CENTRAL UNIVERSITY OF KERALA
केरल केन्द्रीय विश्वविद्यालय

DEPARTMENT OF COMPUTER SCIENCE

SCHOOL OF MATHEMATICAL AND PHYSICAL SCIENCES

Minutes of BOS in Computer Science Held on 09 July 2016 at 11.00 AM

Agenda: To discuss about the Syllabus, feedback of students, previous question papers, evaluation strategies

The following members were present during the meeting.

1. Dr. P. S. Hiremath, Professor, KLE Technological University
2. Dr. Rajesh R.
3. Dr. T.M. Thasleema
4. Mr. Ragesh N.K., Specialist, DSP & Multimedia, Tata Elxsi Ltd., Thiruvananthapuram
5. Mr. Fasil O.K., Software Engineer, NuCore Software Solutions

1) The BOS members have gone through the previous syllabus and current syllabus. The BOS observes the improvement in the curriculum/syllabus. The BOS members also suggested to include some industry related electives. The BOS approved the syllabus.

2) The feedback of 2014-16 batch students and 2015 admitted students were obtained. The BOS members has gone through the measures taken by the Faculty Council and approved the same.

3) The BOS members has gone through the previous question papers. The BOS members also verified (i) whether the question paper covers the entire syllabus, (ii) whether the question papers are upto the mark, (iii) whether the evaluation strategies of the answer papers are good. The BOS members were satisfied with procedures for the same.

Dr. P. S. Hiremath

Dr. Rajesh R.

Dr. T.M. Thasleema

Mr. Ragesh N.K.

Mr. Fasil O.K.

CORE COURSE				
COURSE CODE	**COURSE TITLE**	**CONTACT HRS/WEEK**		**CREDITS**
		LEC	**LAB**	**TUT**	
CSC5204	Data Mining	2	2	1	4

Lec = Lecture, Tut = Tutorial, Lab = Practical

This is an experimental, problem solving, **skill development course**.

Course Objective

   The objective of the course is to provide theoretical and practical aspects of data mining and design business rules for decision support systems.

By completing this course, students will obtain the following course/learning outcomes:
1. Knowledge gained:
    (i)   State-of-art pre and post data processing techniques and algorithms
2. Skill gained:
    (ii)   Extract knowledge using data mining techniques
    (iii) Adapt to new data mining tools
    (iv)   Ability to analyse the real world data mining problems
3. Competency gained:
    (v)   Development of data mining algorithms for real world problems.
    (vi)   Ability to participate in data challenges and to do higher order research

Prerequisites: Basic knowledge in algorithms.

Grading:

Lab experiments and implementation	– 15%
Mini project (individual)	– 10%
Class Test	- 10%
Assignment/Quiz/presentation	– 5%
Final Exam	– 60%

CSC5204 – Data Mining

Module 1
Introduction: Fundamentals of data mining, Data Mining Functionalities, Classification of Data Mining systems, Major issues in Data Mining, Data Preprocessing: Needs Preprocessing the Data, Data Cleaning, Data Integration and Transformation, Data Reduction, Discretization and Concept Hierarchy Generation. Some Considerations in Multi-Source Data Fusion.

Module 2
Data Mining Primitives, Languages, and System Architectures: Data Mining Primitives, Data Mining Query Languages, Architectures of Data Mining Systems. Mining data streams: The Stream Data Model, Sampling Data in a Stream, Filtering Streams, Counting Distinct Elements in a Stream, Estimating Moments, Counting Ones in a Window, Decaying Windows.

Module 3
Mining Association Rules in Large Databases: Association Rule Mining, Mining Single-Dimensional Boolean Association Rules from Transactional Databases, Dynamic Itemset Counting Algorithm, FP-Tree Growth Algorithm, Constraint-Based Association Mining. Handling large datasets in main memory, the limited pass algorithm, Counting frequent item sets in a stream.

Module 4
Classification and Prediction: Issues Regarding Classification and Prediction, Classification by Decision Tree Induction, Other Classification Methods, Prediction, Classifier Accuracy. Uncertain Knowledge Association Through Information Gain. Cluster Analysis Introduction: Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Density-Based Methods, Clustering High-Dimensional data, Constraint-based cluster analysis, Outlier Analysis, Mining Complex Types of Data: Mining Time-Series and Sequence Data, Mining Text Databases, Mining the World Wide Web.

References:
1. Jiawei Han, M. Kamber, Jian Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2nd Ed., 2005.
2. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, *Mining of Massive Data Sets*, Cambridge University Press, Second Edition, 2014.
3. Arun K Pujari, *Data Mining Techniques*, Universities Press, 2nd Ed., 2010.
4. Da Ruan, Guoqing Chen, Etienne E. Kerre, Geert Wets, *Intelligent Data Mining: Techniques and Applications* (Studies in Computational Intelligence), Springer, 1st Ed., 2010.
5. M. Mohammadian, *Intelligent Agents for Data Mining and Information Retrieval*, Idea Group Publishing, 2004.