INFORMATION BOOKLET

PhD Course work

Department of Computer Science & Engineering,
JIS University

81, Nilgunj Road, Agarpara, Kolkata - 700109

Introduction:

The Department of Computer Science and Engineering was established in July 2016. The department offers following programmes: 1. 4-Years B.Tech. programme in Computer Sc. & Engineering 2. 3 Years Bachelor in Computer Application and 3. Ph.D. degree in various specializations of Computer Sc. and Engineering. Our graduates have distinguished themselves in higher studies at the top Universities. They also occupy positions of eminence in the computer industry. Our Alumni remain in constant touch with us and are contributing in the development of the department. Placements for our graduates are the best in the Institute. The faculty members of the department have international experience and training. The departmental research is focused in the areas of Artificial Intelligence, Neuro Computing, Parallel Processing, Software Engineering, Image Processing and Computer Vision, Medical Image Processing, Pattern Recognition, Data mining and Web mining, Biometrics, semantic web, Natural Language Processing (NLP), Machine Learning, and Information Extraction. Besides plan funding, the Department attracts financial inputs through externally funded projects. Programme Objective: The basic objective of the programme offered by the Dept. is to train the students in the area of Computer Science and Engineering and its streams at various levels UG/PG. Apart from learning the concepts of development of skills related to the discipline, the courses are designed in such a way that covers most of the basic as well as advanced courses required for pursuing PhD research in a specific area.

Curriculum Overview

The PhD course work courses are designed in such a way that facilitates the start of the research in following areas of Computer Science and Engineering: Artificial Intelligence, Neuro Computing, Parallel Processing, Software Engineering, Image Processing and Computer Vision, Medical Image Processing, Pattern Recognition, Data mining and Web mining, Biometrics, semantic web, Natural Language Processing (NLP), Machine Learning, and Information Extraction, Computer Networks , and wireless sensor networks etc.

The PhD course work courses are designed in such a way that facilitates the start of the research in following areas of Computer Science and Engineering: Artificial Intelligence, Neuro Computing, Parallel Processing, Software Engineering, Image Processing and Computer Vision, Medical Image Processing, Pattern Recognition, Data mining and Web mining, Biometrics, semantic web, Natural Language Processing (NLP), Machine Learning, and Information Extraction, Computer Networks , and wireless sensor networks etc.

For Ph.D. Computer Science & Engineering course work details and credit allotted were given below which was discussed by the BOS members and approved:

Sl. No.	Course Name	Course Code	Theory /Lab	Credit
1.	Research Methodology	GRM-1	Theory	4
2.	Cloud Computing (Compulsory)	CC-2	Theory	4
3.	Data Warehousing and Mining (Elective-1)	DWDM-3	Theory	4
4.	Literature Review Survey	LRS-4	Theory	Non-Credit (Compulsory)
5.	Data Science & Analytics (Elective -2)	DSA -5	Theory	4

Systematic Literature Review in Computer Science

This work aims to provide a practical guide to assist students of Computer Science courses and related fields to conduct a systematic literature review. The steps proposed in this paper to conduct a systematic review were extracted from a technical report published by the researcher Bárbara Kitchenham and arranged in a more objective format, in order to make information more accessible and practical, especially for those who are having their first contact with this technique. The target audience for this work are undergraduate, master's and doctoral students that are in the initial phase of their bibliographic research.

Introduction

This document aims to provide a practical guide to assist students in conducting a systematic literature (SLR). It is interesting to use this guide with examples of systematic reviews already published, so that the researcher can view the steps outlined. This guide is based on the technical report proposed by Barbara Kitchenham in 2007. Before starting a systematic review, it is interesting to see if there is already a review covering the defined scope. Considering there is no systematic review published covering the defined scope, the following steps should be considered:

1. Defining the main research questions of the literature review Here come the questions that guide what the author wants to figure out with the research. (For example: What technique is being used to solve a particular problem on Software Engineering or Computer Graphics?).

2. Defining keywords. There are several techniques that can be used to find these keywords (the idea is to use one or more of the items below):

a. Analyze the research questions and extract the first keywords;

b. Use papers from the area of interest to find new words and synonyms of the words already found;

c. Define the PICOC [4] to better delineate the scope/ aims of the systematic review (optional).

