DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SYLLABUS FOR Ph.DCOURSE WORK

RAJIV GANDHI UNIVERSITY, RONO HILLS, DOIMUKH

(YEAR 2021-22 ONWARDS)

Ph.D Course WorkCourse Structure

Paper Code	Title	Credit (L-T-P)	Mark Distribution				Examination time (in hours) (T-P-S)
			End Semester		In Semester	Total	
			Theory	Practical	Sessional		
PHDCS-801	Research Methodology	3-1-0	80	0	20	100	3-0-1½
PHDCS-802	Research and Publication Ethics	1-0-1	25	15	10	50	1-1-1
PHDCS-83X	Elective – I (***Any one from the list***)	3-1-0	80	0	20	100	3-0-1½
	PHDCS-831: Data Science						
	PHDCS- 832: Speech and Natural Language Processing						
	PHDCS- 833: Wireless Access Technologies						
	PHDCS-834: Machine Learning						
PHDCS-84X	Elective – II (***Any one from the list***)(Practical only)	0-0-2	0	25	25	50	0-2-2
	PHDCS841: Data Analytics with Python						
	PHDCS842: R for Data Science						
	PHDCS843: Programming with MATLAB						

PROGRAM OBJECTIVES:

- To develop an understanding of various research designs and techniques.
- To provide students an understanding of the expectations from research work.
- To improve research skills of engineering students
- To bridge the skill gaps and make students research ready
- To provide an opportunity to students to develop inter-disciplinary skills.
- The programfocusses on research skill development and more than 50% of the time is spent on practical training and problem solving, to provide the requisite understanding towards application of academic topics from Computer Engineering disciplines into real world engineering research projects.

PHDCS-801: RESEARCH METHODOLOGY (3-1-0)

COURSE OBJECTIVES:

- To develop understanding of the basic framework of research process, publication and patent scopes.
- To develop an understanding of various research designs and techniques.
- To identify various sources of information for literature review and data collection.
- To develop an understanding of the ethical dimensions of conducting applied research.
- Appreciate the components of scholarly writing and evaluate its quality.

LEARNING OUTCOME :

Students who successfully complete this course will be able to:

- Explain key research concepts and issues
- Read, comprehend, and explain research articles in their academic discipline.

Contents:

Unit 1: Meaning of research problem, Sources of research problem, Criteria Characteristics of a good research problem, Errors in selecting a research problem, Scope and objectives of research problem.Approaches of investigation of solutions for research problem, data collection, analysis, interpretation, Necessary instrumentations

Unit 2: Effective literature studies approaches, Analysis plagiarism, Research ethics,

Unit 3: Effective technical writing, how to write report, PaperDeveloping a Research Proposal, Format of research proposal, a presentation and assessment by a review committee

Unit 4: Nature of Intellectual Property: Patents, Designs, Trade and Copyright. Process of Patenting and Development: technological research, innovation, patenting, development. International Scenario: International cooperation on Intellectual Property. Procedure for grants of patents, Patenting under PCT.

Unit 5: Patent Rights: Scope of Patent Rights. Licensing and transfer of technology. Patent information and databases. Geographical Indications.

Unit 6: New Developments in IPR: Administration of Patent System. New developments in IPR; IPR of Biological Systems, Computer Software etc.

References:

1. Stuart Melville and Wayne Goddard, "Research methodology: an introduction for science & engineering students'"
2. Wayne Goddard and Stuart Melville, "Research Methodology: An Introduction"
3. Ranjit Kumar, 2nd Edition, "Research Methodology: A Step by Step Guide for beginners"
4. Halbert, "Resisting Intellectual Property", Taylor & Francis Ltd ,2007.
5. Mayall, "Industrial Design", McGraw Hill, 1992.
6. Niebel, "Product Design", McGraw Hill, 1974.
7. Asimov, "Introduction to Design", Prentice Hall, 1962.
8. Robert P. Merges, Peter S. Menell, Mark A. Lemley, "Intellectual Property in New Technological Age", 2016.
9. T. Ramappa, "Intellectual Property Rights Under WTO", S. Chand, 2008

PHDCS-802: RESEARCH AND PUBLICATION ETHICS (2-0-0)

COURSE OBJECTIVES:

- Provide students with the fundamental knowledge of basics of philosophy of science andethics, research integrity, publication ethics.
- Hands-on sessions are designed to identify research misconduct and predatory publications.
- Indexing and citation databases, open access publications, research metrics (citations, *h*index, Impact Factor etc).
- Guide and mentor students in presenting plagiarism tools for a valid and ethical research report.

