Following is the list of subjects along with the respected syllabi, floated as the Course Work for PhD in Computer Science and Engineering (Session 2020-21).

Core Courses:

Course Code	Course Title	Credits	Year of Offering Course
SET801F	Research and Publication Ethics	2	2020
SET802F	Research Methodology for Engineering and Technology	4	2020

Departmental electives:

Course Code	Course Title	Credits	Year of Offering Course
CSE801E	Machine Learning	2	2020
CSE802E	Deep Learning	2	2020
CSE803E	Soft Computing	2	2020
CSE804E	Data Science	2	2020
CSE805E	Advanced Computational Design of Ethics	2	2020
CSE806E	Advanced Computer Architecture	2	2020
CSE807E	Graph Theory	2	2020

Research and Publication Ethics
SET801F

- RPE 01: PHILOSOPHY AND ETHICS (3 hrs.)

1. Introduction to philosophy: definition, nature and scope, concept, branches
2. Ethics: definition, moral philosophy, nature of moral judgements and reactions

- RPE 02: SCIENTIFICCONDUCT (5hrs.)

1. Ethics with respect to science and research
2. Intellectual honesty and research integrity
3. Scientific misconducts: Falsification, Fabrication, and Plagiarism (FFP)
4. Redundant publications: duplicate and overlapping publications, salami slicing
5. Selective reporting and misrepresentation of data

- RPE 03: PUBLICATION ETHICS (7 hrs.)

1. Publication ethics: definition, introduction and importance
2. Best practices / standards setting initiatives and guidelines: COPE, WAME, etc.
3. Conflicts of interest
4. Publication misconduct: definition, concept, problems that lead to unethical behavior and vice versa, types
5. Violation of publication ethics, authorship and contributorship
6. Identification of publication misconduct, complaints and appeals
7. 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 ofScience. Routledge.

MacIntyre, Alasdair (1967) A Short Histopy ofEthics. London.

P. Chaddah, (2018) Ethics in Competitive Research: Do not get scooped; do not get plagiarized, ISBN:9789387480865

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 ofEnvironmental Health Sciences, 1—10. Retrieved from https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfin

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/pdffEthics Book.pdf

Research Methodology for Engineering & Technology

SET802F

Unit-I (12 hours) Science and Scientific research

Knowledge and the epistemology of knowledge, deductive and inductive inference. A brief history of scientific ideas, important thinkers, scientists and scientific advancements. Principles of effective research, aspects of research - self-development and the creative process, the problem-solver and the problem-creator; Finding and solving research problems, Literature survey, developing a research plan, research proposals.

Unit-II (12 hours) Computational thinking and statistical analysis:

Measures of central tendency, Data and its nature, data representation, curve plotting using MS-Excel and Origin. Hypothesis testing concept of p-value. Student's t-test and F-test. ANOVA(one way and two way), transformation of data. Tests of significance, non-parametric tests, simple, partial and multiple correlations. Basic principles of Statistical Computation using various softwares; design of experiments and analysis of results using various softwares (MATLAB, ORIGIN, SPSS, Design Expert, etc)

Unit-III (10 hours): Scientific Writing:

The research report, Steps in writing a report, Layout of the research report, Writing references and bibliography. Presentations: Importance of effective presentation, Planning a good presentation. Scientific Papers: How to write good papers, models of the paper writing process, identifying ideas and telling the story, The benefits of targeting good journals, Peer review, How to respond to reviewer comments, reviewing a paper. Identification of publication misconduct, complaints and appeals Predatory publishers and journals

Unit-IV (10 hours):

Technical and Scientific documentation using Latex: What is LATEX?, A short history of TEX, Main attractions of LATEX: Automatic Styling according to Journal requirements, Cross references, Writing Complex Maths. The LATEX document, Typical Input Files, Post-processed look, The Edit/Format/Preview Process, Embedding References in the Document, Bibliography management using BIBTEX, Presentations using Beamer, Introduction to Overleaf, Hands-on sessions on LATEX.

Unit-V (16 hours):

Modelling, Simulation and Data analysis using software tools: Introduction, Variable types, arrays and matrices, multidimensional arrays, subarrays, operations, functions, using functions with array inputs, Displaying output data. Two Dimensional Plotting, multiple plots, subplots, Logarithmic and Polar plots. Control Flow/Branching Statements: Logical data type, Relational operators, if – else, switch – case, Loops– While and For loops, nested loops. User defined functions. Additional Data and Plot Types: Complex Data, Complex numbers with relational operators, Complex functions, String functions, 3D line plots, 3D surface, mesh and contour plots, Introduction to Optimal Design using software tools.