3. Defining search string The basic rule is to separate the keywords. For each separated word, find its synonyms and concatenate them with the OR connector. After the definition of the groups of words with their synonyms, concatenate them with AND to end the string. See below an example of a search string extracted from [5]: (pragmatic OR pragmatics OR pragmatism) AND (interoperability OR interoperate OR interoperable OR interoperation OR similarity OR integrate OR integration) AND (solution OR method OR technique OR model OR tool OR framework OR architecture OR infrastructure OR approach) AND (computational OR system OR application OR software)

4. Defining search engines This definition depends on the area of systematic review. In general, considering the main electronic search engines that include the researches produced in the area of Computer Science, the revision should cover part (if not all) of the following bases:

a. IEEE Xplore (www.ieeexplore.com.br)

b. Scopus (www.scopus.com)

c. ScienceDirect (www.sciencedirect.com)

d. Springer (www.springerlink.com)

e. ACM (www.portal.acm.org)

f. Compendex (www.engineeringvillage.com)

5. String refinement The idea is to test the string defined in the previous item in one of the search engines (for example, Scopus). Once the string is performed, verify if the returned papers appear to be relevant. Known papers that are potential candidates for primary studies of this review (which exist on these engines) should appear in the search. If no relevant results appear, the search string must be parsed again to be calibrated. The refinement of the search criteria in each database (search by title, abstract and keywords or complete, limitation or not per year, limitation or not by area, etc.) should be analyzed on a case-by-case basis.

6. Search string execution Once the search string is defined, it must be adapted to each of the search engines, since the syntax used by each one is different. It is always interesting to keep in a document what string was used, how many articles each base returned, and the date of execution. If it is necessary to run the search again (either by the author or by a researcher who intends to

continue the work), you can execute the complete string or restrict to the period that the initial execution succeeds, just complementing the results.

7. Download and store search results All major search engines allow you to explore the set of selected results in various formats. The main formats are BIB (preferably) and CSV. These formats will be used as input for reference management tools.

There are several reference management tools that can be used such as JabRef and Mendeley. For a systematic review, we use these tools primarily to manage downloaded references and remove duplicate papers returned from different search engines. 8. Define inclusion and exclusion criteria In this step the criteria to decide which papers will go to the next stage of the review will be defined. The criteria should be defined from the research questions. Other criteria that may be used are: Restriction by language (only articles in English), restriction by area, restriction by primary articles (excluding secondary studies such as other systematic reviews, area mapping and surveys), etc.

The criteria, especially those of exclusion, may contain restrictions regarding the type of research that one wishes to do. These criteria should be carefully defined and reviewed by your advisor before they are implemented in order to align them with the research aims.

Selection of papers - First stage - Analysis by title and abstract In this stage, only the title and abstract of the selected papers will be analyzed. The suggestion is to export a CVS file (Jabref has this functionality) with the articles' data (with the duplicates removed) and to import it into a worksheet (for example, from Google Drive). The data that will be analyzed in this step (stored in the generated worksheet) is only title and abstract. It is interesting to add a status column so that each researcher involved in this step can assign one of the following values to the analyzed article: Included, excluded or doubtful (According to the defined inclusion / exclusion criteria.

8. Each of the researchers will have their own column to put the result of their analysis. It is important to highlight which analysis should be based on the inclusion and exclusion criteria defined previously. This step is quite subjective.

9. Articles that are marked "doubtful" should be discussed by team members, as well as possible disagreements. If those involved do not reach agreement, a diagonal reading of the paper is suggested to resolve the doubt. It is suggested that the first author does the first screening, followed by the others involved. All papers selected in this step should be downloaded. It is suggested to make the papers available to anyone involved in any cloud service (e.g. Dropbox). 10. Selection of papers - Second stage - Analysis by Introduction and Conclusion (Optional) Depending on the number of papers selected in the previous step, an intermediate step is required to refine the selection. In this stage, the introduction and conclusion of the selected papers will be analyzed. The selection criterion is the same as in the previous step. A copy of the spreadsheet tab used with the selected articles is created and the researchers involved re-mark the papers as included, excluded or doubtful.

