1

00:00:00,630 --> 00:00:08,040

[Auto-generated transcript. Edits may have been applied for clarity.]

Now let's talk about logistic regression. It's a variant of linear regression and one of the most basic classification algorithms.

2

00:00:08,430 --> 00:00:15,300

Linear regression is used for regression tasks, where the goal is to predict continuous values such as height or salary.

3

00:00:15,810 --> 00:00:25,080

Logistic regression, despite its name, is actually used for classification tasks, while the name includes regression, and we'll see why later.

4

00:00:25,290 --> 00:00:28,740

Its primary use is for predicting categorical outputs.

5

00:00:29,430 --> 00:00:36,030

Imagine we want to predict whether a fruit is an apple or an orange based on two features weight and color.

6

00:00:36,690 --> 00:00:40,830

Let's say color is represented numerically. Lower scores for green.

7

00:00:40,920 --> 00:00:49,080

Higher scores for red. You have a data set with color, weight and a class label apple or orange based on historical data.

8

00:00:49,560 --> 00:00:53,580

As new data comes in, the model should classify the fruit correctly.

9

00:00:54,360 --> 00:01:00,900

Using linear regression wouldn't help here because we're not predicting a continuous value, we're predicting a class.

10

00:01:01,290 --> 00:01:09,360

That's where logistic regression comes in. Instead of fitting a straight line, logistic regression fits a sigmoid function.

11

00:01:09,660 --> 00:01:13,440

Suppose the output is zero for orange and one for apple.

12

00:01:13,830 --> 00:01:21,600

In a simplified one dimensional example, using only the color feature, a straight line wouldn't effectively separate the two classes.

13

00:01:22,110 --> 00:01:23,610

There might be overlap.

14

00:01:23,760 --> 00:01:31,590

Some oranges and apples may share similar color values, making it difficult to draw a straight line that separates them cleanly.

15

00:01:31,950 --> 00:01:36,690

The sigmoid function, which has an S shaped curve, is better suited for this.

16

00:01:37,140 --> 00:01:43,920

It's bounded between 0 and 1, and outputs a probability that a given data point belongs to a certain class.

17

00:01:44,280 --> 00:01:54,870

For example, if a fruit falls at a point on the sigmoid curve that corresponds to 0.8, the model would assign an 80% probability that it's an apple.

18

00:01:55,260 --> 00:02:00,300

A threshold is then applied if the predicted probability is greater than 0.5.

19

00:02:01,140 --> 00:02:06,240

Classify it as Apple one. If less, classify it as orange zero.

20

00:02:06,900 --> 00:02:11,190

This threshold can be adjusted depending on the distribution of the historical data.

21

00:02:11,670 --> 00:02:18,570

For instance, if the data shows less overlap between classes, the threshold might be raised or lowered.

22

00:02:18,960 --> 00:02:23,400

For example, predicting Apple only for values greater than 0.7.

23

00:02:23,850 --> 00:02:32,010

The parameters of the sigmoid curve and the threshold, if tuned, are learned through a training algorithm that minimizes a loss function.

24

00:02:32,610 --> 00:02:41,730

In this classification context, the loss function often used is the mean squared error between the predicted probability and the actual class label.

25

00:02:42,180 --> 00:02:50,610

For example, if the true label is one apple and the model predicts 0.3, the error is one -0.3.

26

00:02:50,640 --> 00:02:56,130

The training algorithm adjusts the model to minimize such losses across all training examples.

27

00:02:56,520 --> 00:03:01,920

That's in summary. Logistic regression is used for classification, not regression.

28

00:03:02,370 --> 00:03:07,530

It uses a sigmoid function to map inputs to probabilities between 0 and 1.