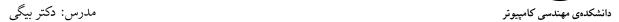
نظریهی یادگیری ماشین

نيمسال دوم ۱۴۰۱ _ ۱۴۰۰



تمرین سری اول موعد تحویل: ۲۰ اسفند

مسئلهی ۱.

$$\mathbb{P}\left[\exists h \in \mathcal{H} \text{ s.t. } L_{\left(\overline{\mathcal{D}}_m,f\right)}(h) > \epsilon \text{ and } L_{(S,f)}(h) = 0\right] \leqslant |\mathcal{H}|e^{-\epsilon m}$$

مسئلهي ٢.

نشان دهید که برای هر توزیع $\mathcal D$ ، دسته بند بهینه ی بیز که آن را با $f_{\mathcal D}$ نشان می دهیم بهینه است، به این معنا که برای هر دسته بند $g:\mathcal X\to 0,1$ داشته باشیم:

$$L_{\mathcal{D}}(f_{\mathcal{D}}) \leqslant L_{\mathcal{D}}(g)$$

راهنمایی: در تعریف دسته بند بهینه بیز داریم، برای هر توزیع احتمالاتی \mathcal{D} بر روی $0,1 \to \mathcal{X}$ ، بهترین تابع پیشبینی کننده برچسب از \mathcal{X} به 0,1 به شکل زیر خواهد بود:

$$f_{\mathcal{D}}(x) = \begin{cases} 1 & \text{if } \mathbb{P}[y=1 \mid x] \geqslant 1/2 \\ 0 & \text{otherwise} \end{cases}$$

مسئلهي ٣.

با فرض وجود الگوریتم یادگیری احتمالا تقریبا درست برای کلاس کانسپت C بر روی \mathbb{R}^2 که از تقاطع نیم فضاهای همتراز با محور شکل گرفته است:

$$\left\{ (x,y) : x \geqslant c_x, y \geqslant c_y \right\},$$

$$\left\{ (x,y) : x \geqslant c_x, y \leqslant c_y \right\},$$

$$\left\{ (x,y) : x \leqslant c_x, y \geqslant c_y \right\},$$

$$\left\{ (x,y) : x \leqslant c_x, y \leqslant c_y \right\},$$

به طوری که \mathbb{R} بدست آوردید. پیچیدگی نمونه را برای $c_x, c_y \in \mathbb{R}$ بدست

مسئلهي ۴.

یک نوع مدل PAC دو توزیعه وجود دارد که الگوریتم یادگیری آن ممکن است صراحتا نمونه های مثبت و منفی را درخواست کند اما باید فرضیهای را بیابد که بر روی هر دو توزیع دادههای مثبت و منفی به خوبی عمل کند. زمانی میگوییم یک مدل PAC دو توزیعه با الگوریتم A به صورت احتمالا تقریبا صحیح کلاس فرضیه \mathcal{H} را یاد میگیرد که اگر برای هر کانسپت هدف \mathcal{H} ، توزیع \mathcal{D}^+ بر روی دادههای با برچسب مثبت، توزیع \mathcal{D}^+ بر روی دادههای با برچسب منفی، برای هر (ϵ,δ) و اگر به اندازه کافی (اما محدود) داده مثبت و منفی به صورت i.i.d. از دو توزیع \mathcal{D}^+ و اگر به اندازه کافی f داده شود، سپس f فرضیه f را خروجی بدهد به طوری که با احتمال حداقل f داشته باشیم:

$$\Pr_{x \sim \mathcal{D}^+}[h(x) = 0] \leqslant \epsilon \text{ and } \Pr_{x \sim \mathcal{D}^-}[h(x) = 1] \leqslant \epsilon$$

الف) ثابت کنید اگر \mathcal{H} با استفاده از مدل پایه ی یک توزیعه ی PAC قابل یادگیری باشد، با استفاده از مدل دو توزیعه ی PAC نیز قابل یادگیری خواهد بود.

ب) (امتیازی) فرض کنید h_0 تابعی است که همیشه 0 خروجی میدهد و h_1 تابعی است که همیشه 1 خروجی میدهد. ثابت کنید اگر یک کلاس فرضیه \mathcal{H} با استفاده از مدل دوتوزیعه ی PAC قابل یادگیری باشد، کلاس فرضیه \mathcal{H} نیز بر روی مدل یک توزیعه ی PAC قابل یادگیری است.

راهنمایی: می توانید از Chernoff bound استفاده کنید.

موفق باشيد