Machine Learning

Dr Changjiang He, Dr Kuo-Ming Zhao
Computer Science| School of Art
University of Roehampton

Lesson 3.1

Logistic Regression

Contents

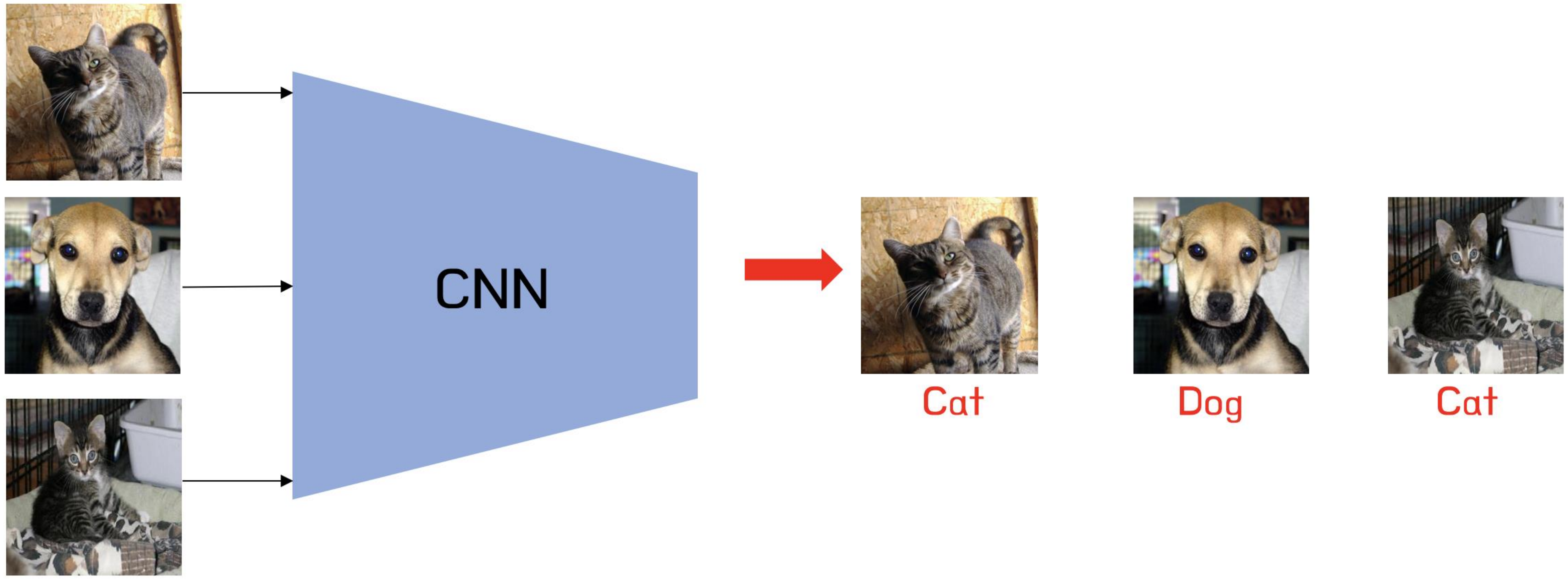
University of Roehampton London

- What is classification
- Review of logistic function
- What is the basic idea of logistic regression

Classification

What is a classifier?

A classifier is a machine learning model that is used to discriminate different objects based on certain features.



Logistic Regression

University of Roehampton London

- In this lesson, we will a learn a simple classification method, Logistic Regression.
- Logistic regression is a supervised learning algorithm which is mostly used to solve binary classification.
- Customer churn, spam email, website or ad click predictions are some examples of using logistic regression.
- It is even used as an activation function for neural network layers.

