Lecture 15 - 2: Generative vs Discriminative models

Tabaré Pérez

May 6, 2020

Today, we will talk about generative models.

So let me start by first giving you a general idea, what are the generative models and how they are different from what we've done in the past, which was discriminative model.

So you remember in this class, whenever we talked about classification, the pictures that you had in mind is that your classifier had a training data, which consisted of some, let's say, positive and negative instances. And the job of the discriminative model was to find a separator that discriminates between these positive and negative examples.

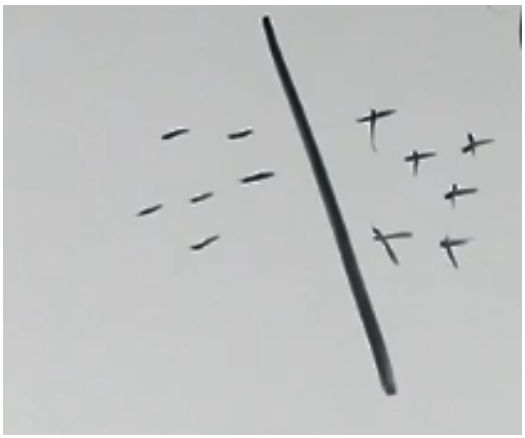

Figure 1: Classification: Linear separator

And you can easily imagine to yourself that for different clouds of these positive and negative points, you may have exactly the same separator even the structure of these clouds it's totally different. So for instance, even here I would have some different structure of positive and negative examples, the separator will be still the same.

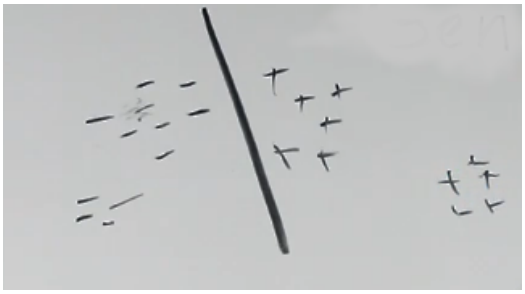

Figure 2: Classification: Linear separator

So the approach that we will be taking is actually trying to understand what is the structure of these positive and negative classes and the way we're going to be thinking about it is saying, if I can understand in some probabilistic terms the structure of positive and negative examples, maybe I can do discrimination in a different way that we've done when we're primarily looking at the separator.

And we sort of start doing that when we were looking at k-means classifiers.

We kind of said, ok, the structure of our data would be just clusters.

But in that particular case, we assume very specific type of structures that we can impose on the data, and it wasn't really probabilistic.

So now, we are going to move in two directions. We're going to look at much broader set of distribution that can be fitted for our data and we're also going to add a probabilistic component to it.

And today, we would look at two classes of generative models which are very commonly used:

- Multinomials
- Gaussians.

And for both of these classes of models, we will look at them within the same lens.

We're going to ask two questions here:

- The first question that we will ask, and this is a question that we will ask about any generative model, is how do we estimate this model? Because we have particular form of probability distribution, how do we fit it to our data? So the first question will be estimation question.
- And the second question's, when you estimated the model, how we can actually do prediction?

And let me give you some very high-level idea how we can use these kind of probabilistic models for prediction, because I will start today talking actually about estimation but later, we will get to prediction.

But you need to see kind of, you know, our ultimate destination, because we do want to use these models to do prediction.

So what we will do, given our training data the same way as we've seen before, our negative and positive points, we will fit probability distributions for the negative class and for the positive class.

Now, given a new point, you can actually compute how likely is it it was generated by the negative class or by the positive class.

And by comparing these probabilities, we can actually induce what is the right label.

But again, before we can go to the question of prediction, we need to start by thinking, how can I find the right type of a distribution to describe each class?

And the first thing that I will do, I will start the road here with one class of models which are called **MULTINOMIALS**.