LEARNING OUTCOMES :

- To be able to describe and apply theories and methods in ethics and research ethics
- To acquire an overview of important issues in research ethics, like responsibility for research, ethical vetting, and scientific misconduct.
- To acquire skills of presenting arguments and results of ethical inquiries.

Contents:

Unit-1: Philosophy and Ethics: Introduction to Philosophy: definition, nature, scope, concept, branches Ethics: definition, moral philosophy, nature of moral judgment and reactions

Unit -2: Scientific Conduct: Ethics with respect to science and research; Intellectual honesty and research integrity, copyright, Scientific misconduct: falsification, fabrication and Plagiarism (FFP); Redundant Publication: duplication and overlapping publication, salami slicing; Selective reporting and misrepresentation of data

Unit – 3: Publication Ethics:

Publication Ethics: definition, introduction and importance Best practice/standard setting initiative and guidelines: COPE, WAME, etc. Conflict and interest

Publication misconduct: definition, concept, problems that leads to unethical behaviour and vice versa, type Violation of publication ethics, authorship and contributorship Identification of publication misconduct, complaint and appeals Predatory publisher and journals Avoiding Plagiarism. Preparing documents for MoUs, Confidentiality Agreements

Unit – 4: Openaccesspublishing: Open access publication and initiatives SHERPA/RoMEO online resource to check publisher copyright and self-archiving policies Software tool to identify predatory publication developed by SPPU Journal finder/journal suggestion tools viz. JANE, Elsevier Journal finder, Springer, Journal Suggester, etc.

Unit – 5:

A. Group Discussion: Subject Specific Ethical Issues FFP, authorship Conflict interest, Complaints and appeals: examples and fraud from India and abroad
B. Software tools: Use of plagiarism software like turnitin, Urkund and other open source software tools

Unit – 6:

A. Database: Indexing database, Citation database: web of science, scopus, etc.
B. Research metrics: Impact factor of Journal as per journal citation report, SNIP,SJR,IPP, Cite Score Metrics: h-index, g-index, i-10 index, altmetrics

(Note : Practical based on topics mentioned in Unit 4 5, and 6)

References:

1. Nicholas H. Steneck. Introduction to the Responsible Conduct of Research. Office of Research Integrity. 2007. Available at: https://ori.hhs.gov/sites/default/files/rcrintro.pdf
2. The Student's Guide to Research Ethics By Paul Oliver Open University Press, 2003
3. Responsible Conduct of Research By Adil E. Shamoo; David B. Resnik Oxford University Press, 2003
4. Ethics in Science Education, Research and Governance Edited by KambadurMuralidhar, Amit Ghosh Ashok Kumar Singhvi. Indian National Science Academy, 2019.
5. Anderson B.H., Dursaton, and Poole M.: Thesis and assignment writing, Wiley Eastern 1997.
6. BijornGustavii: How to write and illustrate scientific papers? Cambridge University Press.
7. Bordens K.S. and Abbott, B.b.: Research Design and Methods, Mc Graw Hill, 2008.
8. Graziano, A., M., and Raulin, M.,L.: Research Methods – A Process of Inquiry, Sixth Edition, Pearson, 2007.

PHDCS-831: DATA SCIENCE (3-1-0)

COURSE OBJECTIVES:

- Help students learn, understand, and practice big data Analytics and machine learning approaches, which include the study of modern computing Big data technologies.
- Scaling up machine learning techniques focusing on industry applications.

LEARNING OUTCOMES :

- Conceptualization and summarization of big data and machine learning, trivial data versus big data, big data computing technologies, machine learning techniques, and scaling up machine learning approaches.

