A Brief Introduction to Generative and Discriminative Models

Zhaokun Zhang*

March 16, 2019

Abstract

There are a wide and growing variety of tasks in computer vision for which machine learning methods based on probability are being successfully applied. In this article I will give a brief introduction to generative model and discriminative model which are typical in statistical classification. And I'll also discuss their basic differences. Finally I will list some advantages as well as disadvantages for the models.

1 Definition

In statistical classification, especially in supervised machine learning, there are two main approaches called the generative approach and the discriminative approach. Analogously, a classifier based on the generative model is a generative classifier, while a classifier based on the discriminative model is a discriminative classifier, though this term also refers to classifiers that are not based on a model.

Generative classifiers learns a model of the joint probability, $p(x, y)$, of the inputs x and the label y, and makes their predictions using Bayers rules to calculate $p(y|x)$, and then picking the most likely label y. However, unlike the generative modeling, the discriminative model, also referred to as conditional models, studies the $p(y|x)$ directly, or learns a direct map from inputs x to the class labels. Note that there is still distinction between the conditional model and the discriminative model, though more often they are simply categorized as discriminative model[1]. Conditional Model models the conditional probability distribution while the traditional discriminative model aims to optimize on mapping the input around the most similar trained samples. A discriminative classifier tries to model by just depending on the observed data while learning how to do the classification from the given statistics.

In addition to precise definition, the terminology is because the generative model can be used to "generate" outcomes, either of an observation and target (x, y), or of an observation x given a target value y, that is, generative model allows us to do more than just making predictions of y given x. We can generate new data points $\{(x_i, y_i)\}$ by drawing from $\hat{p}(x, y)$. And the discriminative model or discriminative classifier(without a model) can be used to "discriminate" the value of the target variable given x, meaning that it does nothing more than classify inputs.

2 Differences between the two models

At the beginning, we can use a simple example to further elaborate the definitions and help to explain the differences. Assume that given a piece of audio, you need to identify the language it belongs to. And now you have two choice to accomplish it:

1. You spend time in studying all kinds of languages and fully understand them, then as long as a person starts, you can know which language it belongs to;

2. You focus on the differences of various language models rather than take time to learn every language, and categorize them. So when you' re listening, you just need to find the characteristics based on the differences.

* *WISE,Xiamen University(Student ID: 15220162202515)*

[1]Source: Wikipedia, https://en.wikipedia.org/wiki/Discriminative_model#Contrast_with_Generative_Model

Thus, in machine learning, the first is the generative model, while the second is the discriminative model. The fundamental difference between these models is that generative model models the distribution of individual classes, while disriminative models learn the boundaries between classes[2]. After learning the joint distribution $p(x, y)$, generative model can reflect the similarity of similar data itself from the statistical point of view, but it doesn't care about the boundary of the various types like discriminative model. Besides, we can get discriminative model from generative model as $p(y|x) = p(x, y)/p(x)$, that is, generative algorithms have discriminative properties, but it's more difficult to get generative model as discriminative model gets $p(y|x)$ directly. When you're computing $p(y|x)$ in logistic regression, you're not computing these two things: $p(x, y)$ and $p(x)$, you're just applying a logistic function to a dot product.

3 Advantages and disadvantages of models

3.1 Generative model

3.1.1 Advantages:

- The actual information we get in generative model is far more than what we need to make a prediction. For example, apart from joint distribution, we can get a marginal distribution ;
- The generative model can deal with the situation where there exists hidden variables or missing data;
- It converges to the real model more quickly, and it's more flexible when dealing with single problem.

3.1.2 Disadvantages

- Since the generative model doesn't care about the boundaries, it may cause misclassification;
- Although it can bring us with more information, it also needs more samples and more complex calculation, which may cause the waste of resources.

3.2 Discriminative model

3.2.1 Advantages

- Comparing with the generative model, discriminative model makes fewer assumptions and requires less training samples which would save calculation resources;
- Higher accuracy, which mostly leads to better learning modeling, and it also tends to generate lower asymptotic errors;
- It has a more flexible framework that could easily cooperate with other needs of application.

3.2.2 Disadvantages

- Discriminative model cannot reflect the natural characteristics of training samples, that is, it can tell you which category it belongs to, but cannot describe the total environment;
- The relationship between variables is difficult to find out and invisible;
- similarly, by the definition, the discriminative model will need the combination of multiple subtasks for a solving complex real-world problem.

Although there holds a belief that discriminative classiers are almost always to be preferred, you should choose from them based on what you wish to achieve from your analysis as these methods have different advantages and specific applications.

[2]Source: Devesh Batra, https://deveshbatra.github.io/Generative-vs-Discriminative-models/

References

[1] Ng, Andrew Y.; Jordan, Michael I.(2002). "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes". Advances in neural information processing systems.

[2] Kelm, B. Michael.(2018). "Combining Generative and Discriminative Methods for Pixel Classification with Multi-Conditional Learning".

[3] Devesh Batra's blog. "Discriminative vs Generative model". https://deveshbatra.github.io/Generative-vs-Discriminative-models/

[4] Wikipedia."Discriminative model". https://en.wikipedia.org/wiki/

[5] Wikipedia."Generative model". https://en.wikipedia.org/wiki/