

به نام ایزد منان

تمرین اول تئوری درس مبانی هوش محاسباتی، «شبکه های عصبی»



استاد درس: دکتر عبادزاده



پاییز ۱۴۰۰ - دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر

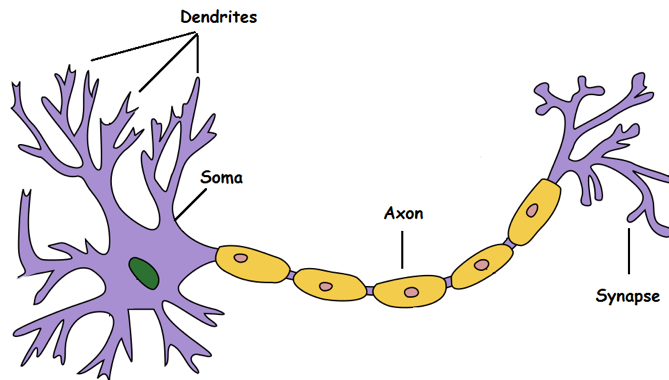
نکاتی در مورد این تمرین نیاز به توجه و دقت دوستان دارد.

- ۱- هرگونه کپی کردن باعث عدم تعلق نمره به تمامی افراد مشارکت کننده در آن می شود.
- ۲- آخرین مهلت ارسال تمرین، ساعت ۲۳:۵۵ دقیقه روز **جمعه ۱۴ آبان** می باشد. این زمان با توجه به جمع بندی های صورت گرفته، شرایط و با توجه به سایر تمرین ها در نظر گرفته شده است و قابل تمدید نمی باشد.
- ۳- دوستان فایل ارسالی خود را به صورت فشرده و به صورت «شماره دانشجویی_HW1» مانند HW1_97310000 نام گذاری کنید.
- ۴- در صورت هرگونه سوال یا مشکل می توانید با تدریس یاران درس از طریق ایمیل در ارتباط باشید.

ci.1400fall@gmail.com

سوال ۱. به سوالات زیر پاسخ دهید.

الف. ساختار نورون (سلول) عصبی انسان مانند شکل ۱ است. نقش هر یک از قسمت‌های مشخص شده را بیان کنید و توضیح دهید هر کدام در یک پرسپترون چگونه معادل شده اند.



شکل ۱- ساختار و بخش‌های مختلف یک نورون عصبی

ب. سه مدل اصلی تقسیم بندی مسائل یادگیری ماشین (یادگیری با نظارت، ...) را شرح دهید و کارایی و نحوه استفاده شبکه‌های عصبی را در این مسائل توضیح دهید.

سوال ۲. به سوالات زیر در رابطه با شبکه پرسپترون چندلایه پاسخ دهید.

الف. چرا در این شبکه از توابع فعالیت (*Activation Function*) استفاده می‌شود و در صورت نبود آنها به چه مشکلی برمی‌خوریم؟

ب. با ذکر دلیل بیان کنید چرا افزودن بایاس به یک نورون عملکرد آن را بهبود می‌بخشد؟

پ. افزایش عمق (تعداد لایه‌ها) و همین‌طور تعداد نورون‌های هر لایه چه تاثیری بر شبکه دارد؟ چه مشکلاتی می‌تواند از آن نتیجه شود؟

ت. یکی از مشکلات روش گرادیان (*Gradient Descent*) نوسان حول نقطه مینیمم است. چگونه با استفاده از ضریب یادگیری این مشکل را برطرف خواهیم کرد؟

سوال ۳. به سوالات زیر پاسخ دهید.

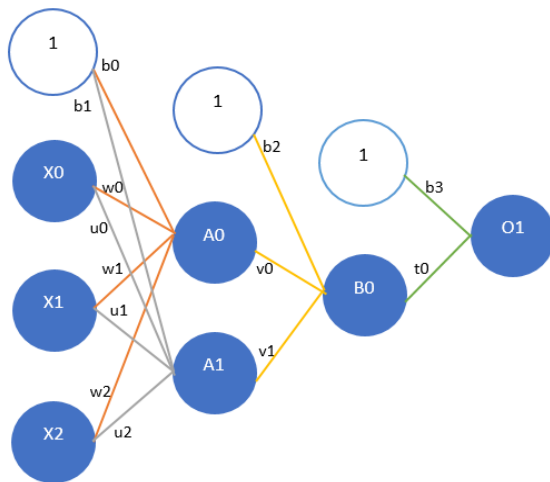
الف. بیش برازش (*Overfitting*) و بیش برازش (*Underfitting*) را در شبکه عصبی توضیح دهید.

