


# MIS710 Machine Learning in Business

## Lab 1

Associate Professor Lemai Nguyen





**You should complete Lab 0 Intro Exercises  
PRIOR TO  
coming to the class in Week 1!**

# Objectives

1. To set up and familiarise with the Google Colab environment for MIS710
2. Build up Python programming basic knowledge and skills
  - Data values and variables
  - Data types
  - Data structures
  - Use built-in functions
  - Write for loop
  - Write conditional statements
  - Write comments

# Useful Python libraries

- NumPy is a Python library to work with multi-dimensional arrays and matrices, and mathematical functions to operate on arrays

The NumPy array is the fundamental data structure in scikit-learn.

Data inside an array must be of the same type.

- pandas is a Python library for data manipulation and analysis on DataFrame tables.

DataFrame columns can be of different types.

- matplotlib is a Python library with functions to produce visualisations such as line charts, histograms, scatter plots...
- seaborn is Python library based on matplotlib with a high-level interface for drawing attractive and informative statistical graphics
- **LabelEncoder** we will use it to fit label encoder and return encoded labels for data preparation

# Lab 1: Python exercises

- Access Google Colab <https://colab.research.google.com/>
- Select Google drive to work with Colab Notebooks
- Create a folder MIS710 under Colab Notebooks folder
- Download MIS710 Lab 1.ipynb and the titanic dataset titanic\_train.csv from the unit site in the CloudDeakin system
- Upload them to the MIS710 folder on your Google drive
- Open MIS710 Lab 1.ipynb
- Follow the instructions
- Don't stay behind!

# **MIS710**

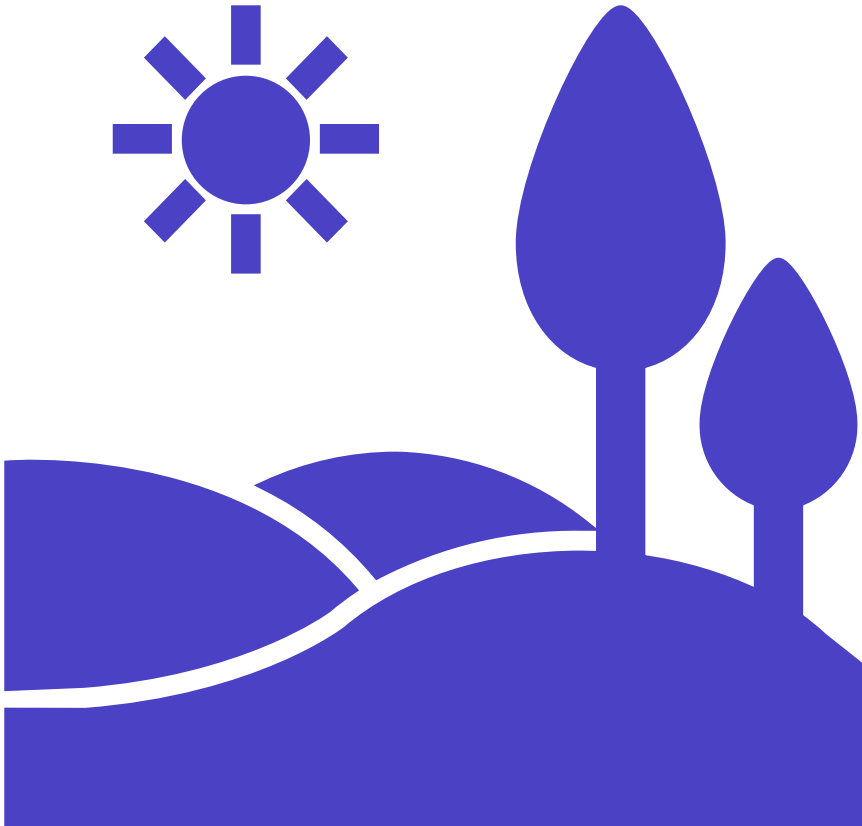
# **Machine Learning**

# **in Business**





# Acknowledgement of Country



We acknowledge the Traditional Custodians of  
our lands and waterways.  
We pay respects to Elders past, present and  
emerging future

# Introduction



Thuc Nguyen  
[thuc.nguyen@deakin.edu.au](mailto:thuc.nguyen@deakin.edu.au)



# Agenda

---

Overview

---

Objectives and How to succeed

---

Machine Learning Steps

---

Why Python?

