The Hong Kong University of Science and Technology (Guangzhou)

UG Course Syllabus

[Course Title] Machine Learning

[Course Code] DSAA 2011

[No. of Credits] 3

[Any pre-/co-requisites] DSAA 1001 or AIAA 2205

Name: Jiaheng WEI

Email: jiahengwei@hkust-gz.edu.cn

Office Hours: To be determined

Course Description

Machine learning is an exciting and fast-growing field that leverages data to build models which can make predictions or decisions. This is an introductory machine learning course that covers fundamental topics in model assessment and selection, supervised learning (e.g., linear regression, logistic regression, neural networks, deep learning, support vector machines, Bayes classifiers, decision trees, ensemble methods); unsupervised learning (e.g., clustering, dimensionality reduction); and reinforcement learning. Students will also gain practical programming skills in machine learning to tackle real-world problems.

Basic knowledge on mathematics (e.g., basic of probability theory, linear algebra, calculus and optimization), programming (e.g., Python/C++/Matlab) and data science are essential and will benefit the study of this course.

Intended Learning Outcomes (ILOs)

By the end of this course, students should be able to:

1. Understand a wide variety of machine learning algorithms and contrast different machine learning paradigms.

2. Understand how to evaluate machine learning models.

3. Explain the principles underlying a variety of machine learning algorithms.

4. Apply a variety of machine learning algorithms to solve real-world problems.

5. Evaluate and compare the performance of different machine learning algorithms.

Assessment and Grading

This course will be assessed using criterion-referencing and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided below, outlining the criteria used for evaluation.

Assessments:

This course will be assessed using criterion-referencing and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided below, outlining the criteria used for evaluation.

Assessment Task	Contribution to Overall Course grade (%)	Due date
Mid-Term	20%	To be determined
Written assignment	40%	To be determined
Final examination	40%	To be determined

* Assessment marks for individual assessed tasks will be released within two weeks of the due date.

Mapping of Course ILOs to Assessment Tasks

[add to/delete table as appropriate]

Assessed Task	Mapped ILOs	Explanation
Mid-Term Final examination	[ILO1, ILO2, ILO3, ILO4, ILO5]	These tasks test students' understanding of the taxonomy of machine learning algorithms (ILO 1) as well as the design of basic machine learning algorithms (ILO 3), and assess students' ability to choose the proper algorithms, apply them to tack real-life problems (ILO 4), and evaluate their performance with proper metrics and by comparison with other algorithms (ILO 2, ILO5).
Written assignment	[ILO2, ILO4, ILO5]	These tasks assess students' ability to choose the proper algorithms, apply them to tack real-life problems (ILO 4), and evaluate their performance with proper metrics and by comparison with other algorithms (ILO 2, ILO5). They can also improve their collaboration skills.

Grading Rubrics

Assessed Task	Rubrics
Mid-Term Final examination	Students need to show their understanding of the taxonomy of machine learning algorithms, the design of basic machine learning algorithms and evaluation of algorithms.
Written assignment	Students need to (i) finish individual assignment by themselves; (ii) collaborate to decompose the project into multiple sub-projects, formulate each sub-project as a proper machine learning problem, tackle the problems and evaluate the applied algorithms.

Final Grade Descriptors:

Grades	Short Description	Elaboration on subject grading description
A	Excellent Performance	Demonstrates a comprehensive grasp of machine learning, expertise in problem-solving, and significant creativity in thinking. Exhibits a high capacity for scholarship and collaboration, going beyond core requirements to achieve learning goals.
B	Good Performance	Shows good knowledge and understanding of machine learning, competence in problem-solving, and the ability to apply and evaluate machine learning algorithms. Displays high motivation to learn and the ability to work effectively with others.
C	Satisfactory Performance	Possesses adequate knowledge of core machine learning algorithms, competence in dealing with familiar problems, and some capacity for analysis and critical thinking. Shows persistence and effort to achieve broadly defined learning goals.
D	Marginal Pass	Has threshold knowledge of machine learning, potential to achieve key professional skills, and the ability to make basic evaluation of algorithms. Benefits from the course and has the potential to develop in the discipline.
F	Fail	Demonstrates insufficient understanding of machine learning and lacks the necessary problem-solving skills. Shows limited ability to think critically or analytically and exhibits minimal effort towards achieving learning goals. Does not meet the threshold requirements for professional practice or development in the discipline.

Course AI Policy

Students are allowed to use AI to help learn the knowledge but need to complete assessment tasks themselves.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on assignments will include [specific details, e.g., strengths, areas for improvement]. Students who have further questions about the feedback including marks should consult the instructor within five working days after the feedback is received.

Resubmission Policy

Assessment tasks are not allowed for resubmission after deadline for the fairness of all students.

Required Texts and Materials

There are no required materials. Students are encouraged to attend lectures and read some recommended materials.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST(GZ)'s Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to Regulations for Academic Integrity and Student Conduct for the University's definition of plagiarism and ways to avoid cheating and plagiarism.

[Optional] Additional Resources

Students are encouraged to familiarize themselves with following materials to obtain a good understanding of machine learning:

1. Book: 'Machine Learning' by Tom Mitchell (McGraw Hill, 1997)
2. `Pattern Recognition and Machine Learning' by Christopher M. Bishop
3. `Statistical Learning Methods' by Hang Li (《统计学习方法》-李航)
4. Book: 'Machine Learning' by Zhihua Zhou (2016) (《机器学习》-周志华)
5. Tutorial notes: CS189 provided by UC Berkeley
6. Tutorial notes: CS229 provided by Stanford University

You may need help with mathematical tools with following materials:

1. `The Matrix Cookbook' [http://matrixcookbook.com] by Kaare Brandt Petersen and Michael Syskind Pedersen
2. `A First Course in Probability' by Sheldon Ross
3. `Probability and Statistics' by Morris H. DeGroot and Mark J. Schervish
4. `Probability Theory' by Achim Klenke
5. `Introduction to Linear Algebra' by Gilbert Strang
6. `Convex Optimization' by Stephen Boyd and Lieven Vandenberghe
7. `Multivariable Calculus' by James Stewart

You may get help with some important Python packages:

1. Numpy: https://numpy.org/doc/stable/reference/
2. Pandas: https://pandas.pydata.org/docs/reference/index.html
3. Scikit: https://scikit-learn.org/stable/user_guide.html

You may need help with certain topics:

1. SVM

https://ocw.mit.edu/courses/6-867-machine-learning-fall-2006/resources/lec4/

2. model selection

Section 9.3 in https://cs229.stanford.edu/lectures-spring2022/main_notes.pdf