



مقدمه‌ای بر یادگیری ماشین

پاییز ۱۴۰۱

اساتید: علی شریفی، بهروز آذرخلیلی

دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

تاریخ برگزاری: ۱ آذر

کوئیز اول

سوالات (۱۰۰ نمره)

۱. (۲۰ نمره) موارد زیر را محاسبه کنید.

• (L1-norm)

$$A = \begin{bmatrix} 2 & -8 \\ 3 & 1 \end{bmatrix}$$

• (infinity-norm)

$$A = \begin{bmatrix} 3 & 6 & -1 \\ 3 & 1 & 0 \\ 2 & 4 & -7 \end{bmatrix}$$

۲. (۴۰ نمره) درستی یا نادرستی گزاره‌های زیر را با ذکر توضیح مختصر مشخص کنید:

الف) در Soft Margin SVM کاهش مقدار پارامتر C حاشیه‌ی جداسازی دو دسته کاهش می‌یابد و خطای آموزش افزایش می‌یابد.

ب) با افزایش مقدار K در روش K-fold cross validation میانگین خطای تخمین زده شده کاهش می‌یابد.

ج) الگوریتم K-means به مقدار دهی اولیه حساس است اما به طور تضمینی حتما همگرا می‌شود.

د) با توجه به داده‌های موجود در شکل زیر، در صورتی که از یک کرنل چند جمله‌ای درجه ۲ به بالا به جای یک کرنل خطی استفاده کنیم، بردارهای پشتیبان تغییر نخواهد کرد.



ه) فرض کنید می‌خواهیم سن افراد را بر اساس اسکن مغز آنها با استفاده از Regression تخمین بزنیم. تعداد افرادی که اسکن مغز آنها در دسترس است ۱۰ نفر است و یک بردار ویژگی ۲۰۰۰۰ تایی برای هر فرد بعد از اسکن مغز در دسترس است.

در این حالت استفاده از روش L2 Regularization را به روش L1 Regularization ترجیح می‌دهیم.

۳. (۲۰ نمره) داده‌های دو بعدی زیر را در نظر بگیرید.

$$D = \{(-2, 1), (-1, 3), (-1, -2)\}$$

برای این داده‌ها اولین مولفه‌ی اصلی را بدست آورید.

۴. (۲۰ نمره - امتیازی) ماتریس $A \in \mathbb{R}^{p \times p}$ را یک ماتریس متقارن و مثبت نیمه معین در نظر بگیرید. اثبات کنید

$$k(x, y) = x^T A y$$

یک تابع کرنل معتبر است.