

Lingnan University
Department of Computing and Decision Sciences
Course Syllabus

Course Title	:	Machine Learning
Course Code	:	CDS3003
Recommended Study Year	:	3
No. of Credits/Term	:	3
Mode of Tuition	:	Sectional Approach
Class Contact Hours	:	3 hours per week
Category in Major Prog.	:	Required
Prerequisite(s)	:	(a) CDS2001 Probability and Statistics 2, (b) CDS2002 Introduction to Artificial Intelligence, and (c) SSC2113 Linear Algebra
Co-requisite	:	Nil
Exclusion	:	Nil
Exemption Requirement	:	Nil

Brief Course Description:

This course is designed to introduce methods in machine learning (i.e., the study of how to create a machine that is capable of improving its performance in some tasks, such as perception, cognition and action, by learning from experience and data). This makes computers more intelligent. Machine learning is a powerful tool for recognising hidden and potentially useful relationships in data. It is essential in many organisations, as they can generate huge amounts of data in their daily operations and machine learning enables them to discover important knowledge from this data. The students will learn some powerful software packages for performing machine learning including generative artificial intelligence (AI) models and apply these packages to handle different kinds of problems.

Aims:

This course aims at delivering the concepts and applications of machine learning. The course discusses different categories of machine learning algorithms: Evolutionary; Connectionist; Symbolic; Bayesian; and Analogy approaches. Genetic programming, learning classifier systems, neural networks, support-vector machines, rule learning, inductive logic programming, probabilistic graphic models, and case-based learning will be discussed. The course will cover the supervised and unsupervised learning paradigms including generative AI models (e.g., generative adversarial networks, diffusion models, language models) as well as some representative algorithms (e.g. logistic regression, decision tree, and PCA). The basic concepts and applications of reinforcement learning will be discussed. The course also aims at introducing some Python packages for machine learning, such as scikit-learn, and enables students to apply different learning algorithms available in these packages. Moreover, the theories and techniques for model selection will be introduced throughout the course. The topic includes confusion matrix and bias-variance trade-off.

Learning Outcomes (LOs):

Upon the successful completion of this course, the student will be able to:

1. Identify different categories of machine learning algorithms;
2. Illustrate and explain supervised learning, unsupervised learning, and reinforcement learning (PLO5);
3. Describe and apply the theories and techniques for model selection;
4. Develop machine learning programs to solve some real-world problems (PLO9).

Indicative Contents:

Machine Learning

Basic concepts of machine learning

Different categories of learning algorithms

Evolutionary approaches: Genetic Programming, learning classifier systems;

Connectionist approaches: neural networks, self-organization maps

Symbolic approaches: Inductive Logic Programming

Bayesian approaches: Bayesian networks

Analogy approaches: Support Vector Machines, case-based learning

Supervised Learning

The concepts of supervised learning

Decision tree, logistic regression

Unsupervised Learning

The concepts of unsupervised learning including generative artificial intelligence
generative adversarial networks, diffusion models, language models

Reinforcement Learning

The basic concepts of reinforcement learning

Applications of reinforcement learning

Model Selection

Confusion matrix, F-measure, regularization, bias-variance trade-off

Teaching Method:

There are different teaching and learning activities including lectures, laboratories, and tutorials. The concepts of machine learning and different categories will be discussed in lectures. Students will also learn various supervised learning, unsupervised learning including generative AI models, reinforcement learning, and model selection methods in lectures. The implementations of some learning algorithms in Python, Java, R, or C++ will be discussed in laboratories. Moreover, students will also understand the implementations of different algorithms in some machine learning software packages such as scikit-learn, WEKA, ECJ, and Aleph. The detailed elaborations and computations of some learning algorithms will be discussed in tutorials.

Measurement of Learning Outcomes:

	Class Attendance and Participation	Assignments	Mid-term Examination	Final Examination
1. Identify different categories of machine learning algorithms			x	x
2. Illustrate and explain supervised learning, unsupervised learning, and reinforcement learning	x	x	x	x
3. Describe and apply the theories and techniques for model selection			x	x
4. Develop machine learning programs to solve some real-world problems		x		

1. There are a number of classroom activities to evaluate if the students can describe supervised learning, unsupervised learning, reinforcement learning, and model selection methods (LO2, and PLO5).
2. Written assignments require students to derive some parts of mathematical theory about the machine learning methods and/or perform related calculations (LO2 and PLO5).
3. Programming assignments require students to implement a prototype of machine learning system to solve some problems using computers. Students will acquire the IT skills related to machine learning (LO4 and PLO9).
4. Understanding of different types of machine learning methods and model selection techniques is comprehensively assessed in the examinations (LO1-3 and PLO5).

