سوالات تشریحی

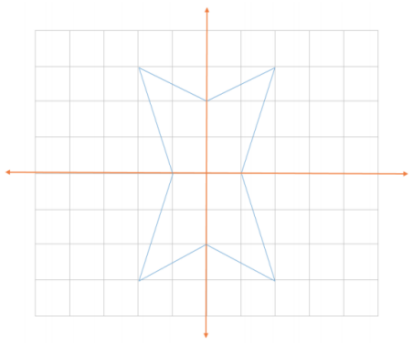
۱. در این بخش می خواهیم با استفاده از نورون هایی که تابع activation آن ها تابع sigmoid است تعدادی از توابع منطقی را ایجاد کنیم. در واقع اگر نورونی n ورودی با مقادیر xn , ... , x1 داشته باشد که با وزن های wn ,..., w1 متصل باشند، ، خروجی نورون به شکل زیر خواهد بود:

$$y = \frac{1}{1 + \exp\left(-w_0 - \sum_{i=1}^{n} x_i w_i\right)}$$

ورودی های شبکه به شکل ۰ و ۱ هستند و در صورتی که خروجی تابع بزرگتر از ۰٫۵ باشد به عنوان ۱ منطقی و در غیر این صورت به عنوان ۰ منطقی برداشت می شود. در ادامه منظور از تعداد لایه ها، تعداد تمام لایه ها به غیر از لایه ی ورودی است. بنابراین یک شبکه ی دو لایه در واقع یک لایه ی نهان داشته و یک شبکه ی یک لایه، لایه ی نهانی ندارد.

الف. با استفاده از یک شبکه ی دو لایه تابع $y_4 = (x_1 \vee \neg x_2) \oplus (\neg x_3 \vee \neg x_4)$ را مدلسازی کنید. مقادیر وزن ها و بایاس در این شبکه را مشخص کنید.

ب. آیا می توان شبکه ای دو لایه با دو ورودی x و y یعنی تنها با یک لایه ی نهان طراحی کرد طوری که برای ورودی ای که داخل ناحیه شکل ۱ قرار بگیرد خروجی ۱ و درغیراینصورت ۰ بدهد . در صورت امکان پذیر بودن ساختار شبکه به همراه وزن ها و بایاس های آن را مشخص کنید. در غیر اینصورت عدم امکان پذیری را ثابت کنید. (توجه: خطوط قرمز معادل واحدهای مختصات است)



۲. به سوالات زیر به صورت تحلیلی پاسخ دهید.

الف. از توابع فعال ساز Sigmoid , tanh در شبکه های عصبی مخصوصا شبکه های عمیق به ندرت استفاده می شود و از تابع Relu استفاده بیشتری می شود. به نظر شما علت این امر چیست؟ (راهنمایی: از طراحی شبکه های LSTM کمک بگیرید.)

ب. مزایا و معایب روش های L1 و L2 برای regularization را بیان کرده و به طور کلی این دو روش را با یکدیگر مقایسه کنید.

ج. تاثیر momentum بر آموزش شبکه عصبی را توضیح دهید.

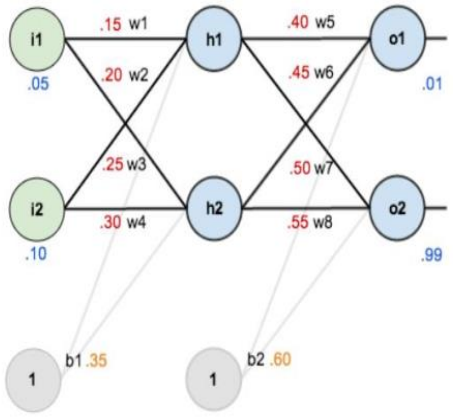
۳. در آموزش شبکه های عصبی سعی در بهینه کردن تابع هدفی داریم که بهینه سازی وزن های آن می تواند از طریق الگوریتم back propagation انجام شود. شبکه زیر را درنظر بگیرید. فرض کنید تابع فعال ساز نورون ها تابع sigmoid می باشد. تابع خطا نیز mean squared error می باشد و به صورت زیر تعریف می شود:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$$

با توجه به توضیحات بالا و مقدارهای ورودی، خروجی، وزن های اولیه و بایاس ها محاسبات زیر را انجام دهید:

الف. خطای شبکه را پس از یک مرحله feed forward محاسبه کنید.

