

تکلیف پنجم درس شناسایی الگو

وحید ملکی
شماره دانشجویی: ۴۰۳۱۳۰۰۴

۲۴ آبان ۱۴۰۴

سؤال ۹

در این مسئله، هدف طراحی یک طبقه‌بند برای پیش‌بینی جنسیت فرد (G) بر اساس دو ویژگی وزن (W) و وضعیت فارغ‌التحصیلی (S) است. فرض شده است:

- W و S از یکدیگر مستقل هستند.
- واریانس توزیع $P(W | G = \text{female})$ و $P(W | G = \text{male})$ برابر است.
- $G \in \{\text{male}, \text{female}\}$
- W : متغیر پیوسته
- $S \in \{0, 1\}$: متغیر باینری

الف) آیا استفاده از طبقه‌بند Bayes Naïve مناسب است؟

بله، استفاده از طبقه‌بند Bayes Naïve در این مسئله کاملاً مناسب و منطقی است.
دلیل: طبقه‌بند Bayes Naïve بر پایه فرض استقلال شرطی ویژگی‌ها نسبت به کلاس بنا شده است، یعنی:

$$P(W, S | G = g) = P(W | G = g) \cdot P(S | G = g)$$

در این مسئله صراحتاً گفته شده که W و S از یکدیگر مستقل هستند. این فرض دقیقاً با اصل اساسی Bayes Naïve هم‌خوانی دارد. بنابراین، این روش نه تنها مجاز، بلکه انتخاب بهینه و کارآمد است.
قانون طبقه‌بندی به صورت زیر است:

$$\hat{G} = \arg \max_{g \in \{\text{male}, \text{female}\}} P(G = g | W, S)$$

با استفاده از قضیه بیز:

$$P(G = g | W, S) \propto P(W, S | G = g) \cdot P(G = g)$$

و با اعمال فرض استقلال (naïve):

$$P(W, S | G = g) = P(W | G = g) \cdot P(S | G = g)$$

ب) توزیع‌های احتمالی مورد نیاز و تعداد پارامترها

از آنجا که استفاده از Bayes Naïve مناسب است، توزیع‌های زیر باید از داده‌های آموزشی تخمین زده شوند:

۰۱. توزیع پیشین $P(G)$

• $P(G = \text{male})$

• $P(G = \text{female}) = 1 - P(G = \text{male})$

تعداد پارامتر: 1

۰۲. توزیع شرطی وزن $P(W | G)$

متغیر W پیوسته است □ مدل‌سازی با توزیع گوسی برای هر کلاس:

$$P(W | G = \text{male}) \sim \mathcal{N}(\mu_{\text{male}}, \sigma_W^2), \quad P(W | G = \text{female}) \sim \mathcal{N}(\mu_{\text{female}}, \sigma_W^2)$$

(واریانس مشترک σ_W^2 طبق فرض مسئله)
پارامترها:

• μ_{male}

• μ_{female}

• σ_W^2 (مشترک)

تعداد پارامتر: 3

۰۳. توزیع شرطی وضعیت فارغ‌التحصیلی $P(S | G)$

متغیر S باینری است □ مدل‌سازی با توزیع برنولی برای هر کلاس:

• $\theta_{S|\text{male}} = P(S = 1 | G = \text{male})$

• $\theta_{S|\text{female}} = P(S = 1 | G = \text{female})$

(احتمال $S = 0$ مکمل است)

تعداد پارامتر: 2

جمع کل پارامترها

$$1 + 3 + 2 = 6$$

تعداد	پارامترها	توزیع
۱	$P(G = \text{male})$	$P(G)$
۳	$\mu_{\text{male}}, \mu_{\text{female}}, \sigma_W^2$	$P(W G)$
۲	$\theta_{S \text{male}}, \theta_{S \text{female}}$	$P(S G)$
۶	مجموع	

در نتیجه، طبقه‌بند Bayes Naïve با ۶ پارامتر کاملاً قابل آموزش است و با فرضیات مسئله سازگار است.