

به نام ایزد منان

تمرین تئوری اول درس مبانی هوش محاسباتی: «شبکه‌های عصبی»



استاد درس: دکتر عبادزاده



پاییز ۱۴۰۱ - 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 می‌شود.

۲- آخرین مهلت ارسال تمرین، ساعت ۲۳:۵۵ دقیقه روز دوشنبه ۳۰ آبان است.

۳- فایل ارسالی خود را به صورت فشرده و به صورت «شماره دانشجویی_HW1» مانند HW1_97310000 نامگذاری کنید.

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

ci.fall.1401@gmail.com

سوال ۱. به سوال‌های زیر پاسخ دهید.

الف) با ذکر دلیل بیان کنید چرا افزودن بایاس به یک نورون عملکرد آن را بهبود می بخشد؟

ب) با ذکر مثال و انجام محاسبات توضیح دهید که در صورت عدم استفاده از توابع فعالیت و یا استفاده از توابع فعالیت خطی برای

همه‌ی لایه‌ها در یک شبکه‌ی پرسپترون چند لایه، چه اتفاقی می‌افتد.

پ) از توابع فعالیت معروف می‌توان به sigmoid و relu اشاره کرد. این دو تابع را با هم مقایسه کنید و نقاط ضعف هریک را بیان

کنید.

ت) مفاهیم dropout و regularization و کاربرد آن‌ها را در شبکه عصبی را توضیح دهید.

سوال ۲. باتوجه به مفاهیم شبکه عصبی پیچشی و شبکه RBF به سوالات زیر پاسخ دهید.

الف) در یک شبکه عصبی پیچشی، با فرض ورودی $3 \times 128 \times 128$ و استفاده از ۱۰ فیلتر 7×7 و padding برابر ۲ و اندازه Stride

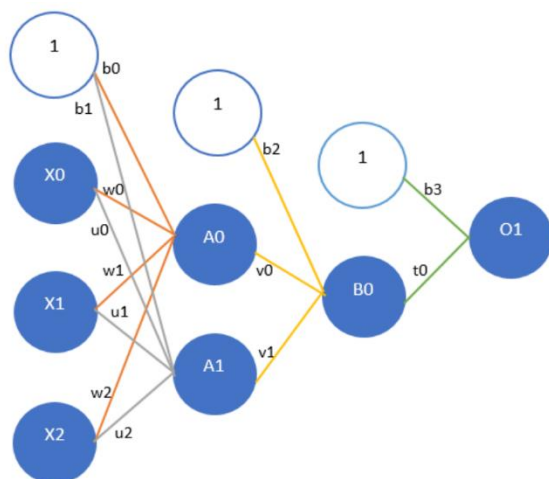
برابر با ۳، اندازه سائز خروجی و تعداد پارامترهای آن را بدست آوردید.

ب) با استفاده از توابع پایه‌ی شعاعی و فرضیات مناسب، تابع xor را مدل کنید و برای هر یک از چهار حالت ورودی‌های x_1 و

x_2 درستی مدل را نشان دهید (توضیحات کافی را برای هر یک از مراحل کار ارائه کنید)

x_1	x_2
0	0
0	1
1	0
1	1

سوال ۳. با توجه با شکل زیر و اطلاعات داده شده، به سوالات پاسخ دهید.



$$X = \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix} \quad A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \quad B = [B_0]$$

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

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

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

$$T = [t_0]$$

$$A_0 = \text{sigmoid}(WX + b_0), \quad A_1 = \text{sigmoid}(UX + b_0), \quad B_0 = \text{sigmoid}(AV + b_2)$$

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

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

الف) مشتق Cost نسبت به u_0 را بدست آورید. (از قاعده زنجیره ای استفاده کنید و برای نورون های میانی نیز

تابع Sigmoid را به عنوان تابع فعالیت در نظر بگیرید.)

ب) اگر مقدار اولیه وزن ها برابر باشند با:

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

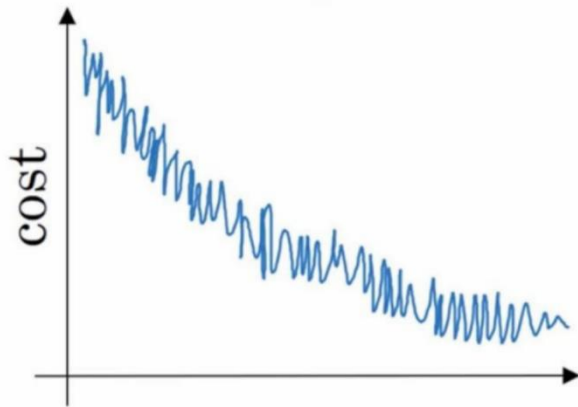
خروجی مدل را در صورتی که مقادیر ورودی برابر باشند با:

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

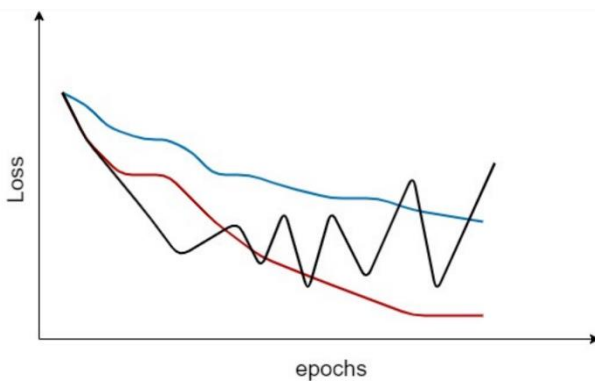
بدست آورید و خطا را محاسبه کنید. (تمام مراحل محاسبه خروجی ذکر شود و y_t را ۰ در نظر بگیرید).

(امتیازی) سوال ۴. به سوالات زیر پاسخ مناسب دهید.

الف) در صورتی که نمودار تابع هزینه یک شبکه عصبی بر حسب تعداد epoch به شکل زیر باشد، با فرض مناسب بودن ضریب یادگیری، علت نوسانات در نمودار را توضیح دهید.



ب) چنانچه در آموزش یک شبکه عصبی، مقدار نرخ یادگیری برابر X باشد نمودار خطای شبکه مشابه خط قرمز میشود. اگر مقدار نرخ یادگیری را برابر Y قرار دهیم به نمودار آبی و اگر به مقدار Z تغییر دهیم به نمودار مشکی تبدیل می شود. اگر قرار باشد مقادیری را برای X و Y و Z متصور شویم، لطفا این مقادیر را با هم مقایسه کنید.



پ) کاهش تعداد لایه ها و نورون های یک شبکه چه تاثیری در بایاس و واریانس و احتمال بیش برازش در آن شبکه دارد؟

ت) برای پیشبینی قیمت یک کالا از یک شبکه عصبی استفاده کرده ایم. توابع فعالیتی که میتوانیم در لایه ی آخر ان استفاده کنیم را از بین توابع فالیت (خطی - سیگموید - رلو - \tanh) انتخاب کنید و دلالتان را هم توضیح دهید.

موفق باشید - تیم تدریسیاری