Department of Computer Science & Engineering, GIT

Pre Ph.D Courses

(w.e.f from 2018-19 admitted batch)

ERC901: RESEARCH METHODOLOGY

Module I

Objectives and Types of Research: Motivation and objectives – Research methods vs Methodology, Types of research – Descriptive vs Analytical, Applied vs Fundamental, Quantitative vs Qualitative vs Empirical

Module II

Research Formulation: Defining and formulating the research problem – Selecting the problem – Necessity of defining the problem – Importance of literature review in defining a problem – Literature review – Primary and secondary sources – reviews, treatise, monographs, patents – web as a source – searching the web – Critical literature review – Identifying gap areas from literature review – Development of work hypothesis

Module III

Research Design and Methods: Research Design – Basic Principles – Need of research design – Features of good design – Important concepts relating to research design – Observation and Facts, Laws and Theories, Prediction and explanation, Induction, Deduction. Development of models, Developing a research plan – Exploration, Description, Diagnosis, Experimentation, Determining experimental and sample designs.

Module IV

Data Collection and Analysis: Execution of Research – Observation and Collection of data – Methods of data collection – Sampling methods – data processing and Analysis strategies – Data analysis with statistical Packages – Hypothesis – testing – Generalization and Interpretation.

Module V

Report and Thesis Writing: Structure and components of scientific reports – Types of report – Technical reports and thesis – Significance – Different steps in the preparation – Layout, structure and Language of typical reports – Illustrations and tables – Bibliography, referencing and footnotes – Oral Presentation – Planning _ Preparation Practice – Making presentation – Use of visual aids, Importance of effective communication.

Text Books

1. C.R.Kothari, Research Methodology – Methods and Techniques, New Age International Publishers, 2004.
2. William M. Trochim, James P. Donnelly, Research Methods, 2/e, Cengage India, 2016.

EID902: DATAMINING

Module I

Introduction to Data Mining: What is Data Mining, Motivating Challenges, The origins of Data Mining, Data Mining Tasks. Data: Types of Data, Data quality, Data Preprocessing, Measures of Similarity and Dissimilarity.

Module II

Classification: Basic Concepts, Decision Trees, and Model Evaluation Preliminaries, General Approach to solving a classification Problem, Decision Tree Induction, Model Overfitting, Evaluating the performance of a classifier. **Classification Alternate Techniques** - Rule-based Classifier, Nearest-Neighbor Classifiers, Bayesian Classifiers, Deep Learning, SVM.

Module III

Association Analysis: Basic Concepts and Algorithms, Problem Definition, Frequent Itemset Generation, Compact Representation of Frequent Itemsets, Alternative Methods for generating Frequent Itemsets, Evaluation of Association Patterns, Handling a Concept Hierarchy, Sequential Patterns.

Module IV

Cluster Analysis: Basic Concepts and Algorithms, Overview, K-Means, Agglomerative Hierarchical Clustering, DBSCAN, Cluster Evaluation, Scalable Clustering Algorithm, Which Clustering Algorithm?

Module V

Anomaly Detection: Characteristics of Anomaly Detection Problems, Characteristics of Anomaly Detection Methods, Statistical Approaches, Proximity-based Approaches, Clustering-based Approaches, Reconstruction-based Approaches, One-class Classification, Information Theoretic Approaches, Evaluation of Anomaly Detection.

Text Book(s)

1. Tan, Steinbach, Vipin Kumar, Introduction to Data Mining, Pearson Education,2006

2. Jiawei Han, MichelineKamber, Data Mining Concepts and Techniques,Morgan Kaufman Publications.

References

1. Margaret H Dunhan, Data Mining Introductory and Advanced Topics,Pearson Education.

2. Ian H. Witten Eibe Frank, Data Mining, Morgan Kaufman Publications