

## به نام ایزد منان

تمرین سوم درس مبانی هوش محاسباتی، «شبکه‌های عصبی»



استاد درس: دکتر عبادزاده

پاییز ۹۹ – دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر

نکاتی در مورد این تمرین نیاز به توجه و دقت دوستان دارد.

۱- در جدول زیر نحوه اعمال جریمه تاخیر در ارسال تمرین‌ها ذکر شده است.

میزان جریمه	میزان تاخیر (روز)
هر روز ۵٪	۱ الی ۲ روز
هر روز ۱۰٪	۲ الی ۶ روز

در صورتی که بین ۷ تا ۱۴ روز تاخیر داشته باشید، نمره شما از ۵۰٪ محاسبه می‌شود و پس از این بازه با توجه به سایر

تمرین‌ها و زمان تحویل، به تمرین ارسالی نمره‌ای تعلق نمی‌گیرد.

۲- هرگونه کپی کردن باعث عدم تعلق نمره به تمامی افراد مشارکت کننده در آن می‌شود.

۳- آخرین مهلت ارسال تمرین، ساعت ۲۳:۵۵ دقیقه روز **چهارشنبه ۱ بهمن** می‌باشد. این زمان با توجه به

جمع‌بندی‌های صورت گرفته، شرایط و با توجه به سایر تمرین‌ها در نظر گرفته شده است و قابل تمدید نمی‌باشد.

۴- دوستان فایل ارسالی خود را به صورت فشرده و به صورت «شماره دانشجویی\_HW3P\_» مانند HW3P\_9731000

نام گذاری کنید. در این فایل باید مواردی نظیر کدها، فایل پی‌دی‌اف گزارش و سایر موارد مورد نیاز در هنگام بررسی وجود

داشته باشد و صرفاً این فایل در روز ارائه در نظر گرفته می‌شود.

۵- کدهای خود را به صورت مناسب کامنت گذاری کنید. به صورتی که بتوان روال اجرا و موارد مورد نیاز را درک کرد.

۶- سعی کنید ابتدا تمامی سوالات و بخش‌ها را مطالعه کنید.

۷- در صورت هرگونه سوال یا مشکل می‌توانید با تدریس‌یاران درس از طریق ایمیل در ارتباط باشید.

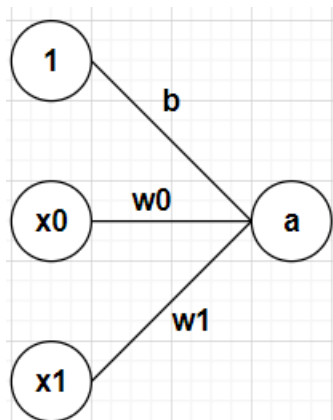
[ci.fall2020.tas@gmail.com](mailto:ci.fall2020.tas@gmail.com)

## بخش دوم – مسائل برنامه‌نویسی و پیاده‌سازی

در این بخش علاوه بر کد و نمودارها، پاسخ تشریحی و تحلیل شما حائز اهمیت است.

- ۱- دیتاست مربوط به این سوال در یک فایل CSV قرار داده شده است. ستون اول و دوم feature ها و ستون سوم مربوط به کلاس هر داده است. این فایل را در برنامه خود وارد کنید. سپس داده‌ها را به دو قسمت آموزش (۷۵ درصد) و تست (۲۵ درصد) تقسیم کنید. داده‌ها را در قالب نمودار scatter رسم کنید. (داده‌ها با کلاس‌های مختلف با رنگ‌های مختلف نشان داده شوند). نمودارها را در گزارش خود قرار دهید.

