



COMP8107

Statistical and Machine Learning Methods

Session 2, In person-scheduled-weekday, North Ryde 2026

School of Computing

Contents

<u>General Information</u>	2
<u>Learning Outcomes</u>	3
<u>General Assessment Information</u>	3
<u>Assessment Tasks</u>	4
<u>Delivery and Resources</u>	6
<u>Unit Schedule</u>	8
<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

Unit Convenor & Lecturer

Iris Jiang

iris.jiang@mq.edu.au

12WW 610

By appointment only. See iLearn for details.

Credit points

10

Prerequisites

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

Corequisites

Co-badged status

STAT8107

Unit description

This unit provides a comprehensive exploration of statistical learning, equipping students with both theoretical foundations and practical skills for analysing complex data. With a strong emphasis on the entire predictive modelling workflow, students learn to build, evaluate, and refine models using R, the tidyverse, and the tidy modelling framework. Key statistical learning concepts, such as the bias-variance trade-off and empirical risk minimisation, are thoroughly examined to develop a critical understanding of model performance and generalisation. Students gain hands-on experience in model assessment using resampling techniques and implementing diverse models for regression and classification. The unit also covers advanced non-linear methods, including decision trees, random forests, and boosting, alongside unsupervised learning techniques like principal component analysis, k-means, and hierarchical clustering. Through theoretical insights and real-world case studies, students are equipped to tackle complex data-driven challenges with confidence and rigour.

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 empirical risk minimisation and its applications in regression and classification.

ULO2: Apply effective resampling techniques for model development.

ULO3: Understand the impact of underfitting and overfitting on model generalisability to unseen data, and analyse strategies for mitigating overfitting through regularised regression and effective hyperparameter tuning.

ULO4: Implement advanced ensemble methods, including random forests and boosting, to enhance predictive performance in complex scenarios.

ULO5: Apply unsupervised learning techniques such as clustering and dimension reduction to analyse 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.

Participation

We strongly encourage all students to actively participate in all learning activities. Regular engagement is crucial for your success in this unit, as these activities provide opportunities to deepen your understanding of the material, collaborate with peers, and receive valuable feedback from instructors, to assist in completing the unit assessments. Your active participation not only enhances your own learning experience but also contributes to a vibrant and dynamic learning environment for everyone.

Release

- Assignment – To be released no later than 21st August.
- Case study – To be released no later than 2nd October.
- Final Exam – To be released on the scheduled final examination date.

Late Assessment Submission Penalty

Unless a Special Consideration request has been submitted and approved, a 5% penalty (of the total possible mark of the task) will be applied for each day a written report or presentation 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. The submission time for all uploaded assessments is **11:55 pm**. A 1-hour grace period will be provided to students who experience technical concerns. For any late submission of time-sensitive tasks, such as scheduled tests/exams, performance assessments/presentations, and/or scheduled practical assessments/labs, please apply for [Special Consideration](#). For example, if the assignment is worth 8 marks (of the entire unit) and your submission is late by 19 hours (or 23 hours 59 minutes 59 seconds), 0.4 marks (5% of 8 marks) will be deducted. If your submission is late by 24 hours (or 47 hours 59 minutes 59 seconds), 0.8 marks (10% of 8 marks) will be deducted, and so on.

Assessments where Late Submissions will be accepted

- Assignment – YES, Standard Late Penalty applies
- Case study – 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 assessments in this unit on time, please inform the convenor and submit a Special Consideration request through <http://connect.mq.edu.au/>.

Short Extension

For some assessment types, students are allowed to request an extra 3 calendar days to complete eligible assessments. No reason or evidence is required.

