

SUMMARY OF INFORMATION ON EACH COURSE

| | | | | | | | | | | |
|-----|--|---|--|-----|-----|-----------|-----|------------|---------|-----|
| 1. | Name of Course | Data Mining and Machine Learning Systems | | | | | | | | |
| 2. | Course Code | TMM 3341 | | | | | | | | |
| 3. | Status of Course [Applies to (cohort)] | Specialisation core for B.Sc (Hons) Bioinformatics | | | | | | | | |
| 4. | MQF Level/Stage | Bachelor – MQF Level 6 | | | | | | | | |
| 5. | Version (State the date of the Senate approval – history of previous and current approval date) | Date of previous version: June 2014 Date of current version: November 2015 | | | | | | | | |
| 6. | Pre-Requisite | TMA1211 Discrete Mathematics and Probability | | | | | | | | |
| 7. | Name(s) of academic/teaching staff | Kalaiarasi SMA Tan Shing Chiang Cheah Wooi Ping | | | | | | | | |
| 8. | Semester and Year offered | Trimester 1, Year 3 | | | | | | | | |
| 9. | Objective of the course in the programme : With the unprecedented rate at which data is being collected today in almost all fields of human endeavor, 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. | Justification for including the course in the programme : 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 text mining. Students can further utilize the knowledge to classify biometric patterns, IDS ruleset, etc | | | | | | | | | |
| 11. | Course Learning Outcomes : | | Domain | | | Level | | | | |
| | 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 | | | | |
| 12. | LO4 | | Design a simple data mining application. | | | Cognitive | | | Level 6 | |
| | Mapping of Learning Outcomes to Programme Outcomes : | | | | | | | | | |
| | Learning Outcomes | PO1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 | PO8 | PO9 |
| | LO1 | X | | | | | | X | | |
| | LO2 | X | | | | | | X | | |
| LO3 | | | | | | | X | X | X | |
| LO4 | | | | | | | | X | X | |
| 13. | Assessment Methods and Types : | | | | | | | | | |
| | Method and Type | | Description/Details | | | | | Percentage | | |
| | 1 Mid Test | | Written test | | | | | 20 | | |
| | 2 Assignment | | Research papers and report | | | | | 20 | | |

SUMMARY OF INFORMATION ON EACH COURSE

| | | | | | | |
|-----|--|---------------------|-----|------------------|-----|-----|
| | 3 Quizzes | Written quizzes | | | | 10 |
| | 4 Final Examination | Written examination | | | | 50 |
| 14. | Mapping of assessment components to learning outcome (LOs) | | | | | |
| | Assessment Components | % | LO1 | LO2 | LO3 | LO4 |
| | 1 Mid Test | 20 | 20 | 29 | 20 | |
| | 2 Assignment | 20 | 20 | | 20 | 67 |
| | 3 Quizzes | 10 | 10 | | 10 | 33 |
| | 4 Final Examination | 50 | 50 | 71 | 50 | |
| 15. | Details of Course | | | | | |
| | Topics | | | Mode of Delivery | | |
| | | | | Lecture | Lab | |
| | Introduction Data Mining Concepts, Input: Concepts, Instances, Attributes and Output: Knowledge Representation; Probability and Statistics. | | | 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. | | | 2 | 2 | |
| | Data preparation Data cleaning; data integration and transformation; data reduction; discretization and concept hierarchy generation. | | | 2 | 2 | |
| | Mining association rules Associations, maximal frequent item sets (ASCI); 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 | 3 | |

SUMMARY OF INFORMATION ON EACH COURSE

| | | | |
|---|---|---------------------------------------|-----------------------------|
| 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, Genetic algorithms, Prediction; Classifier accuracy. | | 5 | 3 |
| 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 | 3 |
| 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 data; Mining the World Wide Web. | | 5 | 3 |
| Data mining applications and trends in data mining | | | |
| Massive Datasets/Text mining, Case Studies, Agent -Based Mining. | | 2 | 2 |
| Total | | 28 | 20 |
| | | | |
| | Total Student Learning Time (SLT) | Face to Face / Guided Learning | Independent Learning |
| | Lecture | 28 | 28 |
| | Tutorials | - | - |
| | Laboratory/Practical | 20 | 10 |
| | Presentation | 0 | 0 |
| | Assignment | 0 | 10 |
| | Mid Term Test | 1 | 3 |
| | Final Exam | 2 | 16 |
| | Quizzes | 2 times | 2 |
| | Sub Total | 51 | 69 |
| | Total SLT | 120 | |
| 16. | Credit Value | 3 | |
| 17. | Reading Materials : | | |
| | Textbooks | | |
| | Jiawei Han and Micheline Kamber, "Data Mining: Concepts and Techniques", 3rd Ed. Morgan Kaufmann Publishers, 2011 (ISBN-10: 0123814790) | | |

SUMMARY OF INFORMATION ON EACH COURSE

| | | |
|----------------|---|---|
| | | |
| | | Reference Material (including 'Statutes' for Law) |
| | | Ian H. Witten, Eibe Frank, Mark Hall. "Data Mining: practical machine learning tools and techniques", 3rd Edition, Morgan Kaufmann Publishers, 2011. |
| | | Dorian Pyle, "Data Preparation for Data Mining", Morgan Kaufmann, 1999. |
| | | Gordon Linoff and Michael J. A. Berry. "Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management", 3rd Edition, Wiley, 2011 |
| | Appendix (to be compiled when submitting the complete syllabus for the programme) : | |
| | <ol style="list-style-type: none"> 1. Mission and Vision of the University and Faculty 2. Programme Objectives or Programme Educational Objectives 3. Programme Outcomes (POs) 4. Mapping of POs to the 8 MQF domain 5. Summary of the Bloom's Taxonomy's Domain Coverage in all the Los in the format below : | |
| | Learning Outcomes (please state the learning Outcomes) | Bloom's Taxonomy Domain |
| Subject | | Affective Cognitive Psychomotor |
| TMM 3341 | Define data mining, knowledge discovery process and advance mining concepts | Level 1 |
| | Explain the data preparation process | Level 5 |
| | Compare different type of data mining and machine learning algorithms | Level 6 |
| | Design a simple data mining application | Level 6 |
| | <ol style="list-style-type: none"> 6. Summary of LO to PO measurement 7. Measurement and Tabulation of result for LO achievement 8. Measurement Tabulation of result for PO achievement | |