

## ضمیمه

### نمادگذاری

در زیر خلاصه‌ای از نمادگذاری این کتاب آورده شده است:

$(a, b]$ : گروه و پرانتز برای نشان دادن بازه‌ها به کار می‌روند، گروه‌ها نشان‌دهنده‌ی این است که بازه شامل مقدار مرز است و پرانتز نشان‌دهنده‌ی این است که بازه مرز را شامل نمی‌شود.

برای مثال  $(1, 3]$  یعنی  $1 < x \leq 3$ .

مجموع  $\sum_{i=1}^n x_i$ :  $x_1 + x_2 + \dots + x_n$ .

حاصل ضرب  $\prod_{i=1}^n x_i$ :  $x_1 \cdot x_2 \cdot \dots \cdot x_n$ .

$\vdash$ : نماد نتیجه‌گیری منطقی. برای مثال  $A \vdash B$  یعنی که  $B$  از  $A$  نتیجه‌گیری منطقی می‌شود.

$>_g$ : نماد کلی‌تر بودن. برای مثال  $h_i >_g h_j$  نشان می‌دهد که  $h_i$  از  $h_j$  کلی‌تر است.

$\operatorname{argmax}_{x \in X} f(x)$ : مقدار  $x$  است که  $f(x)$  در آن ماکزیمم می‌شود. برای مثال،

$$\operatorname{argmax}_{x \in \{1, 2, -3\}} x^2 = -3$$

$\hat{f}(x)$ : تابعی که تابع  $f(x)$  را تخمین می‌زند.

$\delta$ : در یادگیری PAC، مرزی بر روی احتمال شکست است. در شبکه‌های عصبی نیز، جمله‌ی خطای مربوطه‌ی خروجی یک تک واحد است.

$\epsilon$ : مرز خطا فرضیه (در یادگیری PAC).

$\eta$ : ضریب یادگیری در شبکه‌های عصبی و متدهای یادگیری مربوطه.

$\mu$ : میانگین توزیع احتمال.

$\sigma$ : انحراف معیار توزیع احتمال.

$\nabla E(\vec{w})$ : گرادیان  $E$  نسبت به بردار  $\vec{w}$ .

$C$ : کلاس توابع هدف ممکن

$D$ : داده‌های آموزشی

$\mathcal{D}$ : توزیع احتمال روی فضای نمونه‌ای.

$E[x]$ : مقدار امید  $x$ .

$E(\vec{w})$ : مجموع خطاهای مربعی شبکه‌ی عصبی با بردار وزن‌های  $\vec{w}$ .

Error: خطای فرضیه گسسته مقدار یا پیش‌بینی.

$H$ : فضای فرضیه‌ای.

$h(x)$ : پیش‌بینی فرضیه‌ی  $h$  برای نمونه‌ی  $x$ .

$P(x)$ : احتمال  $x$ .

$\text{Pr}(x)$ : احتمال اتفاق  $x$ .

$p(x)$ : چگالی توزیع احتمال  $x$ .

$Q(s, a)$ : تابع  $Q$  در یادگیری تقویتی.

$\mathcal{K}$ : مجموعه‌ی اعداد حقیقی.

$VC(H)$ : بعد Vapnik-Chervonenkis فضای فرضیه‌ای  $H$ .

$V S_{H,D}$ : فضای ویژه؛ مجموعه‌ای از فرضیه‌هایی از  $H$  که با  $D$  سازگارند.

$w_{ij}$ : در شبکه‌های عصبی وزن از گره  $i$  به گره  $j$  است.

$X$ : فضای نمونه‌ای.