Subject Description Form

Subject Code	COMP5541
Subject Title	Machine Learning and Data Analytics
Credit Value	3
Level	5
Pre-requisite/ Co-requisite/ Exclusion	Nil
Objectives	The objectives of this subject are to: 1. present the principles, concepts and models of modern machine learning; 2. introduce analytics skills to analyze data and get insight from it.
Intended Learning Outcomes	Upon completion of the subject, students will be able to: a. understand the essential concepts of machine learning models and algorithms. b. develop insights into the effectiveness of machine learning techniques. c. gain knowledge of applying machine learning techniques to various cutting-edge applications. d. design machine learning solutions to solve new challenging problems in practice by considering different requirements and issues. e. participate in team work, presentation and technical writing.
Subject Synopsis/ Indicative Syllabus	• **Supervised Learning** Definitions, training data, dimensionality of input space, concept learning, supervised learning steps, choice of algorithms, bias-variance dilemma, chaining, conflict resolution, noise, variable types, SVM, k-NN, linear discriminant analysis, naïve Bayes, decision trees, neural networks and multi-layer perceptron, evaluation measures. • **Unsupervised Learning** Data types, dimensionality reduction, feature selection, k-means, mixture models, hierarchical clustering, anomaly detection, neural network-based approaches, Hebbian learning, deep belief networks, self-organizing map, latent variable models. • **Semi-supervised Learning** Supervised learning and unlabeled training data, continuity and manifold assumptions, generative models, graph-based methods, heuristic approaches and low-density separation. • **Reinforcement Learning** Definitions, algorithm for control learning, criterion of optimality, value function, directed policy search, deep reinforcement learning, inverse reinforcement learning, apprenticeship learning. • **Data Analytics Applications** Data types and pattern discovery in data using machine learning, selective applications in topic modeling, genomics, prediction, etc., selected issues like scalability, interpretability, legal/social/ethical issues.

<table>
<tr><td>Teaching/Learning Methodology</td><td>Lectures teach students on the main concepts of the course, together with comprehensive examples, and class questions and answers for easy understanding.
Tutorials and lab sessions help students to review the learned concepts, master the practical techniques and necessary tools for effective system/application development.
Group project offers the opportunity to students to develop analytical and problem solving skills through system implementation and interpersonal communication.</td></tr>
<tr><td>Assessment Methods in Alignment with Intended Learning Outcomes</td><td>
<table>
<tr><td rowspan="2">Specific assessment methods/tasks</td><td rowspan="2">% weighting</td><td colspan="5">Intended subject learning outcomes to be assessed (Please tick as appropriate)</td></tr>
<tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td></tr>
<tr><td>1. Assignments</td><td rowspan="3">55</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td></td></tr>
<tr><td>2. Project</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr>
<tr><td>3. Quiz</td><td>✓</td><td>✓</td><td>✓</td><td></td><td></td></tr>
<tr><td>4. Examination</td><td>45</td><td>✓</td><td>✓</td><td></td><td>✓</td><td></td></tr>
<tr><td>Total</td><td>100</td><td colspan="5"></td></tr>
</table>
Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes:

Continuous assessments consist of project, assignments, lab exercises, and quizzes, which are designed to facilitate students to achieve intended learning outcomes. Lab exercise is designed to encourage students to acquire good understanding of the relevant knowledge, practice in order to enrich their hands-on experience with various software tools. The project is designed to enhance students' ability to acquire the understanding and using different knowledge, principles, techniques, tools to solve a real problem through team. Quizzes are to ensure the students understand the concepts.

Examination will evaluate student's understanding and usage of machine learning and data analytics techniques.</td></tr>
<tr><td>Student Study Effort Expected</td><td>
<table>
<tr><td>Class contact:</td><td></td></tr>
<tr><td>▪ Class activities</td><td>39 Hrs.</td></tr>
<tr><td>Other student study effort:</td><td></td></tr>
<tr><td>▪ Assignments, Projects, Quizzes, Examination</td><td>66 Hrs.</td></tr>
<tr><td>Total student study effort</td><td>105 Hrs.</td></tr>
</table>
</td></tr>
</table>

Reading List and References	1. Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2006. 2. Ian Goodfellow, Deep Learning, MIT, 2016. 3. Jared Dean, Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners. Wiley, 2014. 4. EMC Education Services (Editor), Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, Wiley, 2015. 5. Leskovec, Rajaraman, Ullman, Mining of Massive Datasets, 2nd Ed., Cambridge University Press, 2014. 6. Han, J., and Kamber, M., 2011, Data Mining: Concepts and Techniques, 3rd Ed., Morgan Kaufmann, San Francisco, CA. 7. Tan, P.N., Steinbach, M., Kumar V., 2014, Introduction to Data Mining, 2nd Ed, Addison Wesley.