STAT7107

Statistical Learning

Session 1, In person-scheduled-weekday, North Ryde 2024

School of Mathematical and Physical Sciences

Contents

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, Lecturer
Benoit Liquet-Weiland
benoit.liquet-weiland@mq.edu.au
Contact via email
630

Lecturer
Iris Jiang
iris.jiang@mq.edu.au
Contact via email

Credit points
10

Prerequisites
Admission to MRes

Corequisites

Co-badged status
STAT8107

Unit description
This unit introduces statistical learning techniques and machine learning (ML) algorithms for data analysis. The unit covers a wide range of topics, starting with the basics of loss functions, maximum likelihood, and linear models. Students will also learn about binary and multi-classification, performance measures, and optimisation procedures, including the notion of convexity, gradient descent, and stochastic gradient descent. Students will also explore neural networks, including both shallow and deep neural networks. The unit then moves on to penalised regression, including ridge regression and the Lasso model. Students will also learn about unsupervised learning techniques, including clustering using K-means, image segmentation, and dimension reduction via principal components analysis. Further, partial least square regression, decision tree learning and support vector machine are covered. Non-parametric regression includes kernel regression and spline regression are presented. The unit concludes with case studies. Students will develop a strong foundation in statistical learning techniques and gain practical experience in applying ML algorithms and statistical methods to solve real-world problems.

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 collinearity 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

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

General Faculty Policy on assessment submission deadlines and late submissions: .

- For any late submission of time-sensitive tasks, such as scheduled tests/exams, performance as- sessments/presentations, and/or scheduled practical assessments/labs, students need to submit an application for Special Consideration.
- Should these assessments be missed due to illness or misadventure, students should apply for Special Consideration.
- 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.

Assessments where Late Submissions will be accepted

- Assignment 1 - Yes, Standard Late Penalty applies
- Project - Yes, Standard Late Penalty applies
- Examination - NO, unless Special Consideration is granted

Special Consideration Policy:

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 ask.mq.edu.au.

Assessment Tasks

Name	Weighting	Hurdle	Due
Assignment 1	25%	No	Week 6
Project	35%	No	Week 10
Final Exam	40%	No	Formal Examination Period

Assignment 1

Assessment Type [1]: Quantitative analysis task
Indicative Time on Task [2]: 20 hours
Due: **Week 6**
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.

Project

Assessment Type [1]: Project
Indicative Time on Task [2]: 20 hours
Due: **Week 10**
Weighting: **35%**

An authentic project with a report aimed at a non-technical audience and a presentation to peers

On successful completion you will be able to:

- Analyse the impact of collinearity 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 collinearity 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 in and attend one two-hour class per week.

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

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

Suggested textbooks

Students should obtain the lecture overheads from iLearn prior to the lecture. The lecture overheads are available module by module.

The following are recommended reading books for this unit:

- *The Mathematical Engineering of Deep Learning.* Liquet B., Moka S. and Nazarathy Y. CRC Press, 2024 (https://deeplearningmath.org/)
- *Pattern Recognition and Machine Learning*, Bishop, Christopher M. 2006.
- *Machine Learning: A Probabilistic Perspective,* Kevin Murphy, MIT Press, 2012.
- Local regression and likelihood, C. Loader, Springer-Verlag, 1999. QA276.8 .L6/1999.

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/

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

Week 1: Statistics and ML (part1). Loss function, Maximum Likelihood, Linear model.

Week 2: Statistics and ML (part2). Logistic Model, Binary Classification, performance measure.

Week 3: Optimisation procedures. Convexity, Gradient Descent, Stochastic Gradient Descent

Week 4: Statistics and ML (part3). Multi-Classification, Softmax.

Week 5: Beyond linearity and overfitting. Polynomial model, Train/Validation/Test, cross validation

Week 6: Neural Network (part 1). Shallow neural Network and Deep neural network

Week 7: Neural Network (part 2). Deep neural network

Week 8: Penalized Regression. Collinearity, Ridge regression, Lasso Model.

Week 9: Unsupervised Learning. Clustering using Kmeans, Image segmentation, Hierarchical Clustering

Week 10: Unsupervised Learning. Dimension Reduction via Principal Components Analysis and auto-encoder

Week 11: Partial Least Square. PLS regression and PLS discriminant Analysis

Week 12: Decision tree learning. Classification tree and regression tree, Random Forest and Gradient Boosting tree

Week 13: Revision.

Policies and Procedures

Macquarie University policies and procedures are accessible from Policy Central (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). 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) 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 ask.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/

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 AskMQ, 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 is the first offering of this unit. 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 2024.02 of the Handbook