



تمرین پنجم - بخش دوم

شماره دانشجویی: ۹۸۱۰۱۰۷۴

محمدجواد هزاره

سوال ۱

(آ) با این روش اطلاعات پیشین خود از پارامترها را به تخمین اضافه می‌کنیم. به این صورت که تعدادی داده‌ی مصنوعی به نمونه اضافه می‌کنیم که از توزیع پیشین مورد نظر ما تبعیت می‌کنند و سپس با داده‌های جدید و این داده‌ها تخمین را انجام می‌دهیم. با این کار به نحوی از بیش‌برازش (overfitting) می‌توان جلوگیری کرد.

(ب) جداول CPT به صورت زیر خواهد بود.

X_i	Y	$\mathbb{P}(X_1 Y)$	$\mathbb{P}(X_2 Y)$	$\mathbb{P}(X_3 Y)$	$\mathbb{P}(X_4 Y)$
0	Stop	$\frac{1}{4}$	$\frac{1}{4}$	1	$\frac{1}{4}$
1	Stop	$\frac{3}{4}$	$\frac{3}{4}$	0	$\frac{3}{4}$
0	Go	1	$\frac{2}{5}$	0	$\frac{3}{5}$
1	Go	0	$\frac{3}{5}$	1	$\frac{2}{5}$

Y	$\mathbb{P}(Y)$
Stop	$\frac{4}{9}$
Go	$\frac{5}{9}$

بنابراین چون احتمال صفر داریم بهتر است از Laplace Smoothing استفاده کنیم. با توجه به تعداد داده‌های مشاهده شده، از این روش با $k = 2$ استفاده می‌کنیم. بنابراین جدول‌ها به صورت زیر خواهند شد.

X_i	Y	$\mathbb{P}(X_1 Y)$	$\mathbb{P}(X_2 Y)$	$\mathbb{P}(X_3 Y)$	$\mathbb{P}(X_4 Y)$
0	Stop	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{6}{8}$	$\frac{3}{8}$
1	Stop	$\frac{5}{8}$	$\frac{5}{8}$	$\frac{2}{8}$	$\frac{5}{8}$
0	Go	$\frac{7}{9}$	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{5}{9}$
1	Go	$\frac{2}{9}$	$\frac{5}{9}$	$\frac{7}{9}$	$\frac{4}{9}$

Y	$\mathbb{P}(Y)$
Stop	$\frac{4}{9}$
Go	$\frac{5}{9}$

حال برای احتمال برچسب داده‌ی جدید داریم:

$$\begin{cases} \mathbb{P}(\text{Stop} | 1, 1, 1, 0) = \frac{4}{9} \times \frac{5}{8} \times \frac{5}{8} \times \frac{2}{8} \times \frac{3}{8} \approx 0.016 \\ \mathbb{P}(\text{Go} | 1, 1, 1, 0) = \frac{5}{9} \times \frac{2}{9} \times \frac{5}{9} \times \frac{7}{9} \times \frac{5}{9} \approx 0.030 \end{cases}$$

بنابراین برچسب Go به این داده تخصیص خواهد یافت.