 07. Data Science |
| 03. Information Retrieval | |
| 04. Soft Computing | |
| 05. Image processing | |

List of Electives

CSE....01 Cryptography and Network Security

Introduction to Cryptography, Security Attacks and Services, Mathematical Background, Number Theory, Substitutions and Permutations, Modular Arithmetic, Classical Cryptosystems, Cryptanalysis of Classical Cryptosystems

Symmetric Encryption Algorithms: Theory of Block Cipher Design, Feistel Cipher Network Structures, DES and Triple DES, Modes of Operation, Modern Symmetric Encryption Algorithms, Cryptanalysis of Symmetric Key Ciphers

Public Key Cryptography: Testing for Primality, Factoring Large Numbers, RSA, Diffie-Hellman, ElGamal, Key Exchange Algorithms, Public-Key Cryptography Standards.

Hashes and Message Digests, Cryptanalysis of Asymmetric Key Ciphers, Modern Trends in Asymmetric Key Cryptography

Digital Signature Standard (DSS and DSA), Authentication of Systems, X.509 Authentication Service, Elliptic Curve Cryptosystems, Digital Watermarking and Steganography.

Network Security: Secret Sharing Schemes, Pretty Good Privacy (PGP), Secure Socket Layer (SSL), Intruders and Viruses, Firewalls, IPSEC, Web Security, privilege management infrastructure (PMI) and Access Control, Security in e-commerce & smart cards, Bitcoin, Ad-hoc Networks Security Mechanisms

Suggested Readings:

1. "Cryptography and network security: principles and practices." William Stallings, Pearson Education India, 2006.
2. "Applied cryptography: protocols, algorithms, and source code in C." Bruce Schneier, John Wiley & Sons, 2007 "
3. "A Handbook of Applied Cryptography", Alfred J. Menezes, Paul C. Van Oorschot and Scott A. Vanstone, CRC Press Series

CSE.....02:Wireless & Mobile Technology

UNIT – I

Fundamentals of Wireless Communication Technology, Evolution of mobile systems:GSM,CDMA,3G,4G; Handoff , Location management

UNIT – II

Wireless MAC ,Mobile Network Layer: Mobile IP; Mobile Transport Layer : Traditional TCP, Indirect TCP, Snooping TCP, Mobile TCP

UNIT – III

Mobile Ad-hoc Networks(MANET):Layered architecture of MANET, Ad hoc network routing protocols, MAC layer issues of MANET, Routing In MANET:DSR,AODV,DSDV,OLSR, Transport layer issues, Security issues in MANET ; Vehicular Ad-hoc Networks

UNIT – IV

Wireless Sensor Networks(WSN):Introduction, MAC layer protocols for WSN,Routing Protocols for WSN, Transport Control protocols and Middlewares, ,Operating Systems for WSN; Architecture and Design issues of Wireless Mesh Networks(WMNs),MAC layer issues, Routing issues in WMN, Security issues in WMN

Suggested Readings:

1. “Mobile Communications” ,J. Schiller, Addison-Wesley
2. “Fundamentals of Mobile and Pervasive Computing”, Frank Adelstein, S.K.S. Gupta, Golden G. Richard and Loren Schwiebert , McGraw-Hill Professional.
3. “Ad-hoc and Sensor Networks: Theory and Applications” ,C.D.M. Cordeiro & D.P. Agarwal , World Scientific Publishing Company
4. “Ad Hoc Wireless Networks: Architectures and Protocols”, C. Siva Ram Murthy and B.S. Manoj, Pearson.

CSE.....03: Information Retrieval

UNIT – I

Introduction: Goals and history of IR. The impact of the web on IR. Basic IR Models: Boolean and vector-space retrieval models; ranked retrieval; text-similarity metrics; TF-IDF (term frequency/inverse document frequency) weighting; cosine similarity.

UNIT – II

Basic Tokenizing, Indexing, and Implementation of Vector-Space Retrieval: Simple tokenizing, stop-word removal, and stemming; inverted indices; efficient processing with sparse vectors; Java implementation. Experimental Evaluation of IR: Performance metrics: recall, precision, and F-measure; Evaluations on benchmark text collections.

UNIT – III

Query Operations and Languages: Relevance feedback; Query expansion; Query languages. Text Representation: Word statistics; Zipf's law; Porter stemmer; morphology; index term selection; using thesauri. Metadata and markup languages (SGML, HTML, XML). Text Categorization: Categorization algorithms: Rocchio, nearest neighbor, and naive Bayes. Applications to information filtering and organization.

UNIT – IV

Web Search: Search engines; Spidering; Metacrawlers; directed Spidering; link analysis (e.g. hubs and authorities, Google PageRank); shopping agents. Language-Model Based Retrieval : Using naïve Bayes text classification for ad hoc retrieval. Improved smoothing for document retrieval.

UNIT – V

Text Clustering: Clustering algorithms: agglomerative clustering; k-means; expectation maximization (EM). Applications to web search and information organization. Information Extraction and Integration: Extracting data from text; semantic web; collecting and integrating specialized information on the web.

