1

00:00:00,620 --> 00:00:08,330

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

Another classifier is what we call Naive Bayes. It belongs to a family of Bayes classifiers, which are probabilistic classifiers.

2

00:00:08,630 --> 00:00:12,980

Like logistic regression, they output the probability of a class label.

3

00:00:13,550 --> 00:00:17,030

The Naive Bayes algorithm is the simplest in this family.

4

00:00:17,180 --> 00:00:20,600

Though there are other Bayes classifiers with different assumptions,

5

00:00:21,020 --> 00:00:26,360

these classifiers are based on Bayes theorem, a statistical principle from the 1800s.

6

00:00:26,780 --> 00:00:31,250

Bayes theorem calculates the probability of a hypothesis given some evidence.

7

00:00:31,790 --> 00:00:38,870

For example, it can tell us the probability that someone is sick given that their Covid 19 test result is positive.

8

00:00:39,440 --> 00:00:42,470

This is typically explained using a probability tree.

9

00:00:42,800 --> 00:00:48,830

A person may be sick or healthy, and then they take a test which may come back positive or negative.

10

00:00:48,890 --> 00:00:52,250

If a sick person tests negative, that's a false negative.

11

00:00:52,640 --> 00:00:57,020

A healthy person may also receive a positive result, which is a false positive.

12

00:00:57,470 --> 00:01:00,470

What we observe is the test result, the evidence.

13

00:01:00,740 --> 00:01:04,730

But what we're interested in is the underlying condition. The hypothesis.

14

00:01:05,240 --> 00:01:13,190

Bayes's theorem lets us compute the probability of the hypothesis being sick given the observed evidence a positive test result.

15

00:01:13,430 --> 00:01:17,000

In machine learning, this idea is applied to spam detection.

16

00:01:17,360 --> 00:01:26,660

The hypothesis is whether an email is spam or not, and the evidence is the presence of specific words in the email, such as prize or lottery.

17

00:01:26,720 --> 00:01:33,460

You may not directly observe whether an email is spam, but based on the evidence, the presence of certain words.

18

00:01:33,500 --> 00:01:37,160

Bayes theorem provides a probability, for example,

19

00:01:37,280 --> 00:01:48,680

in email that includes the words winner and prize might have a 92.8% probability of being spam, meaning a 7.2% probability of not being spam.

20

00:01:49,010 --> 00:01:54,560

The term naive refers to the assumption that all pieces of evidence are independent of each other.

21

00:01:54,950 --> 00:02:03,530

For instance, the algorithm assumes that the presence of the word lottery is independent of the word million, even though they often co-occur.

22

00:02:04,040 --> 00:02:13,130

This independence assumption is rarely true in real life, but it simplifies computation, and the algorithm often performs well regardless.

23

00:02:13,700 --> 00:02:18,500

Naive Bayes is simple, efficient, and includes a learning phase.

24

00:02:18,830 --> 00:02:25,340

Here's how the learning works. You're given a data set of emails where each email is labeled as spam or not spam,

25

00:02:25,580 --> 00:02:31,310

and each column represents the presence of a particular word lottery, prize winner, etc.

26

00:02:31,700 --> 00:02:35,510

The class label might be zero for spam and one for not spam.

27

00:02:35,930 --> 00:02:39,050

From this, probabilities are computed from the data set.

28

00:02:39,380 --> 00:02:47,060

During prediction, the algorithm multiplies the learned probabilities for each piece of evidence to compute the final classification.

29

00:02:47,480 --> 00:02:52,070

While Naive Bayes does not minimize a loss function like some other algorithms,

30

00:02:52,520 --> 00:02:57,080

there's still a learning phase in which probabilities are estimated from data.