Machine Learning
CSE-801E

L-P
2-0

UNIT I

Review of basic concepts of linear algebra, probability and optimization. Definition of learning systems. Goals and applications of machine learning. Aspects of developing a learning system: training data, concept representation, function approximation. Classification and regression. Types of learning. Basic Machine learning pipeline – training, testing and validation. Evaluation Metrics – accuracy, precision, recall, ROC curve. Linear Regression, Logistic regression, Multi-class Regression, SoftMax Regression. Regularization Theory, Overfitting, Underfitting, bias-variance tradeoff.

UNIT II

Discriminative Learning: Perceptron, Maximum margin classifiers and Support Vector Machines: Hard and soft margin, Higher dimensional space and Kernel trick. Decision trees – Concept of pure and impure nodes and the measure of impurity using entropy and Gini. ID3, CART, C4.5 Neural Networks and Backpropagation.

UNIT III

Generative Learning: Bayes Classifier and Naïve Bayes Classifier. Maximum Likelihood Estimation, Maximum a Posteriori Estimation. Non-Parameterized Density Estimation: Parzen window and KNN Density estimation. Clustering: Agglomerative clustering, K-means, Gaussian Mixture models and Expectation Maximization algorithm.

UNIT IV

Computational Learning Theory: Computational learning theory, PAC learning model, Sample complexity, VC Dimension. Dimensionality Reduction. The Curse of Dimensionality. PCA, SVD, LDA, ICA. tSNE

UNIT V

Sequence Modelling and Ensemble Methods: Hidden Markov Models, Conditional Random Fields. Bagging & boosting and its impact on bias and variance C5.0 boosting, Random forest, AdaBoost, Gradient Boosting Machines and XGBoost.

References

1. *Machine Learning. Tom Mitchell*. First Edition, McGraw- Hill, 1997.
2. *Introduction to Machine Learning*, Edition 2, Ethem Alpaydin
3. *Pattern Recognition and Machine Learning*. Bishop, Christopher M. springer, 2006.
4. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Géron, Aurélien. O'Reilly Media, 2019.

Deep Learning
CSE-802E

L-P
2-0

UNIT I

History of Deep Learning, Deep Learning Success Stories, McCulloch Pitts Neuron, Thresholding Logic, Perceptrons, Perceptron Learning Algorithm, Multilayer Perceptrons (MLPs), Representation Power of MLPs, Sigmoid Neurons, Gradient Descent, Feedforward Neural Networks, Representation Power of Feedforward Neural Networks, Feed-Forward Neural Networks, Backpropagation, Gradient Descent (GD), Momentum Based GD, Nesterov Accelerated GD, Stochastic GD, AdaGrad, RMSProp, Adam.

UNIT II

Eigenvalues and eigenvectors, Eigenvalue Decomposition, Basis, Principal Component Analysis and its interpretations, Singular Value Decomposition. Autoencoders and relation to PCA, Regularization in autoencoders, Denoising autoencoders, sparse autoencoders, Contractive autoencoders.

UNIT III

Regularization: Bias Variance Tradeoff, L2 regularization, Early stopping, Dataset augmentation, Parameter sharing and tying, Injecting noise at input, Ensemble methods, Dropout, Greedy Layerwise Pre-training, Better activation functions, Better weight initialization methods, Batch Normalization, Learning Vectorial Representations Of Words, Convolutional Neural Networks, LeNet, AlexNet, ZF-Net, VGGNet, GoogLeNet, ResNet

UNIT IV

Recurrent Neural Networks, Backpropagation through time (BPTT), Vanishing and Exploding Gradients, Truncated BPTT, GRU, LSTMs, Encoder Decoder Models, Attention Mechanism, Attention over images.

UNIT V

Generative Models: Generative adversarial networks, Variational Autoencoders. Transformers, BERT, GPT.