Contents:

Unit 1: Introduction to core concepts and technologies: Introduction, Terminology, data science process, data science toolkit, Types of data, Example applications.

Unit 2: 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 3: 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 4: Data visualization: Introduction, Types of data visualization, Data forvisualization, Data types, Data encodings, Retinal variables, Mapping variables to encodings, Visual encodings.

Unit 5: Applications of Data Science,Technologies for visualization, Bokeh

Unit 6: Recent trends in various data collection and analysis techniques, various visualization techniques, application development methods of used in data science.

References:

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

PHDCS- 832: SPEECH AND NATURAL LANGUAGE PROCESSING (3-1-0)

COURSE OBJECTIVES:

- Teach students the leading trends and systems in natural language processing.
- Make them understand the concepts of morphology, syntax, semantics and pragmatics of the language and that they are able to give the appropriate examples that will illustrate the above mentioned concepts.
- Teach them to recognize the significance of pragmatics for natural language understanding.
- Enable students to be capable to describe the application based on natural language processing and to show the points of syntactic, semantic and pragmatic processing.

LEARNING OUTCOMES :

- Understand approaches to syntax and semantics in NLP.
- Understand approaches to discourse, generation, dialogue and summarization within filters.
- Understand current methods for statistical approaches to machine translation.
- Understand machine learning techniques used in NLP, including Hidden Markov Models and probabilistic context-free grammars, clustering and unsupervised methods, log-linear and discriminative models, and the EM algorithm as applied within NLP

Contents:

Unit – 1: Discrete Signals and Systems: A Review – Introduction to DFT – Properties of DFT – Circular Convolution – Filtering methods based on DFT – FFT Algorithm – Decimation in time algorithm, Decimation in frequency algorithm – z-transform – Filter Design : FIR and IIR.

Unit – 2: Analogy and Physiology of Speech Production mechanism: Lungs, Larynx, vocal Tract – Categorization of Speech Sounds – Elements of a language: vowels, Nasals, Fricatives, plosives. Speech perception; Voiced and Unvoiced sounds

Unit – 3: Spectral Analysis of Speech: Formant Frequency, Fundamental Frequency, Power spectral analysis, PSD, Energy Spectra – Analysis and Synthesis of Pole-Zero Speech Models, Linear Prediction of Speech Signal – Cepstral Frequency, Mel-cepstra, Homomorphic signal Processing-Homomorphic Filtering

Unit – 4: Speech Coding: Optimum scalar and vector quantization, waveform coding, vecoders – Speech recognition – Dynamic time warping, Hidden Markov Model – Speaker verification and Identification – Speech synthesis: Text-to-speech system, synthesizer techniques, HMM-based synthesis, sine-wave synthesis

Unit- 5: Brief Review of Regular Expressions and Automata; Finite State Transducers; Word level Morphology and Computational Phonology; Basic Text to Speech; Part of Speech Tagging; Parsing with CFGs;

Unit – 6: Probabilistic Parsing. Representation of Meaning; Semantic Analysis; Lexical Semantics; Word Sense; Disambiguation; Discourse understanding; Natural Language Generation; Techniques of Machine Translation; Indian Language case studies.

References:

1. John G. Proakis, Dimitris K Manolakis, Digital Signal Processing, Pearson Education
2. S.Salivahanan, A. Vallavaraj, C. Gnanapriya, DigitalSignal Processing, TMH
3. Lawrence Rabiner and Biing-Hwang Juang, Fundamentals of Speech Recognition, Pearson Education.
4. T.E.Quatieri, Speech Signal Processing, Pearson Education
5. Ben Gold , Nelson Morgan and Dan Ellis, Speech and Audio Signal Processing: Processing and Perception of Speech and Music, WILEY Publication

PHDCS- 833: WIRELESS ACCESS TECHNOLOGIES (3-1-0)

COURSE OBJECTIVES:

- To study the evolving wireless technologies and standards
- To understand the architectures of various access technologies.
- To understand various protocols and services provided by next generation networks.

LEARNING OUTCOMES :

After successfully completing the course student will be able to

- Keep the student updated on latest wireless technologies and trends in the communication field
- Understand the transmission of data through various networks.

