GENERATIVE & DISCRIMINATIVE LEARNING

- GENERATIVE MODELS ATTEMPT TO ESTIMATE A DISTRIBUTION OVER ALL VARIABLES (INPUT AND OUTPUTS) IN A SYSTEM. THIS IS INEFFICIENT, SINCE WE ONLY NEED CONDITIONAL DISTRIBUTIONS OF OUTPUT GIVEN INPUT TO PERFORM CLASSIFICATION OR PREDICTION. THIS CAN BE WASTEFUL AND NON-ROBUST.

- USING A MINIMALIST APPROACH, DISCRIMINATIVE LEARNING ONLY CONSIDERS INPUT-OUTPUT MAPPINGS FOR CLASSIFICATION AND REGRESSION; AND ESCHEWS THE MODELING OF THE UNDERLYING DISTRIBUTIONS. (PROVIDES MODEL FOR TARGET VARIABLES CONDITIONED ON THE OBSERVED VARIABLE.)

EXAMPLES

1. GENERATIVE LEARNING
   - MAXIMUM LIKELIHOOD
   - MAXIMUM A POSTERIORI
   - BAYESIAN INFERENCE
   - NAIVE BAYES
   - HIDDEN MARKOV MODEL

2. DISCRIMINATIVE LEARNING
   - EMPIRICAL RISK MINIMIZATION
   - SUPPORT VECTOR MACHINES
     REGULARIZATION THEORY
   - MAXIMUM ENTROPY DISCRIMINATION
   - BAYES' POINT MACHINES
   - LINEAR REGRESSION
   - NEURAL NETWORKS

PROBABILISTIC MODELS FOR CLASSIFICATION

x = DATA VECTOR

y = LABEL/CLASS OF DATA VECTOR

GOAL: TO MODEL $p(y|x)$

TWO APPROACHES

1. GENERATIVE MODEL
2. DISCRIMINATIVE MODEL

1. GENERATIVE MODEL: MODEL $p(x,y)$

ASSUME THAT THE PRIOR $p(y)$ IS GIVEN.

OBTAIN $p(y|x)$ AS:

$$p(y|x) = \frac{p(x,y)}{p(x)} = \frac{p(x|y)\,p(y)}{\sum_y p(x,y)} \quad ; \quad p(x) > 0$$