References

1. *Deep Learning, Ian Goodfellow, Yoshua Bengio, Aaron Courville, MIT Press.*
2. *Pattern Recognition and Machine Learning, Christopher Bishop, Springer.*
3. *Neural Networks and Deep Learning: A Textbook, Charu C Aggarwal*

Soft Computing
CSE-803E

L-P
2-0

UNIT I
Soft Computing: Introduction, soft computing vs. hard computing, various types of soft computing techniques, applications of soft computing.
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.

UNIT II
Fuzzy Arithmetic: Fuzzy Numbers, Linguistic Variables, Arithmetic Operations on Intervals & Numbers, Lattice of Fuzzy Numbers, Fuzzy Equations.
Fuzzy Proposition: Classical Logic, Multivalued Logics, Fuzzy Propositions, Fuzzy Qualifiers, Fuzzy Decision Making, Fuzzy Control Systems. Fuzzyfications & Defuzzificataions, Applications of fuzzy logic, Industrial applications.

UNIT III
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 IV
Genetic Algorithm: Concept of Genetics and Evolution and its application to probabilistic search techniques. Basic GA framework and different GA architectures. GA operators: Encoding, Crossover, Selection, Mutation, etc. Solving single-objective optimization problems using GAs.

UNIT V
Introduction of Hybrid Systems: Neuro-Fuzzy Hybrid Systems, Genetic Neuro Hybrid Systems, Genetic Fuzzy Hybrid and Fuzzy Genetic Hybrid Systems.
Applications: Medicine, Economics etc.

Text Books:

1. J.S.R.Jang, C.T. Sun And E.Mizutani, "Neuro-Fuzzy And Soft Computing", PHI / Pearson Education, 04.
2. S.N.Sivanandam & S.N.Deepa, "Principles of Soft Computing", Wiley India Pvt Ltd, 11.
3. F. O. Karray and C. de Silva, "Soft computing & Intelligent System Design", Pearson, 09.

Reference Books:

1. S.Rajasekaran & G.A.Vijayalakshmi Pai, "Neural Networks, Fuzzy Logic And Genetic Algorithm: Synthesis & Applications", Prentice-Hall Of India Pvt. Ltd., 2006.
2. Freeman J.A. & D.M. Skapura, "Neural Networks: Algorithms, Applications and Programming Techniques", Addison Wesley, Reading, Mass, 1992.
3. S. Haykin. "Neural Networks-A Comprehensive Foundations", Prentice-Hall Int., 1999.

DATA SCIENCE
CSE-804E

L-P
2-0

UNIT I

Introduction, Basic statistical descriptions of data, Data visualization, Measuring data similarity and dissimilarity, Data preprocessing. Data Warehousing.

UNIT II

Frequent pattern mining: Market basket analysis, Frequent itemset mining algorithms, Mining closed and maximal patterns, Association rule mining, Correlation analysis, Pattern evaluation measures, Advanced Pattern Mining.

UNIT III

Classification: Decision tree induction, Bayes' theorem, Naive Bayesian classification, Rule-based classification, Metrics for evaluating classifier performance, Ensemble methods, Classification by Backpropagation (Neural network-based classification), Classification Using Frequent Patterns, Support vector machines (SVMs), Lazy Learners, Other classification methods.

UNIT IV

Clustering: Clustering Techniques — Partitioning methods, Hierarchical methods, Density-based methods, Grid-based methods, Probabilistic model-based methods, Clustering high-dimensional data, Clustering graph and network data, Clustering with constraints, Evaluation of clustering, Other cluster analysis methods.

UNIT V

Data Mining Trends and Research Frontiers: Mining Complex Data Types, Mining sequence data (time series, symbolic sequences, biological sequences), Mining data streams, Mining text data, Web mining, Graph mining, Mining spatial data, Social network analysis, Visual and audio data mining, Privacy preserving data mining, Outlier detection. Introduction to data mining in R and Python.

UNIT VI

Big data analytics: Big data, Big data technologies, Processing big data, Hadoop MapReduce, Big data tools and techniques.

REFERENCES:

1. Jiawei Han, Micheline Kamber and Jian Pei, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, Elsevier.
2. Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Introduction to Data Mining", Addison-Wesley, Pearson.
3. Mohammed J. Zaki and Wagner Meira Jr., "Data Mining and Analysis: Fundamental Concepts and Algorithms", Cambridge University Press.
4. Charu C. Aggarwal, "Data Mining: The Textbook", Springer.
5. Ambiga Dhiraj, Michael Minelli and Michehe Chambers, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business", Wiley.
6. Arvind Sathi, "Big Data Analytics: Disruptive Technologies for Changing Game", McP.
7. Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", Wiley.
8. Tom White, "Hadoop: The Definitive Guide", O'Reilly.