11. Selection of papers - Third stage - Complete reading and quality checklist In this last stage, quality criteria (checklist) should be defined to verify the suitability of the papers analyzed. Besides the checklist, it is interesting to set a cut-off point so that less-qualified papers (according to the defined checklist) can be deleted. The lead author should do all the initial work (reading all articles, applying the checklist, deleting the articles below the cutoff point). This stage should be guided by the advisors.

12. Extraction of answers related to research questions At this stage, the research questions should be answered by analyzing the papers selected in the previous step. You can create a spreadsheet with the papers (title or ID) in the rows and the query questions in the columns. Then, when proceeding with the reading of the articles, the possible answers extracted may be posted directly to the spreadsheet. The synthesis of the extracted data can be presented in different forms. Usually, tables, graphs and other artifacts are used to facilitate the visualization of this information. This step can be performed in conjunction with the previous step. Conclusion Systematic literature reviews have been extensively used to identify, evaluate and interpret the studies published in the literature. A systematic review allows the researcher to collect evidence in order to identify research opportunities in a given field of study. Thus, systematic literature reviews are highly recommended for students who are starting their research and wish to evaluate effectively a particular area and clearly understand how their proposal may contribute considering what has already been published. The main contribution of this work is to assist in the construction of systematic literature review focused on the Computer Science. To achieve this goal, we developed a practical and concise guide that assists in simple and didactic steps the conduction of a systematic review in accordance with the technique discussed by Barbara in Kitchenham. We hope that with the use of this guide, the implementation of this valuable technique becomes easier, more accessible and be encouraged as an important step towards the production of quality research.

Research Methodology

Subject Code :- GRM - 1

Credit :- 4 (Theory)

Syllabus

Definition, Types & Methods of Research in Health • Literature search strategy • Using reference manager/ endnote • Research/study design options • Survey Techniques & Data collection methods • Design of survey and data collection instruments – relevance to study objectives; development and types of questions; length order, layout and coding of survey instrument • Instrument adaptation and validation • Introduction to data quality assessment • Data triangulation • Double method of data collection and management • Data base manipulations • Research on Diagnostic Tests • Critical appraisal of Journal Article and Writing a Research Paper • Sample Size Consideration in Medical Research • Facility based Vs population based health statistics • Statistical Methods in Research Qualitative research methods • Approaches, main qualitative methods (unstructured observation, structured observation, unstructured and semi-structured interviewing, systematic interviewing, multiple informant interviewing (FGD, PRA etc) • Representativeness, reliability and validity • Qualitative methods for community health need assessment - Community health needs assessment in the context of population-based methods, • Management and analysis of qualitative data Ethics in Research • Nuremberg code, Helsinki declaration, ICMR guidelines, accepted ethical principles concerning research on human subjects, informed consent, confidentiality, obtaining communal consent for field trials • Online research ethics certification course • Report writing, Protocol Development • Citation ,referencing and bibliography • Data driven management / program implementation.

References

1. Research Methodology Practice – P. Philominathan – Shri A.V.V.M. Pushpam College – Poondi –Thanjavar

2. Research Methodology – Methods & Techniques 2 ed, Kothari C. R. – Vishwa Prakashan – New Delhi 1990.

3. An Introduction to Research Procedure in Social Sciences – Gopal M. A. – Asia Publishing House - Bombay

4.Darlington, Richard B., Sharon Weinberg, and Herbert Walberg (1973). Canonical variate analysis and related techniques. Review of Educational Research, 453-454.

5. Gorsuch, Richard L. (1983) Factor Analysis. Hillsdale, NJ: Erlbaum

6. Morrison, Donald F. (1990) Multivariate Statistical Methods. New York: McGraw-Hill.

7.Rubenstein, Amy S. (1986). An item-level analysis of questionnairetype measures of intellectual curiosity. Cornell University Ph. D. thesis.

BOOKS:

1. Argyris C. "Personality and Organisation: The Conflict Between System and Individuals "Harper and Row, New York, 1995.

