

هوش مصنوعی پاییز ۱۳۹۹ استاد: محمدحسین رهبان

دانشگاه صنعتی شریف دانشکدهی مهندسی کامپیوتر

مهلت ارسال: \_

## رگرسیون و پرسپترون

پاسخ تمرین ششم، بخش اول

## سوالات نظری (۴۰ نمره)

## ۱. (۲۰ نمره)

اگر داده سوم را به تابع فرضیه اول بدهیم، عبارت w(1)x(1)+w(1)x(1)+w(1) مستقل از برچسب این داده برابر v(1)x(1)+v(1) مستقل از برچسب واقعی اش، به هر شیوهای که وزن های مدل را مقداردهی کنیم برچسب پیشبینی شده برای این داده یکسان خواهد بود.

اما در فرضیه ی دوم مقدار عبارت  $w(\cdot) + w(\cdot) + w(\cdot) + w(\cdot) + w(\cdot) + w(\cdot)$  برابر با  $w(\cdot) + w(\cdot) + w(\cdot) + w(\cdot)$  می شود که مدل می قواند این وزن را برحسب برچسب داده ی سوم یاد بگیرد و برچسب درستی را برای آن پیش بینی کند. پس مدل فرضیه ی دوم قابلیت پیش بینی کردن این داده را برحسب برچسب واقعی اش، برخلاف فرضیه ی اول، دارد.

## ۲. (۲۰ نمره) در پرسپترون می دانیم:

predict according to the current model

$$\hat{y}_i = \begin{cases} +1 & \text{if } \mathbf{w}^T \mathbf{x}_i > 0 \\ -1 & \text{if } \mathbf{w}^T \mathbf{x}_i \le 0 \end{cases}$$

- if  $y_i = \hat{y}_i$ , no change
- else,  $\mathbf{w} \leftarrow \mathbf{w} + y_i \mathbf{x}_i$

به ترتیب برای دادهها مراحل زیر را طی میکنیم: داده اول:

$$W^T X_1 = \cdot \implies \hat{y} = -1 \implies \hat{y_1} \neq y_1$$
 $W_1 = W_1 + (1, 1)^T = (1, 1)^T$ 
 $X_1 + X_2 = \cdot :$ معادله خط جدا کننده:

داده دوم:

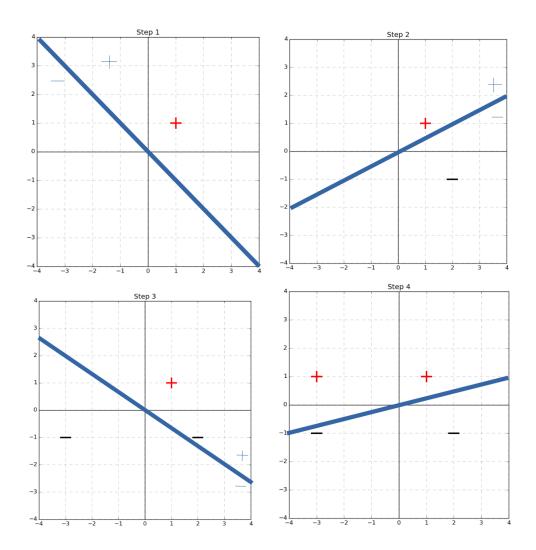
$$W^T X_{\mathsf{Y}} = (\mathsf{1}, \mathsf{1})(\mathsf{Y}, -\mathsf{1})^T = \mathsf{1} \implies \hat{y_{\mathsf{Y}}} = \mathsf{1} \implies \hat{y_{\mathsf{Y}}} \neq y_{\mathsf{Y}}$$
 $W_{\mathsf{Y}} = W_{\mathsf{1}} - (\mathsf{Y}, -\mathsf{1})^T = (-\mathsf{1}, \mathsf{Y})$ 
معادله خط جدا کننده:

داده سوم:

$$W^T X_{\mathbf{r}} = (-\mathbf{1}, \mathbf{1})(-\mathbf{r}, -\mathbf{1})^T = \mathbf{1} \implies \hat{y_{\mathbf{r}}} = \mathbf{1} \implies \hat{y_{\mathbf{r}}} \neq y_{\mathbf{r}}$$
 $W_{\mathbf{r}} = W_{\mathbf{1}} - (-\mathbf{r}, -\mathbf{1})^T = (\mathbf{1}, \mathbf{r})$ 
معادله خط جدا کننده:

داده چهارم:

$$W^T X_{\mathfrak{f}} = (\mathfrak{T}, \mathfrak{T})(-\mathfrak{T}, \mathfrak{I})^T = -\mathfrak{T} \implies \hat{y_{\mathfrak{f}}} = -\mathfrak{I} \implies \hat{y_{\mathfrak{f}}} \neq y_{\mathfrak{f}}$$
 $W_{\mathfrak{f}} = W_{\mathfrak{T}} + (-\mathfrak{T}, \mathfrak{I})^T = (-\mathfrak{I}, \mathfrak{f})$ 
معادله خط جدا کننده:



قسمت ب: مشخصا خطی که حاشیه بین مرز تصمیم و داده آموزش را بیشینه کند، خط y=0 است ولی در پرسپترون ما اصلا به این عامل توجه نمی کنیم و هیچ پنالتی برای کم بودن این فاصله در نظر نمی گیریم که پرسپترون بخواهد این فاصله را بیشینه کند. در واقع تنها کافی است که آخرین داده ای که الگوریتم با آن مواجه می شود، درست دسته بندی شود.