MA4554: STATISTICAL FOUNDATIONS OF DATA SCIENCE

Effective Term

Semester A 2025/26

Part I Course Overview

Course Title

Statistical Foundations of Data Science

Subject Code

MA - Mathematics

Course Number

4554

Academic Unit

Mathematics (MA)

College/School

College of Science (SI)

Course Duration

One Semester

Credit Units

3

Level

B1, B2, B3, B4 - Bachelor's Degree

Medium of Instruction

English

Medium of Assessment

English

Prerequisites

MA2510 Probability and Statistics, and
MA3515 Introduction to Optimization

Precursors

Nil

Equivalent Courses

Nil

Exclusive Courses

Nil

Part II Course Details

Abstract

This course introduces the area of machine learning from a statistical perspective. Understanding modern machine learning algorithms is of huge importance, and requires a deep understanding of statistical theory and algorithms. The course will present the necessary mathematical tools required which are divided primarily into variational and (mostly) statistical methodologies. Examples of this include optimisation, which is based on the notion of convexity, where the students will be introduced to stochastic gradient methods. For the latter, the main focus of the statistical component is on the discussion of Markov chains and how one can use these entities to sample from probability distributions. Once the students become aware of this they will then discuss topics in statistical machine learning such as neural networks, deep learning, statistical learning theory and empirical risk minimization. Overall the course provides a solid introduction of statistics that underpins machine learning, from an analytical, methodological and application viewpoint.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Identify what is machine learning, and how mathematics & statistics underpins its foundations.	20	x	x	
2	Develop an understanding of variational and optimisation schemes that are fundamental, while being able to implement such algorithms.	20	x	x	
3	Acquire proficiency in sampling algorithms such as MCMC which are based on the theory of Markov chains.	20	x	x	x
4	Understand a range of applications in machine learning, such as classification and clustering, as well as in Bayesian context.	20	x	x	x
5	Utilize the statistical and variational perspectives for modern machine learning topics, which include neural networks, statistical learning theory and empirical risk minimisation and related topics.	20	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	Lecture	Learning through teaching is primarily based on lectures.	1, 2, 3, 4, 5	39 hours in total

2	Take-home assignments	Students will be recommended to solve simple maths exercises after each course, to understand in depth the mathematical techniques involved in the course.	1, 2, 3, 4, 5	after-class
3	Online applications	Students will be required to perform online tests of the main algorithms being taught, on images of their own.	1, 2, 3, 4, 5	after-class
4	Math Help Centre	Learning activities in Math Help Centre provides students extra help.	1, 2, 3, 4, 5	after-class

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks ("-" for nil entry)	Allow Use of GenAI?
1	Test	1, 2, 3, 4, 5	30	Questions are designed for the first part of the course to see how well the students have learned the basic algorithms, their theory, and link them to applications.	No
2	Experimental reports on online experiments	1, 2, 3, 4, 5	10	Delivery of an experimental report on the main algorithms presented in the course. Experiments will be performed online and the report will be handed in as a Python notebook.	Yes
3	Formative take-home maths exercises	1, 2, 3, 4, 5	0	The goal is to write down quick solutions of exercises given in the lecture notes. Students may be required to present orally their solution during the course.	Yes

Continuous Assessment (%)

40

Examination (%)
60

Examination Duration (Hours)
2

Minimum Examination Passing Requirement (%)
30

Additional Information for ATs
40% Coursework
60% Examination (Duration: 2 hours, at the end of the semester)
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. Test

Criterion
1.1 Understanding of the mathematical proofs

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
1. Test

Criterion
1.2 Ability to describe an image processing algorithm

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
1. Test

Criterion
1.3 Ability to comment the visual effects of an image processing algorithm

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
2. Experimental reports on online experiments

Criterion
2.1 Ability to choose adequate images for experiments

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
2. Experimental reports on online experiments

Criterion
2.2 Ability to explore quickly a new algorithm by adequate experiments

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
2. Experimental reports on online experiments

Criterion
2.3 Ability to detect and comment on visual defaults caused by algorithms

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
3. Examination

Criterion
3.1 Ability to make a variant of mathematical arguments seen in the course processing effect

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
3. Examination

Criterion
3.2 Ability to conceive and describe precisely an algorithm with a prescribed image processing effect

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
3. Examination

Criterion
3.3 Ability to analyse the visual content of an image and to link it to mathematical operators

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
4. Formative take-home maths exercises

Criterion
4.1 Pedagogical ability to describe orally the solution of an exercise to peers

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Part III Other Information

Keyword Syllabus

Fundamentals of machine learning will be discussed emphasising the importance of data, which will lead to brief discussion on supervised, unsupervised learning and common methods. These will include classification, clustering and regressions.

Introduction to optimisation theory: notion of convex, strong-convexity, KKT conditions etc.

Stochastic optimisation methods, motivated from the GD method. Algorithms that will be discussed include SGD, the heavy-ball method and convergence properties.

Sampling from probability distributions with Bayesian statistics:

-Bayesian statistics: understand the fundamental concepts of being Bayesian and its advantages.

- Monte Carlo algorithms: overview on MCMC algorithms such as the RWMH, MALA and the Gibbs sampler.

Understand key concepts of statistical deep learning:

-Neural networks: discussion of deep neural networks, Bayesian neural networks and theoretical properties such as the Universal Approximation Theorem.

- Statistical learning theory: Cover topics in VC dimension, robustness, statistical learning theory, and diffusion processes.

Reading List

Compulsory Readings

	Title
1	Complete lecture notes provided by the lecturer (can be updated during the course)
2	Complete slides provided by the lecturer (can be updated during the course)

Additional Readings

	Title
1	J.H. Friedman, R. Tibshirani and T. Hastie. The Elements of Statistical Learning, Springer, 2001.
2	G. Casella and C. Robert. Monte Carlo Statistical Methods, Springer, 1999.

3	L. Vandenberghe and S. P. Boyd. Convex Optimisation, CUP, 2004.
4	I. Goodfellow, Y., Bengio and A. Courville. Deep Learning, MIT Press, 2015.