Generative and Discriminative Models

Suchismit Mahapatra

Introduction

General Setup

- Fixed, unknown distribution $\mathcal{D}$ over $\mathcal{X} \times \mathcal{Y}$, to be learned.
  - $\mathcal{X}$ is the instance space and $\mathcal{Y}$ is the label space.
- Given a dataset $\mathcal{S} = \{x_i, y_i\}_{i=1,2\ldots N}$.
- Learning :-
  - Identify a hypothesis space $\mathcal{H}$ and a loss function $\mathcal{L}$.
  - Minimize average loss over training data.
- Guarantee :-
  - If we can find a hypothesis $h \in \mathcal{H}$ which minimizes loss over observed data, statistical learning theory guarantees good generalization performance (as a function of $\mathcal{H}$).

Discriminative Model

Overview

- Goal is to learn directly how to make predictions.
- Discovers patterns in the data by looking at positive/negative examples.
- Uses the above to define a "prediction rule".
- Assumptions are made in the form of hypothesis class $\mathcal{H}$ i.e. model type and complexity.
- Approximates $h : \mathcal{X} \rightarrow \mathcal{Y}$ as estimating $P(Y|X)$.

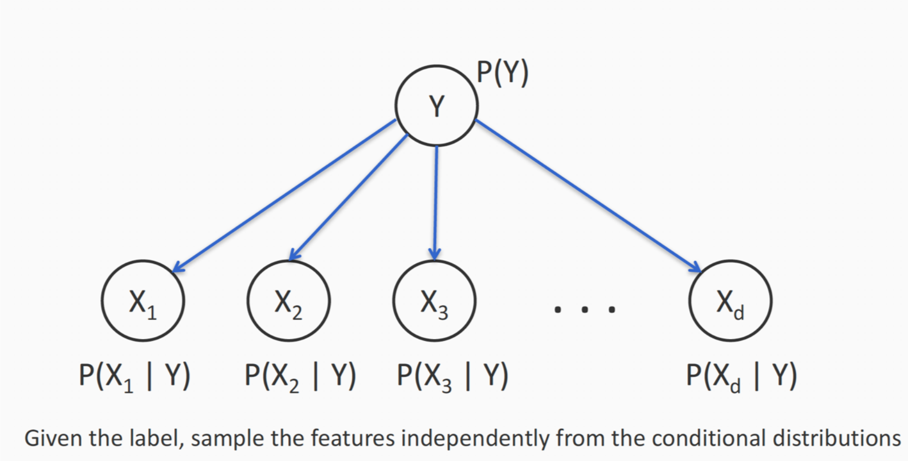
Generative Model

Overview

- Explicitly models how instances from each class/category are generated.
- Estimates the class conditional feature distributions $P(X|Y)$ and class priors $P(Y)$.
- Assumptions are made in functional forms of both above.
- We have seen this for Naïve Bayes earlier.

Generative Model

Naïve Bayes



[Credit: Vivek Srikumar]

Why do we need Generative Models ?

Motivation

- Access to (potentially) infinite data ?
- Feature de-noising
- Improves robustness to predictions.
- Anomaly detection.
- Structured prediction i.e. speech to text etc.

Generative vs Discriminative

Summary

- Generative Model
  - Learns $P(X, Y)$.
  - Uses the capacity of the model to characterize how the data is generated.
  - E.g. Naïve Bayes, Hidden Markov Models
- Discriminative Model
  - Learns $P(Y|X)$.
  - Uses the capacity of the model to characterize the decision boundary only.
  - E.g. Logistic Regression, SVM, Boosted Decision Trees

Generative vs Discriminative

Summary