

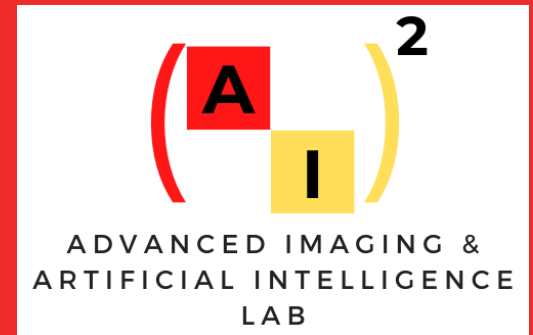
# Course Overview

---

## ENEN645/ENSF617 – Introduction to Machine Learning and Data Mining

Roberto Souza  
Associate Professor  
Electrical and Computer Engineering  
Schulich School of Engineering

W2026



# Course Delivery

---

- Synchronous and in-person
  - MF – 9:30 am to 10:45 am
  - Room ST135
- A Teams channel will be created for discussion and questions.

# Course Syllabus

---

**#1** Types of data mining: **classification, clustering, association, prediction.** **Processes:** data preparation, model building. Techniques: decision tree, **neural network**, evolutionary computing, Bayesian network. Applications: multi-media, text and web mining.

**#2** Overview of the historical context that allowed deep learning to flourish. Different types of **neural networks**; how to **train and deploy** them in various problems; Fine-tuning pre-trained models to achieve state-of-the-art results in **relevant applications**; overview of the current trends in deep learning.

# Course Syllabus (main topics)

---

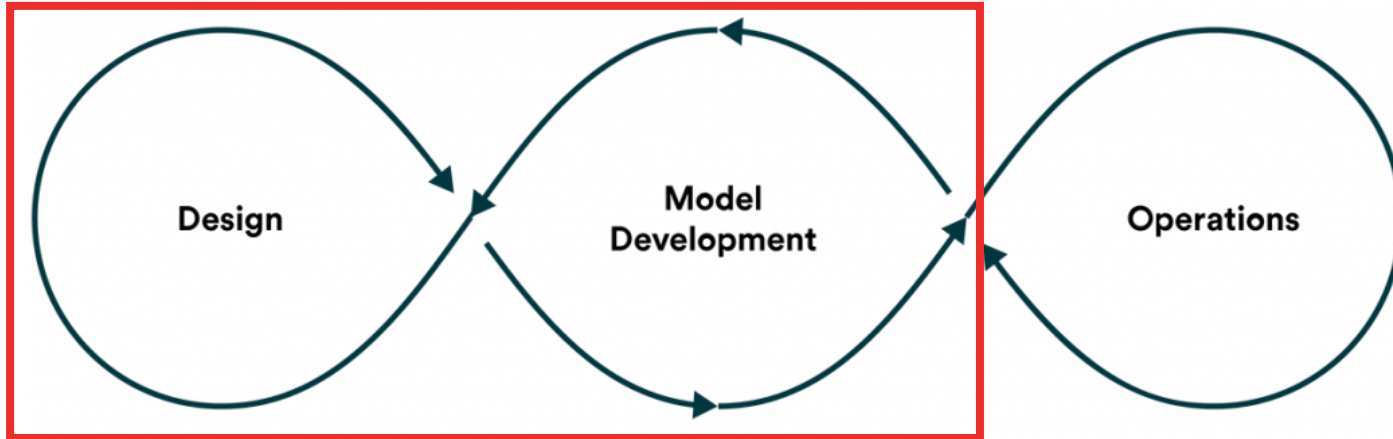
- Python Bootcamp and machine learning concepts
- Data preparation and pre-processing
- Regularization techniques
- Traditional machine learning models (Decision Trees, Random Forests, ...)
- Neural Networks
- Transfer Learning and Domain Adaptation
- Generative models
- Self-supervised learning
- Physics informed neural networks

# Learning Outcomes

---

1. Design and develop data mining and machine learning solutions for relevant problems
2. Select appropriate experimental setups and metrics for evaluating machine learning models
3. Select appropriate machine learning models for different types of problems
4. Have a comprehensive overview of current trends in machine learning
5. Acquire hands-on experience with machine learning programming frameworks