Text Books:

1. Introduction to Information Retrieval- Manning, Prabhakar, Schütze, Cambridge University Press (2008)
2. Modern Information Retrieval– Ricardo Baeza-Yate, Berthier Ribeiro-Neto, Pearson Education Asia (2005)

Reference Books:

1. Introduction to Modern Information Retrieval(2nd Edition) , G.G. Chowdhury, Neal-Schuman Publishers (2003)
2. Information Retrieval: Algorithms, and Heuristics, David A. Grossman, Ophir Frieder, Academic Press (2000).
3. Text Information Retrieval Systems , Charles T. Meadow, Bert R. Boyce, Donald H. Kraft, Academic Press(2000).

CSE04: Soft Computing

UNIT – I

Neural Networks: History, overview of biological Neuro-system, Mathematical Models of Neurons, ANN architecture, Learning rules, Learning Paradigms-Supervised, Unsupervised and reinforcement Learning, ANN training Algorithms-perceptions, Training rules, Delta, Back Propagation Algorithm, Multilayer Perceptron Model, Hopfield Networks, Associative Memories, Applications of Artificial Neural Networks

UNIT – II

Fuzzy Logic: Introduction to Fuzzy Logic, Classical and Fuzzy Sets: Overview of Classical Sets, Membership Function, Fuzzy rule generation. Operations on Fuzzy Sets: Compliment, Intersections, Unions, Combinations of Operations, Aggregation Operations. Fuzzy Arithmetic: Fuzzy Numbers, Linguistic Variables, Arithmetic Operations on Intervals & Numbers, Lattice of Fuzzy Numbers, Fuzzy Equations

UNIT – III

Fuzzy Logic: Classical Logic, Multivalued Logics, Fuzzy Propositions, Fuzzy Qualifiers, Linguistic Hedges. Uncertainty based Information: Information & Uncertainty, Non specificity of Fuzzy & Crisp Sets, Fuzziness of Fuzzy Sets.

UNIT – IV

Introduction of Neuro-Fuzzy Systems, Architecture of Neuro Fuzzy Networks Application of Fuzzy Logic: Medicine, Economics etc. Genetic algorithms(GAs), Evolution strategies(ESs), Evolutionary programming(EP), Genetic Programming(GP), Selecting, crossover, mutation, schema analysis, analysis of selection algorithms; convergence; Markov & other stochastic models.

UNIT – V

Soft computing approaches: Simulated Annealing, Tabu Search, Ant colony based optimization, etc. A fusion Approach of multispectral images with SAR, Optimization of travelling Salesman problem using genetic algorithm approach, Soft computing based hybrid fuzzy controllers.

Text Books:

1. Neuro-Fuzzy and Soft computing: A Computational Approach to Learning and Machine Intelligence- Jang, Sun, Mizutani, Pearson Education (2004)
2. Principles of Soft Computing – S.N. Sivanandam and S. N. Deepa, Wiley India Pvt Limited (2011)

Reference Books:

1. Neural Networks, Fuzzy Logic and Genetic Algorithm: Synthesis and Applications – S. Rajasekaran and G.A Vijayalakshmi pai, Prentice-Hall of India Pvt Limited (2006)
2. Fuzzy Set Theory: Foundations and Applications- George J. Klir, Ute St. Clair, Bo Yuan, Prentice Hall(1997).
3. Neural Networks: Algorithms, Applications and Programming Techniques- Freeman J.A. & D.M.

CSE05: Image Processing

UNIT – I

Introduction: Background, Digital Image Representation, Fundamental steps in Image Processing, Elements of Digital Image Processing - Image Acquisition, Storage, Processing, Communication, Display. Digital Image Formation : A Simple Image Model, Geometric Model- Basic Transformation (Translation, Scaling, Rotation), Perspective Projection, Sampling & Quantization - Uniform & Non uniform.

UNIT – II

Mathematical Preliminaries: Neighbour of pixels, Connectivity, Relations, Equivalence & Transitive Closure; Distance Measures, Arithmetic/Logic Operations, Fourier Transformation, Properties of The Two Dimensional Fourier Transform, Discrete Fourier Transform, Discrete Cosine & Sine Transform.

UNIT – III

Image Enhancement : Spatial Domain Method, Frequency Domain Method, Contrast Enhancement - Linear & Nonlinear Stretching, Histogram Processing; Smoothing - Image Averaging, Mean Filter, Low-pass Filtering; Image Sharpening. High-pass Filtering, High-boost Filtering, Derivative Filtering, Homomorphic Filtering; Enhancement in the frequency domain - Low pass filtering, High pass filtering.

UNIT – IV

Image Restoration: Degradation Model, Discrete Formulation, Algebraic Approach to Restoration - Unconstrained & Constrained; Constrained Least Square Restoration, Restoration by Homomorphic Filtering, Geometric Transformation - Spatial Transformation, Gray Level Interpolation.

UNIT – V

Image Segmentation: Point Detection, Line Detection, Edge detection, Combined detection, Edge Linking & Boundary Detection - Local Processing, Global Processing via The Hough Transform; Thresholding - Foundation, Simple Global Thresholding, Optimal Thresholding; Region Oriented Segmentation - Basic Formulation, Region Growing by Pixel Aggregation, Region Splitting & Merging.

