DARBHANGA COLLEGE OF ENGINEERING

DARBHANGA

COURSE FILE OF DATA MINING
(05 1822)

DEPARTMENT
OF
COMPUTER SCIENCE AND ENGINEERING

FACULTY NAME
MR. SUNIL KR. SAHU
(Assistant Professor)

Institute/College Name:	Darbhanga College of Engineering
Program Name:	B.Tech (CSE, 8th semester)
Course Code:	051822
Course Name:	Data Mining
Lecture/Tutorial(per week):	3/1
Course Credits:	5
Course Co-coordinator Name:	Mr. Sunil Kumar Sahu

1. Scope and Objective of Course

Before starting a data mining project, it is essential to determine the tasks to be performed and properly manage allocation of tasks among individuals involved in the data analysis. Hence, planning is important as it results in effective data analysis.

1. Be familiar with the algorithms of data mining,
2. Be acquainted with the tools and techniques used for Knowledge Discovery in Databases.
3. Be exposed to web mining and text mining.

2. Textbooks

TB1: Data Mining and Data Warehouse by Udit Agarwal.
TB2: Data Mining. A tutorial-based Primer by Roiger, Michael W. Geatz and Pearson Education.
TB3: Data Mining Introductory & advanced topic by Margaret H. Dunham, Pearson Education

3. Reference Books

1. Data Mining: Next Generation Challenges and Future Direction by Kargupta, et al, PHI.
2. Data Warehousing, Data Mining & OLAP by Alex Berson Stephen J.Smith.

Other readings and relevant websites

S. No.	Link of journals, Magazines, websites and Research papers
1.	https://www.youtube.com/watch?v=Z6f9ckEElsU&list=PL8751DA481F0F0D17
2.	https://www.tutorialspoint.com/software_engineering/index.htm
3.	https://www.youtube.com/watch?v=m5c27rQtD2E
4.	https://www.youtube.com/watch?v=0hnqxIsXcy4

Course plans

Lecture No.	Date of Lecture	Topics	Web Links for Videos Lecture	Text Books/Reference books/Reading Materials	Page No. of Text Books
1-3	29/01/18 to 06/0218	**Introduction** Motivated Data Mining Data Mining on what kind of Data, Data Mining Functionalities, Classification of Data Mining System, Major issues in Data Mining.	https://www.youtube.com/watch?v=0hnqxlsXcy4	(TB2) Data Mining. A tutorial-based Primer by Roiger, Michael W. Geatz and Pearson Education.	(TB2)1-26
		Laboratory-1			
4-8	12/02/18 to 14/02/18	**Data Warehouse and OLAP Technology for Data Mining** Data Warehouse, Data Warehouse Architecture, Data Warehouse Implementation, Development of Data cube technology, Data Warehousing to Data Mining.	https://www.youtube.com/watch?v=m-aKj5ovDfg	(TB2) Data Mining. A tutorial-based Primer by Roiger, Michael W. Geatz and Pearson Education.	(Tb2)30-52
		Laboratory-2			
9-13		**Data Preprocessing** Data cleaning, Data Integration and Transformation, Data Reduction, Discrimination and concept Hierarchy Generation.	https://www.youtube.com/watch?v=m5c27rQtD2E	TB1: Data Mining and Data Warehouse by Udit Agarwal.	(Tb1)121-130
		Assignment-1, Laboratory-3			
14-18		**Data Mining Primitives, Primitives, Languages and System Architectures** Data Mining Primitives, Data Mining query language, Designing GUI on a Data Mining query language, Architectures of Data Mining System	https://www.youtube.com/watch?v=meWQLWq7QSE	TB1: Data Mining and Data Warehouse by Udit Agarwal.	(Tb1)131-174
		Laboratory-4			

19-25		**Mining Association rules in large database** Association rules mining, Mining single-dimensional Boolean Association rules from transaction database, mining multilevel Association rules from transaction database, Mining multidimensional Association rules from relational databases and Data warehouses, Association mining to correlation analysis, Constraint based association mining.	https://www.youtube.com/watch?v=itRkLa2kq6w	TB1: Data Mining and Data Warehouse by Udit Agarwal.	(Tb1)121-130
		Assignment-2, Laboratory-5			
26-31		**Classification and Prediction** What is classification and prediction, Issues regarding classification and prediction, Classification by decision tree Induction, Bayesian Classification, Classification by Back propagation, Classification based on concepts from association rule mining, Prediction, Classification accuracy.	https://www.youtube.com/watch?v=vuc93jbO2Dw	TB1: Data Mining and Data Warehouse by Udit Agarwal.	(Tb1)175-201
		Laboratory-6, 7 and 8.			
32-37		**Cluster Analysis** What is cluster analysis, Types of data in cluster analysis, A categorization of major clustering methods, Partitioning methods, Hierarchical Methods, Density based methods, Grid based methods, and Model based clustering methods.	https://www.youtube.com/watch?v=vuc93jbO2Dw	TB1: Data Mining and Data Warehouse by Udit Agarwal.	(Tb1)202-247
		Laboratory-9,10and 11.			
38-41		**Applications and trends in Data Mining** Data mining applications, Social impacts of Data Mining, Trends in Data Mining.	https://www.youtube.com/watch?v=8PJ24SrQqy8&list=PLYihddLF-CgYuWNL55Wg8ALkm6u8U7gps&index=7	TB1: Data Mining and Data Warehouse by Udit Agarwal.	(Tb1)115-120
		Assignment-3, Laboratory-12			

