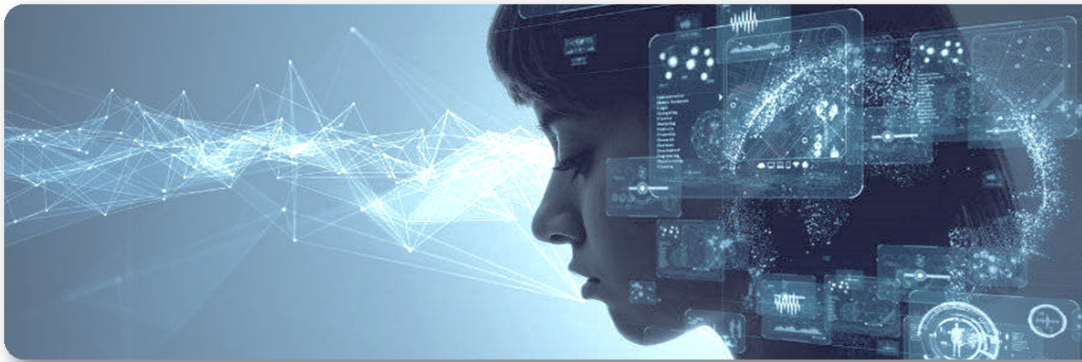

COMP 2211 Exploring Artificial Intelligence

Naïve Bayes Classifier - Why $P(e_i|B_i)$ can be found by substituting $x = e_i$ to $f(x)$?

Dr. Desmond Tsoi

Department of Computer Science & Engineering
HKUST, Hong Kong SAR, China

Explanation

- In continuous probabilities, the probability of getting precisely any given outcome is 0, and this is why densities are used instead.
- Therefore, we do not deal with expressions such as $P(X = x)$ but with $P(|X - x| < \Delta(x))$, which stands for the probability of X being close to x.
- Let's simplify the notation and write $P(X \sim x)$ for $P(|X - x| < \Delta(x))$.
- If we apply the Bayes' rule here, we will get

$$P(X \sim x|W \sim w) = \frac{P(W \sim w|X \sim x)P(X \sim x)}{P(W \sim w)}$$

- Because we are dealing with probabilities. If we now introduce densities:

$$pdf(x|w)\Delta(x) = \frac{pdf(w|x)\Delta(w)pdf(x)\Delta(x)}{pdf(w)\Delta(w)}$$

Explanation

- Since probability = density $\times$ neighborhood-size
- Also, since all $\Delta(\cdot)$ cancel out in the expression above, we get

$$pdf(x|w) = \frac{pdf(w|x)pdf(x)}{pdf(w)}$$

  which is the Bayes rule for densities.
- The conclusion is that, given that the Bayes rule also holds for densities, it is legitimate to use the same methods replacing probabilities with densities when dealing with continuous random variables.

Based on various resources in the Internet.