Text Books:

1. Digital Image Processing (Third Edition)- Rafael C. Gonzalez, Pearson(2007)
2. Digital Image Processing(Seventh Edition)- Bernd Jahne, Springer(2017)
3. Digital Image Processing & Analysis(Second Edition)- Scott E Umbraugh, CRC press(2010)

Reference Books:

1. Fundamentals of Digital Image Processing–Anil K. Jain, Pearson(1988).
2. Image Processing, Analysis & Machine Vision- Milan Sonka, Vaclav Hlavac and Roger Boyle, Cengage Learning(2014).

UNIT – I

Introduction to NLP: Definition, issues and strategies, application domain, tools for NLP, Linguistic organisation of NLP, NLP vs PLP. Word Classes: Review of Regular Expressions, CFG and different parsing techniques Morphology: Inflectional, derivational, parsing and parsing with FST, Combinational Rules Phonology: Speech sounds, phonetic transcription, phoneme and phonological rules, optimality theory, machine learning of phonological rules, phonological aspects of prosody and speech synthesis. Pronunciation.

UNIT – II

Spelling and N-grams: Spelling errors, detection and elimination using probabilistic models, pronunciation variation (lexical, allophonic, dialect), decision tree model, counting words in Corpora, simple N-grams, smoothing (Add One, Written-Bell, Good-Turing), N-grams for spelling and pronunciation.

UNIT – III

Syntax: POS Tagging: Tagsets, concept of HMM tagger, rule based and stochastic POST, algorithm for HMM tagging, transformation based tagging Sentence level construction & unification: Noun phrase, co-ordination, sub-categorization, concept of feature structure and unification.

UNIT – IV

Semantics: Representing Meaning: Unambiguous representation, canonical form, expressiveness, meaning structure of language, basics of FOPC Semantic Analysis: Syntax driven, attachment & integration, robustness Lexical Semantics: Lexemes (homonymy, polysemy, synonymy, hyponymy), WordNet, internal structure of words, metaphor and metonymy and their computational approaches Word Sense Disambiguation: Selectional restriction based, machine learning based and dictionary based approaches.

UNIT – V

Pragmatics: Discourse: Reference resolution and phenomena, syntactic and semantic constraints on Coreference, pronoun resolution algorithm, text coherence, discourse structure. Dialogues: Turns and utterances, grounding, dialogue acts and structures. Natural Language Generation: Introduction to language generation, architecture, discourse planning (text schemata, rhetorical relations).

Text Books:

3. Speech and Language Processing: An introduction to Language processing, Computational Linguistics, and Speech Recognition- D. Jurafsky & J. H. Martin (Second Edition), Prentice Hall (2008).
4. Foundation of Statistical Natural Language Processing (Kindle edition)- Christopher D. Manning, Hinrich Schutze, The MIT Press (1999).
5. Machine Translation- Pushpak Bhattacharyya, Chapman and Hall/CRC publication (2015).

Reference Books:

1. Natural Language Understanding (Second Edition) – James Allen, Pearson (1994).
2. Natural Language Processing: A Paninian Perspective (Eastern Economy Edition)- Bharathi, A. Vineet Chaitanya and Rajeev Sangal Prentice Hall India.
3. Statistical Language Learning- Eugene Charniak MIT Press (1993).

CSE.....07: Data Science

UNIT I:

Introduction to core concepts and technologies: Introduction, Terminology, data science process, data science toolkit, Types of data, Example applications.

UNIT II:

Data collection and management: Introduction, Sources of data, Data collection and APIs, Exploring and fixing data, Data storage and management, Using multiple data sources

UNIT III:

Data analysis: Introduction, Terminology and concepts, Introduction to statistics, Central tendencies and distributions, Variance, Distribution properties and arithmetic, Samples/CLT, Basic machine learning algorithms, Linear regression, SVM, Naive Bayes.

UNIT IV:

Data visualization: Introduction, Types of data visualization, Data for visualization: Data types, Data encodings, Retinal variables, Mapping variables to encodings, Visual encodings.

UNIT V:

Applications of Data Science, Technologies for visualization, Bokeh (Python)

UNIT VI:

Recent trends in various data collection and analysis techniques, various visualization techniques, application development methods of used in data science.

Text Books:

1. Cathy O’Neil and Rachel Schutt. Doing Data Science, Straight Talk From The Frontline. O’Reilly. 2014
2. Jure Leskovek, AnandRajaraman and Jeffrey Ullman. Mining of Massive Datasets. v2.1, Cambridge University Press. 2014. (free online)

Reference books:

1. Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. ISBN 0262018020. 2013.
2. Foster Provost and Tom Fawcett. Data Science for Business: What You Need to Know about Data Mining and Data-analytic Thinking. ISBN 1449361323. 2013.
3. Trevor Hastie, Robert Tibshirani and Jerome Friedman. Elements of Statistical Learning, Second Edition. ISBN 0387952845. 2009. (free online)