

# COMP3420 — Artificial Intelligence for Text and Vision

## Week 01 Lecture 1: Introduction and Overview

Diego Mollá

Department of Computer Science  
Macquarie University

COMP3420 2023H1

# Acknowledgement of Country

I acknowledge the traditional custodians of the Macquarie University land, the Wallumattagal clan of the Dharug nation, whose cultures and customs have nurtured, and continue to nurture, this land, since the Dreamtime. We pay our respects to Elders past, present and future.

# Welcome to COMP3420!

## Artificial Intelligence for Text and Vision

Availability of digital data in increasingly larger volumes, both as text and images, has enabled machine learning to provide effective solutions to applications that require intelligent processing of text and images. This unit explores the use of Artificial Intelligence techniques, in particular deep learning techniques, for tasks related to the processing of text and computer vision. Application areas include text search, sentiment analysis, information extraction, and image recognition.

<https://coursehandbook.mq.edu.au/2023/units/COMP3420/?year=2023>

# Welcome to COMP3420!

## Artificial Intelligence for Text and Vision

Availability of digital data in increasingly larger volumes, both as text and images, has enabled machine learning to provide effective solutions to applications that require intelligent processing of text and images. **This unit explores the use of Artificial Intelligence techniques, in particular deep learning techniques, for tasks related to the processing of text and computer vision.** Application areas include text search, sentiment analysis, information extraction, and image recognition.

<https://coursehandbook.mq.edu.au/2023/units/COMP3420/?year=2023>

# Programme

- 1 Artificial Intelligence for Text and Vision
- 2 Example Applications
- 3 Unit Practicalities

## Reading

- Lecture Notes
- Unit guide

# Programme

- 1 Artificial Intelligence for Text and Vision
- 2 Example Applications
- 3 Unit Practicalities

# Why AI for Text and Vision?

## Information Overload

- A great deal of digital information is available as text and images.
- People can understand text and images easily.
- But it's very hard to process by machines!



# Why Machine Learning for Text and Vision?

- This unit will focus on **machine learning** applications for text and vision.
- This is so because recent advances in machine learning have led to spectacular results in these fields.
- Our everyday life is affected by AI applications on text and vision.
- We will learn what is “under the hood” of these successful systems.



# Programme

- 1 Artificial Intelligence for Text and Vision
- 2 Example Applications
- 3 Unit Practicalities

# Output of ChatGPT (16 February 2023) I

“List and briefly describe 6 applications of AI for text”

Here are six applications of AI for text:

- 1 Sentiment Analysis: AI can be used to analyze large volumes of text, such as customer reviews or social media posts, to determine the sentiment of the content. This is useful for businesses to understand how customers feel about their products or services.
- 2 Language Translation: AI can be used to automatically translate text from one language to another, making it easier for people to communicate across different languages.

# Output of ChatGPT (16 February 2023) II

“List and briefly describe 6 applications of AI for text”

- 3 Chatbots: AI can be used to create chatbots that can interact with customers in natural language. This is useful for customer support, as chatbots can answer frequently asked questions and help customers find the information they need.
- 4 Text Summarization: AI can be used to summarize large volumes of text, such as news articles or research papers, to extract the most important information. This is useful for researchers and journalists who need to quickly digest large amounts of information.

# Output of ChatGPT (16 February 2023) III

“List and briefly describe 6 applications of AI for text”

- 5 Natural Language Processing: AI can be used to analyze and understand natural language text, such as emails or chat messages. This can be used to identify the intent of the message and extract important information, such as dates or locations.
- 6 Content Generation: AI can be used to generate content, such as news articles or marketing copy, based on a set of parameters or inputs. This can be useful for businesses that need to create large volumes of content quickly, or for individuals who want to create content but may not have the time or expertise to do so.

# Output of ChatGPT (16 February 2023) I

“List and briefly describe 6 applications of AI for vision”

Here are six applications of AI for vision:

- 1 Object Recognition: AI can be used to recognize and classify objects in images and videos, such as vehicles, people, and animals. This is useful for applications such as security cameras or self-driving cars.
- 2 Facial Recognition: AI can be used to recognize and identify individuals based on their facial features, which is useful for security and law enforcement applications.
- 3 Medical Image Analysis: AI can be used to analyze medical images, such as X-rays or MRI scans, to identify abnormalities or assist with diagnosis. This is useful for medical professionals who need to quickly analyze large volumes of images.

# Output of ChatGPT (16 February 2023) II

“List and briefly describe 6 applications of AI for vision”

- ④ Autonomous Vehicles: AI can be used to help autonomous vehicles navigate and avoid obstacles in real time, making them safer and more efficient.
- ⑤ Augmented Reality: AI can be used to create immersive augmented reality experiences by recognizing and tracking objects in real time. This is useful for applications such as gaming and marketing.
- ⑥ Quality Control: AI can be used to detect defects in manufacturing processes by analyzing images of products or parts. This is useful for improving product quality and reducing waste.

# Programme

- 1 Artificial Intelligence for Text and Vision
- 2 Example Applications
- 3 Unit Practicalities

# What This Unit is About

- COMP3420 explores AI technology for text and vision.
  - Text: Emphasis on text search and the use of machine learning for classification.
  - Vision: Emphasis on the use of machine learning for classification, segmentation, object recognition.
- Programming language: Python.
- This unit has the following prerequisite:
  - COMP2200 Data Science.



