

# Data Analysis

---

## *Syllabus Information*

### **CS 4100 - Data Analysis**

**Associated Term:** 2024/25 Academic Session

#### **Learning Objectives:**

The module focuses on presenting and explaining machine learning methods for analysing data. The first part of the course consists of a quick overview of pre-processing methods and basic statistical concepts, a detailed presentation of simple algorithms for supervised learning, e.g. linear and logistic regression models, the K-Nearest Neighbours algorithm, and Shallow Neural Networks, and a description of how to provide a statistically sound evaluation of the output of machine learning models. The second part of the course explores more advanced techniques, e.g. ensemble methods, and introduces unsupervised learning approaches, e.g. clustering. Compared with CS5100, CS4100 has optional lab sessions and reduced CW assignments. Learning Outcomes: 1. Extract value and insight from data. 2. Understand, develop, validate, evaluate, and use effectively machine learning models and statistical models. 3. Work with and implement methods and techniques such as clustering, classification, regression, decision trees, neural networks, and ensemble methods.

**Required Materials:** [Click here for the reading list system](#)

**Technical Requirements:** The total number of notional learning hours associated with the course are 150. **These will normally be broken down as follows:** 33 hour(s) of Lecture(s) across 11 week(s) 11 hour(s) of Laboratory across 11 week(s) 106 hours of Guided Independent Study **Summative Assessment:** Examination (70%) 2 hours Program and Report (30%) 20 hours