AINT351 - Revision

Three types of learning

Imagine a machine experiences a sequence of sensory inputs `x1, x2, ... xn`

Supervised learning:

- The machine is also given `y1, y2, ..yn` and its goal is to learn and reproduce them from the inputs
- Learning by examples, input and output is given so it knows how to reproduce the output from the input

Unsupervised learning:

- The machine should build a representation of x that can be used for decision making, prediction
- There is no desired output, you are given inputs and after some iterations you start to categorise data based on some criteria

Reinforcement learning:

- The machine can generate actions `a1, a2,.. an` that affects its environment and receives a reward or punishment based on them. Its goal is to learn actions that maximise long term reward
- Learning based on rewards for actions so that it learns to maximise long term reward

Goals of supervised learning

Classify input data:

- In this case the desired outputs `y1,y2,...yn` are discrete class labels and the goal is to **classify** new output correctly from the new input
- have an image of a digit and want to know what digit it is based on previous examples of that digit

Goals of unsupervised learning

Regression

- In this case the desired outputs `y1, y2, ...yn` are continuous values and the goal is to **predict** new output correctly from new input
- Have the data from babies and can try to predict its weight given its height

We wish to find useful representations of data. This can involve

- Finding clusters
- Dimensionality reduction

- Finding the hidden cause of the surface phenomena
- Modelling the data probability density
- Data compression

Probability

Types of data:

Discrete data: only certain values

- Dice value = {1,2,3,4,5,6}
- Flip a coin = {H,T}

Continuous data: any value

- Length measurement
- Weight measurement

Probability functions

- A probability function maps possible values of a variable to its respective probabilities
  - e.g. if value is x we can write its possible probabilities as p(x)
- Probability functions have the following properties
  - P(x) is a number with a value between 0 to 1.0
  - The area under a probability function is always unity

The addition law of probability

- If two events A and B are mutually exclusive then
- `P(A U B)` = the probability event A **OR** B occurs
- `P(A U B) = P(A) + P(B)`
- If two events A and B are **NOT** mutually exclusive then
- `P(A U B) = P(A) + P(B) - P(A n B)`
- You have to subtract the intersect as it is where both events happen

Probability distributions

Bernoulli distribution:

- The probability of a success of failure, heads or tails, 1 or 0
- `n` is the number of times that the experiment is repeated

Discrete distribution:

- A finite amount of probabilities all of which have equal probability of occurring
- A dice throw, each outcome has a probability = 1/6

Cumulative probability:

- The probability of this event happening **ASWELL AS** all the previous events
- A dice landing on 6 as well as all the chances of it landing on 1,2,3,4 and 5 = 6/6

Binomial distribution:

- 2 outcomes
  - Heads or tails
- What is the probability of getting exactly 3 heads in 5 coin tosses
- HHHTT
  - `(1/2)^3 x (1/2)^2`
- THHHT
  - `(1/2)^1 x (1/2)^3 x (1/2)^1`
- All equal = `(1/2)^3 x (1/2)^2`
- therefore the overall probability =
  - `N x (1/2)^3 x (1/2)^2`
  - where N = number of unique arrangements
- There are exactly 10 ways to get 3 heads in 5 coin tosses
  - N = 10
- `10 x (1/2)^3 x (1/2)^2 = 0.3125`

Uniform distribution:

- A distribution that has constant probability
- 0.5 of values 0 and 1.0

Continuous data distributions:

- A continuous random variable is a random variable with a set of possible values that is infinite or uncountable
- looks like Gaussian distribution

Represi