ب. چگونه می‌توان با استفاده از نمودار میزان خطا برحسب زمان دو مورد بالا را متوجه شد؟

پ. چگونه می‌توان مشکل بیش برازش را برطرف کرد؟

ت. از تکنیک‌های برطرف کردن بیش برازش، *Dropout* و *Regularization* را تشریح کنید.

سوال ۴. با توجه با شکل زیر و اطلاعات داده شده، به سوالات پاسخ دهید.



$$X = \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix}$$

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

$$B = [B_0]$$

$$W = [W_0 \quad W_1 \quad W_2]$$

$$U = [u_0 \quad u_1 \quad u_2]$$

$$V = [v_1 \quad v_2]$$

$$T = [t_0]$$

$$A_0 = \text{sigmoid}(WX + b_0),$$

$$A_1 = \text{sigmoid}(UX + b_0),$$

$$B_0 = \text{sigmoid}(AV + b_2)$$

$$O_1 = \text{sigmoid}(BT + b_3), \quad \text{cost func} = (O_1 - y_t)^2, \quad y_t \text{ برچسب داده ورودی است}$$

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

الف. مشتق Cost نسبت به u_0 را بدست آورید. (از قاعده زنجیره ای استفاده کنید و برای نوروں‌های میانی نیز تابع *Sigmoid* را به عنوان *Activation Function* در نظر بگیرید).

ب. اگر مقدار اولیه وزن‌ها برابر باشند با:

$w_2 = 0.7$	$w_1 = 0.5$	$w_0 = 0.1$	$b_0 = -0.3$
$u_2 = 0.1$	$u_1 = 0.9$	$u_0 = 0.6$	$b_1 = 0.5$
	$v_1 = 0.6$	$v_0 = 0.2$	$b_2 = -0.4$
		$t_0 = 0.8$	$b_3 = 0.2$

خروجی مدل را در صورتی که مقادیر ورودی برابر باشند با:

$$X_2 = 1 \quad X_1 = 0 \quad X_0 = 1$$

بدست آورید و خطا را محاسبه کنید. (تمام مراحل محاسبه خروجی ذکر شود و y_t را \bullet در نظر بگیرید).

سوال ۵. به سوالات زیر پاسخ دهید در رابطه با شبکه عصبی پیچشی (CNN) پاسخ دهید.

الف. مفهوم اتصال محلی (*Local connectivity*) و به اشتراک گذاری پارامترها در شبکه عصبی پیچشی را توضیح دهید.

ب. با فرض ورودی $3 \times 64 \times 64$ و استفاده از ۶ فیلتر 5×5 و لایه گذاری (*Padding*) ۲ و اندازه گام (*Stride*) ۳، اندازه سائز

خروجی را بدست آوردید.

پ. تعداد پارامترهای قسمت (ب) را بدست آورید.

سوال ۶. در مورد تابع هزینه ماشین بردار پشتیبانی (SVM) به سوالات زیر پاسخ دهید.

الف. مینیمم و ماکسیمم که این تابع می تواند اختیار کند چه قدر است.

ب. در صورتی تمامی پرسپترون های خروجی شبکه عصبی مقادیر نزدیک به صفر اختیار کنند، در صورت وجود N کلاس داده

برای طبقه بندی (*Classification*)، خروجی تابع را بدست آورید.

پ. میدانیم مقدار هزینه SVM برای هر پرسپترون خروجی برابر است با:

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

در صورتی که سیگما روی تمامی کلاس ها (به همراه $j=y_i$) باشد، چه تفاوتی در خروجی L_i حاصل می شود؟ آیا این تغییر

روی یادگیری موثر خواهد بود؟