Generative vs Discriminative Models in Supervised Learning

Ira A. Fulton Schools of Engineering
Arizona State University

Objective

Objective

Differentiate between generative and discriminative models of supervised learning

Objective

Discuss challenges in Bayesian learning

Supervised Learning

- **The set-up:** the given training data consist of <sample, label> pairs, or ($\mathbf{x}$, y); the objective of learning is to figure out a way to predict label y for any new sample $\mathbf{x}$.
  - E.g., Given n pairs $<\mathbf{x}^{(i)}, y^{(i)}>$, $i=1, \ldots, n$; $\mathbf{x}^{(i)}$: i-th sample represented as d-dimensional vectors; $y^{(i)}$: corresponding labels.

- Equivalently, to find $P(y|\mathbf{x})$

Two Types of Models

Generative Model

- $P(\mathrm{y}|\mathbf{x}) \propto P(\mathrm{y})\, p(\mathbf{x}|\mathrm{y})$

➔ To learn $P(\mathrm{y})$ and $p(\mathbf{x}|\mathrm{y})$.

Discriminative Model

- Directly learn $P(\mathrm{y}|\mathbf{x})$
- No assumption made on $p(\mathbf{x}|\mathrm{y})$

Two Types of Models

Generative Model

- $P(y|\mathbf{x}) \propto P(y)\, p(\mathbf{x}|y)$

➔ To learn $P(y)$ and $p(\mathbf{x}|y)$.

➔ Bayesian learning, Bayes classifiers.

- Example: Naïve Bayes Classifier

Discriminative Model

- Directly learn $P(y|\mathbf{x})$
- No assumption made on $p(\mathbf{x}|y)$

- Example: Logistic Regression

Practical Difficulty of Bayesian Learning

Consider doing Bayesian learning without making simplifying assumptions.

- Given n training pairs $<\mathbf{x}^{(i)}, y^{(i)}>$, $i=1, \ldots, n$. Each $\mathbf{x}^{(i)}$ is d-dimensional.
- We need to learn $P(y)$ and $p(\mathbf{x}|y)$

➔ $p(\mathbf{x}|y)$ can be very difficult to estimate:

➔ Consider a very simple case: binary features, and y is also binary. How many probabilities do we need to estimate?