# What this course is about?



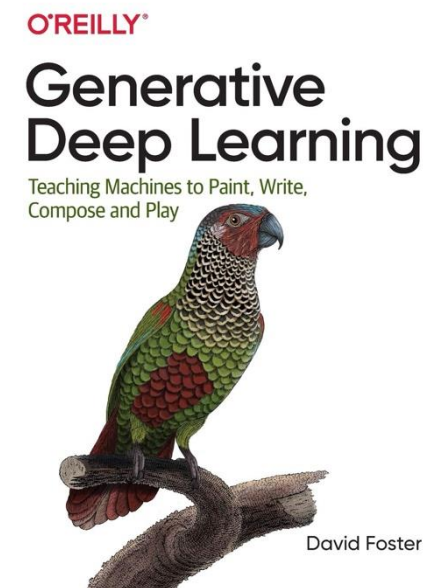
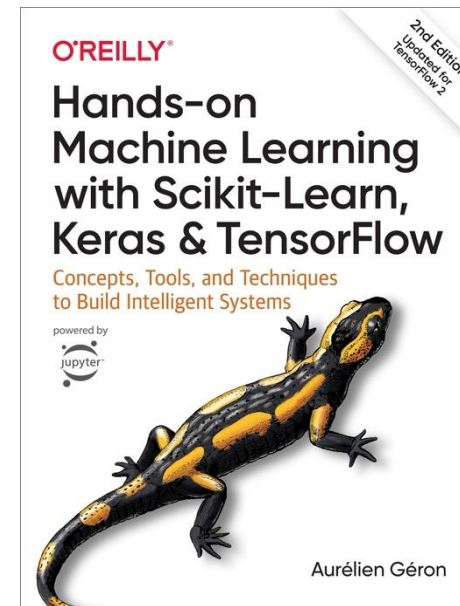
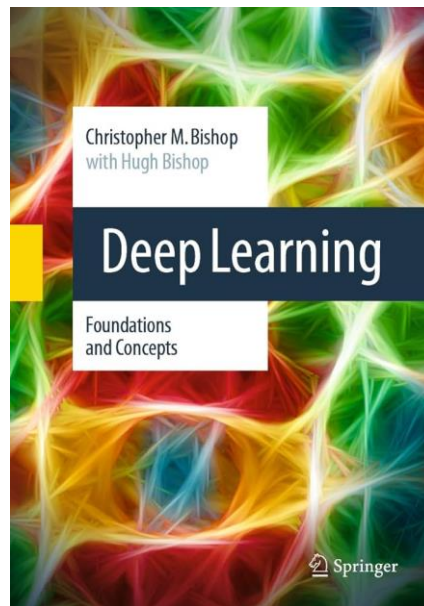
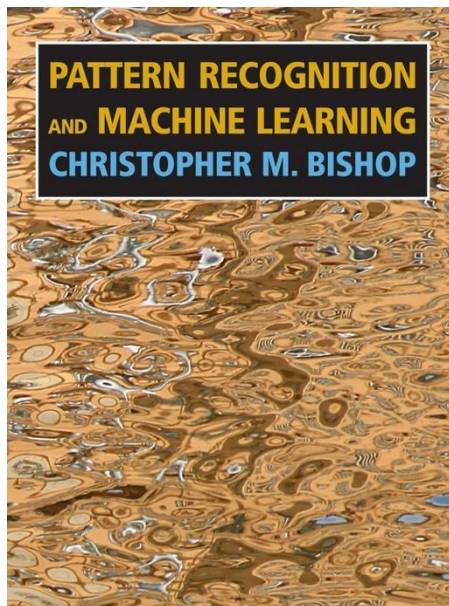
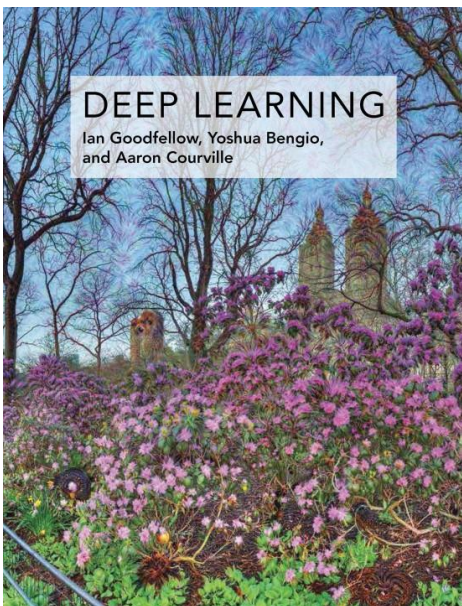
- Model deployment
- CI/CD pipelines
- Monitoring & triggering

**This course is about designing and developing machine learning models to achieve the best quantitative metrics\* to the problems being modelled.**

# Textbook

---

- No mandatory textbook for this course



# Course Assessment

---

Component	Learning Outcomes	Weight
Participation	1, 2, 3, 4	5%
Assignments (2)	1, 2, 3, 4, 5	20%
Midterm	1, 2, 3, 4	30%
Final Project	1, 2, 3, 4, 5	45%

