MA6630: INTRODUCTION TO STATISTICAL LEARNING

Effective Term

Semester B 2024/25

Part I Course Overview

Course Title

Introduction to Statistical Learning

Subject Code

MA - Mathematics

Course Number

6630

Academic Unit

Mathematics (MA)

College/School

College of Science (SI)

Course Duration

One Semester

Credit Units

3

Level

P5, P6 - Postgraduate Degree

Medium of Instruction

English

Medium of Assessment

English

Prerequisites

MA5617 Statistical Data Analysis;
OR a pass in all of the following courses
PH5102 Introduction to Biostatistics in One Health
BIOS5800 Probability
BIOS5801 Statistical Computing

Precursors

Nil

Equivalent Courses

Nil

Exclusive Courses

Nil

Part II Course Details

Abstract

Statistical learning is a new interdisciplinary area, which has connections to a variety of subjects including statistics, applied mathematics and computer sciences. It has been successfully applied in pattern recognition, signal processing, data mining, bioinformatics and financial engineering, etc. This course presents an overview of many cutting-edge techniques and algorithms in statistical learning. The covered topics include linear and nonlinear classification and regression, support vector machine, kernel methods, model averaging, boosting, as well as high-dimensional data. This course will provide the students the fundamental ideas and intuition behind modern statistical learning methods.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Understand basic concepts of statistical learning, and classic algorithms.	25	x	x	
2	Explain the fundamental principles and theories underlying the learning algorithms and generalize to various new applications.	25	x	x	
3	Implement the computer code and programming the statistical learning method and apply to real problems.	25	x	x	
4	Conduct mathematical derivation of the algorithms in the course.	25	x	x	x

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: Accomplishments
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

Learning and Teaching Activities (LTAs)

	LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Teaching	Learning through teaching is primarily based on lectures	1, 2, 3, 4	3 hours/week
2	Take-home assignments	Learning through take-home assignments helps students implement advanced theory for better understanding	1, 2, 3, 4	After-class

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Hand-in assignments	1, 2, 3, 4	20	
2	Project	1, 2, 3, 4	20	

Continuous Assessment (%)

40

Examination (%)

60

Examination Duration (Hours)

2

Additional Information for ATs

For a student to pass the course, at least 30% of the maximum mark for the examination must be obtained.

Assessment Rubrics (AR)

Assessment Task

1. Hand-in assignments (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

Problem solving based on comprehensive understanding

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of statistical learning models and algorithms and can always apply them to solve complex real problems

Good

(B+, B, B-) Adequately demonstrates an understanding of statistical learning models and algorithms and can usually apply them to solve real problems

Fair

(C+, C, C-) Demonstrates some understanding of statistical learning models and algorithms and can sometimes apply them to solve simple problems

Marginal

(D) Demonstrates limited understanding of statistical learning models and algorithms and can seldom apply them to solve simple problems

Failure

(F) Demonstrates little understanding of statistical learning models and algorithms and can rarely or never apply them to solve simple problems

---

Assessment Task

2. Project

Criterion

Real data analytic ability and software usage based on comprehensive understanding

Excellent

(A+, A, A-) Demonstrates a comprehensive understanding of statistical learning models and algorithms and has strong ability in applying R to solve complex problems

Good

(B+, B, B-) Demonstrates an understanding of statistical learning models and algorithms and has ability in applying R to solve relevant problems

Fair

(C+, C, C-) Demonstrates some understanding of statistical learning models and algorithms and has some ability to solve simple problems using R

Marginal

(D) Demonstrates limited understanding of statistical learning models and algorithms and has limited ability to solve simple problems using R

Failure

(F) Demonstrates little understanding of statistical learning models and algorithms and has little ability to solve simple problems using R

Assessment Task

3. Examination (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

Creativity and problem solving ability based on comprehensive understanding

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of statistical learning models and algorithms and can always apply them to solve complex real problems

Good

(B+, B, B-) Adequately demonstrates an understanding of statistical learning models and algorithms and can usually apply them to solve real problems

Fair

(C+, C, C-) Demonstrates some understanding of statistical learning models and algorithms and can sometimes apply them to solve simple problems

Marginal

(D) Demonstrates limited understanding of statistical learning models and algorithms and can seldom apply them to solve simple problems

Failure

(F) Demonstrates little understanding of statistical learning models and algorithms and can rarely or never apply them to solve simple problems

---

Assessment Task

1. Hand-in assignments (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

Problem solving based on comprehensive understanding

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of statistical learning models and algorithms and can always apply them to solve complex real problems

Good

(B+, B) Adequately demonstrates an understanding of statistical learning models and algorithms and can usually apply them to solve real problems

Marginal

(B-, C+, C) Demonstrates limited understanding of statistical learning models and algorithms and can seldom apply them to solve simple problems

Failure

(F) Demonstrates little understanding of statistical learning models and algorithms and can rarely or never apply them to solve simple problems

---

Assessment Task

2. Project (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

Real data analytic ability and software usage based on comprehensive understanding

Excellent

(A+, A, A-) Demonstrates a comprehensive understanding of statistical learning models and algorithms and has strong ability in applying R to solve complex problems

Good

(B+, B) Demonstrates an understanding of statistical learning models and algorithms and has ability in applying R to solve relevant problems

Marginal

(B-, C+, C) Demonstrates limited understanding of statistical learning models and algorithms and has limited ability to solve simple problems using R

Failure

(F) Demonstrates little understanding of statistical learning models and algorithms and has little ability to solve simple problems using R

Assessment Task

3. Examination (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

Creativity and problem solving ability based on comprehensive understanding

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of statistical learning models and algorithms and can always apply them to solve complex real problems

Good

(B+, B) Adequately demonstrates an understanding of statistical learning models and algorithms and can usually apply them to solve real problems

Marginal

(B-, C+, C) Demonstrates limited understanding of statistical learning models and algorithms and can seldom apply them to solve simple problems

Failure

(F) Demonstrates little understanding of statistical learning models and algorithms and can rarely or never apply them to solve simple problems

---

Part III Other Information

Keyword Syllabus

Concepts of learning; bias-variance trade-off; linear regression, ridge regression, Lasso regression; Logistic regression; Bayes rule; tree methods; boosting; cross-validation; regularisation; model averaging and bagging

Reading List

Compulsory Readings

	Title
1	An Introduction to Statistical Learning: with Application in R by James, Springer; 1st ed. 2013, Corr. 7th printing 2017 edition

Additional Readings

	Title
1	The Element of Statistical Learning by Hastie, Tibshirani and Friedman, Springer; 2nd edition
2	Pattern Recognition and Machine Learning by Bishop, Springer, 2006