

به نام ایزد منان

تمرین سوم درس مبانی هوش محاسباتی، «شبکه های 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 تمرین ها ذکر شده است.

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

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

تمرین ها و زمان تحویل، به تمرین ارسالی نمره ای تعلق نمی گیرد.

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

۳- آخرین مهلت ارسال تمرین، ساعت ۲۳:۵۵ دقیقه روز **دوشنبه ۲۲ دی** 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.

۴- دوستان فایل ارسالی خود را به صورت فشرده و به صورت «شماره دانشجویی_HW3» مانند HW3_97310000

نام گذاری کنید. در این فایل باید مواردی نظیر کدها، فایل پی دی اف گزارش و سایر موارد مورد نیاز در هنگام بررسی وجود داشته باشد و صرفاً این فایل در روز ارائه در نظر گرفته می شود.

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 بخش ها را مطالعه کنید.

۷- در صورت هرگونه سوال یا مشکل می توانید با تدریس یاران درس از طریق ایمیل در ارتباط باشید.

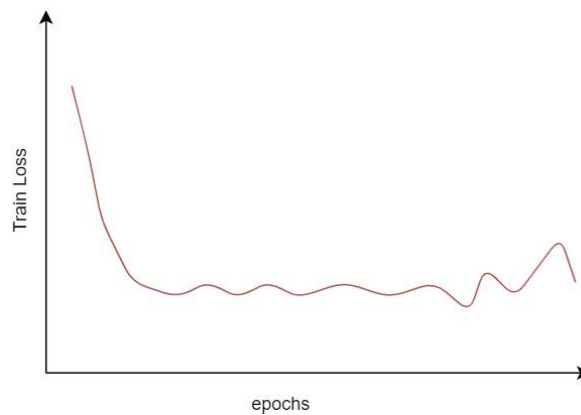
ci.fall2020.tas@gmail.com

بخش اول – مباحث تئوری و مسائل تشریحی

۱- گفته می‌شود که ایده شبکه های مصنوعی عصبی از نحوه عملکرد سلول های عصبی مغز الهام گرفته شده است. اما نورون های مدل سازی شده با نورون های مغز تفاوت های اساسی دارند.

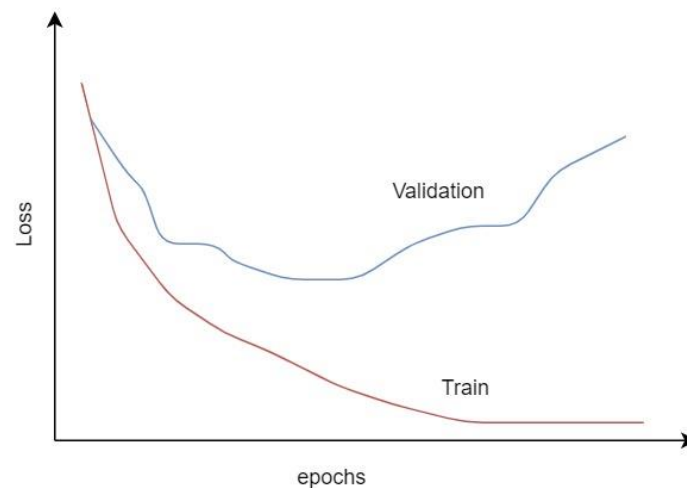
- چه چیزی نورون های بیولوژیکی مغز را متفاوت می‌کند؟ مهم ترین تفاوت ها را نام ببرید.
- در دهه های اخیر، با پیشرفت توان پردازشی کامپیوترها، ساخت و آموزش شبکه های عصبی مصنوعی در ابعاد هرچه بزرگتر میسر شده است. (برای مثال مدل GPT-3 دارای ۱۷۵ میلیارد پارامتر آموزش داده شده است) با کمی تحقیق این عدد را با تعداد سیناپس های مغز انسان مقایسه کنید. به نظر شما رسیدن به تعداد نورون های یک مغز واقعی ، عملکرد در سطح آن را هم تضمین می‌کند؟
- در بین انواع روش های نوین یادگیری ماشین، **یادگیری تقویتی** شباهت زیادی با نحوه تعامل موجودات زنده با محیط را دارد. در خصوص نحوه کار این روش توضیح دهید.
- الگوریتم **backpropagation** عامل اصلی یادگیری در شبکه های عصبی محسوب می‌شود. اما شواهدی دال بر وجود چنین الگوریتمی در مغز وجود ندارد. به نظر شما دلیل آن چیست؟