2. Arthur G., Beclein and Glueck W.G., Management., Hold sounders international edition, 1983.

3. Argyris C., "Interpersonal Competence and organizational Effectiveness", Erwin Dorsey Press, Moscow, 1992.

4. Batten.D. "Tough Minded Management" American Management Association, New York, 1963.

5. Beach S. "personnel – the management f people at work", Macmillan publishing co. Inc., New York.

6. Cellerman S.W. "Management by otivation", American Management Association, New York, 1968.

7. Chopra B.S., et.al., "Leadership for the Indian Managers, A human relation approach" , The Times research foundation.

8. Calhoon R.P. "Personal Management and supervision", Application Century crafts, New York, 1967.

9. Chatterjee N.N. "Management of Personal in Indian enterprises : Concept, Practices and emerging trends", Allied book agency, Calcutta.

10.Calhoon R.P. "Personnel Management and Supervision" Appleton Century, New York, 1964.
11.Desai K.G. "Human Problem in Indian Industries", : Sindhu, Bombay, 1969.

12.Devis Keith "Human Relation at work, McGrow Hill book company, Inc., New York, 1957.
13.Dunette M.D., Ca,pbell J.P. Hakel M.D. "Factors Contributing to job satisfaction and job dissatisfaction six occupational group, organizational Behaviour and Human Performance" 1967, p.no. 143-174.

14.Davis K. "Human behavior at work", Tata McGrow Hill, New Delhi – 1977.

15.Dwivedi R.S., "Supervisor Personality" Samilya Publication, Bombay, 1978.

16.Flippo E.B. “Principles of personnel Management”, McGrow Hill, New York, 1931.

17.Ford Martin E., “Motivating Human”, Sage Publication, New Delhi, 1995.

18.Ghosh “Personnel Administration in India”, Sudha Publication New Delhi, 1969.

19.Gongali H.C. “Structure and Process of organization” – Asia Publishing Huse, Bombay, 1964.

20.Gonguli S, Chaudhari S. and Guha S.C., “A motivation test for measuring entrepreneurial traits”, Management and Labour Studies, 1980, 44-49.

21.Gonguli T., “An Experimental Study of Workers morale and productivity Industrial Health in India”, Series No.77 – All India Institute of Hugience and public health, Calcutta.

22.Guha T.N., “Job Satisfaction among shoe factory workers.” Productivity 1965, P.no. 89 to 94.
23.Gonguli H.C., “Structure and Process of Organisation”, Asia Publishing House, Bombay, 1964.

24.Gupta, “Evelopment of Bank of rural development”, Deep & Deep Publication.

25.Herzberge F., “Job Work and the nature of man”, world publishing co., New York, 1966.
26.Hunt J.W., “managing People at Work”, McGrow Hill – 1986.

27.Jacob Mankidy – “Human Resource Management in Bank Contemporary Issue” – National Institute of Bank Management, Pune.

28.Jacob Mankidy – “Study of Profile and Report”, National Institute of Bank Management, Pune, 1996.

29.Jacob ankidy & Zarin Dastar. “Organisagtion in Conflict – A stricke in Andhra Bank”, National Institute of Bank Management Mumbai, 1976.

30.Jacob K.K. “Personnel Management in India” – Himanshu Publication, Udaipur.

31.James A.F., Stonner and Charles Wankel, “Management”, Printice Hall of India, New Delhi, 1986.

32.Koontz H. and O’Donnel C., “Principle of Management”, McGrow Hill, New York, 1959.
33.Luthan F., “Organisational behavior” third edition, McGrow Hill, 1985.

34.McGregor D. “The human side of Enterprise”, McGrow Hill New York, 1960.

Cloud Computing

Subject Code :- CC - 2

Credit :- 4 (Theory)

PRE-REQUISITES:: Computer networks, operating system, algorithms

OBJECTIVE:: This course covers a series of current cloud computing technologies, including technologies for Infrastructure as a Service, Platform as a Service, Software as a Service, and Physical Systems as a Service. The course is also highly project oriented, involving hand-on exploration of existing technologies as well as development of new technologies.

