Generative Models and Latent Variable Models

Advanced Statistical Inference

Simone Rossi

Naive Bayes

1. Consider the following training set with two classes and two features:

$$(\boldsymbol{x}_1, y_1) = ((0,1)^\top, 0), \qquad (\boldsymbol{x}_2, y_2) = ((2,3)^\top, 0),$$

$$(\boldsymbol{x}_3, y_3) = ((4,1)^\top, 1), \qquad (\boldsymbol{x}_4, y_4) = ((6,3)^\top, 1).$$

   Following the Gaussian Naive Bayes construction in the slides, compute:

   - the class priors $p(y=0)$ and $p(y=1)$;
   - the empirical means μ_{01}, μ_{02}, μ_{11}, μ_{12};
   - the empirical variances σ_{01}^2, σ_{02}^2, σ_{11}^2, σ_{12}^2.

2. Using the parameters from the previous exercise and the Gaussian Naive Bayes likelihood

$$p(\boldsymbol{x}_\star \mid y_\star = k) = \prod_{d=1}^{2} \mathcal{N}(x_{\star d} \mid \mu_{kd}, \sigma_{kd}^2),$$

   classify the test point

$$\boldsymbol{x}_\star = (1,2)^\top.$$

   Compute:

   - $p(\boldsymbol{x}_\star \mid y=0)$ and $p(\boldsymbol{x}_\star \mid y=1)$;
   - the unnormalized scores $p(\boldsymbol{x}_\star \mid y=k)p(y=k)$;
   - the posterior probabilities $p(y=0 \mid \boldsymbol{x}_\star)$ and $p(y=1 \mid \boldsymbol{x}_\star)$;
   - the predicted class.

3. Repeat the same computation for

$$\boldsymbol{x}_\star = (5,2)^\top.$$

   Compute the posterior probabilities and the predicted class.

4. Using the same Gaussian Naive Bayes parameters, consider the test point

$$\boldsymbol{x}_\star = (3, 2)^\top.$$

Compute the posterior probabilities in the following two cases:

- uniform prior: $p(y = 0) = p(y = 1) = 0.5$;
- imbalanced prior: $p(y = 0) = 0.2$, $p(y = 1) = 0.8$. In each case, give the predicted class.

Gaussian Mixture Models and EM

1. A two-component Gaussian mixture model has weights

$$\pi_1 = 0.4, \qquad \pi_2 = 0.6.$$

For a point $x = 1$, suppose the Gaussian densities evaluate to

$$\mathcal{N}(x; \mu_1, \sigma_1^2) = 0.30, \qquad \mathcal{N}(x; \mu_2, \sigma_2^2) = 0.10.$$

Compute:

- the joint terms $p(x, z_1 = 1)$ and $p(x, z_2 = 1)$;
- the marginal density $p(x)$;
- the responsibilities $\gamma(z_1)$ and $\gamma(z_2)$.

2. For another point $x = 3$, suppose the Gaussian densities are

$$\mathcal{N}(x; \mu_1, \sigma_1^2) = 0.05, \qquad \mathcal{N}(x; \mu_2, \sigma_2^2) = 0.25.$$

Using the same mixture weights, compute the two responsibilities.

3. Consider the dataset

$$x_1 = 1, \qquad x_2 = 3.$$

Using the same mixture parameters as in the two previous exercises, compute the log-likelihood

$$\log p(\boldsymbol{X} \mid \pi_1, \pi_2, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \sum_{n=1}^{2} \log \Big(\pi_1 \mathcal{N}(x_n; \mu_1, \sigma_1^2) + \pi_2 \mathcal{N}(x_n; \mu_2, \sigma_2^2) \Big).$$

4. In one dimension, an E-step on the dataset

$$x_1 = 0, \qquad x_2 = 2, \qquad x_3 = 4$$

returns the responsibilities

$$\gamma_{11} = 1, \qquad \gamma_{21} = 0.5, \qquad \gamma_{31} = 0,$$

$$\gamma_{12} = 0, \qquad \gamma_{22} = 0.5, \qquad \gamma_{32} = 1.$$

Compute the M-step updates:

- N_1 and N_2;
- π_1 and π_2;
- μ_1 and μ_2;
- the scalar variances σ_1^2 and σ_2^2.

5. Consider now a three-component mixture model with

$$\pi_1 = 0.2, \qquad \pi_2 = 0.5, \qquad \pi_3 = 0.3.$$

For a point $x = 2$, suppose

$$\mathcal{N}(x; \mu_1, \sigma_1^2) = 0.10, \qquad \mathcal{N}(x; \mu_2, \sigma_2^2) = 0.40, \qquad \mathcal{N}(x; \mu_3, \sigma_3^2) = 0.20.$$

Compute:

- the marginal density $p(x)$;
- the three responsibilities $\gamma(z_1)$, $\gamma(z_2)$, $\gamma(z_3)$;
- which component has the largest posterior responsibility.