۲- علی می‌خواهد با داده‌های مربوط به ساعات عبور و مرور افراد در آزمایشگاه، یک مدل شبکه عصبی پرسپترون ۲ لایه‌ای را آموزش دهد تا بتواند در هر ساعت تعداد افراد حاضر در آزمایشگاه را پیش‌بینی کند. او پس از آموزش روی داده‌های *train* نمودار خطا را بعد از هر *epoch* رسم کرد و شکل زیر حاصل شد:



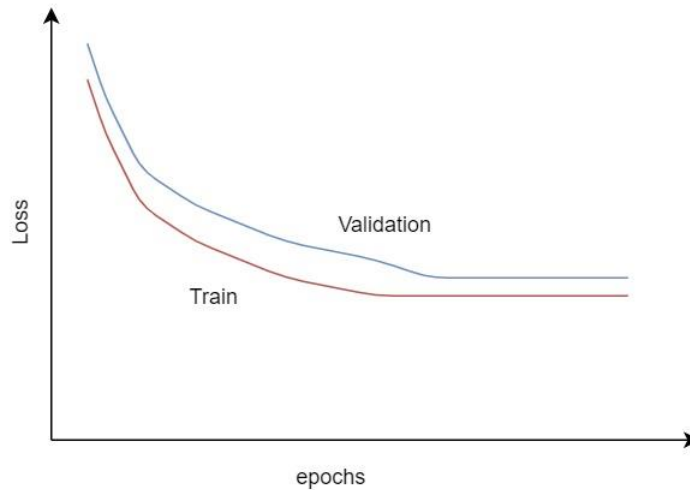
- چه مشکلی در هایپرپارامترهای مدل وجود دارد؟ لاقط یک پیشنهاد برای رفع مشکل بدهید.

علی پس از تلاش زیاد توانست نمودار خطا روی داده‌های آموزش را نزدیک به صفر برساند. اما خوشحالی‌اش با رسم نمودار خطای داده‌های *validation* فروکش کرد:



- مشکل مدل علی چه چیزی است؟ برای رفع این مشکل چه پیشنهادی دارید؟ (لاقط ۳ مورد)

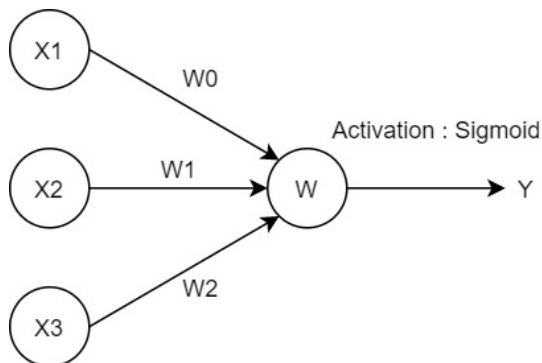
علی پس از مشورت با استاد درس هوش محاسباتی خود، تصمیم گرفت تغییراتی در مدل خود انجام دهد. به طور دقیق‌تر او تعداد نورون‌ها و لایه‌های مدل را تغییر داد. نمودار زیر خطای *Train* و *validation* را بعد از تغییرات علی را نشان می‌دهد.



- به نظر شما علی چه تغییری در مدل اعمال کرده بود؟ دلیل خود را بیان کنید.

۳- تا چند سال پیش تابع *sigmoid* معمول ترین گزینه برای اعمال رفتار غیرخطی در خروجی نورون ها بود، اما در سال های اخیر و با عمیق تر شدن مدل ها، جایش را به توابعی مثل *relu* و *gelu* داده است. با رسم شکل هر کدام از این توابع، دلیل این جایگزینی را توضیح دهید.

۴- می خواهیم یک شبکه پرسپترون تک لایه با یک نورون را آموزش دهیم.



اگر تابع هزینه ما

$$cost(x) = (sigmoid(Wx + b) - y_t)^2$$

(W بردار وزن ها ، b عدد بایاس و y_t برچسب داده ورودی است)

موارد زیر را به دست آورید

- ابتدا مشتق تابع $cost$ نسبت به w_0 را به دست آورید. (از قاعده زنجیره ای کمک بگیرید)
 - فرض کنید مقدار اولیه وزن ها به صورت $w_0 = 0.5, w_1 = 0.2, w_2 = 0.3, b = 2$ باشد. درین صورت مقدار خروجی مدل ($feed\ forward$) به ازای ورودی $\{0, 1, 0\}$ و برجسب : ۲، و خطای به دست آمده را حساب کنید.
 - با استفاده از مشتقات به دست آمده در قسمت اول و داده های قسمت دوم، یک بار وزن w_0 را به روزرسانی کنید.
- (طبق روش $gradient\ decent$ و نرخ یادگیری $= 0.1$)

موفق باشید

تیم تدریسیاری