

Which Type of Model Should It Use?

Qi Zhao

July 26, 2018

It has established that there are two different types of model that relate the world state and the observed data, the discriminative model [4] and the generative model [2]. But when should it use each type of model? There is no definitive answer to this question, but some considerations are:

1. Inference is generally simpler with discriminative models. They directly model the conditional probability distribution [1] of the world $\Pr(w|x)$ given the data. In contrast, generative models calculate this probability via Bayes rule [5], and sometimes this requires a computationally expensive algorithm.

2. Generative methods build probability models $\Pr(x|w)$ over the data whereas discriminative models just build a probability model $\Pr(w|x)$ over the world state. The data (usually an image) are generally of much higher dimension than the world state (some aspect of a scene), and modeling it is costly. Moreover, there may be many aspects of the data which do not influence the state; it might devote parameters to describing data configuration although they both imply the same world state (Figure 1).

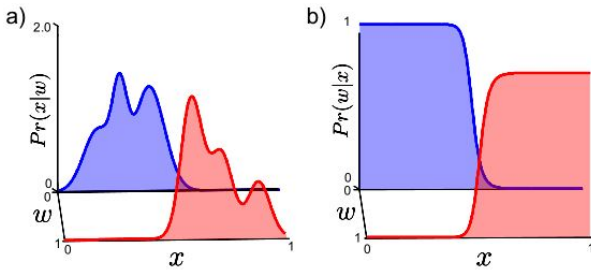


Figure 1. Generative vs. discriminative models. a) Generative approach: we separately model the probability $\Pr(x|w)$ for each class. This may require a complex model with many parameters. b) Posterior probability distribution $\Pr(w|x)$ computed via Bayes rule with a uniform prior. Notice that the complicated structure of the individual class conditional density functions isn't reflected in the posterior: here, it would have been more efficient to take a discriminative approach and model this posterior directly.

3. Modeling the likelihood $\Pr(x|w)$ [3] mirrors the actu-

al way that the data were created; the state of the world did create the observed data through some physical process (usually light being emitted from a source, interacting with the object and being captured by a camera). If wishing to build information about the generation process into our model, then this approach is desirable. For example, it can account for phenomena such as perspective projection and occlusion.

4. Sometimes parts of the training or test data vector x may be missing. Here, generative models are preferred. They model the joint distribution over all of the data dimensions and can effectively interpolate the missing elements.

5. A fundamental property of the generative approach is that it allows incorporation of expert knowledge in the form of a prior. It is harder to impose prior knowledge in a principled way in discriminative models.

It is notable that generative models are more common in vision applications.

References

- [1] D. Brook and D. Evans. An approach to the probability distribution of CUSUM run length. *Biometrika*, 59(3):539–549, 1972. 1
- [2] M. R. Sabuncu, B. T. Yeo, K. Van Leemput, B. Fischl, and P. Golland. A generative model for image segmentation based on label fusion. *IEEE Transactions on Medical Imaging*, 29(10):1714–1729, 2010. 1
- [3] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1(2):113–122, 1982. 1
- [4] X. Tan, Y. Li, J. Liu, and L. Jiang. Face liveness detection from a single image with sparse low rank bilinear discriminative model. In *European Conference on Computer Vision*, pages 504–517, 2010. 1
- [5] R. A. Waller and D. B. Duncan. A Bayes' rule for the symmetric multiple comparisons problem. *Journal of the American Statistical Association*, 64(328):1484–1503, 1969. 1