Contents:

Unit 1: Necessity for wireless terminals connectivity and networking. Wireless networking advantages and disadvantages, Overview of wireless access technologies. Narrowband and broadband networks, fixed and nomadic networks. Wireless local loop (WLL), Public Switched Telephone Network (PSTN) interfaces.

Unit 2: Fixed wireless access (FWA) networks, frequency bands for different networks. Criterions for frequency bands allocation, Network topologies, hotspot networks. Communication links: point-to-point (PTP), point- to-multipoint (PMP), multipoint-to-multipoint (MTM).

Unit 3: Standards for most frequently used wireless access networks: WPAN (802.15, Bluetooth, DECT, IrDA), UWB (Ultra-Wideband), WLAN (802.11, Wi-Fi, HIPERLAN, IrDA), WMAN (802.16, WiMAX, HIPERMAN, HIPERACCESS), WWAN (802.20), Other technologies for broadband wireless access, Local Multipoint Distribution Service (LMDS), Multichannel Multipoint Distribution Service (MMDS). Ad Hoc networks, Network services. Services types based on carrier frequency and bandwidth.

Unit 4: Wireless access networks planning, design and installation. Services provision, legislative and technical aspects, Technical and economical factors for network planning: expenses, coverage, link capacity, network complexity and carrier-to-interference ratio (C/I). Base station or access point allocation. Base station and access point equipment. Terminal mobility issues regarding wireless access to Internet. Wireless networking security issues.

Unit 5: Example of laptop or handheld PC wireless connection in real environment. PC wireless interface equipment. Wireless access network exploitation and management, software requirements, link quality control. Business model, wireless network services market, market research and marketing, service providers, wireless data application service providers (WDASP) and their role on public telecommunication services market, billing systems.

Unit 6: Recent trends in wireless networking and various access mechanism, new standards of wireless communication.

References:

M. P. Clark, Wireless Access Networks: Fixed Wireless Access and WLL networks -- Design and Operation, John Wiley & Sons, Chichester

D. H. Morais, Fixed Broadband Wireless Communications: Principles and Practical Applications, Prentice Hall, Upper Saddle River

R. Pandya, Introduction to WLLs: Application and Deployment for Fixed and Broadband Services, IEEE Press, Piscataway

PHDCS-834: MACHINE LEARNING (3-1-0)

COURSE OBJECTIVES:

- To introduce students to the basic concepts and techniques of Machine Learning.
- To become familiar with regression methods, classification methods, clustering methods.
- To become familiar with Dimensionality reduction Techniques.

LEARNING OUTCOMES :

- Gain knowledge about basic concepts of Machine Learning
- Identify machine learning techniques suitable for a given problem
- Solve the problems using various machine learning techniques
- Apply Dimensionality reduction techniques.
- Design application using machine learning techniques.

Contents:

Unit 1: Supervised Learning (Regression/Classification), **Basic methods:** Distance-based methods, Nearest-Neighbors, Decision Trees, Nave Bayes, **Linear models:** Linear Regression, Logistic Regression, Generalized, Support Vector Machines, Nonlinearity and Kernel Methods, Beyond Binary Classification: Multi-class/Structured Outputs, Ranking

Unit 2: Unsupervised Learning: Clustering: K-means/Kernel K-means, Dimensionality Reduction: PCA and kernel PCA, Matrix Factorization and Matrix Completion , Generative Models (mixture models and latent factor models)

Unit 3: Evaluating Machine Learning algorithms and Model Selection, Introduction to Statistical Learning Theory, Ensemble Methods (Boosting, Bagging, Random Forests)

Unit 4 : Sparse Modeling and Estimation, Modeling Sequence/Time-Series Data, DeepLearning and Feature Representation Learning

Unit 5 : Scalable Machine Learning (Online and Distributed Learning): A selection from some other advanced topics, e.g., Semi-supervised Learning, Active Learning, Reinforcement Learning, Inference in Graphical Models, Introduction to Bayesian Learning and Inference

Unit 6: Recent trends in various learning techniques of machine learning and classification methods for IOT applications. Various models for IOT applications.

References:

1. Kevin Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012
2. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Springer 2009 (freely available online)
3. Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007.