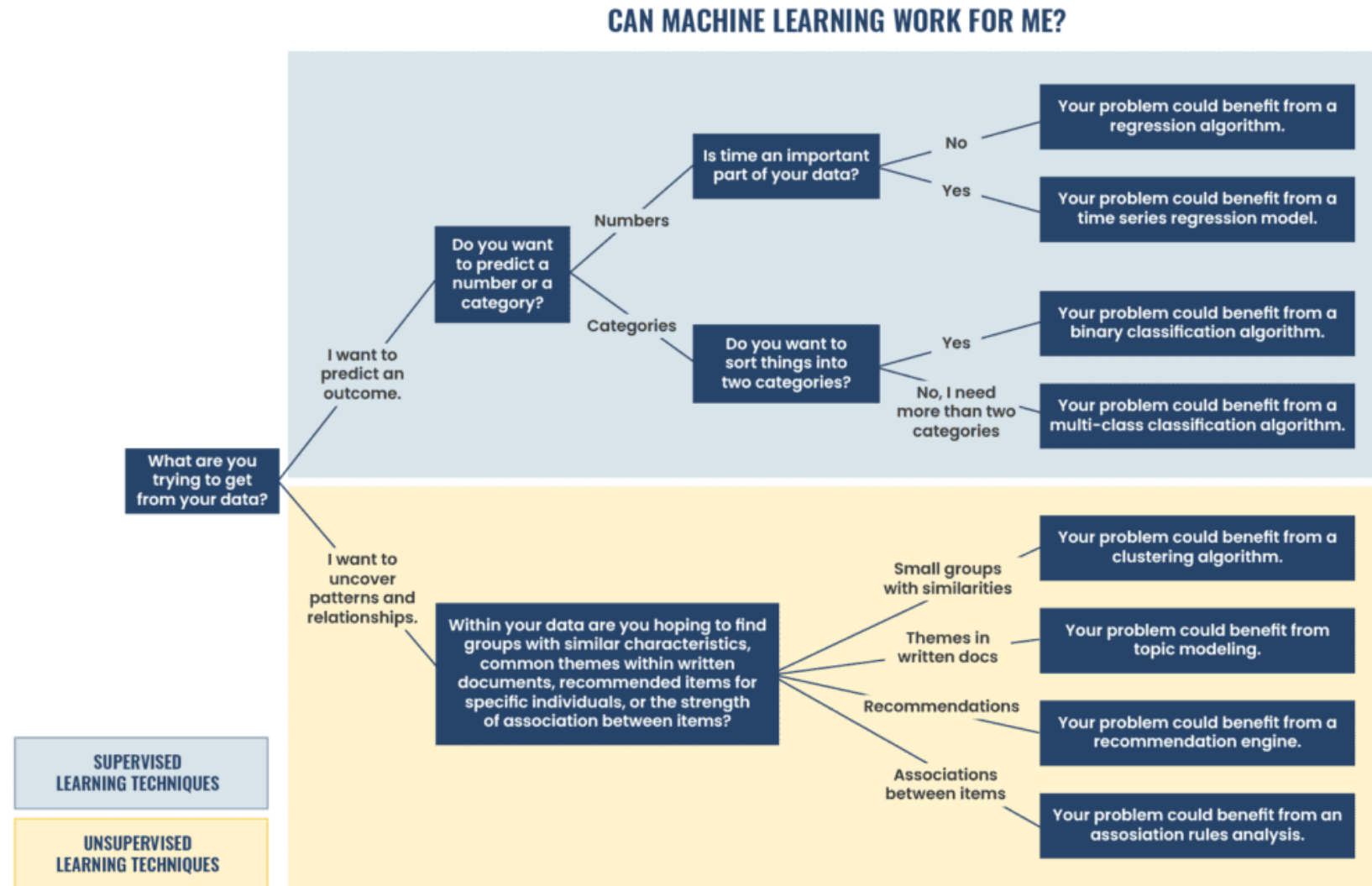
---

Useful Libraries

---

Have Fun!

