



COMP8107

Statistical and Machine Learning Methods

Session 1, In person-scheduled-weekday, North Ryde 2025

School of Computing

Contents

<u>General Information</u>	2
<u>Learning Outcomes</u>	2
<u>General Assessment Information</u>	3
<u>Assessment Tasks</u>	4
<u>Delivery and Resources</u>	6
<u>Unit Schedule</u>	7
<u>Policies and Procedures</u>	8
<u>Changes from Previous Offering</u>	10

Disclaimer

Macquarie University has taken all reasonable measures to ensure the information in this publication is accurate and up-to-date. However, the information may change or become out-dated as a result of change in University policies, procedures or rules. The University reserves the right to make changes to any information in this publication without notice. Users of this publication are advised to check the website version of this publication [or the relevant faculty or department] before acting on any information in this publication.

General Information

Unit convenor and teaching staff

Convenor

Iris Jiang

iris.jiang@mq.edu.au

Credit points

10

Prerequisites

STAT6110 or STAT6191 or STAT8310 or (Admission to GradCertResFSE or GradDipResFSE)

Corequisites

Co-badged status

STAT8107

Unit description

This unit covers statistical learning techniques and machine learning (ML) algorithms for data analysis. Topics include loss functions, maximum likelihood, linear models, binary and multi-class classification, performance measures, and optimisation methods such as convexity, gradient descent, and stochastic gradient descent. Students will explore neural networks (shallow and deep), penalised regression (ridge and Lasso), and unsupervised learning techniques like K-means clustering, image segmentation, and principal components analysis. Other topics include partial least squares regression, decision trees, support vector machines, and non-parametric regression methods such as kernel and spline regression. The unit concludes with case studies, giving students a solid foundation in statistical learning and practical experience in applying ML algorithms to real-world problems.

Learning in this unit enhances student understanding of global challenges identified by the United Nations Sustainable Development Goals ([UNSDGs](#)) Industry, Innovation and Infrastructure

Important Academic Dates

Information about important academic dates including deadlines for withdrawing from units are available at <https://www.mq.edu.au/study/calendar-of-dates>

Learning Outcomes

On successful completion of this unit, you will be able to:

ULO1: Demonstrate an understanding of loss functions, maximum likelihood, linear models, and their applications in regression and classification tasks.

ULO2: Apply optimisation procedures such as gradient descent, and stochastic gradient descent to solve optimisation problems in machine learning

ULO3: Analyse the impact of colinearity and overfitting in linear models and implement penalised regression techniques such as ridge regression and the Lasso model

ULO4: Develop a practical understanding of neural networks, including both shallow and deep neural networks

ULO5: Apply unsupervised learning techniques such as clustering and dimension reduction to analyze complex datasets and develop insights into underlying patterns and relationships

ULO6: Utilise statistical learning techniques and state-of-the-art software tools to solve real-world problems.

General Assessment Information

Requirements to Pass this Unit

To pass this unit you must:

- Achieve a total mark equal to or greater than 50%.

Hurdle Assessments

There is no Hurdle Assessment.

Late Assessment Submission Penalty

Unless a Special Consideration request has been submitted and approved, a 5% penalty (of the total possible mark) will be applied each day a written assessment is not submitted, up until the 7th day (including weekends). After the 7th day, a grade of 0 will be awarded even if the assessment is submitted. Submission time for all written assessments is set at 11:55 pm. A 1-hour grace period is provided to students who experience a technical concern.

For any late submission of time-sensitive tasks, such as scheduled tests/exams, performance assessments/presentations, and/or scheduled practical assessments/labs, students need to submit an application for Special Consideration.

Assessments where Late Submissions will be accepted.

In this unit late submissions will be accepted as follows:

- Assignment 1 – YES, Standard Late Penalty applies;
- Case study/analysis – YES, Standard Late Penalty applies.
- Final Exam – NO, unless Special Consideration is granted.

Special Consideration

The Special Consideration Policy aims to support students who have been impacted by short-term circumstances or events that are serious, unavoidable and significantly disruptive, and which may affect their performance in assessment. If you experience circumstances or events that affect your ability to complete the written assessments in this unit on time, please inform the convenor and submit a Special Consideration request through <http://connect.mq.edu.au>.

