

In-Class-4

Madhu Peduri

1. Give the formula for the posterior numerator for each variety, e.g., posterior numerator (Iris-setosa).

$$\begin{aligned} P(\text{Iris-setosa} \mid X) = & p(\text{Iris-setosa}) * \\ & p(\text{sepal-length} \mid \text{Iris-setosa}) * \\ & p(\text{sepal-width} \mid \text{Iris-setosa}) * \\ & p(\text{petal-length} \mid \text{Iris-setosa}) * \\ & p(\text{petal-width} \mid \text{Iris-setosa}) \end{aligned}$$

$$\begin{aligned} P(\text{Iris-versicolor} \mid X) = & p(\text{Iris-versicolor}) * \\ & p(\text{sepal-length} \mid \text{Iris-versicolor}) * \\ & p(\text{sepal-width} \mid \text{Iris-versicolor}) * \\ & p(\text{petal-length} \mid \text{Iris-versicolor}) * \\ & p(\text{petal-width} \mid \text{Iris-versicolor}) \end{aligned}$$

$$\begin{aligned} P(\text{Iris-virginica} \mid X) = & p(\text{Iris-virginica}) * \\ & p(\text{sepal-length} \mid \text{Iris-virginica}) * \\ & p(\text{sepal-width} \mid \text{Iris-virginica}) * \\ & p(\text{petal-length} \mid \text{Iris-virginica}) * \\ & p(\text{petal-width} \mid \text{Iris-virginica}) \end{aligned}$$

2. Calculate P for each variety, e.g., $P(\text{Iris-setosa})$

$$P(\text{Iris-setosa}) = 50/150 = 1/3$$

$$P(\text{Iris-versicolor}) = 50/150 = 1/3$$

$$P(\text{Iris-virginica}) = 50/150 = 1/3$$

3. Give the formula for $p(\text{sepal-length} \mid \text{Iris-setosa})$, if the mean value and variance of sepal-length for Iris-setosa is 5.0 and 0.12, respectively. Substitute the values for x , μ , and σ^2 into the formula.

$$x = 5.9$$

$$\mu = \mu = 5.0$$

$$\sigma^2 = \text{variance} = 0.12$$

$$P(\text{sepal-length} \mid \text{Iris-setosa}) = (1/\sqrt{2 * \pi * 0.12}) * e ^ { -(5.9 - 5.0)^2/(2 * 0.12) } \\ = 0.04$$

4. How many conditional probabilities will the Naïve Bayesian Classifier need to calculate to classify the test sample?

Number of features = 4

Number of classes = 3

Total number of conditional probabilities = $4 * 3 = 12$

5. If posterior numerator(Iris-setosa) = 0.005, posterior numerator(Iris-versicolor) = 0.002, and posterior numerator(Iris-virginica) = 0.003, which variety did the Naïve Bayesian Classifier predict the test sample to be?

Iris -setosa, as it has highest posterior probability out of three classes.