Lec8- Classification

Recap

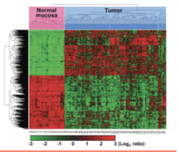
Classification

What is classification?

Given a collection of records (training set)

Each record contains a set of attributes, one of the attributes is the class

Find a method to assign the class of previously unseen records based on their other attributes and the training set as accurately as possible



Person	Height(m)	Weight(kg)	Gender
P1	1.79	75	M
P2	1.64	54	F
P3	1.70	63	M
P4	1.88	78	M
P5	1.75	70	??

How to do classification?

Training data with class

Classification method

Data to be classified

K-nearest neighbor classification: The standard procedure of KNN

Suppose we have chosen distance metric and K

Normalization

Compute distances

Identify the K most similar data

Take their class out and find the mode class

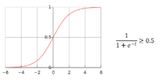
Clustering VS Classification

	Clustering	Classification
Goal	Find similarity (clusters) in the data	Assign class to the new data
Data	Data without class	Training data with class and testing data without class
Classes	Unknown number of classes	Knows number of classes
Output	The cluster index for each point	The class assignment of the testing data
Algorithm	One phase	Two phases (training and application)

Logistic Regression

$$w_h H + w_w W + w_0 \geq 0.5$$

$$\frac{1}{1+e^{-(w_h H + w_w W + w_0)}} \geq 0.5$$



$$\frac{1}{1+e^{-z}} \geq 0.5$$

Training: fit the Training data, to get Wh, Ww, W0

Testing: run the formula

Logistic Regression model training

How to train

To get Wh, Ww, W0

To make the model fit the training data

$$\frac{1}{1+e^{-(w_h H + w_w W + w_0)}} \geq 0.5$$

correct for the training data

$$Y^{output} = \frac{1}{1+e^{-(w_h H + w_w W + w_0)}}$$

1 for male, 0 for female

Loss function

$$L = (Y^{output} - Y)^2$$

Y: the true label we have for training data

find out the minimum value of loss function

Gradient descent algorithm

$(Y^{output} - Y)^2$ **is a function of W s**

for each w, find a value to make the function value smallest

calculate derivative

find out direction

repeat

find when the derivative is very small, close to 0

To get the formula

Initialize Wh, Ww, W0 with random values

Calculate the output Y^{output}

Update the weights

$w_i = w_i + \Delta w_i$

$\Delta w_i = 2 * \alpha (Y - Y^{output}) \frac{\partial y^{output}}{\partial w_i}$

repeat, until no more to update

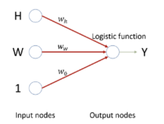
Good Wh, Ww, W0 make the Loss function smallest

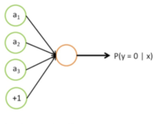
From logistic regression to neural networks

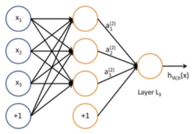
The simplest neural network: Logistic Regression

$$Y^{output} = \frac{1}{1+e^{-(w_h H + w_w W + w_0)}}$$

visualize



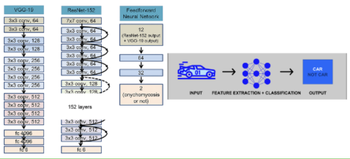




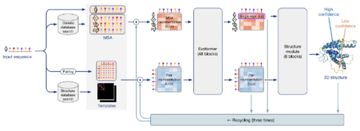
From Logistic Regression to Neural Networks

- Fast prediction
- Successful in real-life problems
- High tolerance to noisy data
- Long training time
- Poor interpretability

From neural network to deep learning



Alpha Fold: the most successful deep learning application



Presented with xmind