Generative vs discriminative models

Wednesday, October 31, 2018 10:20 AM

695

Let's say you have input data x and you want to classify the data into labels y. A generative model learns the **joint** probability distribution `p(x,y)` and a discriminative model learns the **conditional** probability distribution `p(y|x)` - which you should read as *"the probability of y given x"*.

Here's a really simple example. Suppose you have the following data in the form (x,y):

`(1,0), (1,0), (2,0), (2, 1)`

`p(x,y)` is

```
      y=0   y=1
     -----------
x=1 | 1/2   0
x=2 | 1/4   1/4
```

`p(y|x)` is

```
      y=0   y=1
     -----------
x=1 | 1     0
x=2 | 1/2   1/2
```

If you take a few minutes to stare at those two matrices, you will understand the difference between the two probability distributions.

The distribution `p(y|x)` is the natural distribution for classifying a given example `x` into a class `y`, which is why algorithms that model this directly are called discriminative algorithms. Generative algorithms model `p(x,y)`, which can be transformed into `p(y|x)` by applying Bayes rule and then used for classification. However, the distribution `p(x,y)` can also be used for other purposes. For example, you could use `p(x,y)` to *generate* likely `(x,y)` pairs.

From the description above, you might be thinking that generative models are more generally useful and therefore better, but it's not as simple as that. This paper is a very popular reference on the subject of discriminative vs. generative classifiers, but it's pretty heavy going. The overall gist is that discriminative models generally outperform generative models in classification tasks

From <https://stackoverflow.com/questions/879432/what-is-the-difference-between-a-generative-and-discriminative-algorithm>

A **generative algorithm** models how the data was generated in order to categorize a signal. It asks the question: based on my generation assumptions, which category is most likely to generate this signal?

A **discriminative algorithm** does not care about how the data was generated, it simply categorizes a given signal.

- **Discriminative models** learn the (hard or soft) **boundary** between classes
- **Generative models** model the **distribution** of individual classes

1: Discriminative algorithms allow you to classify points, without providing a model of how the points are actually generated. So these could be either:

- **probabilistic** algorithms try to learn P(Y|X)P(Y|X) (e.g., logistic regression);
- or **non-probabilistic** algorithms that try to learn the mappings directly from the points to the classes (e.g., perceptron and SVMs simply give you a separating hyperplane, but no model of generating new points).

So yes, discriminative classifiers are any classifiers that aren't generative.
Another way of thinking about this is that **generative algorithms make some kind of structure assumptions on your model**, but discriminative algorithms make fewer assumptions. For example, Naive Bayes assumes conditional independence of your features, while logistic regression (the discriminative "counterpart" of Naive Bayes) does not.

From <https://stats.stackexchange.com/questions/12421/generative-vs-discriminative>

Yes, Naive Bayes is generative because it captures $P(X|Y)$P(X|Y) and $P(Y)$P(Y). For example, if we know that $P(Y=English)=0.7$P(Y=English)=0.7 and $P(Y=French)=0.3$P(Y=French)=0.3, along with English and French word probabilities, then we can now generate a new document by first choosing the language of the document (English with probability 0.7, French with probability 0.3), and then generating words according to the chosen language's word probabilities.

From <https://stats.stackexchange.com/questions/12421/generative-vs-discriminative>

CRF is discriminative
Hidde Markov is generative

SVMs and decision trees are discriminative because they learn explicit boundaries between classes

From <https://stats.stackexchange.com/questions/12421/generative-vs-discriminative>

From <https://stackoverflow.com/questions/879432/what-is-the-difference-between-a-generative-and-discriminative-algorithm>