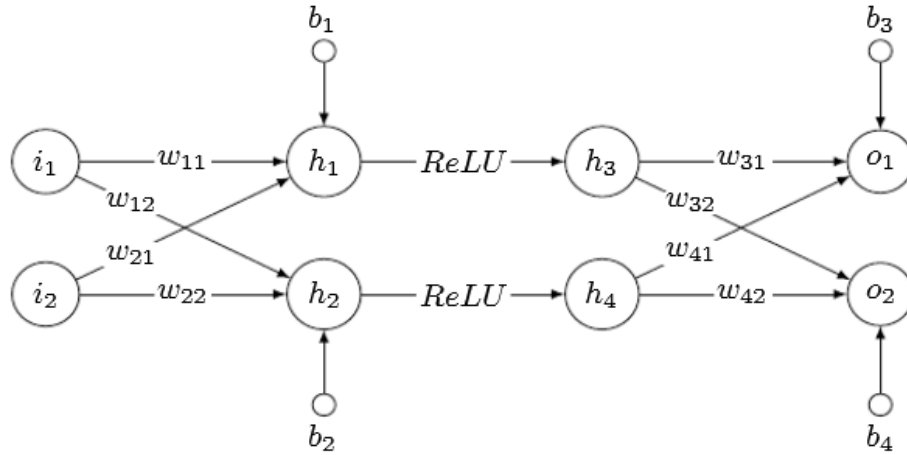


(۱) شبکه زیر را که fully connected است و از تابع فعالسازی Relu استفاده می کند در نظر بگیرید. ورودی ها (i_1, i_2) لایه های مخفی (h_1, h_2) و (h_3, h_4) می باشند. واحدهای خروجی (o_1, o_2) و مقادیر واقعی خروجی (t_1, t_2) می باشند.



مقادیر متغیرها در جدول زیر دیده می شود:

Variable	i_1	i_2	w_{11}	w_{12}	w_{21}	w_{22}	w_{31}	w_{32}	w_{41}	w_{42}	b_1	b_2	b_3	b_4	t_1	t_2
Value	2.0	-1.0	1.0	-0.5	0.5	-1.0	0.5	-1.0	-0.5	1.0	0.5	-0.5	-1.0	0.5	1.0	0.5

(الف) خروجی های شبکه (o_1, o_2) را با در نظر گرفتن ورودی های (i_1, i_2) که در جدول بالا مشخص شده است محاسبه کنید.

(ب) خطای میانگین مربعات (mean squared error) خروجی (o_1, o_2) با در نظر گرفتن داده های واقعی (t_1, t_2) را محاسبه کنید.

(ج) وزن w_{21} را با الگوریتم نزول گرادیانی (gradient descent) و نرخ یادگیری (learning rate) 0.1 به کمک مقدار loss محاسبه شده در (ب) بدست آورید.

(۲) سه روش از روشهای optimization ارائه شده در درس را توضیح داده و با هم مقایسه کنید.

(۳) اگر یک شبکه عمیق را آموزش داده و validation error از training error خیلی کمتر شود چه دلایلی برای آن می تواند وجود داشته باشد (دو دلیل).

(۴) توضیح دهید که چرا به activation function در شبکه های عمیق نیاز وجود دارد.

(۵) وقتی یک شبکه عمیق را پیاده سازی می کنیم معمولا یک تابع 'forward' و یک تابع 'backward' برای هر لایه پیاده سازی می کنیم. توضیح دهید این توابع چه کاری انجام می دهند به چه متغیرهایی نیاز دارند چه آرگومانهایی می گیرند و چه آرگومانهایی برمی گردانند.

سوالات چند گزینه ای:

۶) برای جلوگیری از overfitting می توان:

(الف) اندازه شبکه را افزایش داد

(ب) از data augmentation استفاده کرد

(ج) یادگیری را زودتر متوقف کرد

۷) کدام یک از روشهای optimization ممنتم مرتبه اول (first order momentum) را استفاده می کند

(الف) Stochastic Gradient Descent

(ب) Adam

(ج) RMSProp

(د) Gauss-Newton

۸) عمیق کردن شبکه با اضافه کردن لایه ها منجر به کدام یک از موارد زیر می شود:

(الف) کاهش سرعت یادگیری

(ب) کاهش training loss

(ج) بهبود کارایی بر روی داده های دیده نشده

۹) رنج خروجی تابع فعالسازی تانژانت هایپربولیک چقدر است؟

(الف) ۰ تا ۱

(ب) ۱- تا ۱

(ج) ۰ تا بی نهایت

(د) ۱- تا بی نهایت

۱۰) لطفا کد نوشته شده در لینک زیر را با کمک ابزار keras در tensorflow بازنویسی کنید و در حد حداکثر یک صفحه

مدلهای مختلف استفاده شده و نتایج آنها را توضیح دهید (تا ۱۲ شب)

https://colab.research.google.com/drive/15UbFTQnh_wEcnqUPZ8k04u4ZW59c36_u?usp=sharing#scrollTo=109NtaGNhjA3