# Staff

**Diego Molla:** Unit convenor, Lecturer  
([diego.molla-aliod@mq.edu.au](mailto:diego.molla-aliod@mq.edu.au)).

**Greg Baker:** Lecturer ([gregory.baker2@hdr.mq.edu.au](mailto:gregory.baker2@hdr.mq.edu.au)).

**Abid Ali:** Tutor ([abid.ali@hdr.mq.edu.au](mailto:abid.ali@hdr.mq.edu.au)).

# Delivery

- Lectures:**
- On campus sessions on Monday 3-5pm: 14SCO T5.
  - Recordings will be available in iLearn / Echo360.

- Practicals:**
- Register to your 2-hour block.
  - These are in-campus sessions.
  - See [timetables.mq.edu.au/2023/](https://timetables.mq.edu.au/2023/).

## Please Note

- Practicals start from this week.
- Currently, only one practical class is open (Tuesday 4-6pm, 4RPD G12).

# Web Resources

- The unit is available in iLearn (<http://ilearn.mq.edu.au>).
- All the administrative material presented in this lecture is also available at this site.
  - Unit Outline.
  - Administrative Information.
  - Lecture Notes and recordings.
  - Pointers to Reading.
  - Other Useful Stuff.
- You are expected to keep up-to-date by using iLearn for:
  - Relevant news and information.
  - Discussions.
  - [Submission of assignments](#).

# GitHub

- Some of the material of this unit is available in a collection of GitHub repositories.
- <https://github.com/COMP3420-2023S1/repositories>
  - `public_material_2023S1`: Lectures, practicals, code.
  - `practical-ml-vision-book`: Fork of GitHub repository of Vision textbook.
- If you know how to use git, this will be the best way to make sure you have the latest versions.
- You can use the github browser interface to download individual files.
- We will also experiment with GitHub Classroom to submit and automark practical exercises and assignments.

# Textbooks

- Most of the contents will use these books:
  - “[NLTK Book](#)”: Steven Bird, Ewan Klein, Edward Loper. Natural Language Processing with Python — Analyzing Text with the Natural Language Toolkit.  
<http://www.nltk.org/book>
  - “[Deep Learning Book](#)”: François Chollet. Deep Learning with Python. (available in the library).
  - Dan Jurafsky, James H. Martin. Speech and Language Processing. 3rd ed. draft.  
<https://web.stanford.edu/~jurafsky/slp3/>
  - “[Vision Book](#)”: Valliappa Lakshmanan, Martin Görner, Ryan Gillard (2021), Practical Machine Learning for Computer Vision. O-Reilly. (Available in the library).
- Every week there will be [assigned readings](#); these readings are essential.
- The iLearn page also has pointers to online resources.
  - Recommendations for additions are welcome.

# Assessment

## Assessment Components

- Assignment 1: 10%, due Week 3.
- Assignment 2: 15%, due 2nd week of recess.
- Assignment 3: 25%, due Week 12.
- Exam: 50%, during the examination period.

## Final Assessment

- Your final mark and grade are entirely determined by the sum of marks of the individual assessment tasks.
- To pass the unit, the sum of marks must be at least 50% of the total assessment marks.
- This unit does not have hurdle assessments.

# Practical Assignments

- ① **Simple Computer Vision** (10%, due Week 3)
  - Use of pre-packaged tools.
  - Can be used as a **diagnostic test** (results will be out before census date).
- ② **Practical Computer Vision** (15%, due 2nd week of recess)
  - Use of techniques used in commercial and research applications.
  - Use of real (messy) image data.
- ③ **Text Processing** (20%, due Week 12)
  - Practical use of text processing with real (messy) text data.

# Submitting your Assignment

- Read the assignment specifications.
- Submit in iLearn (we will also experiment with GitHub Classroom).
- Late submission policy:
  - Late submissions will incur a penalty unless a Special Consideration request has been submitted and approved.
  - <https://students.mq.edu.au/study/assessment-exams/special-consideration>

## Plagiarism

- You may discuss but not write together.
- Read the Academic Integrity Policy.  
<https://policies.mq.edu.au/document/view.php?id=3>



# Tentative Lecture Schedule — Diego

- 1 Unit introduction; Python
- 2 Machine Learning for image classification
- 3 Convolutional networks for image classification
- 4 Popular convnet architectures
- 5 Object detection and image segmentation
- 6 Prelude to Advanced AI for Computer Vision

# Lecture Schedule — Greg

- 7 Simple text processing
- 8 Text search
- 9 Simple machine learning for text classification
- 10 Deep learning for text classification
- 11 Processing text sequences
- 12 Language models and text generation
- 13 Exam review

# Important Things To Do

- Download the lecture notes **before** attending the lecture.
- Read the practical exercises **before** attending the session.
  - time in the sessions is gold.
- Read the online Unit Guide; this is your “contract”.
- Schedule an average of 10 hours per week for working on this unit:
  - As in every 10-credit-point unit.
  - This includes the mid-semester break.
  - The total work load of all 10-credit points units is 150 hours.

# What's Next

## Week 1

- Image Processing in Python
- Workshop: Image Processing in Python