ب. با در نظر گرفتن learning rate برابر با ۰,۳ وزن های آپدیت شده لایه دوم (w5,...,w8) را پس از یک مرحله back propagation محاسبه نمایید. **(امتیازی: محاسبه وزن های لایه اول (w1,...,w4)**



پیاده سازی

۱. در این سوال هدف استفاده از شبکه fully connected برای پیش بینی باران در روز آتی به صورت دو کلاسه به معنی بارانی بودن یا نبودن هوا طبق تعدادی ویژگی می باشد. مجموعه داده مورد استفاده در این بخش دارای اطلاعات مشاهدات آب و هوا در تعداد ایستگاه هواشناسی می باشد. مجموعه داده را می توانید از این لینک دریافت کنید.

گام های حل مسئله:

https://drive.google.com/file/d/167biTSSVRfDq0K8QWw0T3aKnTGAbMROs/view

الف. پیش پردازش ها و نرمال سازی هایی که از نظر شما برای ورودی ها مورد نیاز است انجام دهید. همچنین داده ها را به بخش آموزش و اعتبارسنجی و آزمایش تقسیم کنید. (راهنمایی: ویژگی های categorical باید encode شوند.)

ب. شبکه multilayer perceptron با تعداد لایه های نهان و دیگر هایپر پارامترهای دلخواه تعریف کنید. مقدار هایپر پارامتر ها را منطقی انتخاب کنید و مقادیر را حتما در گزارش خود ذکر کنید. به عنوان راهنمایی برای loss استفاده از binary-cross-entropy یکی از گزینه هاست. در این بخش برای راحتی کار می توانید از sklearn استفاده کنید.

ج. پس از compile مدل تعریف شده در قسمت قبل، در این قسمت به آموزش مدل بپردازید.

د. جهت ارزیابی accuracy روی داده های آموزش و آزمایش را گزارش کنید و نتایج را طبق گام های انجام شده تحلیل نمایید.

۲. هر مجموعه داده ای که به صورت متوالی و پی در پی اطلاعاتی را نشان دهد می تواند به صورت توالی مورد پردازش قرار گیرد. پیش بینی قیمت سهام های بورس طبق پیشینه قیمت آن ها، پیش بینی آلودگی هوا توسط معیارهای آلودگی در سیکل های زمانی گذشته و روند تغییرات آن ها، از این دست مسائل هستند. برای حل این مسائل شبکه هایی مانند LSTM که feedback connection دارند کاربرد فراوانی دارند که در این بخش به بررسی یک نمونه ساده از شبکه LSTM می پردازیم. شبکه های بازگشتی در هر گام، به تحلیل اطلاعات پیشین و انتقال آن ها به قسمت بعد جهت پردازش می پردازند.

مجموعه داده مورد نظر شامل دمای میانگین روزانه در چند سال متوالی در یک منطقه می باشد که می توانید آن را از این لینک دریافت کنید. این داده ها از نوع univariate time series هستند به این معنی که فقط یک فاکتور در نظر گرفته می شود و برای پیش بینی مقدار آن در آینده ، فارغ از عوامل تاثیرگذار احتمالی دیگر در نظر گرفته می شوند. در مجموعه داده ستون date نشان دهنده تاریخ و ستون temp نشان دهنده دمای میانگین متناظر آن تاریخ می باشد. هدف پیش بینی دمای یک روز طبق ۵ روز گذشته آن است. استفاده از keras،tensorflow یا pytorch در پیاده سازی این بخش به انتخاب شماست.

گام های حل مسئله:

الف. داده های ورودی و خروجی مدل را بسازید و هم چنین آن ها را به بخش های آموزش، اعتبار سنجی و آزمایش تقسیم کنید.

(راهنمایی: جهت ساخت ورودی مدل ابتدا همه ستون temp را به صورت یک لیست در نظر بگیرید و با ماسکی به طول ۵، یکی یکی در طول آن حرکت کنید. در هر شیفت یک اندیسی، ۵ عدد وجود خواهد داشت که این لیست نمونه ای از ورودی می باشد. اندیس بعدی یعنی عدد ششم به عنوان خروجی یا target در نظر گرفته می شود.)

ب. در گام بعدی به تعریف مدل LSTM با تعداد دلخواه لایه LSTM، سپس لایه Dense در ادامه بپردازید. loss مدل rmse باشد و هایپر پارامتر های دیگر به دلخواه انتخاب شوند.

ج. مدل را آموزش دهید. (تابع fit در اکثر framework های مطرح این حوزه برای آموزش مدل تعریف شده است.)

https://docs.google.com/spreadsheets/d/1IVhGpP3iXgWpoVpmir0b5fdcWo7MagTx/edit#gid=983681963

د. مقدار loss را برای داده های آموزش و آزمایش گزارش کرده و نتایج را تحلیل کنید.