# Participation (5%)

---

- Student participation will account for 5% of the final grade. How will participation be measured?
  - Students helping with proposed activities during class
  - Students' questions and answers during class
  - Students' participation on the Teams channel discussion board
  - Students helping each other during class

# Assignments (20%)

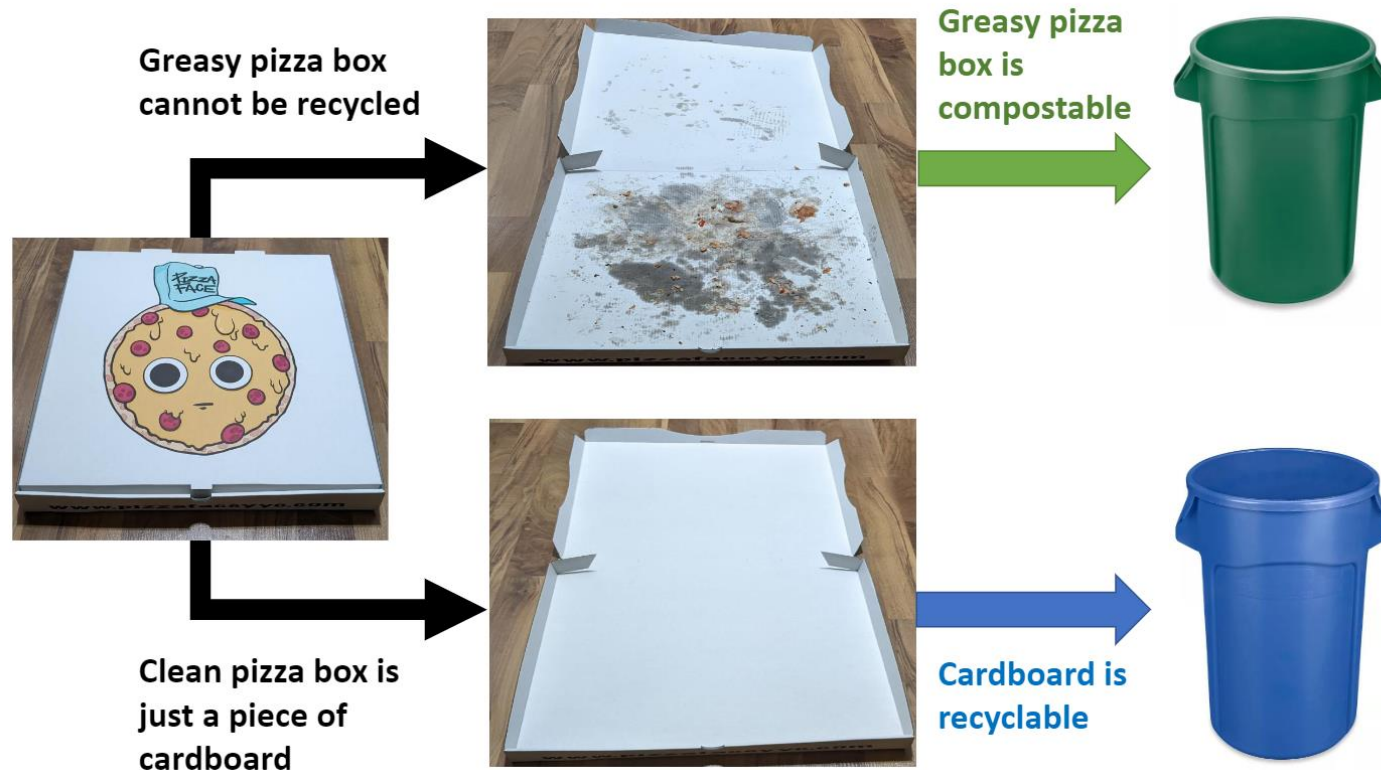
---

- Team-based – up to 4 people per team
- **Assignment 01 (10%):**
  - Proposing a garbage classification system based on images and natural language
  - **Due:** 30 January 2026 (midnight) | **Delivery method:** D2L dropbox
- **Assignment 02 (10%):**
  - Propose a solution for the problem of assignment 01
  - **Due:** 13 February 2026 (midnight) | **Delivery method:** D2L dropbox

