

الف) مدل های پارامتریک دارای تعداد ثابت و محدودی پارامتر مستقل از اندازه دیتا ست هستند در نتیجه :  
 Linear SVM primal problem دارای  $D+1$  پارامتر  $(w, b)$  هستند و میتوان به عنوان مدل پارامتریک در نظر گرفت در حالیکه dual problem را میتوان مدل غیر پارامتریک دید که پارامترهای به اندازه دیتاست بستگی دارد.

ب) درست، در عمل، دلیل اینکه SVM ها در برابر برازش بیش از حد مقاوم هستند، حتی در مواردی که تعداد ویژگی ها بیشتر از تعداد مشاهدات است، این است که از منظم سازی استفاده می کنند. آنها برای جلوگیری از برازش بیش از حد، در تنظیم دقیق پارامتر تنظیم،  $C$ ، و در مورد SVM های غیر خطی، انتخاب دقیق هسته و تنظیم پارامترهای هسته کلیدی هستند.

ج) توانایی SVM برای مقابله با نویز به قدرت نویز و هسته استفاده شده بستگی دارد، برای هسته های با بایاس بالا مانند خطی یا چند جمله ای با نویز نباید مشکل ایجاد شود، برای بایاس کم مانند RBF - طبقه بندی را تحت تاثیر قرار می دهد.

د) اگر داده های مجموعه آموزشی را نتوان با خطی جدا کرد، خیر

ترکیبی از نوع خاصی از طبقه بندی کننده های ضعیفی که ما استفاده می کنیم. (برای مثال XOR را به عنوان طبقه کننده ضعیف در نظر بگیرید) مهم نیست که چند بار تکرار انجام شود، خطای آموزشی صفر حاصل نمی شود.

ه) درست، وزن در الگوریتم AdaBoost بر اساس دقت طبقه بندی کننده به هر طبقه بندی کننده اختصاص می یابد و دقت (accuracy) همواره نامنفی است

و) درست، الگوریتم AdaBoost نسبت به داده های نویزی یا پرت بسیار حساس است زیرا تابع جریمه ان نمایی است و هربار دسته بندی وزن مرتبا افزایش می یابد تا زمانی که به کلاس درست تعلق بگیرد.



$(2, 0), (1, 1), (1, 0), (0, 1) \rightarrow$  are support vectors (ب)

منظ برای SV ها مقدار  $0 \neq a_i$  است :

$$L(\alpha) = -\frac{1}{r} [(-\alpha_1 [1] - \alpha_2 [1] + \alpha_3 [2] + \alpha_4 [1]) \cdot (-\alpha_1 [1] - \alpha_2 [1] + \alpha_3 [2] + \alpha_4 [1])] + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4$$

$$\Rightarrow L(\alpha) = -\frac{1}{r} [ \alpha_1^2 + \alpha_2^2 - 2\alpha_1\alpha_2 + \alpha_3^2 + \alpha_4^2 - 2\alpha_2\alpha_4 + 4\alpha_3^2 + 4\alpha_3[\alpha_4 - \alpha_2] ] + 2\alpha_3 + 2\alpha_4$$

$$\sum a_i = 0 \Rightarrow L(\alpha) = -\frac{1}{r} [(\alpha_1 - \alpha_4)^2 + (\alpha_1 + \alpha_3)^2] + 2\alpha_3 + 2\alpha_4$$

در سبب  $a_i$  متغیر می باشد :

$$\frac{dL}{d\alpha_3} = 0 \rightarrow \underline{\alpha_1 + \alpha_3 = r}$$

$$\frac{dL}{d\alpha_4} = 0 \rightarrow \underline{\alpha_1 - \alpha_4 = -r}$$

$$\frac{dL}{d\alpha_1} = 0 \rightarrow r(\alpha_1 - \alpha_4) + r(\alpha_1 + \alpha_3) = 0$$

Aman

$$w = \sum \alpha_i y_i x_i \Rightarrow w = -\alpha_1 [1] - \alpha_2 [1] + \alpha_3 [2] +$$

$$\alpha_4 [1] = \begin{bmatrix} -\alpha_2 + 2\alpha_3 + \alpha_4 \\ -\alpha_1 + \alpha_4 \end{bmatrix} \xrightarrow{\text{مقدارهای صغیر}} \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$y_i (w x_i + b) + 1 = 0 \rightarrow r + b = 1 \rightarrow \underline{b = -r}$$

$$\text{در دلتا: } w n + b = 0 \Rightarrow r n_1 + r n_r - r = 0 \rightarrow n_1 + n_r = \frac{r}{r}$$

ب) اگر نقطه داده یک بردار پشتیبان نباشد، حذف آن تاثیری بر مدل ندارد. از طرف دیگر، حذف بردارهای پشتیبان موقعیت ابرصفحه را تغییر می دهد. بعد ابرصفحه به تعداد ویژگی ها (بردارها) بستگی دارد.

ج) فرض کنید سه بردار پشتیبان داریم که  $(0,0)$  برای کلاس منفی و  $(1,1), (1,-1)$  برای کلاس مثبت و  $\text{margin} = 1$  است با حذف بردار  $(1,-1)$  margin ما به  $\sqrt{2}$  افزایش می یابد در حالت کلی یا تغییر نمی کند یا افزایش می یابد.

-3

الف) خیر

page: ( )

سوال ۳ ب)

Subject:

Year:

Month:

Day:

$$L(w, a) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n a_i (y_i (w x_i + b) - 1)$$

$$a_i \geq 0, \forall i$$

$$\frac{dL}{dw} = 0 \rightarrow w = \sum_{i=1}^n a_i y_i x_i$$

$$\frac{dL}{db} = 0 \rightarrow b = \sum_{i=1}^n a_i y_i = 0$$

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = a_1 y_1 x_1 + a_2 y_2 x_2 + a_3 y_3 x_3$$

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - a_2 \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix} - a_3 \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix} \rightarrow$$

جابجایی در  $\phi(x)$  و ابعاد  $y_i$

$$\Rightarrow \begin{cases} w_1 = a_1 - a_2 - a_3 \\ w_2 = \sqrt{2} a_2 - \sqrt{2} a_3 \\ w_3 = -a_2 - a_3 \end{cases}$$

از معادله نسبت به  $b$  می رانیم :  $a_1 = a_2 + a_3$  در نتیجه  $w_1 = 0$  است

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} (1 \ 0 \ 0) + b = 1 = 0 \rightarrow \underline{b = 1}$$

$$\begin{cases} y_2 (w x_2 + 1) - 1 = 0 \\ y_3 (w x_3 + 1) - 1 = 0 \end{cases} \Rightarrow \begin{cases} -\sqrt{2} w_2 + w_3 = -2 \\ +\sqrt{2} w_2 + w_3 = -2 \end{cases}$$

$$\Rightarrow \underline{w_2 = 0}, \underline{w_3 = -2}$$

$$\text{معادله آخر را حل می کنیم: } -2x_3 = -1 \Rightarrow x_3 = \frac{1}{2}$$

$$w = [0 \ 0 \ -2]$$

Arman

-4

الف) 4

ب) 3

ج) 2

د) 1

ه) 6

5- تنها یک داده misclassify داریم و وزن هر داده  $1/8$  است در نتیجه :

$$\text{Error} = 1/8$$

$$\alpha = \frac{1}{2}(\ln(7/8)/(1/8)) = \frac{1}{2} \ln(7) = 0.9729$$