# Overview



<https://www.phdata.io/blog/can-ai-or-ml-solve-your-business-problem/>

# Examples

---

Customer segmentation

---

Predictive analytics

---

Fraud detection

---

Supply chain optimisation

---

Sentiment analysis

---

Personalised recommendation

# Strategies for Success

---

Active participation

---

Regular practice

---

Consistency

---

Ask questions

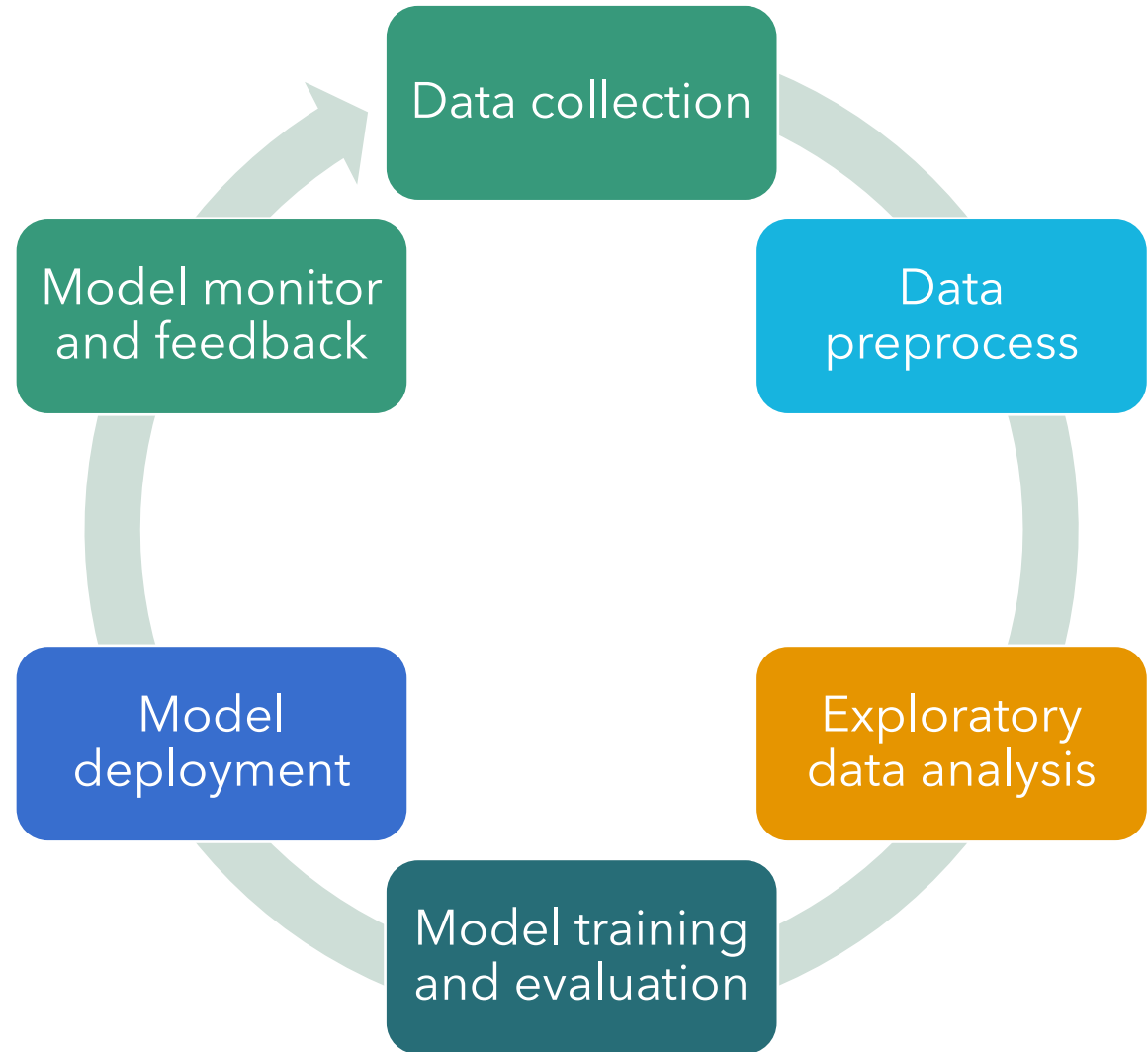
---

Engage in discussion

---

Problem-solving

# Machine Learning Pipeline



# Tools and Techniques

Programing language (Python)	BI Tools
Open source	Closed source
Free	Expensive
Powerful and flexible	Limited function
Steep learning curve	Easy to learn



# Useful libraries

---

numpy

---

pandas

---

matplotlib

---

seaborn

---

scikit-learn

---

scipy

**Let's have some fun!**



**ARE WE HAVING FUN YET ?**