Release Dates

1. Assignment 1 – To be released no later than Friday of Week 8 at 11:55pm AEDT
2. Case study/analysis – To be released no later than Friday of Week 12 at 11:55pm AEDT
3. Final Exam – To be released on Result Publication Date

Assessment Tasks

Name	Weighting	Hurdle	Due
Assignment 1	25%	No	Friday of Week 6 at 11:55pm AEDT
Case study	35%	No	Friday of Week 10 at 11:55pm AEDT
Final Exam	40%	No	Formal Examination Period

Assignment 1

Assessment Type [1](#): Qualitative analysis task

Indicative Time on Task [2](#): 20 hours

Due: **Friday of Week 6 at 11:55pm AEDT**

Weighting: **25%**

Written Report

On successful completion you will be able to:

- Demonstrate an understanding of loss functions, maximum likelihood, linear models, and their applications in regression and classification tasks.
- Apply optimisation procedures such as gradient descent, and stochastic gradient descent to solve optimisation problems in machine learning
- Utilise statistical learning techniques and state-of-the-art software tools to solve real-world problems.

Case study

Assessment Type ¹: Case study/analysis

Indicative Time on Task ²: 20 hours

Due: **Friday of Week 10 at 11:55pm AEDT**

Weighting: **35%**

An authentic case study with a scientific magazine-style article aimed at non-technical audience.

On successful completion you will be able to:

- Analyse the impact of colinearity and overfitting in linear models and implement penalised regression techniques such as ridge regression and the Lasso model
- Develop a practical understanding of neural networks, including both shallow and deep neural networks
- Apply unsupervised learning techniques such as clustering and dimension reduction to analyze complex datasets and develop insights into underlying patterns and relationships
- Utilise statistical learning techniques and state-of-the-art software tools to solve real-world problems.

Final Exam

Assessment Type ¹: Examination

Indicative Time on Task ²: 2 hours

Due: **Formal Examination Period**

Weighting: **40%**

An invigilated exam is to be scheduled in the university exam period.

On successful completion you will be able to:

- Demonstrate an understanding of loss functions, maximum likelihood, linear models, and their applications in regression and classification tasks.
- Apply optimisation procedures such as gradient descent, and stochastic gradient descent to solve optimisation problems in machine learning
- Analyse the impact of colinearity and overfitting in linear models and implement penalised regression techniques such as ridge regression and the Lasso model

- Develop a practical understanding of neural networks, including both shallow and deep neural networks
- Apply unsupervised learning techniques such as clustering and dimension reduction to analyze complex datasets and develop insights into underlying patterns and relationships

¹ If you need help with your assignment, please contact:

- the academic teaching staff in your unit for guidance in understanding or completing this type of assessment
- the [Writing Centre](#) for academic skills support.

² Indicative time-on-task is an estimate of the time required for completion of the assessment task and is subject to individual variation

Delivery and Resources

Classes

Lectures (beginning in Week 1): There is one two-hour lectures each week.

SGTA classes (beginning in Week 2): Students must register for one one-hour class per week.

The timetable for classes can be found on the University website at: <https://publish.mq.edu.au>

Enrolment can be managed using eStudent at: <https://students.mq.edu.au/support/technology/systems/estudent>

Suggested textbooks

The following book is useful as supplementary resources, for additional questions and explanations.

1. James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer.
2. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
3. Kuhn, M., & Silge, J. (2022). *Tidy modeling with R*. O'Reilly Media, Inc.

Technology Used and Required

This subject requires the use of the following computer software:

- R: R is a free statistical software package. Access and installation instructions may be found at: <https://www.r-project.org/>
- RStudio: RStudio is an open source tool that is used to manage and present work

performed using R. Access and installation instructions may be found at <https://rstudio.com/products/rstudio/download/>

- LaTeX: LaTeX is a free mathematical typesetting program. You should use this to help to typeset your assignment. Access and installation instructions may be found at: <https://www.latex-project.org/get/>

Communication

We will communicate with you via your university email or through announcements on iLearn. Queries to convenors can either be placed on the iLearn discussion forum or sent to your lecturers from your university email address.

COVID Information

For the latest information on the University's response to COVID-19, please refer to the Coronavirus infection page on the Macquarie website: <https://www.mq.edu.au/about/coronavirus-faqs>. Remember to check this page regularly in case the information and requirements change during semester. If there are any changes to this unit in relation to COVID, these will be communicated via iLearn.

Unit Schedule

This is a draft schedule and is subjected to change.

