منبع: جزوه ؟

اسرا آذرنوش ۹۸۱۱۷۰۹۹۸

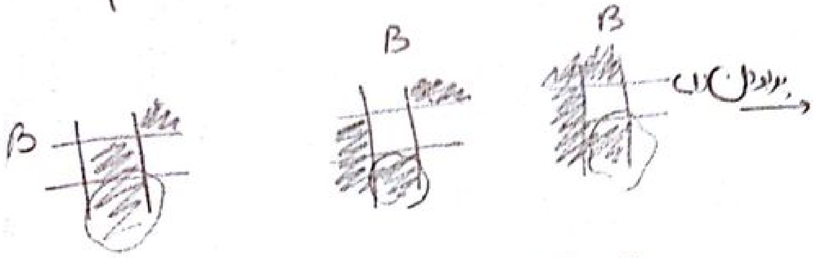
۱) هر حرف ویژن یک ماتریس ۳×۳ است که یا مقدارهای ۱ (سیاه) دارد و یا ندارد A

پر بودن (۱) →

۱/۸	۲/۲	۱/۳
۰	۱	۱
۱/۲	۰	۰

Laplace →

۲/۶	۳/۶	۲/۶
۱/۶	۴/۶	۴/۶
۲/۶	۱/۶	۱/۶

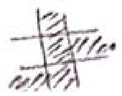


پر بودن (۱) →

۱/۲	۱/۲	۲/۲
۲/۲	۱/۲	۰
۲/۲	۱	۰

Laplace →

۱/۶	۳/۶	۳/۶
۴/۶	۴/۶	۱/۶
۳/۶	۳/۶	۱/۶



A or B →

۰	۱	۰
۰	۱	۱
۱	۱	۰

احتمال هر ۲ را محاسبه و هر کدام بیشتر بود به احتمال بیشتر مال همان است

$$P(A|x) \rightarrow \frac{3}{6}\times\frac{3}{6}\times\frac{2}{6}\times\frac{4}{6}\times\frac{4}{6}\times\frac{4}{6}\times\frac{2}{6}\times\frac{1}{6}\times\frac{2}{6} = \frac{4^3\times 2^5\times 3}{6^9}$$

$$P(B|x) \rightarrow \frac{4}{6}\times\frac{2}{6}\times\frac{3}{6}\times\frac{2}{6}\times\frac{1}{6}\times\frac{1}{6}\times\frac{3}{6}\times\frac{4}{6}\times\frac{4}{6} = \frac{4^3\times 3^2\times 2^2}{6^9}$$

۳ > ۱ ⟶ احتمال A است

احتمال پر بودن هر خانه برای هر حرف هر ۳ را بدست می آوریم و احتمال نبود نیز بدست می آید (نبودن = ۱ − بودن)

پس با توجه به رنگی بودن یا نبودن حرفی که باید بدانیم احتمال را برای هر ۲ محاسبه می کنیم و آنکه مقدار بیشتری دارد به احتمال بیشتر آن حرف است

۱) ساده ← مثال ← رنگی بودن یا نبودن

← کدام ← feature

۲) برای محاسبه feature ها اطلاعات را محاسبه می کنیم و آنکه بیشترین IG را دارد در نظر می گیریم

ج

$w = 0$

اگر اطلاعات ترینینگ قابل جداسازی باشد آنگاه الگوریتم Perceptron training در نهایت Converge خواهد شد

با فرض اینکه داده‌ها به صورت خطی قابل تفکیک هستند حاشیه را به صورت زیر تعریف می‌کنیم

$$\delta = \max_{w} \min_{(x_i, y_i)} [y_i w^T x_i] \quad \text{به طوری که} \quad \|w\| = 1$$

$\hat{w}$ را hyperplane قابل جداسازی با حاشیه δ در نظر بگیریم

$$\begin{cases} \hat{w}^T w^{(k)} = \hat{w}^T (w^{(k-1)} + y_i x_i) \\ = \hat{w}^T w^{(k-1)} + \hat{w}^T y_i x_i \\ \geq \hat{w}^T w^{(k-1)} + \delta \\ \geq k\delta \end{cases}$$

$$\|w^{(k)}\| \geq k\delta$$

$$\begin{cases} \|w^{(k)}\|^2 = \|w^{(k-1)} + y_i x_i\|^2 \\ \leq \|w^{(k-1)}\|^2 + \|y_i x_i\|^2 \\ \leq \|w^{(k-1)}\|^2 + 1 \\ \leq k \end{cases} \qquad y_i w^{(k-1)T} x_i < 0$$

$$\|w^{(k)}\| \leq \sqrt{k}$$

$$k\delta \leq \|w^{(k)}\| \leq \sqrt{k} \longrightarrow k \leq 1/\delta^2$$

اگر دسته‌ها جداپذیر نباشند الگوریتم هرگز تمام نمی‌شود

$$\|w^k\|_2 \|w\|_2 \cos\theta \geq \sqrt{k}\delta$$

$$\cos\theta \geq \sqrt{k}\delta \longrightarrow k \leq 1/\delta^2 \longrightarrow \cos\theta \leq 1$$

بنابراین در تعداد متناهی گام برای دسته‌هایی که خطی جداپذیرند Converge می‌کند