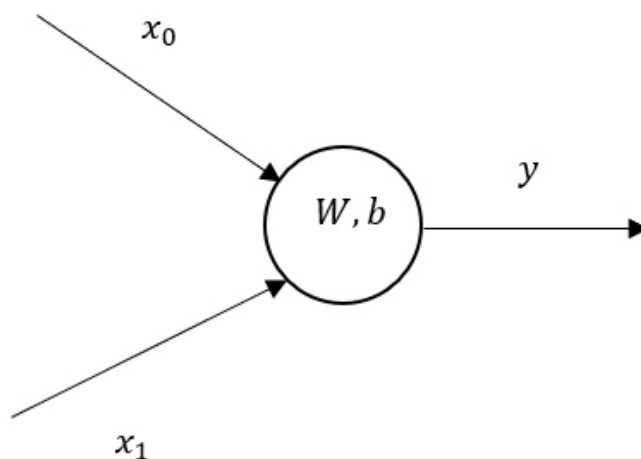


## تمرین سوم درس هوش محاسباتی

۱

فایلی برای داده‌ها همراه این تمرین آپلود شده است. داده‌ها را در یک نمودار scatter نشان دهید و نقطه‌های هر دسته را با رنگ‌های جداگانه نشان دهید. آن را به دو دسته train و test جدا کنید. (ابتدا shuffle کنید) داده‌ها نقاط دو بعدی هستند که برچسب صفر یا یک دارند.

شبکه عصبی زیر را در نظر بگیرید.



$$y = S(WX + b), W = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}, b \in R, X = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

$$cost = \frac{1}{2}(y - y_0)^2$$

$$X, y_0 \in data$$

$$S(x) = \frac{1}{1 + e^{-x}}, (Sigmoid), \frac{dS(x)}{dx} = S(x) * (1 - S(x))$$

در این شبکه مقادیر زیر را بر حسب

$$x_0, x_1, y$$

بدست آورید:

$$\frac{dcost}{dW}, \frac{dcost}{db}$$

راهنمایی: از قاعده زنجیره‌ای مشتق استفاده کنید.

$$\frac{dcost}{dW} = \frac{dcost}{dy} \frac{dy}{dW} = \frac{1}{2} 2(y - y_0) \frac{dy}{dW} = \dots$$

۳

با استفاده از داده‌ها، این شبکه را آموزش دهید. به این صورت که:

```

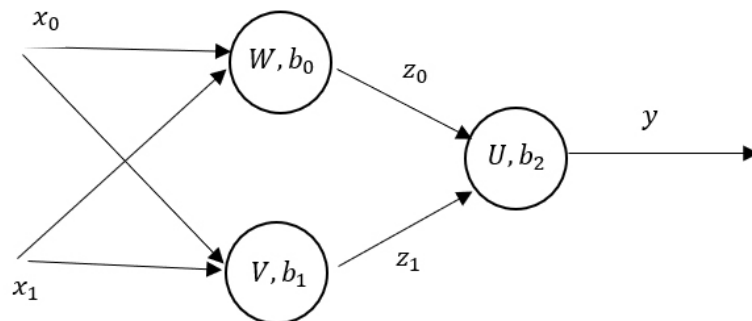
۱ initialize W and B from N(0, 1)
۲ n_epoch = 50
۳ lr = 0.01
۴ for i from 0 to n_epoch:
۵     for each W in variables:
۶         grad[W] = 0
۷         for x0, x1, y0 in train_data:
۸             compute y
۹             compute cost
۱۰            grad[W] += dcost/dW
۱۱     for each W in variables:
۱۲         W = W - lr * grad[w]

```

۴

حالا داده‌های تست را به شبکه بدهید و دقت دسته‌بندی را ارزیابی کنید. نمودار scatter برای داده‌های تست رسم کنید و رنگ هر نقطه بر اساس خروجی شبکه برای آن نقطه باشد.

حالا شبکه زیر را در نظر بگیرید:



$$z_0 = S(XW + b_0)$$

$$z_1 = S(XV + b_1)$$

$$z_2 = S(ZU + b_2)$$

$$Z = \begin{bmatrix} z_0 \\ z_1 \end{bmatrix}, W = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}, V = \begin{bmatrix} v_0 \\ v_1 \end{bmatrix}, U = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix}$$

$$cost = \frac{1}{2}(y - y_0)^2$$

$$X, y_0 \in data$$

$$S(x) = \frac{1}{1 + e^{-x}}, (Sigmoid), \frac{dS(x)}{dx} = S(x) * (1 - S(x))$$

در این شبکه مقادیر زیر را به دست آورید:

$$\frac{dcost}{dW}, \frac{dcost}{dV}, \frac{dcost}{dU}, \frac{dcost}{db_0}, \frac{dcost}{db_1}, \frac{dcost}{db_2}$$

۵

قسمت ۳ و ۴ را برای این شبکه انجام دهید. کارایی این شبکه نسبت به قبلی چه تفاوتی دارد؟ چرا؟

توجه: در این تمرین تحلیل شما نیز حائز اهمیت است.