Paper III
Title: Graph Theory & Data Mining
M. Marks: 100

Unit I
Relations: Relations and their properties, Binary Relations, Equivalence relations, Diagraphs, Computer representation of relations and digraphs, Transitive Closures, Warshall's Algorithm.
Graph Theory: Graph Terminology, Representing Graphs, Connectivity of Graphs: Paths and Circuits, Eularian and Hamiltonian Paths.
Trees: Rooted trees, Application of trees: Binary Search Trees, Decision Trees, Prefix Codes, Tree traversal algorithms, trees and sorting, spanning trees, minimal spanning trees.

Unit II
Data Mining: Introduction to data mining.
Decision Trees: General idea, Where to use Decision Trees, How it works, Strengths and weaknesses.
Neural Networks: General idea, Where to use Neural Networks, How it works, Strengths and weaknesses.

Unit III
Nearest Neighbor and Clustering: General idea, Where to use nearest-neighbor prediction, How Clustering and Nearest–Neighbor Prediction work, Strengths and weaknesses.
Genetic Algorithms: General idea, Where to use them, How Genetic Algorithms work, Strengths and weaknesses.

Unit IV
Social Networks: Small world Phenomenon, Properties of Social Networks.
Web Mining: Web Content mining, Web Structure mining, Web Usage mining.
Applying Social Network Analysis to Web: PageRank Algorithm, Hyperlink Induced Topic Search (HITS).

Unit V
Social Networks as Graphs: Varieties of Social Networks, Graphs with Several Node Types, Clustering of Social-Network Graphs, Distance Measures for Social-Network Graphs, Applying Standard Clustering Methods, Betweenness, The Girvan-Newman Algorithm, Using Betweenness to Find Communities, Discovery of Communities, Finding Cliques, Finding Overlapping Communities, Maximum-Likelihood Estimation, The Affiliation-Graph Model.

REFERENCES:

• Alex Berson, Stephen J. Smith, "Data Warehousing, Data Mining and OLAP", Tata McGraw Hill, 2007.

- Kenneth H. Rosen, "Discrete Mathematics and Its Applications", The Random House.

- Knoke, "Social Network Analysis", Sage, 2nd Edition.

- Scott, J., "Social Network Analysis: A Handbook", Sage.

Paper Title: Recent Advances in Computer Science　　　　　Max Marks: 100
Attempt 5 questions with one question from each section　　　Time: 2.5 hours

UNIT – I:

Big Data: Definition and Explanation. Characteristics of Big Data (Basic, 3V and 5V). Why Big Data is important? Types of Big Data: Structured, Semi-structured, and unstructured.
Big Data Technology Foundation: Physical Infrastructure (Generation, Computation, Communication, and Storage), Security Infrastructure. Current trends and Challenges.

UNIT – II:

Cloud Computing: Overview, Evolution and Characteristics. How Cloud Computing works? Pros and Cons of Cloud Computing. Challenges of Cloud Computing. Comparison with traditional computing architecture (Client/Server). Comparison with other recent computing trends (Grid, Cluster and Distributed Computing).
Virtualization: Introduction to virtualization, types and implementation levels.
Cloud Computing Architecture: Cloud computing stack, Introduction to Cloud Service Models - Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). Introduction to Cloud Deployment Models – Public Cloud, Private Cloud, Hybrid Cloud, and Community Cloud. Services provided at various levels.

UNIT – III:

IoT Definition: Overview, Application, Potential and Challenges, Architecture, M2M vs IoT. Internet vs IoT: Layers, Protocols, Packet-services, Performance parameters of Packet-networks (Web, P2P, Sensor Networks, & Multimedia).

Unit - IV :

Definition of learning systems. Goals and applications of machine learning. Aspects of developing a learning system: training data, concept representation, function approximation. The concept learning task. Concept learning as search through a hypothesis space. General-to-specific ordering of hypotheses. Finding maximally specific hypotheses. The importance of inductive bias.

Unit - V:

Blockchain: Introduction, Advantage over conventional distributed database, Blockchain Network, Mining Mechanism, Distributed Consensus, Merkle Patricia Tree, Gas Limit, Transactions and Fee, Anonymity, Reward, Chain Policy, Life of Blockchain application, Soft & Hard Fork, Private and Public blockchain.

References:

1. "Big Data for Dummies", Judith Hurwith, Alan Nugent, Fern Halper, and Marcia Kaufman, John Wiley & Sons, 2013.

2. "Big Data – Principles and best practices of scalable real-time data systems", Nathan Marz and James Warren, Dreamtech Press, 2016.

3. "Cloud Computing Bible", Barrie Sosinsky, Wiley-India, 2010.

4. "Cloud Computing: Principles and Paradigms", Editors: Rajkumar Buyya, James Broberg, Andrzej M. Goscinski, Wiley, 2011.

