The Hong Kong Polytechnic University

Subject Description Form

Subject Code	CSE6016
Subject Title	Probability, Statistics, and Machine Learning for Engineers
Credit Value	3
Level	6
Pre-requisite / Co-requisite / Exclusion	Nil
Objectives	This subject is intended to introduce a broad range of statistical and machine learning methods for civil engineering applications; and equip students with knowledge of applying the skills learned to quantitatively model and solve real-world problems in civil engineering.
Intended Learning Outcomes	Upon completion of the subject, students will be able to: a. Appreciate the need for a quantitative approach for modelling different applications in civil engineering. b. Understand the conceptual foundations of probabilistic and statistical models, and their broad applications in civil engineering. c. Understand the mathematical foundations of different types of machine learning methods and can differentiate the suitability of each method in various applications. d. Demonstrate the capability to quantitatively model and solve real-world engineering problems using statistical and machine learning methods.
Subject Synopsis / Indicative Syllabus	1. Introduction (1 week) An overview of the statistical and machine learning methods in engineering, history, real-world applications, and future development 2. Probability and statistics (6 weeks) Fundamental concepts of probabilistic/statistical models including random variables, probability distributions, and count and discrete dependent variable models. Further application of the models to a variety of contexts in civil engineering such as risk and reliability analysis for designing infrastructure. 3. Machine learning (6 weeks) The machine learning basics, including the concept of learning (e.g., learnability, Occam's razor), supervised learning (e.g., perceptron, SVM, decision tree, linear regression, logistic regression), unsupervised learning (e.g., k-means, hierarchical clustering), reinforcement learning (e.g., Q-learning), and statistical learning (e.g., graphical models).

<table>
<tr><td>Teaching / Learning Methodology</td><td>A series of lectures and tutorials will be provided to introduce the principles of knowledge of the application of statistics and machine learning. Students will be required to relate the lecture materials with assignments and through homework, lab reports and quizzes. Reading and studying, e-Learning by students, including literature review and information searching, is required to achieve all the intended learning outcomes of the subject.</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="6">Intended subject learning outcomes to be assessed</td></tr>
<tr><td>a</td><td>b</td><td>c</td><td>d</td><td></td><td></td></tr>
<tr><td>Assignments</td><td>30</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td></td><td></td></tr>
<tr><td>Projects</td><td>30</td><td>✓</td><td></td><td></td><td>✓</td><td></td><td></td></tr>
<tr><td>Quizzes</td><td>40</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td></td><td></td></tr>
<tr><td>Total</td><td>100%</td><td colspan="6"></td></tr>
</table>

Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes:

Student will be assessed with continuous assessment. The continuous assessment consists of a set of assignments (30%), project (30%), and quizzes (40%).

Assignments: assessment of the studies with respect to the understanding of the relevant subject matters, including the principles, methodologies, and techniques, by providing answers to the assignment questions.

Projects: assessment of the ability to solve real-world problems by using learned techniques and developing practical solutions.

Quizzes: assessment of students' understanding of the relevant subject matters including the principles, concepts, methodologies, and techniques by proving answers to the questions.

A letter grading system will be used to assess students' performance. Students must attain at least a grade of D in the final grade to achieve a passing grade.</td></tr>
<tr><td>Student Study Effort Expected</td><td>
<table>
<tr><td>Class contact:</td><td></td></tr>
<tr><td>▪ Lectures / Tutorials</td><td>39 Hrs.</td></tr>
<tr><td>Other student study effort:</td><td></td></tr>
<tr><td>▪ Course work</td><td>39 Hrs.</td></tr>
<tr><td>▪ Reading and studying, e-Learning</td><td>39 Hrs.</td></tr>
<tr><td>Total student study effort</td><td>117 Hrs.</td></tr>
</table>
</td></tr>
<tr><td>Reading List and References</td><td>Reference Textbooks:
(1) Ang, A. and Tang, W.H. (2007) Probability Concepts in Engineering – Emphasis on Applications to Civil and Environmental Engineering, 2nd Edition, John Wiley & Sons.
(2) Washington, S.P., Karlatftis, M.D., Mannering, F.L., Anastasopoulos, P. (2020) Statistical and Econometric Methods for Transportation Data Analysis. 3rd Edition,</td></tr>
</table>

	Chapman & Hall/CRC. (3) Mitchell, T. (1997) Machine Learning. McGraw Hill. http://www.cs.cmu.edu/~tom/mlbook.html (4) Bishop, B. (2006) Pattern Recognition and Machine Learning. Springer.