Assessment:

Class Attendance and Participation	5%
Assignments	45%
Mid-term Examination	20%
Final Examination	30%
Total	100%

Required/Essential Readings:

1. Domingos, Pedro. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books, 2015.
2. Shalev-Shwartz, S. and Ben-David, S. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge university press, 2014.
3. Foster, David. *Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play*. O'Reilly Media, 2019.

Recommended/Supplementary Readings:

1. Bishop, Christopher M., *Pattern Recognition and Machine Learning*, Springer, 2008.
2. Geron, Aurelien, *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, O'Reilly Media, 2017.
3. Hamel, Lutz H., *Knowledge Discovery with Support Vector Machines*, Wiley, 2009.
4. James, Gareth, Witten, Daniela, Hastie, Trevor, and Tibshirani, R., *An Introduction to Statistical Learning: with Applications in R*, Springer, 2013.
5. Kelleher, John D., Mac Namee, Brian, and D'Arcy, Aoife. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*, MIT Press, 2015.
6. Mitchell, Tom M., *Machine Learning*, McGraw-Hill, 1997.
7. Murphy, Kevin P., *Machine Learning: A Probabilistic Perspective*, MIT press, 2012.
8. Raschka, Sebastian and Mirjalili, Vahid, *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow, 2nd Edition*, Packt Publishing, 2017.
9. Sutton, Richard S. and Barto, Andrew G., *Reinforcement Learning: An Introduction*, MIT Press, 1998.
10. Brown, Tom B., et al. "Language Models are Few-Shot Learners." In Proceedings of the 2020 Conference on Neural Information Processing Systems (NeurIPS 2020).

Important Notes:

- (1) Students are expected to spend a total of 9 hours (i.e. 3 hours of class contact and 6 hours of personal study) per week to achieve the course learning outcomes.
- (2) Students shall be aware of the University regulations about dishonest practice in course work, tests and examinations, and the possible consequences as stipulated in the Regulations Governing University Examinations. In particular, plagiarism, being a kind of dishonest practice, is "the presentation of another person's work without proper acknowledgement of the source, including exact phrases, or summarised ideas, or even footnotes/citations, whether protected by copyright or not, as the student's own work". Students are required to strictly follow university regulations governing academic integrity and honesty.
- (3) Students are required to submit writing assignment(s) using Turnitin.
- (4) To enhance students' understanding of plagiarism, a mini-course "Online Tutorial on Plagiarism Awareness" is available on <https://pla.ln.edu.hk/>.

Rubric for Final Examination of CDS3003 - Machine Learning

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Identify different categories of machine learning algorithms (LO1)	Four or Five of the following points are achieved <ul style="list-style-type: none"> ▪ Identify Evolutionary approaches ▪ Identify Connectionist approaches ▪ Identify Symbolic approaches ▪ Identify Bayesian approaches ▪ Identify `Analogy approaches 	Two or Three of the following points are achieved <ul style="list-style-type: none"> ▪ Identify Evolutionary approaches ▪ Identify Connectionist approaches ▪ Identify Symbolic approaches ▪ Identify Bayesian approaches ▪ Identify Analogy approaches 	None or One of the following points are achieved <ul style="list-style-type: none"> ▪ Identify Evolutionary approaches ▪ Identify Connectionist approaches ▪ Identify Symbolic approaches ▪ Identify Bayesian approaches ▪ Identify Analogy approaches
Illustrate supervised learning, unsupervised learning, and reinforcement learning (LO2)	All of the following points are achieved <ul style="list-style-type: none"> ▪ Illustrate some supervised learning algorithms ▪ Illustrate some unsupervised learning algorithms ▪ Illustrate the basic concepts and some applications of reinforcement 	Two of the following points are achieved <ul style="list-style-type: none"> ▪ Illustrate some supervised learning algorithms ▪ Illustrate some unsupervised learning algorithms ▪ Illustrate the basic concepts and some applications of reinforcement 	None or One of the following points are achieved <ul style="list-style-type: none"> ▪ Illustrate some supervised learning algorithms ▪ Illustrate some unsupervised learning algorithms ▪ Illustrate the basic concepts and some applications of reinforcement
Describe the concepts of regularization and bias-variance trade-off (LO3)	All of the following points are achieved <ul style="list-style-type: none"> ▪ Describe the concept of regularization ▪ Describe the concept of bias-variance trade-off 	One of the following points are achieved <ul style="list-style-type: none"> ▪ Describe the concept of regularization ▪ Describe the concept of bias-variance trade-off 	None of the following points are achieved <ul style="list-style-type: none"> ▪ Describe the concept of regularization ▪ Describe the concept of bias-variance trade-off
Presentation	Content of submission/ presentation is well formatted with a clearly readable layout and no/very few grammatical/ formatting mistakes	Content of submission/ presentation is properly formatted with a reasonable layout and no more than a few grammatical/ formatting mistakes	Content of submission/ presentation is not properly formatted and/or there are more than a few grammatical/ formatting mistakes