$\therefore p(y|x)$ ARE MODELLED

2. DISCRIMINATIVE MODEL

DIRECTLY MODEL $p(y|x)$

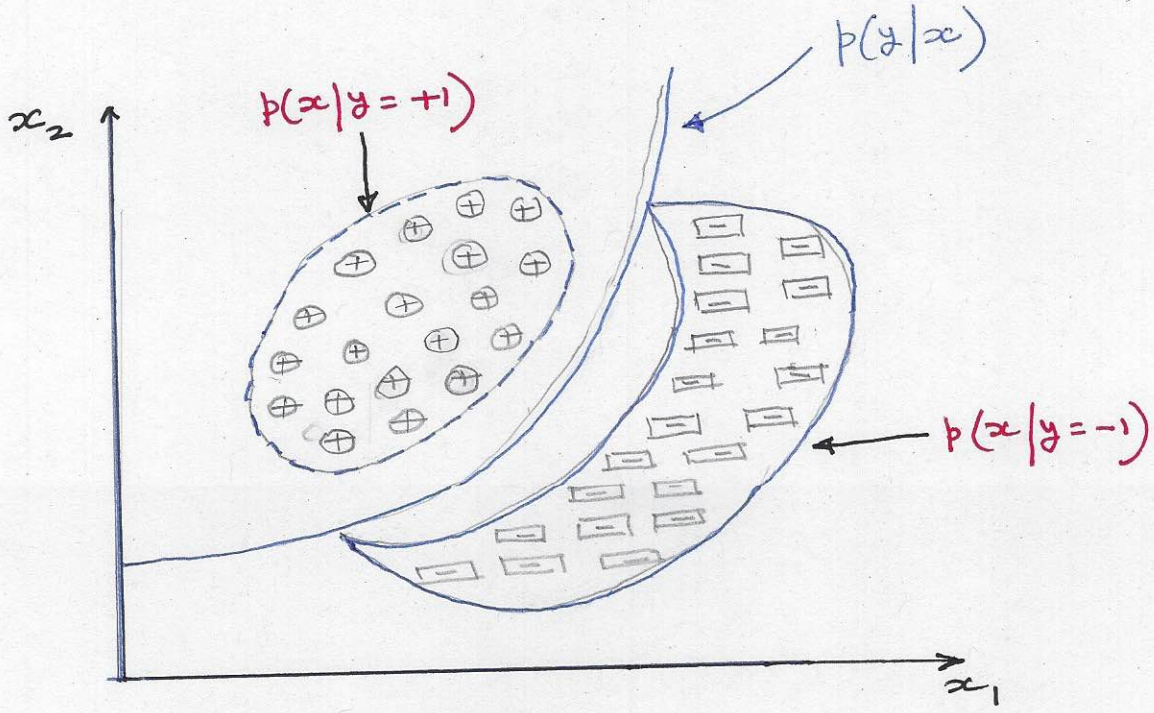
GENERATIVE - DISCRIMINATIVE MODELS

- x = DATA VECTOR
  $y \in \{-1, +1\}$ = LABEL OF DATA VECTOR
  IN A MULTICLASS PROBLEM WITH n CLASSES; $y \in \{1, 2, \ldots, n\}$
- GIVEN INPUT DATA POINT x, A DISCRIMINATIVE MODEL COMPUTES
  $p(y|x)$ = PROBABILITY OF x BEING LABELLED -1 OR $+1$
  NOTE: $p(y=-1|x) = 1 - p(y=+1|x)$
- A GENERATIVE MODEL, OFTEN CAPTURES THE GENERATION OF PROCESS x BY MODELING $p(x|y=+1)$ AND $p(x|y=-1)$
- SEE PICTURE
  - DISCRIMINATIVE MODEL $p(y|x)$ MOSTLY FOCUS ON HOW WELL THEY CAN SEPARATE THE POSITIVES FROM THE NEGATIVES. A SAMPLE FAR FROM THE DECISION BOUNDARY IN THE POSITIVE REGION MAY NOT LOOK LIKE A POSITIVE SAMPLE AT ALL. BUT A DISCRIMINATIVE MODEL WILL GIVE A HIGH PROBABILITY TO IT BEING POSITIVE
  - GENERATIVE MODELS TRY TO UNDERSTAND THE BASIC FORMATION OF INDIVIDUAL CLASSES, AND THUS, CARRY RICHER INFORMATION THAN DISCRIMINATIVE MODELS. THESE MODEL $p(x,y)$.
    GIVEN THE PRIOR $p(y)$, ONE CAN ALWAYS DERIVE DISCRIMINATIVE MODELS $p(y=+1|x)$ FROM THE GENERATIVE MODELS BASED ON BAYES' RULE BY

$$p(y=+1|x) = \frac{p(x|y=+1)\,p(y=+1)}{\sum_{y \in \{-1,1\}} p(x|y)\,p(y)}$$

HOWEVER GENERATIVE MODELS ARE MUCH HARDER TO LEARN THAN DISCRIMINATIVE MODELS, AND OFTEN SIMPLIFIED ASSUMPTIONS ARE MADE ABOUT DATA FORMATION.

- DISCRIMINATIVE MODELS FOCUS ON CLASSIFICATION BOUNDARIES BETWEEN THE POSITIVE AND THE NEGATIVES, WHERE AS GENERATIVE MODELS EMPHASIZE THE DATA GENERATION PROCESS IN EACH INDIVIDUAL CLASS.



GENERATIVE VS. DISCRIMINATIVE MODELS