5. "Designing Internet-of-Things", Adrain McEwen, & Hakim Cassimally, Wiley.

6. "The Internet of Things", Samuel Greengard, MIT Press.

7. "The Silent Intelligence: The Internet of Things", Daniel Kellmereit & Daniel Obodovksi, DND Ventures LLC.

8. "Internet of Things: A hands on approach", Arhdeep Bahga, & Vijay Madisetti, Orient Blackswan.

9. Machine Learning, Tom Mitchell, McGraw Hill

10. Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller and Steven Goldfeder, Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction, Princeton University Press (July 19, 2016).

School of Applied Science and Technology
University of Kashmir

Syllabus

Course Title: Research Methodology   **Course Code: PhD-AST-RM-2023**
Credits: (5L+ 1P)   **Maximum Marks: 150 (Theory - 125 + Practical – 25)**
Examination Duration: Theory: 3 Hours, Practical: 1 Hour

Notes: 1. *To pass, a candidate has to secure a minimum of 50% marks separately in theory and practical examinations.*
2. *For theory examination, two questions shall be asked from each unit. A candidate has to attempt one question form each unit. Each question shall carry 25 marks.*

Unit I: Research Methodology and Design

- Research methodology and statistical reasoning, population and sample.
- Research problem and hypothesis, variables, basic concepts and importance of statistics.
- Introduction to construction of questioners, validity, quantitative and qualitative research methodologies.
- Experimental designs: between subjects or independent groups design, repeated measures or within subjects design, complex/factorial design.
- Non-experimental designs: quasi-experimental or natural groups design, observational methods: types, data analyses of observational and descriptive data, case study, survey research.

Unit II: Basic Statistics

- Introduction to methods of data collection. Steven's levels of measurement: nominal, ordinal, interval, and ratio scale.
- Descriptive statistics: mean, median, mode, range, quartile deviation, variance, standard deviation.
- Correlation and Regression: Pearson's product moment correlation, Spearman's rank order correlation, simple and multiple regression, outliers.
- Sampling methods: simple random, stratified, systematic, cluster, and multistage. sampling errors. Sample size determination.
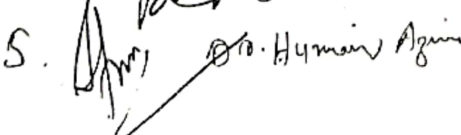
Unit III: Parametric Tests

- The normal curve and its properties (area, skewness, kurtosis, etc.). Tests of normality.
- Inferential statistics: null hypothesis testing, statistical significance testing, one-tailed and two-tailed tests, degrees of freedom, confidence interval, p value.
- Parametric tests: Z-tests, t-tests, analysis of variance (ANOVA).

A

1.   2.   Prof. Mubashir Hussain Masom
3.   2.   Dr. Jawid Iqbal
5.   Dr. Humair Azim   4.   Dr. Adil Elaie
6.

Unit IV: Non-parametric Tests

- Non-parametric tests: descriptions and assumptions, chi-square test, sign test, Wilcoxon's sign rank test, median test, Mann-Whitney U test.

Unit V: Scientific Writing and Research Ethics

- Critical analysis of scientific articles in terms of their importance, consistency, and justifications.
- Journal indexing and research metrics: journal impact factor, h-index, g-index, Eigen factor score, Altmetrics, etc. Publication models for scholarly communications.
- submissions.
- Ethical considerations in research, types of ethical issues, examples of ethical failures. University of Kashmir research policy.

Unit VI: Laboratory Work

- Use of spreadsheets or other software tools (Excel, etc.) for preliminary data analysis and graphical representations.
- Use of software for statistical computations (SPSS, R, etc.) for inferential statistics, parametric and non-parametric tests.
- Use of word processors (Word, Latex, etc.), typing assistant software (Grammarly, etc.) and reference formatting software (Mendeley, Endnote, Zotero, etc.) for manuscript writing.
- Use of plagiarism detection software (iThenticate, Turnitin, URKUND, Plagiarism Checker X, etc) for checking plagiarism of manuscripts.

Books

1. Umesh Kumar B Dubey, D P Kothari. (2022). Research methodology: techniques and trends. CRC Press,.
2. Mongomery, D. C. (2017). Montgomery: design and analysis of experiments. John Willy & Sons.

1. Stewart Jr, C. N. (2023). Research ethics for scientists: A companion for students. John Wiley & Sons.
2. Laake, P., Benestad, H. B., & Olsen, B. R. (Eds.). (2007). Research methodology in the medical and biological sciences. Academic Press.
3. Louis C, Lawrence M, and Keith M. (2007). Research methods in education. 6th ed.. Routledge, London.
4. John W. Creswell. (2014). Research design: Qualitative, Quantitative, and mixed method approaches, Sage Publications.

1.

3.

5.

2

2.

4.

6.