Syllabus

Introduction to Cloud Computing, History of Cloud Computing, Cloud service providers, Pros and Cons of Cloud Computing, Benefits of Cloud Computing, Cloud computing vs. Cluster computing vs. Grid computing, Cloud Computing Architecture, Service Models (XaaS), Infrastructure as a Service(IaaS), Platform as a Service(PaaS), Software as a Service(SaaS)., Deployment Models, Public cloud, Private cloud, Hybrid cloud, Community cloud, Cloud security, Case Study on Open Source & Commercial Clouds: Eucalyptus, Microsoft Azure, Amazon EC2, Selected topics from research papers

References

1. Cloud Computing (Wind) by Dr. Kumar Saurabh, 2nd Edison, Wiley India
2. Cloud Computing: Principles and Paradigms, Editors: Rajkumar Buyya, James Broberg, Andrzej M. Goscinski, Wile, 2011 Cloud Computing: Principles, Systems and Applications, Editors: Nikos Antonopoulos, Lee Gillam, Springer, 2012

OUTCOME OF THE COURSE:: Applying key comparative methodologies to assess the comparative advantages and disadvantages of public vs. private computing clouds Applying relevant methods to assess the important security and sustainability challenges involved in adopting various cloud architectures and making informed decisions for the organizations.

Data Warehousing & Data Mining

Subject Code :- DWDM – 3 (Elective -1)

Credit :- 4 (Theory)

Syllabus

Escalating need for strategies information, Failures of Past Decision-Supporting System, Operational Versus Decision-Supporting System, Data Warehousing- The only Viable Solution, data Warehouse Defined. The Building Blocks: Defining Features, Data Warehouse and Data Marts, Overview of the Components, Metadata in the Data Warehouse. Planning and Planning Management: Planning your Data Warehousing, The Data Warehouse Project, The project team, Project Management Considerations. Defining the Business Requirement: Dimension Analysis, Information Package- A New Concept, Requirements Gathering Methods, Requirements Definition: Scope and content. Requirements as the Driving force for Data Warehousing: Data Design, The Architectural Plan, Data Storage Specification, and Information Delivery Strategy., The Architectural Component: Understanding Data Warehouse Architecture, Distinguishing Characteristics, Architectural framework, Technical Architecture. Infrastructure as the Foundation for Data Warehousing: Infrastructure Support Architecture, Hardware Operational System, Database Software, Collection of Tools. The Significant Role of Metadata: Why Metadata is Important, Metadata Types by Functional Areas, Business Metadata, How to Provide Metadata. Principles of Dimensional Modeling: From Requirement to Data Design, The STAR Schema, STAR Schema keys, Advantages of STAR Schema. Dimensional Modeling: Updates to the Dimensional Tables, Miscellaneous Dimensions, The Snowflake Schema, Aggregate Fact Tables, and Families of STARS. Data Extraction, Transformation, and Loading. OLAP in the Data Warehouse: Demand for Online Analytical Processing, Major Features and Functions, OLAP Models, OLAP Implementation Consideration, Introduction : Data mining, kinds of data mined, kinds of patterns mined, technologies used: statistics, Machine learning, Database systems and Data Warehousing, Information Retrieval, Major issues in Data Mining: Mining methodology, User Interaction, Efficiency and Scalability, Diversity and database types, Data Mining & society, Data Preprocessing: Overview, Data cleaning, Data Integration, Data Reduction, Data Transformation, Data cleaning: Missing Values, Noisy data, Data cleaning as a process. Data Integration: Entity identification problem, Redundancy and Correlation Analysis, Tuple duplication, Data value conflict detection and Resolution. Data Reduction: Overview, wavelet transforms, Principle components Analysis, Attribute subset selection, Regression and log-linear models, Histograms, clustering, sampling, Data cube Aggregation. Data Transformation and Data Discretization by Binning, Discretization by Histogram Analysis, Discretization by cluster, Decision Tree and correlation Analysis, concept Hierarchy generation for Nominal data, Mining Frequent Patterns, Association and Correlations: Basic Concepts,