Logistic Function

The basis of logistic regression is the logistic function, also called the **sigmoid function**, which takes in any real valued number and maps it to a value between 0 and 1.

$$Sigmoid\ Function:\ y = \frac{1}{1 + e^{-x}}$$

Probability

Probability measures the likelihood of an event to occur. For example, if we say “there is a 90% chance that this email is spam”:

$$P(spam) = 0{,}9$$

Odds is the ratio of the probabilities of positive class (email is spam) and negative class (email is not spam).

$$odds = \frac{P\ (spam)}{P\ (not\ spam)}$$

Log Odds

- Log odds is the logarithm of odds.
- In the case of logistic regression, log odds is used

Probability	Odds	Log Odds
0,05	0,05	-1,28
0,1	0,11	-0,95
0,2	0,25	-0,60
0,3	0,43	-0,37
0,4	0,67	-0,18
0,5	1,00	0,00
0,6	1,50	0,18
0,7	2,33	0,37
0,8	4,00	0,60
0,9	9,00	0,95
0,95	19,00	1,28

Log Odds

- Log odds is the logarithm of the ratio between the probability of positive class and negative class.
- Probability of 0.5 means that there is an equal chance, and the log odds is 0.

Probability	Odds	Log Odds
0,05	0,05	-1,28
0,1	0,11	-0,95
0,2	0,25	-0,60
0,3	0,43	-0,37
0,4	0,67	-0,18
0,5	1,00	0,00
0,6	1,50	0,18
0,7	2,33	0,37
0,8	4,00	0,60
0,9	9,00	0,95
0,95	19,00	1,28

Transform of Logistic Function

Let's go back to the sigmoid function and show it in a different way:

$$y = \frac{1}{1+e^{-x}} \Rightarrow 1 + e^{-x} = \frac{1}{y} \Rightarrow e^{-x} = \frac{1-y}{y} \Rightarrow e^{x} = \frac{y}{1-y}$$

Taking the natural log of both sides:

$$\Rightarrow x = \log\left(\frac{y}{1-y}\right) \quad \text{.........} \; \textbf{(1)}$$

Logistic Regression

Now we write down the basic form of linear regression, given multiple input variables x1, x2, …, xn and define the output as z:

$$\mathbf{z} = \beta 0 + \beta 1.x1 + \cdots + \beta n.xn$$

Then we can combine the transformed logistic function for the input decomposition:

$$\beta 0 + \beta 1.x1 + \cdots + \beta n.xn = \log\left(\frac{y}{1-y}\right)$$

Logistic Regression

$$\mathbf{z} = \beta 0 + \beta 1.x1 + \cdots + \beta n.xn$$

$$\beta 0 + \beta 1.x1 + \cdots + \beta n.xn = \log\left(\frac{y}{1-y}\right)$$

Assume y is the probability of positive class. If z is 0, then y is 0.5. For positive values of z, y is higher than 0.5 and for negative values of z, y is less than 0.5. If the probability of positive class is more than 0.5 (i.e. more than 50% chance), we can predict the outcome as a positive class (1). Otherwise, the outcome is a negative class (0).