Rubric for Mid-Term Examination of CDS3003 - Machine Learning

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Identify different categories of machine learning algorithms (LO1)	<p>All of the following points are consistently achieved</p> <ul style="list-style-type: none"> ▪ Identify Evolutionary approaches ▪ Identify Connectionist approaches 	<p>One of the following points are consistently achieved</p> <ul style="list-style-type: none"> ▪ Identify Evolutionary approaches ▪ Identify Connectionist approaches 	<p>None of the following points are achieved</p> <ul style="list-style-type: none"> ▪ Identify Evolutionary approaches ▪ Identify Connectionist approaches
Illustrate supervised learning, unsupervised learning, and reinforcement learning (LO2)	<p>All of the following points are consistently achieved</p> <ul style="list-style-type: none"> ▪ Illustrate some supervised learning algorithms ▪ Illustrate some unsupervised learning algorithms 	<p>One of the following points are consistently achieved</p> <ul style="list-style-type: none"> ▪ Illustrate some supervised learning algorithms ▪ Illustrate some unsupervised learning algorithms 	<p>None of the following points are achieved</p> <ul style="list-style-type: none"> ▪ Illustrate some supervised learning algorithms ▪ Illustrate some unsupervised learning algorithms
Apply the concept of confusion matrix (LO3)	The student can create different confusion matrix for some supervised learning problems. Moreover, all values are correct.	The student can create different confusion matrix for some supervised learning problems. However, some values are incorrect.	The student cannot create different confusion matrix for some supervised learning problems.
Apply the concepts of different measures (LO3)	The student can calculate the values of all measures. More than 90% of these values are correct.	The student can calculate the values of all measures. Between 60%- 89.99% of these values are correct.	The student cannot calculate the values of many measures. Less than 60% of these values are correct.
Presentation	Content of submission/ presentation is well formatted with a clearly readable layout and no/very few grammatical/ formatting mistakes	Content of submission/ presentation is properly formatted with a reasonable layout and no more than a few grammatical/ formatting mistakes	Content of submission/ presentation is not properly formatted and/or there are more than a few grammatical/ formatting mistakes

Rubric for Written Assignments of CDS3003 - Machine Learning

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Perform calculations that are related to the theory of Decision Tree. (LO2)	More than 80% of the calculations are correct.	Between 50% and 80% of the calculations are correct.	Less than 50% of the calculations are correct.
Perform calculations that are related to Logistic Regression. (LO2)	More than 80% of the calculations are correct.	Between 50% and 80% of the calculations are correct.	Less than 50% of the calculations are correct.
Derive some parts of mathematical theory about some unsupervised learning methods. (LO2)	The student correctly derives these parts of mathematical theory.	The student derives these parts of mathematical theory with some errors.	The student cannot derive these parts of mathematical theory.

Rubric for Programming Assignments of CDS3003 - Machine Learning

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Problem definition (demonstrate the understanding of the problem and consider alternative solutions) (LO4)	Clearly states the problem, lists related constrains, and identifies alternative solutions.	The problem is stated but related constrains and alternative solutions are not considered thoroughly.	The problem is marginally defined and with no consideration of constrains and alternative solutions.
Design of the machine learning system (LO4)	The design of the system is strongly related to the problem, The student can clearly explain the approaches and techniques involved. Moreover, the design of the system is very innovative.	The design of the system is related to the problem. The student does not clearly explain the approach and techniques involved. Moreover, the design of the system is not very innovative.	The design of the system is weakly related to the problem. The student does not clearly explain the approach and techniques involved. Moreover, the design of the system is not innovative.
Implementation of the machine learning system (LO4)	The machine learning system can be implemented correctly. The efficiency and effectiveness of the system are excellent.	The machine learning system is implemented with some minor errors. The efficiency and effectiveness of the system are acceptable.	The machine learning system cannot be implemented or the implementation(s) has too many errors. The system cannot be executed, or the efficiency and effectiveness of the system are poor.

Rubric for Class Attendance and Participation of CDS3003 Machine Learning

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Attendance	Full, punctual attendance in class and mandatory seminars	Occasional absences or lateness from class or mandatory seminars	Frequent or recurring absence or lateness from class or mandatory seminars
Class Participation	Active class participation and leadership in group activities.	Passive class participation and active in group activities.	Lack of participation and active disruption of class and group activities.
Explain supervised learning, unsupervised learning, and reinforcement learning (LO2)	The student can explain all of these learning strategies.	The student can explain some or most of these learning strategies. Some explanations are not completely correct and precise.	The student can only explain 0-1 of these learning strategies.