Syllabus

Topics	No. of Lectures	Weightages
Introduction: Motivated Data Mining Data Mining on what kind of Data, Data Mining Functionalities, Classification of Data Mining System, Major issues in Data Mining.	3	8%
Data Warehouse and OLAP Technology for Data Mining : Data Warehouse, Data Warehouse Architecture, Data Warehouse Implementation, Development of Data cube technology, Data Warehousing to Data Mining	5	15%
Data Preprocessing: Data cleaning, Data Integration and Transformation, Data Reduction, Discrimination and concept Hierarchy Generation.	4	10%
Data Mining Primitives, Primitives, Languages and System Architectures: Data Mining Primitives, Data Mining query language, Designing GUI on a Data Mining query language, Architectures of Data Mining System.	5	15%
Mining Association rules in large database : Association rules mining, Mining single-dimensional Boolean Association rules from transaction database, mining multilevel Association rules from transaction database, Mining multidimensional Association rules from relational databases and Data warehouses, Association mining to correlation analysis, Constraint based association mining.	7	12%
Classification and Prediction : What is classification and prediction, Issues regarding classification and prediction, Classification by decision tree Induction, Bayesian Classification, Classification by Back propagation, Classification based on concepts from association rule mining, Prediction, Classification accuracy**.**	6	15%
Cluster Analysis : What is cluster analysis, Types of data in cluster analysis, A categorization of major clustering methods, Partitioning methods, Hierarchical Methods, Density based methods, Grid based methods, Model based clustering methods	6	15%
Applications and trends in Data Mining: Data mining applications, Social impacts of Data Mining, Trends in Data Mining.	5	10%
Total	41	100%

Evaluation and Examination Blue Prints:

Internal assessment is done through quiz tests, presentations, assignments and projects work. Two sets of question paper are asked from each faculty and out of these two, without the knowledge of faculty, one question paper is chose for the concerned examination. Examination rules and regulations are uploaded on the student's portals. Evaluation is a very transparent process and the answer sheets of sessional tests, internal assessment assignments are returned back to the students.

The components of evaluation along with their weightage followed by the university is given below:

Component-1	Sessional test-1	25%
	Sessional test-2	
	Sessional test-3	
Component-2	Assignments, Quiz's, Test, Seminars	05%
Component-3	End Term Examination	70%
Totals		100%

Designation	**Name**	**Signature**
Course Coordinator	Mr. Sunil Kumar Sahu	
H.O.D	Dr._____	
Principal	Dr._____	
Date	…../…./…….	

DM LAB MANUAL

Institute/College Name:	Darbhanga College of Engineering
Program Name:	B.Tech (CSE, 8th semester)
Course Code:	051822
Course Name:	Data mining Lab
Lab (per week)	2
Course Credits:	3
Course Co-coordinator Name:	Mr. Sunil Kumar Sahu

Lab Objective

The Data mining Lab has been developed by keeping in mind the following objectives:

1. Be familiar with the algorithms of data mining,
2. Be acquainted with the tools and techniques used for Knowledge Discovery in Databases.
3. Be exposed to web mining and text mining.

Lab Outcome

At the end of the course, the student should be able to

1. Use Weka tools to demonstrate data mining concepts
2. Illustrate the algorithms of data mining
3. Analyze the performance of tools and techniques used for Knowledge Discovery in Databases.
4. Analyze various web mining and text mining Algorithms
5. Apply the Classification Algorithms for data mining applications
6. Apply the Clustering Algorithms for data mining applications

LAB MANUALS: -

S. No.	Experiment Details	Date	Signature & Date
1.	Creation of a Data Warehouse.		
2.	Implement of Apriori Algorithm.		
3.	Implement of FP-Growth Algorithm		
4.	Implement of K-means clustering.		
5.	One Hierarchical clustering algorithm		
6.	Bayesian Classification.		
7.	Decision Tree.		
8.	Support vector machines Classification		
9.	Applications of classification for web mining		
10.	Case Study on Text Mining or any commercial application		
11.			
12.			

Designation	Name	Signature
Course Coordinator	Mr. Sunil Kumar Sahu	
H.O.D	Dr._____	
Principal	Dr._____	
Date	…../…./…….	