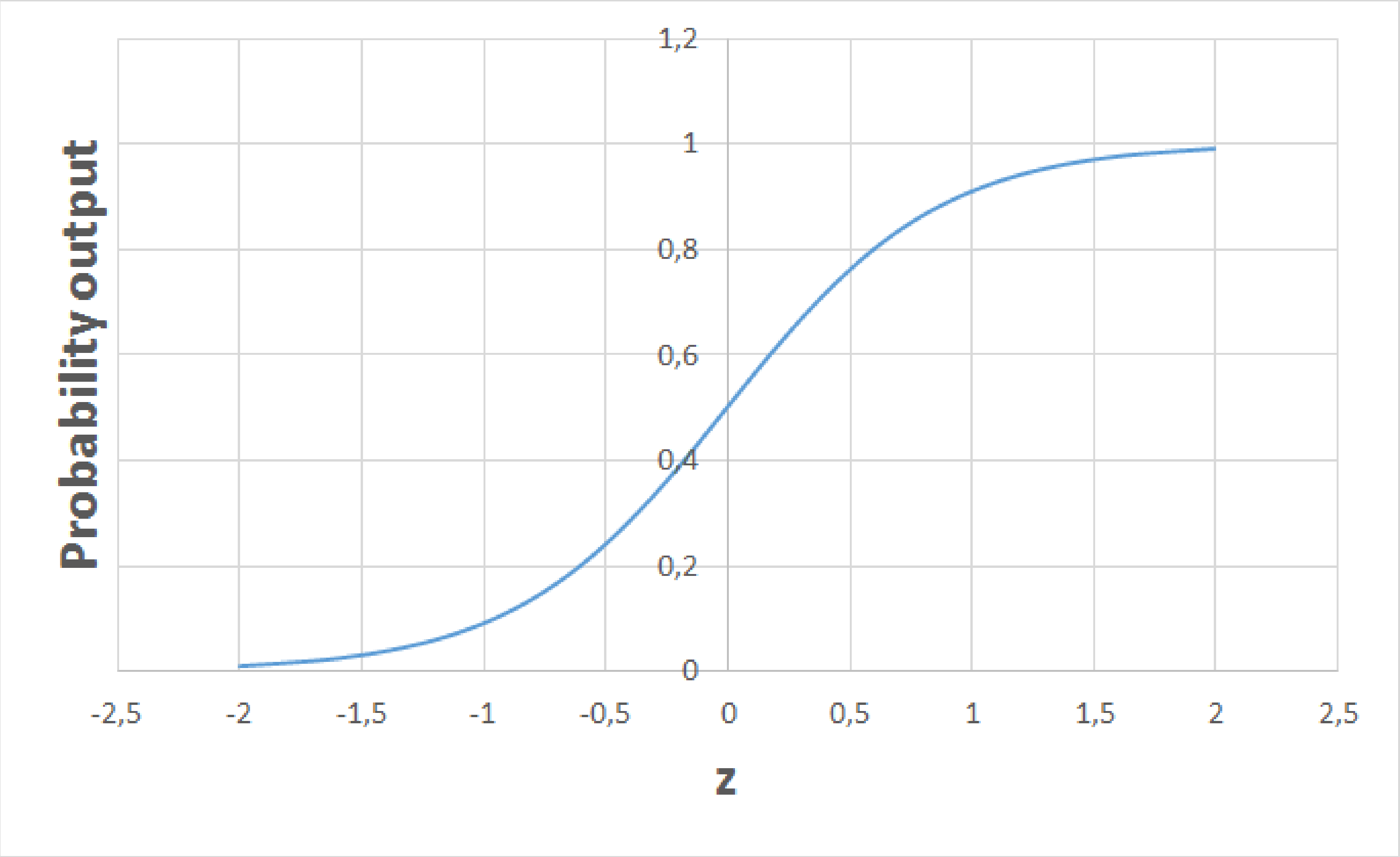
Logistic Regression

The table on the right shows some values of z with corresponding y (probability) values. All real numbers are mapped between 0 and 1.

y	z
0	-infinity
0,01	-4,60
0,1	-2,20
0,2	-1,39
0,3	-0,85
0,4	-0,41
0,5	0,00
0,6	0,41
0,7	0,85
0,8	1,39
0,9	2,20
0,99	4,60
1	+infinity

Logistic Regression

If we plot this function, we will get the famous s shaped graph of logistic regression:



The classification problem comes down to solving a linear equation:

$$\mathbf{0} = \beta 0 + \beta 1.x1 + \cdots + \beta n.xn$$

Parameters of the function are determined in training phase with maximum-likelihood estimation algorithm. Then, for any given values of independent variables (x1, x2, …, xn), the probability of positive class can be calculated.

Types of Logistic Regression

- Binary logistic regression:
  The dependent variable has only two possible outcomes, being either 0 or 1.

- Multinomial logistic regression:
  The dependent variable has three or more possible outcomes.

- Ordinal logistic regression:
  The dependent variable has three or more outcomes in a defined order.

Summary

- Logistic regression is a simple yet very powerful algorithm to solve binary classification problems.

- The logistic function (i.e., sigmoid function) is also commonly used in very complex neural networks as the activation function of output layer.