MULTIMEDIA UNIVERSITY®

COURSE INFORMATION

1 .	**Name of Course**	Data Mining
2 .	**Course Code**	TDS3301
3 .	**Type of Course** (e.g. : Core, major, elective etc.)	Specialisation Core for BCS [DS] and elective for all other specializations
4 .	**Synopsis**	This course aims to develop individuals with skills in exploring huge amount of data to discover out of box patterns and hidden knowledge. The topics cover various state-of-art algorithms for data mining and also evaluation methods to assess the quality of the results yieled using the algorithms.
5 .	**Version** (State the date of theSenate's approval - previous and the current approval date)	Current: January 2018 Previous: June 2016
6 .	**Name(s) of Academic Staff**	Ting Choo Yee, Ho Chin Kuan, Foo Lee Kien, Chua Sook Ling @ Linda Chua, Khor
7 .	**Semester and Year Offered**	Trimester 1 (Delta)
8 .	**Credit Value**	4
9 .	**Pre-Requisite**	TCP1101 Programming Fundamentals

10 . **Objective of the course in the programme:**

To equip students with knowledge of fundamental concepts and techniques of data mining, including data preparation, mining frequent patterns, classification methods and cluster analysis. At the end of this course, students will have the ability to formulate and solve data mining problems.

11 . **Justification for including the course in the programme:**

To prepare students with the skills to perform data analysis and determine whether a real world problem has a data mining solution.

12 .

Course Learning Outcomes (CLO)		Domain	Level
CLO1:	Describe the data mining process and applications of data mining	Cognitive	2
CLO2:	Apply appropriate data pre-processing techniques to improve the quality of data	Cognitive	3
CLO3:	Apply appropriate data mining methods to discover useful patterns	Cognitive	3
CLO4:	Evaluate data mining results	Cognitive	4

13 . **Mapping of the Course Learning Outcomes to the Programme Learning Outcomes, Teaching Methods and Assessment:**

Course Learning Outcomes (CLO) (Must tally with CLOs in item 12)	Programme Learning Outcomes (PLO)												Teaching Methods	Assessment Method
	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8	PLO9	PLO10	PLO11	PLO12		
CLO1							✓						Lecture, Practical	Test
CLO2									✓				Lecture, Practical	Assignment, Lab Test, Final Exam
CLO3									✓				Lecture, Practical	Assignment, Test, Final Exam
CLO4									✓				Lecture, Practical	Assignment
Total							**1**		**3**				*Indicate the relevancy between the CLO and PLO by ticking "✓" the appropriate relevant box (This description must be read together with standards 2.1.2, 2.2.1, and 2.2.2 in Area 2 – pages 16 & 18 of COPPA 2.0)*	

14 . **Transferable Skills:**

Critical Thinking - developed via exploring data in assignment and report writing - assessed via presentation and report writing. Research - developed via empirical studies in assignment.

15 . **Distribution of Student Learning Time (SLT)**

	Course Content Outline	**CLO	Teaching and Learning Activities – Guided Learning (F2F)* *L	*T	*P	*O	Guided Learning (NF2F)*	Independent Learning (NF2F)*	Total SLT
1	**Introduction** Data mining concepts – what is data mining, what motivated data mining, why is data mining important; Types of data repositories on which data mining can be performed; Data mining process; Issues in data mining	**1**	2		2			4	8
2	**Data Preprocessing** Data cleaning – handling missing values and noisy data; Data integration – correlation analysis; Data transformation - normalization; Data reduction – attribute subset selection, dimensionality reduction, numerosity reduction	**2**	8		2		12	10	32
3	**Association Rule Mining** Market basket analysis; Mining frequent itemsets; Method for mining frequent patterns – the Apriori algorithm; Generating association rules; Evaluation of association rules	**3,4**	4		4			8	16
4	**Classification** Classification in data mining and example; Classification methods - decision tree learning, naïve Bayes classifier and artificial neural networks; Metrics for evaluating classifier performance; Evaluating the accuracy of a classifier	**3,4**	8		8			16	32

5	**Clustering** Clustering in data mining; Types of data in cluster analysis; Measures of proximity; Clustering methods - k-means and hierarchical methods; Clustering example; Cluster evaluation and validation; outlier analysis	**3,4**	4		4			8	16
6	**Applications and Trends in Data Mining** Data mining applications – financial, retail, telecommunication, biological data analysis, intrusion detection; Data mining tools; Privacy and social impacts of data mining; Trends in data mining	**1**	2		2			4	8
								Total SLT	**112**

SUMMATIVE ASSESSMENT

1. Continuous Assessment	**Percentage %**	**Total SLT**
Test	20%	5
Lab Test	10%	5
Assignments	30%	16
	Total SLT for Continuous Assessment	**26**

2. Final Assessment	**Percentage %**	**Total SLT**	
		F2F	**ILT**
Final Exam	40%	2	20
	Total SLT for Final Assessment (F2F + NF2F)	**22**	
Grand Total	**100%**	**160**	

****Indicate the CLO based on the CLO's numbering in Item 12.**
***L= Lecture, *T= Tutorial, *P= Practical, *O= Others, F2F*= Face to Face, NF2F*= Non Face to Face**

16 . **Identify Special Requirement to Deliver the Course (e.g., software, nursery, computer lab, simulation room):**
Computer lab

17 . **Main References:**
Ian Witten, Eibe Frank and Mark Hall (2016). Data Mining: Practical Machine Learning Tools and Techniques (4th Edition), Morgan Kaufmann Publishers.

18 . **Additional References:**
Jiawei Han, Micheline Kamber and Jian Pei (2012). Data Mining: Concepts and Techniques (3rd Edition), Morgan Kaufmann Publishers

Note:
Cells shaded light grey contain formulas / fixed values. Edit these formulas only if needed.