-۲



a. شبکه عصبی شکل مقابل را با توجه به روابط داده شده پیاده سازی

کنید. منظور از  $y^{[i]}$  برچسب داده  $i$  ام و  $a^{[i]}$  خروجی شبکه برای

داده‌ی  $i$  ام است.

b. مشتق تابع  $J$  را نسبت به  $w_0, w_1, b$  محاسبه کنید. (در گزارش

بیاورید) (راهنمایی: از قانون زنجیره ای استفاده کنید).

c. با استفاده از Gradient Descent شبکه را آموزش دهید. تعداد

گام‌ها و نرخ آموزش مناسب را پیدا و گزارش کنید.

d. شبکه‌ی آموزش داده شده را بر روی داده‌های تست امتحان کنید و دقت را گزارش کنید. همچنین

داده‌های تست را با توجه به برچسب بدست آمده در نمودار scatter رسم کنید. به نظر شما چه ویژگی

از شبکه باعث شده تا مرز تصمیم‌گیری اینگونه باشد؟

$$\sigma(x) = \frac{1}{e^{-x} + 1}$$

$$\frac{d\sigma(x)}{dx} = \sigma(x) * (1 - \sigma(x))$$

$$a = \sigma(x_0 * w_0 + x_1 * w_1 + b_0)$$

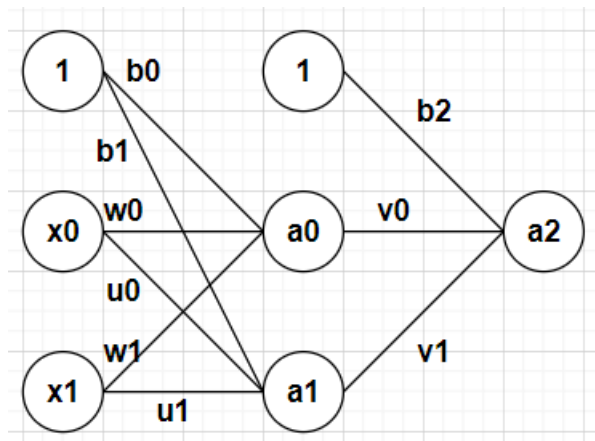
$$J = \frac{1}{m} \sum_{i=1}^m (y^{[i]} - a^{[i]})^2 ; m = \text{number of data samples}$$

a. شبکه عصبی شکل زیر را با توجه به روابط داده شده پیاده سازی کنید. منظور از  $y^{[i]}$  برچسب داده  $i$  ام و  $a_2^{[i]}$  خروجی شبکه برای داده  $i$  ام است.

b. مشتق تابع  $J$  را نسبت به  $w_0, w_1, b_0, u_0, u_1, b_1, v_0, v_1, b_2$  محاسبه کنید. (در گزارش بیاورید) (راهنمایی: از قانون زنجیره ای استفاده کنید).

c. با استفاده از Gradient Descent شبکه را آموزش دهید. تعداد گام ها و نرخ آموزش مناسب را پیدا و گزارش کنید.

d. شبکه‌ی آموزش داده شده را بر روی داده‌های تست امتحان کنید و دقت را گزارش کنید. همچنین داده‌های تست را با توجه به برچسب بدست آمده در نمودار scatter رسم کنید. دقت بدست آمده در این قسمت با قسمت قبل چه تفاوتی دارد؟ به نظر شما دلیل این تفاوت چیست؟



$$\sigma(x) = \frac{1}{e^{-x} + 1}$$

$$\frac{d\sigma(x)}{dx} = \sigma(x) * (1 - \sigma(x))$$

$$a_0 = \sigma(x_0 * w_0 + x_1 * w_1 + b_0)$$

$$a_1 = \sigma(x_0 * u_0 + x_1 * u_1 + b_1)$$

$$a_2 = \sigma(a_0 * v_0 + a_1 * v_1 + b_2)$$

$$J = \frac{1}{m} \sum_{i=1}^m \left( y^{[i]} - a_2^{[i]} \right)^2 ; m = \text{number of data sample}$$

برای آموزش شبکه می‌توانید مطابق شبه کد زیر عمل کنید.

```
Initialize W, U, V from standard normal distribution.
Initialize b = 0
M = number of data_samples
Set learning_rate and number_of_epochs
For i from 0 to number_of_epochs do
    For each param do
        grad[param] = 0
    For each data sample do
        Compute network output
        Compute cost for this data
        For each param do
            grad[param] += dcost/dparam
    For each param do
        param = param - (learning_rate * grad[param] / M)
```

موفق باشید