Week	Topics	
1	Introduction to Statistical Learning	
2	Linear methods for Regression	
3	Resampling and Model Selection	
4	Variable Selection and Regularisation	
5	From Probabilities to Decisions: Logistic Regression & Classification Metrics	
6	LDA: From Linear Decision Rules to Lower Dimensions	Assignment 1 due
7	Beyond Linear Boundaries: QDA & Naïve Bayes	
	Session 1 Break	
8	Trees and Forests	
9	Boosting	
10	Ensemble and Stacking	Case study/analysis due

11	Exploring Unlabeled Data: Principal Component Analysis and Clustering	
12	A Predictive Modelling Case Study	
13	Revision	

Policies and Procedures

Macquarie University policies and procedures are accessible from [Policy Central](https://policies.mq.edu.au) (<https://policies.mq.edu.au>). Students should be aware of the following policies in particular with regard to Learning and Teaching:

- [Academic Appeals Policy](#)
- [Academic Integrity Policy](#)
- [Academic Progression Policy](#)
- [Assessment Policy](#)
- [Fitness to Practice Procedure](#)
- [Assessment Procedure](#)
- [Complaints Resolution Procedure for Students and Members of the Public](#)
- [Special Consideration Policy](#)

Students seeking more policy resources can visit [Student Policies](https://students.mq.edu.au/support/study/policies) (<https://students.mq.edu.au/support/study/policies>). It is your one-stop-shop for the key policies you need to know about throughout your undergraduate student journey.

To find other policies relating to Teaching and Learning, visit [Policy Central](https://policies.mq.edu.au) (<https://policies.mq.edu.au>) and use the [search tool](#).

Student Code of Conduct

Macquarie University students have a responsibility to be familiar with the Student Code of Conduct: <https://students.mq.edu.au/admin/other-resources/student-conduct>

Results

Results published on platform other than [eStudent](#), (eg. iLearn, Coursera etc.) or released directly by your Unit Convenor, are not confirmed as they are subject to final approval by the University. Once approved, final results will be sent to your student email address and will be made available in [eStudent](#). For more information visit connect.mq.edu.au or if you are a Global MBA student contact globalmba.support@mq.edu.au

Academic Integrity

At Macquarie, we believe [academic integrity](#) – honesty, respect, trust, responsibility, fairness and courage – is at the core of learning, teaching and research. We recognise that meeting the expectations required to complete your assessments can be challenging. So, we offer you a range of resources and services to help you reach your potential, including free [online writing an](#)

[d maths support](#), [academic skills development](#) and [wellbeing consultations](#).

Student Support

Macquarie University provides a range of support services for students. For details, visit <http://students.mq.edu.au/support/>

The Writing Centre

[The Writing Centre](#) provides resources to develop your English language proficiency, academic writing, and communication skills.

- [Workshops](#)
- [Chat with a WriteWISE peer writing leader](#)
- [Access StudyWISE](#)
- [Upload an assignment to Studiosity](#)
- [Complete the Academic Integrity Module](#)

The Library provides online and face to face support to help you find and use relevant information resources.

- [Subject and Research Guides](#)
- [Ask a Librarian](#)

Student Services and Support

Macquarie University offers a range of [Student Support Services](#) including:

- [IT Support](#)
- [Accessibility and disability support](#) with study
- Mental health [support](#)
- [Safety support](#) to respond to bullying, harassment, sexual harassment and sexual assault
- [Social support including information about finances, tenancy and legal issues](#)
- [Student Advocacy](#) provides independent advice on MQ policies, procedures, and processes

Student Enquiries

Got a question? Ask us via the [Service Connect Portal](#), or contact [Service Connect](#).

IT Help

For help with University computer systems and technology, visit http://www.mq.edu.au/about_us/offices_and_units/information_technology/help/.

When using the University's IT, you must adhere to the [Acceptable Use of IT Resources Policy](#). The policy applies to all who connect to the MQ network including students.

Changes from Previous Offering

This unit has undergone a comprehensive and significant revision to better align with practical applications in predictive modelling. The content has been completely rewritten to emphasise hands-on experience, focusing on modern statistical learning techniques using *R*. The unit now integrates the *tidyverse* and *tidymodels* frameworks, providing students with a cohesive and streamlined approach to data manipulation, visualisation, and model building. This practical focus ensures that students not only understand theoretical concepts but also gain the skills necessary to apply predictive modelling techniques to real-world data problems.

Unit information based on version 2025.04 of the [Handbook](#)