PHDCS841: DATA ANALYTICS WITH PYTHON (0-0-2)

COURSE OBJECTIVES:

- Import, clean, enrich, transform, visualize and output the analysis of a large datasets, as well as using various visualization tools.

LEARNING OUTCOMES :

- Learn to write, test and debug Python 3 code with confidence, including working with Containers, Conditionals & Loops, Functions & Modules and Error Handling.
- Learn the fundamentals of some of the most widely used Python packages; including NumPy, Pandas and Matplotlib, then apply them to Data Analysis and Data Visualization projects.
- Build and code a Graphical User Interface (GUI) for result demonstration.

Contents:

Introduction to Python and its Basics; Learn the basics of Python language; Learn Regular Expressions in Python; Learn Scientific libraries in Python – NumPy, SciPy, Matplotlib and Pandas; Effective Data Visualization; Learn Scikit-learn and Machine Learning

References:

1. Gowrishankar S, Veena A, "Introduction to Python Programming", 1st Edition, CRC Press/Taylor & Francis, 2018.
2. Jake VanderPlas, "Python Data Science Handbook: Essential Tools for Working with Data", 1st Edition, O'Reilly Media, 2016.
3. AurelienGeron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems", 2nd Edition, O'Reilly Media, 2019.
4. Wesley J Chun, "Core Python Applications Programming", 3rd Edition, Pearson Education India, 2015.
5. Miguel Grinberg, "Flask Web Development: Developing Web Applications with Python", 2nd Edition, O'Reilly Media, 2018.

PHDCS842: R FOR DATA SCIENCE (0-0-2)

COURSE OBJECTIVES:

- Introduce the basics of programming in R for data analytics.
- Introduce good practices of workflows and reproducibility in data science.
- To develop skills independently in programming and data science workflow.
- Introduce the fundamental principles of writing scripts in R and how they are applied in practice for data analytics.

LEARNING OUTCOMES :

- Understand the basics in R programming in terms of constructs, control statements, string functions
- Understand the use of R for Big Data analytics
- Learn to apply R programming for Big data processing.
- Able to appreciate and apply the R programming from a statistical perspective

Contents:

Introducing to R; Matrices, Arrays And Lists; Data Frames; Object Oriented Programming with R; Interfacing R to other languages

References:

1. Norman Matloff, The Art of R Programming: A Tour of Statistical Software Design, No Starch

Press, 2011
2. Jared P. Lander, R for Everyone: Advanced Analytics and Graphics, Addison-Wesley Data & Analytics Series, 2013.
3. Mark Gardener, Beginning R – The Statistical Programming Language, Wiley, 2013
4. Robert Knell, Introductory R: A Beginner's Guide to Data Visualisation, Statistical Analysis and Programming in R, Amazon Digital South Asia Services Inc, 2013

PHDCS843: PROGRAMMING WITH MATLAB (0-0-2)

COURSE OBJECTIVES:

- To enhance programming knowledge in Research and Development with MATLAB software.
- To provide a working introduction to the MATLAB technical computing environment viz. Themes of data analysis, visualization, programming etc.

LEARNING OUTCOMES :

- Learn to apply R MATLAB for breaking a complex task up into smaller, simpler tasks for data processing.
- Able to Tabulate results , analyze and visualize data

Contents:

Introduction to Programming with MATLAB; Programming Environment; Graph Plots; Procedures and Functions; Control Statements; Manipulating Text; GUI Interface; Discrete Linear Systems; Spectral Analysis; Speech Signal Analysis

References:

1. Stormy Attaway, MATLAB: A Practical Introduction to Programming and Problem Solving, 3rd edition, Elsevier, 2013
2. Hunt, Brian R., Ronald L. Lipsman, and Jonathan M. Rosenberg. A guide to MATLAB: for beginners and experienced users. Cambridge university press, 2014.
3. Gilat, A., *MATLAB: An introduction with applications* (4^{th} Ed). New York: Wiley, 2012.
4. Stephen J. Chapman, MATLAB Programming for Engineers, 6E,Cengage Learning India Pvt. Ltd., 2019