04/03 Lecture 1. Introduction

- Read Chapter 1 and Chapter 2 (sec. 2.3 & 2.4) of Dive into Deep Learning
- Linear methods are limited
  - Only work for simple tasks
  - Used in the final stages of more complex methods
- Neural Networks and Deep Learning
  - Focus of the unit
- Supervised Learning
  - Requires human input
- Unsupervised Learning
  - Machine learns from data inputs, only
- Key Components
  1) Data that we can learn from
  2) A model to transform the data
  3) An objective function that quantifies how well (or badly) the model is doing
  4) An algorithmm to adjust the model's parameters to optimise the objective function
- General Framework
  - Define an ML problem
    - Supervised, unsupervised or reinforcement learning
    - Combination of all these types
  - Data collection, cleaning, validation, and storage
  - Modelling, Training and Testing
    - A good representation of the functions: a set of functions that have the following characteristics:
      - Capacity: capable of representing

general complex input-output relationships to achieve good training performance
- Compactness: for good testing performance
- ~~Learnability~~ Learnability: the optimis-ation problem should be solvable

- Vectors
  - Simply a list of scalar values
  - Each scalar value is called an element
  - One-dimensional tensor
  - Can be column or row-oriented
- Matrices
  - A matrix is a list of column or row vectors
  - Each row is a row vector
  - Each column is a column vector
  - A column vector is a ~~mate~~ matrix with one column
  - A row vector is a matrix with one row
- Tensors
  - Generic way of describing n-dimensional arrays with an arbitrary number of axes
  - Scalars are zero-dimensional arrays
  - Vectors are one-dimensional arrays
  - Matrices are two-dimensional arrays
  - A 3-dimensional array is a list of matrices
  - A (n-1)-dimensional array is a list of n dimensional arrays

Summary of Linear Algebra

- Scalors, vectors, matrices and tensors are basic mathematical objects in linear algebra
- Vectors generalise scalars, and matrices generalise vectors
- Scalars, vectors, matrices and tensors have zero, one, two, and an arbitrary number of axes, respectively
- Elementwise multiplication of two matrices is called their Hadamard product. It is not matrix multiplication
- Norms are often used in objective functions

Summary of Calculus

- Derivative: the instantaneous rate of change of a function with respect to its variable and the slope of the tangent line to the curve of the function
- A gradient is a vector whose components are the partial derivatives of a multivariate function with respect to all its variables
- The chain rule enables us to differentia-te composite functions