

1.	Name of Course/Module	Data Mining and Machine Learning Systems
2.	Course Code	TDM3341
3.	Status of Subject	Major for B. Sc Bioinformatics
4.	MQF Level/Stage	Bachelor Degree – MQF Level 6
5.	Version (state the date of the last Senate approval)	June 2012
6.	Requirement for Registration	HEM 2019 Probability and Bio-Statistics
7.	Name(s) of academic/teaching staff	Kalaiarasi SMA G. Mohamed Hanifa
8.	Semester and Year offered	Trimester 2 (Delta level)
9.	Objective of the course/module in the programme :	
	With the unprecedented rate at which data is being collected today in almost all fields of human endeavour, there is an emerging economic and scientific need to extract useful information from it. Data mining and machine learning systems provide to extract patterns, rules, changes, associations and anomalies in massive databases.	
10.	Learning Outcomes :	
	At the completion of the subject, students should be able to:	
	LO1: Define data mining, knowledge discovery process and advance mining concepts (Cognitive, Level 1)	
	LO2: Explain the data preparation process (Cognitive, Level 5)	
	LO3: Compare different type of data mining and machine learning algorithms (Cognitive, Level 6)	
	LO4: Design a simple data mining application (Psychomotor, Level 7)	
11.	Synopsis:	
	This course will provide main topics in data mining and machine learning, including: classification, clustering, association rules, sequence similarity, and so on. Emphasis will be laid on performance and implementation issues, as well as on applications such as web mining.	
	Kursus ini memberikan bahan-bahan pengenalan meliputi topik-topik utama dalam perlombongan data dan pembelajaran mesin, termasuk klasifikasi, pengkelompokan, undang-undang perkaitan, kesamaan jujukan, dan sebagainya. Penekanan diberikan ke atas isu-isu prestasi dan implementasi, dan juga aplikasi-aplikasi seperti perlombongan web.	
12.	Mapping of Subject to Programme Outcomes :	
	Programme Outcomes	% of Contribution
	PO1: Apply soft skills in work and career related activities	22.22
	PO7: Demonstrate knowledge and understanding of essential facts, concepts, principles, and theories relating to bioinformatics	33.33
	PO8: Apply principles and knowledge of bioinformatics in relevant areas	22.22

	PO9: Demonstrate the ability in analysing, modelling, designing, developing and evaluating computing solutions	22.22	
13.	Assessment Methods and Types :		
	Method and Type	Description/Details	Percentage
	Test		20%
	Quiz		10%
	Assignment	Report & Presentation	20%
	Final Exam		50%
14.	Details of Subject		
	Topics	Mode of Delivery	
		Lecture	Lab
	1. Introduction Data Mining Concepts, Input: Concepts, Instances, Attributes and Output: Knowledge Representation; Review of Graph Theory, Lattices, Probability and Statistics.	2	2
	2. Machine learning concepts and approaches Supervised learning: The supervised learning framework, concepts and hypothesis, training and learning, learning by enumeration, learning by construction; Learning Boolean functions: Boolean functions and formulae, monomials, disjunctive normal form and conjunctive normal form, a learning algorithm for monomials, learning disjunctions of small monomials.	2	2
	3. Data preparation Data cleaning; data integration and transformation; data reduction; discretization and concept hierarchy generation.	2	2
	4. Mining association rules Associations, maximal frequent item sets (ASCII); Closed frequent item sets; Covering Algorithms and Association Rules, Association Rules Mining, Linear Models, and Instance-Based Learning . Mining single-dimensional Boolean association rules from transactional databases; Mining multilevel association rules from transaction databases; Mining multidimensional association rules from relational databases and data warehouses; From association mining to correlation analysis; Constraint-based association mining.	5	5
5. Classification and prediction Issues regarding classification and prediction; Classification by decision tree induction; Bayesian classification; Classification by back propagation; Other classification methods, k-nearest neighbour classifiers, Case-based reasoning, Genetic algorithms, Rough set approach, Fuzzy set approaches; Prediction; Classifier accuracy.	5	5	

	6. Cluster analysis Types of data in clustering analysis; A categorization of major clustering methods; Hierarchical methods; Density-based methods; Grid-based methods; Model-based clustering methods; Outlier analysis.		5	5
	7. Mining complex types of data Multidimensional analysis and descriptive mining of complex data objects; Mining Spatial databases; Mining multimedia databases; Mining time-series and sequence data; Mining text databases; Mining the World-Wide Web.		5	5
	8. Data mining applications and trends in data mining Massive Datasets/Text mining, Case Studies, Agent-Based Mining.		2	2
	Total		28	28
15.	Tutorials / Lab			
	Using Matlab / Weka			
	<ul style="list-style-type: none"> • Data preparation • Associations Mining • Classification and prediction • Cluster analysis 			
16.	Total Student Learning Time (SLT)	Face to Face (Hour)	Total Guided and Independent Learning	
	Lecture	28	28	
	Tutorials			
	Laboratory/Practical	28	14	
	Presentation			
	Assignment	-	10	
	Mid Term Test	1	3	
	Final Exam	2	20	
	Quizzes	2 times	2	
	Sub Total	59	77	
	Total SLT	136/40 = 3.4 =>3		
17.	Credit Value	3		
18.	Reading Materials :			
	Textbook		Reference Materials	
	1. Jiawei Han and Micheline Kamber, Data Mining: Concepts and Techniques, 2nd edition, Morgan Kaufmann Publishers, 2005 (ISBN: 1-55860-489-8).		1. Ian H. Witten and Eibe Frank. Data Mining: practical machine learning tools and techniques with Java implementations, Morgan Kaufmann Publishers, San Francisco, CA, (2000).	
			2. Dorian Pyle, Data Preparation for Data Mining, Morgan Kaufmann, 1999 (1558605290).	
			3. Michael J. A. Berry, <i>et al.</i> Data Mining Techniques: For Marketing, Sales, and Customer Support, John Wiley, 1997.	

		<ol style="list-style-type: none"> 4. T.M.Mitchell, Machine Learning, McGraw-Hill, 1997. 5. M.Anthony and N. Biggs, Computational Learning Theory: An Introduction, Cambridge University Press, 1997.
19.	<p>Appendix (to be compiled when submitting the complete syllabus for the programme) :</p> <ol style="list-style-type: none"> 1. Mission and Vision of the University and Faculty 2. Mapping of Programme Objectives to Vision and Mission of Faculty and University 3. Mapping of Programme Outcome to Programme Objectives 4. Programme Objective and Outcomes (Measurement and Descriptions) 	