نظریه یادگیری ماشین

نیم‌سال اول ۱۴۰۳-۱۴۰۴

مدرس: دکتر امیر نجفی

دانشکده‌ی مهندسی کامپیوتر

تمرین دوم

مسئله‌ی ۱. وابستگی پیچیدگی راداماخر به توزیع (۱۰ نمره)

فرض کنید توزیع P را به گونه‌ای انتخاب کرده‌ایم که تک‌نقطه‌ای باشد یعنی $z = \delta(z.)$ و نمونه‌ها از این توزیع می‌آیند. فرض کنید خانواده‌ی $\mathcal{F}$ را به گونه‌ای انتخاب کرده‌ایم که برای هر $x \in \mathbb{R}$ و $f \in \mathcal{F}$، $-۱ \leqslant f(x) \leqslant ۱$.
نشان دهید که:

$$\mathcal{R}_n(\mathcal{F}) \leqslant \frac{۱}{\sqrt{n}}$$

مسئله‌ی ۲. خواص بعد VC (۱۵ نمره)

بعد VC یک خانواده از مجموعه های $\mathcal{S}$ را برابر بعد VC توابع به شکل ۱_S تعریف می‌کنیم که $S \in \mathcal{S}$. فرض کنید $\mathcal{S}, \mathcal{T}$ دو خانواده از مجموعه ها با بعد VC محدود باشند نشان دهید خانواده های زیر نیز بعد VC محدود دارند:

الف)

$$\mathcal{S}^c := \{S^c | S \in \mathcal{S}\}$$

ب)

$$\mathcal{S} \sqcap \mathcal{T} = \{S \cap T | S \in \mathcal{S}, T \in \mathcal{T}\}$$

ج)

$$\mathcal{S} \sqcup \mathcal{T} = \{S \cup T | S \in \mathcal{S}, T \in \mathcal{T}\}$$

مسئله‌ی ۳. خواص پیچیدگی راداماخر (۱۵ نمره)

خواص زیر از پیچیدگی راداماخر را ثابت کنید:

الف)

$$\mathcal{R}_n(\mathcal{F}) = \mathcal{R}_n(\mathrm{conv}(\mathcal{F}))$$

ب)

$$\mathcal{R}_n(\mathcal{F} + \mathcal{G}) \leqslant \mathcal{R}_n(\mathcal{F}) + \mathcal{R}_n(\mathcal{G})$$

و مثالی ارائه دهید که نتوان کران را بهبود داد.

ج) برای تابع ثابت و کراندار g:

$$\mathcal{R}_n(\mathcal{F}+g) \leqslant \mathcal{R}_n(\mathcal{F}) + \frac{||g||_\infty}{\sqrt{n}}$$

مسئله‌ی ۴. محاسبه بعد VC (۱۰ نمره)

بعد VC مجموعه‌های زیر را حساب کنید:

الف) مجموعه دایره‌های در صفحه.

ب) مجموعه مربع‌های در صفحه.

مسئله‌ی ۵. SVM تنک (۱۵ نمره)

یکی دیگر از روش‌های فرموله کردن الگوریتم SVM، بر پایه تنکی (sparsity) بردارهای کمکی است. فرض کنید به جای بیشینه کردن مقدار حاشیه می‌خواهیم تنکی بردارها را بیشینه کنیم. این کار را با کمینه کردن $||\alpha||_p$ که با آن بردار w به دست می‌آید انجام می‌دهیم ($p \geqslant ۱$). ابتدا حالتی را درنظر بگیرید که $p = ۲$. به مساله بهینه سازی زیر می‌رسیم:

$$\min_{\alpha,b} \frac{۱}{۲}\sum_{i=۱}^{m} \alpha_i^۲ + C\sum_{i=۱}^{m}\xi_i$$

$$\text{subject to} \quad y_i\left(\sum_{j=۱}^{m}\alpha_j y_j \mathbf{x}_i\cdot\mathbf{x}_j + b\right) \geqslant ۱ - \xi_i, \quad i\in[m]$$

$$\xi_i, \alpha_i \geqslant ۰, \quad i \in [m].$$

الف) نشان دهید که با در نظر گرفتن قید نامنفی بودن برای α این مسئله را می‌توانیم به یک نمونه از مسئله اولیه SVM تبدیل کنیم.

ب) مساله بهینه سازی دوگان مساله بالا را بنویسید.

ج) اگر قرار دهیم $p = ۱$، به یک α تنک‌تر دست می‌یابیم. مساله دوگان را در این حالت به دست آورید.

مسئله‌ی ۶. قضیه رادون (۱۰ نمره)

برای مجموعه $S = \{\mathbf{x}_۱, \dots, \mathbf{x}_m | \forall i \in [m] : \mathbf{x}_i \in \mathbb{R}^n\}$، پوش محدب S مجموعه‌ای است به صورت:

$$\{\mathbf{z} \in \mathbb{R}^n \mid \exists \lambda_1, \dots, \lambda_m \in [0, 1], \sum_{i=1}^{m} \lambda_i = 1, \mathbf{z} = \sum_{i=1}^{m} \lambda_i \mathbf{x}_i\}.$$

اگر $m \geqslant n + 2$، ثابت کنید S دارای دو زیرمجموعه مجزا S_1 و S_2 است که پوش محدب آن‌ها تداخل دارد.

مسئله‌ی ۷. مثال نقض (۱۰ نمره)

اغلب این‌گونه است که بُعد VC یک کلاس فرضیه برابر (یا از بالا محدود شونده) با تعداد پارامترهایی است که برای تعریف هر فرضیه در کلاس نیاز است. برای مثال، اگر $\mathcal{H}$ کلاس مستطیل‌های هم‌تراز با محور در $\mathbb{R}^d$ باشد، آنگاه $VCdim\mathcal{H} = 2d$ است که برابر با تعداد پارامترهای مورد نیاز برای تعریف مستطیلی در $\mathbb{R}^d$ می‌باشد. در اینجا مثالی وجود دارد که نشان می‌دهد این همیشه درست نیست. ما خواهیم دید که یک کلاس فرضیه می‌تواند بسیار پیچیده باشد و حتی یادگیری‌پذیر نباشد، اگرچه تعداد کمی پارامتر دارد.

دامنه $\mathcal{X} = \mathbb{R}$ و کلاس فرضیه

$$\mathcal{H} = \{x \mapsto \lceil \sin(\theta x) \rceil : \theta \in \mathbb{R}\}$$

را در نظر بگیرید. (اینجا، $\lceil -1 \rceil = 0$ را می‌گیریم). اثبات کنید که $VCdim\mathcal{H} = \infty$.

نکته: راه‌های متعددی برای اثبات این نتیجه وجود دارد. یک راه با اعمال لم زیر است: اگر $0.x_1x_2x_3\dots$ بسط دودویی $x \in (0, 1)$ باشد، آنگاه برای هر عدد طبیعی m، $\lceil \sin(2^m \pi x) \rceil = (1 - x_m)$ به شرطی که $\exists k \geqslant m$ که در آن $x_k = 1$.

مسئله‌ی ۸. رأی اکثریت (۱۵ نمره)

نشان دهید که اگر کلاس فرضیه $\mathcal{H}$ دارای VC-dimension برابر d باشد، آنگاه کلاس $\text{MAJ}_k(\mathcal{H})$ دارای VC-dimension برابر $O(kd \log kd)$ است. منظور از $\text{MAJ}_k(\mathcal{H})$ کلاس توابعی است که با گرفتن رأی اکثریت بر روی k تابع از $\mathcal{H}$ به دست می‌آید.