

# تکلیف پنجم درس شناسایی الگو

وحید ملکی

شماره دانشجویی: ۴۰۳۱۳۰۰۴

۲۴ آبان ۱۴۰۴

## سؤال ۱۳

کدام یک از گزینه‌های زیر درباره‌ی تخمین پارزن صحیح هستند؟

- (آ) با انتخاب بهینه‌ی پارامتر هموارسازی، طبقه‌بند پارزن برای مجموعه‌های آموزشی بسیار بزرگ، طبقه‌بند بیز را تقریب می‌زند.
- (ب) تخمین گر پارزن یک تخمین بی‌اریب (بدون بایاس) از توزیع ارائه می‌دهد.
- (ج) معمولاً از تخمین گر پارزن برای ساخت طبقه‌بندهای ناپارامتری و از تخمین گر kNN برای تخمین تابع چگالی احتمال استفاده می‌شود.
- (د) هرچقدر تعداد نمونه‌ها بیشتر شود، در تخمین توزیع به روش پارزن از کرنل با هموارسازی کمتر استفاده می‌گردد.

بررسی گام به گام

(الف) ✓ صحیح

• با افزایش تعداد نمونه‌ها ( $n \rightarrow \infty$ ) و انتخاب بهینه‌ی پهنای پنجره ( $h \rightarrow 0$  به صورت کنترل‌شده)، تخمین پارزن به توزیع واقعی همگرا می‌شود.

• در این حالت، طبقه‌بند پارزن عملکردی نزدیک به طبقه‌بند بیز بهینه خواهد داشت.

(ب) ✗ نادرست

- تخمین پارزن بایاس دارد (مگر در حالت حدی  $n \rightarrow \infty$  و  $h \rightarrow 0$ ).
- بایاس به دلیل هموارسازی ایجاد می‌شود و به طور کلی بی‌اریب نیست.

(ج) ✗ نادرست

• تخمین پارزن  $\rightarrow$  برای تخمین چگالی احتمال (ناپارامتری)

• kNN  $\rightarrow$  معمولاً برای طبقه‌بندی ناپارامتری

•  $\Rightarrow$  معکوس بیان شده  $\rightarrow$  گزینه نادرست

(د) ✓ صحیح

- با افزایش تعداد نمونه‌ها، اطلاعات بیشتری داریم.
- می‌توان  $h$  (پارامتر هموارسازی) را کوچک‌تر کرد.
- نتیجه: تخمین دقیق‌تر با هموارسازی کمتر

—

خلاصه پاسخ

گزینه	توضیح مختصر	وضعیت
الف	تقریب بیز با $n \uparrow$ ، $h$ بهینه	✓
ب	تخمین پارزن بایاس دارد	×
ج	پارزن برای چگالی، kNN برای طبقه‌بندی	×
د	$n \uparrow \Rightarrow h \downarrow$	✓