# Assignment 1

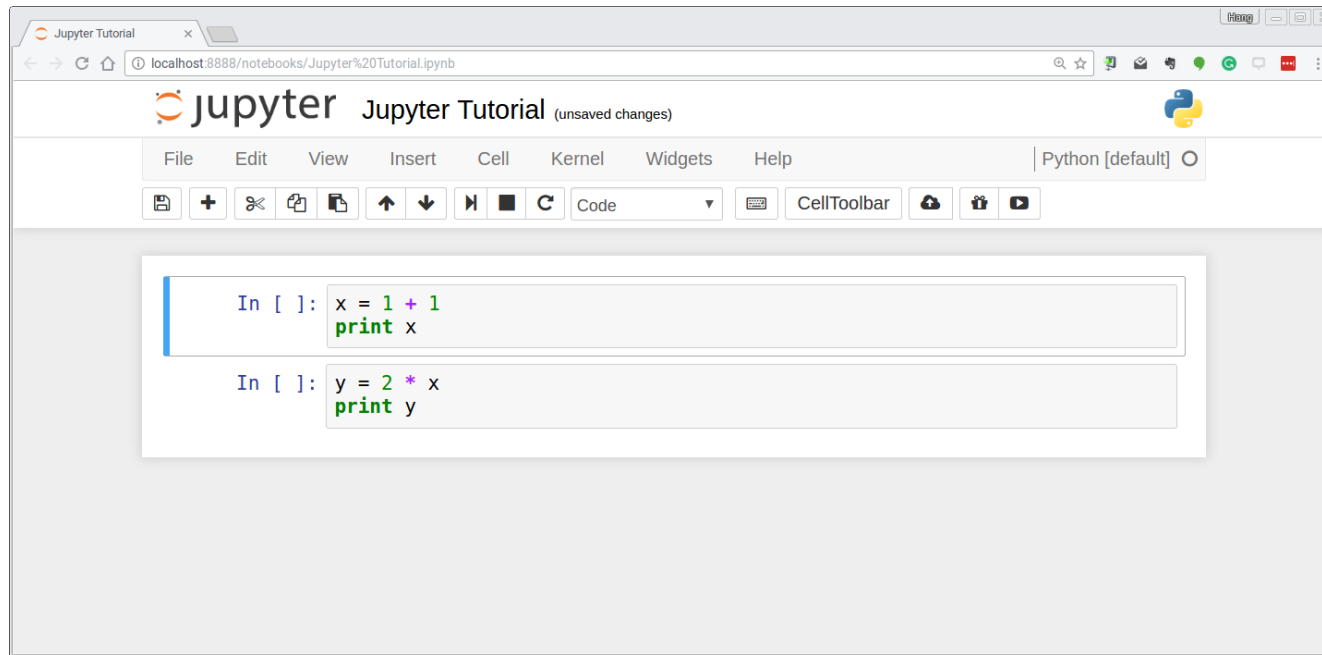
---

- Design a garbage classification system based on images and text
- Use text to add context potentially not available in the image



# Assignment 02

---



- Develop a deep learning solution to the garbage classification problem used in assignment 1, using both the images and text as inputs (data will be provided)

# Midterm (30%)

---

- Quizzes are individual
- Multiple choice and subjective questions
- A sample quiz will be provided for studying
- Content: all topics covered until the class preceding the midterm
  
- **Date:** 23 February 2026 in the classroom
- **Accommodation:** 06 March 2026 in the classroom

# Final Project (45%)

---

- 4-page report + 1 additional page only with references (if necessary)
- The report template in MS Word and Overleaf is available in the class GitHub repository
- Final report due date: **08 April 2026 at midnight**
- Final project presentations: **10 and 13 April 2026**
  - Presentation duration TBD based on number of teams that enroll to present
  - Send slides one day before your presentation
  - Presentations are not graded, but can help raise your final report grade

# Grades

---

Letter Grade	Total Mark (T)
A+	$T \geq 95\%$
A	$90\% \leq T < 95\%$
A-	$85\% \leq T < 90\%$
B+	$80\% \leq T < 85\%$
B	$75\% \leq T < 80\%$
B-	$70\% \leq T < 75\%$
C+	$65\% \leq T < 70\%$
C	$60\% \leq T < 65\%$
C-	$55\% \leq T < 60\%$
D+	$50\% \leq T < 55\%$
D	$45\% \leq T < 50\%$
F	$T < 45\%$

# The Programming Environment (Part 1)

---



<https://colab.research.google.com/>



<https://jupyter.org/>



<https://github.com/rmsouza01/ENSF617-ENEN-645-W2026>



<https://www.overleaf.com/project>



**UNIVERSITY OF CALGARY**  
Research Computing Services

[https://rcs.ucalgary.ca/index.php/RCS\\_Home\\_Page](https://rcs.ucalgary.ca/index.php/RCS_Home_Page)

# The Programming Environment (Part 2)

---

- Python 3
- Python libraries:
  - NumPy
  - Matplotlib
  - Pandas
  - Scikit-learn
  - Tensorflow ( version  $\geq 2.0$ )
  - PyTorch
  - Weight and Bias
- Please have your programming environment in your computer or on Google Colab set up asap

# Deep Learning Frameworks

---



**I hope you enjoy the  
class 😊**

# Questions?