CS 4824/ECE 4424: GNB Decision Surface

Acknowledgement:
Many of these slides are derived from Tom Mitchell, Pascal Poupart, Pieter Abbeel, Eric Eaton, Carlos Guestrin, William Cohen, and Andrew Moore.
Gaussian Naïve Bayes - decision surface

- Assume $Y=$PlayBasketball (boolean) $X_1=$Height $X_2=$Age
  
  $Y^{New} \leftarrow \arg \max_{y_k} P(Y | y_k) \prod_{i} P(X_{i}^{New} | Y = y_k); \text{assume } P(Y=1) = 0.5$
What is the minimum possible error?

- Best case:
  - Conditional independence assumption is satisfied
  - We can perfectly estimate $P(Y)$, $P(X|Y)$ (i.e. infinite training data)
But…

- Naïve Bayes allows estimating $P(Y|X)$ by learning $P(Y)$ and $P(X|Y)$

- Why not learn $P(Y|X)$ directly?