Frequent itemset Mining methods: Apriori Algorithm, Generate Association rules from Frequent itemsets, Improving the efficiency of Apriori, A pattern-growth approach for mining frequent itemsets, using frequent itemset using Vertical data format, Mining closed and max. patterns. Pattern Evaluation Methods, Advanced Pattern Mining: A Road map, Pattern mining in Multilevel, Multidimensional space, Constraint Based Frequent Mining, Classification: Basic Concepts, Decision Tree induction, Bayes Classification Method, Rule based Classification, Model evaluation & selection, techniques to improve classification accuracy. Classification Advanced Methods: Bayesian Belief networks, Classification by Back Propagation, Support Vector Method, Classification using frequent Patterns, lazy learners, other classification methods. Cluster Analysis: Basic Concepts & Methods, Cluster Analysis, partitioning methods, Hierarchical Methods, Density based Methods, Grid based Methods, Evaluation of Clustering. Advanced Cluster Analysis: Probabilistic Model based Clustering, Clustering High Dimensional Data, Clustering Graph & Network data, Clustering & Constraints.

References

1. Data warehousing fundamentals, first edition, paulraj ponniah, Wiley.

2. Data Mining Concept & Techniques, Jiawei Han|Micheline Kamber|Jian Pei, 3rd Edition, M K Publishers.

3. Data Warehousing in the real world, low price edition, Sam Anahory, Dennis Murray, Pearson Education.

4. Data warehousing Tool kit

5. Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management by Gordon S. Linoff and Michael J. Berry (Apr 12, 2011).

6. Data Mining: A Tutorial Based Primer by Richard Roiger and Michael Geatz (Oct 6, 2002).

Literature Review Survey

Subject Code :- LRS - 4

Credit :- Compulsory (No Credit) (Theory)

Recommendations to the Researchers

This work aims to provide a practical guide to assist students of Computer Science courses and related fields to conduct a systematic literature review. The steps proposed in this paper to conduct a systematic review were extracted from a technical report published by the researcher and arranged in a more objective format, in order to make information more accessible and practical, especially for those who are having their first contact with this technique. The target audience for this work is doctoral students that are in the initial phase of their bibliographic research.

Syllabus

Defining the main research questions of the literature review, Defining keywords, Defining search string, Defining search engines, String refinement, Search string execution, Download and store search results, Define inclusion and exclusion criteria, Selection of papers - First stage - Analysis by title and abstract, Selection of papers - Second stage - Analysis by Introduction and Conclusion (Optional), Selection of papers - Third stage - Complete reading and quality checklist, Extraction of answers related to research questions.

Data Science and Analytics

Subject Code :- DSA – 5 (Elective -2)

Credit :- 4 (Theory)

Recommendations to the Researchers

This class is meant to be interesting, and it's meant to help you unveil a completely new area of human knowledge, supporting the basic course on Data Analysis and Data Mining. It gives the opportunity to learn analytical skills and tools instead of only leveling coding skills. To anyone thinking about taking this class I would suggest the following: - Take it only if you are interested in learning something new - Be prepared to work - Be independent, and look for new, unusual solutions. - Do not miss/skip classes and homework. First, homework grades will be responsible for the bulk of your class grade. Second, each class is dedicated to a different area, and you do not want to miss any of them.

Syllabus

Introduction to data science, Exploratory data analysis, Introduction to machine learning, Linear regression and regularization, Model selection and evaluation, Classification: kNN, decision trees, Classification: SVM, Ensemble methods: random forests, Intro to probability: Naïve Bayes and logistic regression, Feature engineering and selection, Clustering: k-means, hierarchical clustering, Dimensionality reduction: PCA and SVD, Text mining and information retrieval, Network Analysis, Recommender systems, Relational databases, SQL, Big data storage and retrieval: noSQL, GraphDB, Big data distributed computing: map reduce, spark rdd, Advanced: neural networks and deep learning, Generalizing lecture, Presentations of final projects.

References :-

1. James, G., Witten, D., Hastie, T., Tibshirani, R. An introduction to statistical learning with applications in R. Springer, 2013.

2. Han, J., Kamber, M., Pei, J. Data mining concepts and techniques. Morgan Kaufmann, 2011.

3. Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning, 2nd edition. — Springer, 2009. 4. Murphy, K. Machine Learning: A Probabilistic Perspective. - MIT Press, 2012.