Advanced Computational Design of Ethics
CSE-805E

L-P
2-0

UNIT I
Fundamentals of ethics, metaethics, normative ethics: virtue theories, deontology, prima facie duties, consequentialism, contractarianism, divine command theory, Islamic ethics, natural law theory, Kant and Categorical Imperatives, applied ethics: normative principles in applied ethics, issues in applied ethics

UNIT II
Introduction to Frankenstein, technoethics: Nanoethics, biotech ethics, computer ethics, Internet ethics, media ethics, Value Sensitive Design (VSD), Universal Declaration of Human Rights, biased computer system, framework for analyzing bias in computers, minimizing bias in computer system design

UNIT III
Machine ethics: nature, importance, and difficulty, why machine ethics?, philosophical concerns with machine ethics, morality of artificial agents, engineering morality, philosophers, engineers and the design of AMAs, top-down, bottom-up and hybrid approaches, ontology agnostic approach

UNIT IV
Introduction to scientific method, intelligent agents: weak AI, strong AI, ethics of artificial intelligence, ethically aligned design: principles, norms, methodologies, personal data, economic and legal issues,

UNIT V
Braitenberg Vehicle, principle of double effect, Asimov's Laws of Robotics, Kantian Machine, SIROCCO, JEREMY, W.D, MedEthEx

Suggested readings
W. Wallach and C. Allen, Moral machines: teaching robots right from wrong. Oxford University Press.
J. Deigh, An Introduction to Ethics. Cambridge University Press.
S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed.
Prentice Hall.

Advanced Computer Architecture

CSE-806E

L-P

2-0

UNIT I

Modern day architectures, Program Execution: compilation, object files, Function call and return, Address space, Data and representation. Computer Organization: Memory, registers, instruction set architecture, Instruction processing. RISC, CISC.

UNIT II

Using processor software approach: Virtual memory, Address translation, paging. Introduction to operating systems, processes, system calls, process management. Single Cycle Implementation.

UNIT III

Using processor hardware approach: pipelined processors, structural, data and control hazards, its impact on programming, static and dynamic instruction scheduling, branch handling and prediction. Multicycle Implementation.

UNIT IV

Cache memory, organization and its impact on programming, Cache coherence, solutions to cache coherence problem, Protection.

UNIT V

Parallel programming, inter process communication, synchronization, parallel architecture (Fyln's texonomy), shared memory vs message passing, and introduction to programming with message passing with MPI.

Text Books:

1. "Computer Architecture: A Quantitative Approach", David A Patterson and John L. Hennessy

2. "Advanced Computer Architecture: Parallelism, Scalability, Programmability", Kai Hwang

Graph Theory

CSE-807E

L-P

2-0

UNIT I
Preliminaries: Graphs, isomorphism, subgraphs, matrix representations, degree, operations on graphs, degree sequences Connected graphs and shortest paths: Walks, trails, paths, connected graphs, distance, cut-vertices, cut-edges, blocks, connectivity, weighted graphs, shortest path algorithms.

UNIT II
Trees: Characterizations, number of trees, minimum spanning trees. Special classes of graphs: Bipartite graphs, line graphs, chordal graphs. Hamilton graphs: Necessary conditions and sufficient conditions.

UNIT III
Independent sets, coverings, matchings: Basic equations, matchings in bipartite graphs, perfect matchings, greedy and approximation algorithms. Vertex colorings: Chromatic number and cliques, greedy coloring algorithm, coloring of chordal graphs, Brook's theorem. Edge colorings: Gupta-Vizing theorem, Class-1 graphs and class-2 graphs, equitable edge-coloring.

UNIT IV
Planar graphs: Basic concepts, Eulers formula, polyhedrons and planar graphs, characterizations, planarity testing, 5-color-theorem.

UNIT V
Directed graphs: Out-degree, in-degree, connectivity, orientation, Eulerian directed graphs, Hamilton directed graphs, tournaments, ER Graphs, Random Graphs.

Text Books:

1. "Introduction to Graph Theory", Douglas Brent West.

2. "Graph Theory", Diestel, R.

3. "Introduction to Graph Theory", Richard J. Trudeau.