Assessment Tasks

Name	Weighting	Hurdle	Due	Groupwork/Individual	Short Extension	AI Approach
Assignment	25%	No	04/09/2026	Individual	Yes	Open
Case study	35%	No	16/10/2026	Individual	Yes	Open
Final Exam	40%	No	Exam Period	Individual	No	Observed

Assignment

Assessment Type ¹: Experiential task

Indicative Time on Task ²: 20 hours

Due: **04/09/2026**

Weighting: **25%**

Groupwork/Individual: Individual

Short extension ³: Yes

AI Approach: Open

Written Report

On successful completion you will be able to:

- Demonstrate an understanding of empirical risk minimisation and its applications in regression and classification.
- Apply effective resampling techniques for model development.
- Understand the impact of underfitting and overfitting on model generalisability to unseen data, and analyse strategies for mitigating overfitting through regularised regression and effective hyperparameter tuning.
- Utilise statistical learning techniques and state-of-the-art software tools to solve real-world problems.

Case study

Assessment Type ¹: Written Submission

Indicative Time on Task ²: 20 hours

Due: **16/10/2026**

Weighting: **35%**

Groupwork/Individual: Individual

Short extension ³: Yes

AI Approach: Open

An authentic case study where students turn real data into a magazine-style article, focusing on clear communication and storytelling for a non-technical audience.

On successful completion you will be able to:

- Demonstrate an understanding of empirical risk minimisation and its applications in regression and classification.
- Apply effective resampling techniques for model development.
- Understand the impact of underfitting and overfitting on model generalisability to unseen data, and analyse strategies for mitigating overfitting through regularised regression and effective hyperparameter tuning.
- Implement advanced ensemble methods, including random forests and boosting, to enhance predictive performance in complex scenarios.
- Apply unsupervised learning techniques such as clustering and dimension reduction to analyse 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: **Exam Period**

Weighting: **40%**

Groupwork/Individual: Individual

Short extension ³: No

AI Approach: Observed

You will undertake a final examination during the formal examination period.

On successful completion you will be able to:

- Demonstrate an understanding of empirical risk minimisation and its applications in regression and classification.
- Apply effective resampling techniques for model development.
- Understand the impact of underfitting and overfitting on model generalisability to unseen data, and analyse strategies for mitigating overfitting through regularised regression and effective hyperparameter tuning.
- Implement advanced ensemble methods, including random forests and boosting, to enhance predictive performance in complex scenarios.
- Apply unsupervised learning techniques such as clustering and dimension reduction to analyse 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
- [Academic Success](#) 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.

³ An automatic short extension is available for some assessments. Apply through the [Service Connect Portal](#).

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. <https://www.statlearning.com/>
2. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://hastie.su.domains/ElemStatLearn/>
3. Kuhn, M., & Silge, J. (2022). *Tidy modeling with R*. O'Reilly Media, Inc. <https://www.tmw.r.org/>

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/>
- **Quarto:** An open-source scientific and technical publishing system, installed by default with the latest release of RStudio.

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 the unit convenor via the contact email on iLearn.

Unit Schedule

Week	Topics	Assesment
1	Introduction to Statistical Learning	
2	Linear methods for Regression	
3	Resampling and Model Selection	
4	Variable Selection and Regularisation	
5	Exploratory Data Analysis & Assignment Q&A	
6	From Probabilities to Decisions: Logistic Regression & Classification Metrics	Assignment due
7	LDA: From Linear Decision Rules to Lower Dimensions	
8	Beyond Linear Boundaries: QDA & Naïve Bayes	
Session 2 Break		
9	Beyond Linear Boundaries: QDA & Naïve Bayes (Continued) & Case Study Q&A	
10	Trees and Tree-based Ensemble Methods	Case Study due
11	Trees and Tree-based Ensemble Methods (Continued)	
12	Exploring Unlabeled Data: Principal Component Analysis and Clustering	

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 and 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/>

Academic Success

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

- [Ask a Study Coach](#)
- [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 <https://students.mq.edu.au/support/technology/service-desk>.

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

We value student feedback to be able to continually improve the way we offer our units. As such we encourage students to provide constructive feedback via student surveys, to the teaching staff directly, or via the FSE Student Experience & Feedback link in the iLearn page.

Unit information based on version 2026.03 of the [Handbook](#)