



Assignment NO.2 Solutions

Neural Networks | Fall 1400 | Dr.Mozayani

Teacher Assistants:

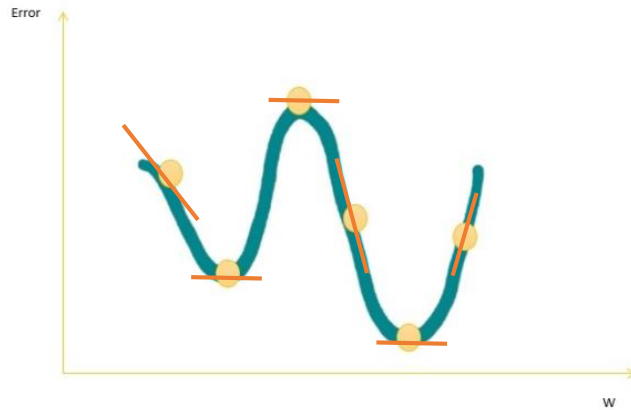
Fatemeh Hajizadeh

Fatemeh Anvari

Student name : **Amin Fathi**

Student id : **400722102**

Problem 1.a



از آنجا که در نقاط ۱ و ۴ مقدار گرادیانت (شیب) مقداری منفی است و لازم است تا به سمت راست نمودار حرکت کنیم ؛ فلذا مقدار وزن را افزایش می دهیم .

همچنین در نقطه ۶ ، از آنجا که مقدار شیب مثبت ایست و لازم است تا به سمت چپ حرکت کنیم ، فلذا مقدار وزن را کاهش می دهیم .

در نقاط ۲ و ۳ و ۵ مقدار شیب صفر می باشد و بنابراین وزن تغییری نمیکند.

Problem 1.b

نقاط ۲ و ۳ به ترتیب مینیمم محلی ، ماکسیمم سراسری (یا محلی) گویند . که روش مناسب برای خروج از گیر کردن الگوریتم در این نقاط استفاده از مامنتوم Momentum می باشد ، در این روش ضریبی از وزن ها قبلی را به وزن های جدید اضافه می کنیم که از گیر کردن الگوریتم پیشگیری میکند ، از دیگر مزایای این روش سرعت بالا و حرکت منسجم تر الگوریتم می باشد .

روابط مامنتم :

$$W_{new} = W_{old} - lr * (\nabla_W L)_{W_{old}}$$

$$V_{new} = \alpha * V_{old} + lr * (\nabla_W L)_{W_{old}}$$

$$W_{new} = W_{old} - (\alpha * V_{old} + lr * (\nabla_W L)_{W_{old}})$$

$$0 < \alpha < 1$$

مقدار مناسب آلفا هم طبق تجربه مقداری نظیر ۰,۹ یا ۰,۹۵ است.

ممانتوم به کمک وزن دهی به تمامی آپدیت های اخیر در مقایسه با آپدیت قبلی محاسبه میشود که به سریع تر شدن همگرایی کمک میکند. Vold همان بخش موثر از آپدیت های اخیر است که به کمک ما می آید.

منابع :

[Improving Vanilla Gradient Descent | by Devin Soni 🏠 | Towards Data Science](#)

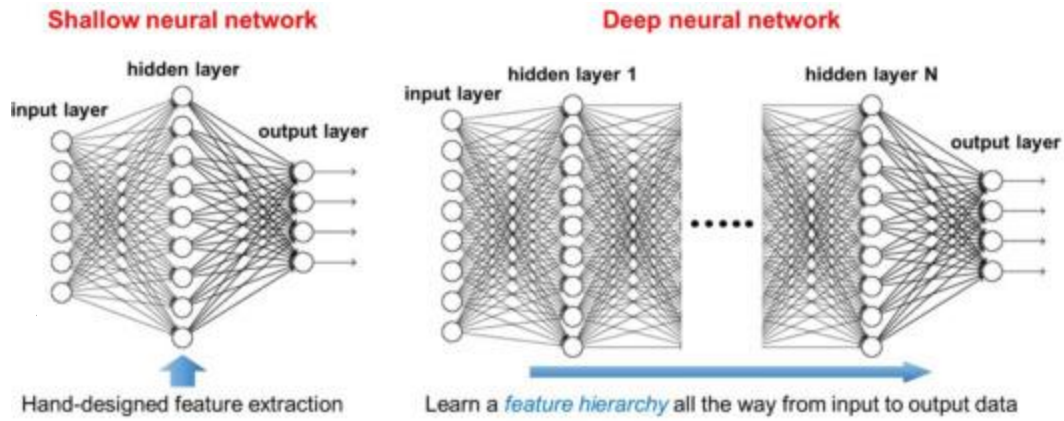
[Overview of various Optimizers in Neural Networks | by Satyam Kumar | Towards Data Science](#)

Problem 2

تفاوت اصلی دو نوع شبکه عمیق و کم عمق در تعداد لایه ها hidden و همچنین تعداد نورون های هر لایه است. با تنظیم این دو پارامتر میتوان عملکرد شبکه را بهبود بخشید. برخی مقادیر تجربی که کارآیی خوبی دارند در هر نوع مختلف از شبکه ها وجود دارد که به عنوان مثال به این که تعداد نورون های هر لایه در شبکه های feed-forward بهتر است توانی از دو باشد، اشاره کرد. یا اینکه تعداد نورون ها در لایه های میانی بهتر است در ابتدا زیاد تر شود و سپس کمتر شود (یا حداقل برای مدتی ثابت بماند). داشتن لایه های میانی متعدد باعث استخراج ویژگی ها به صورت سلسله مراتبی (از ساده به پیچیده) میشود که در حل مساله برای ما کمک کننده خواهد بود. البته تا حدی که بیش از حد پیچیده نشود و دچار بیش برازش نشویم.

تعریف شبکه عصبی عمیق و کم عمق:

شبکه های عصبی به صورت کلی میتوانند به دو دسته کم عمق و عمیق تقسیم شوند. به شبکه ای که ۱ یا ۲ لایه مخفی داشته باشد شبکه عصبی کم عمق گفته میشود که با درک درست ساختار شبکه میتوان از شبکه های کم عمق به بهترین شکل استفاده کرد. شبکه ی عصبی با تعداد لایه های مخفی بیشتر (۳ و بالا تر) با نام شبکه های عصبی عمیق شناخته میشوند. در زیر یک تصویر برای درک بهتر از معماری هر دو نوع شبکه ضمیمه شده است. قابل ذکر است که امروزه به دلیل وجود سخت افزار مناسب و الگوریتم های متعدد و کارا، شبکه های عمیق کاربرد بسیار بیشتر و قدرتمند تری در دنیای واقعی دارند. البته حتما باید مراحل انتخاب مدل مناسب از مدل ساده تر به مدل پیچیده تر انجام شود تا پیچیدگی اضافی نداشته باشیم. و صد البته نیاز به داده یادگیری مناسب هم یکی از مسائل مهم در یادگیری شبکه های عمیق است.



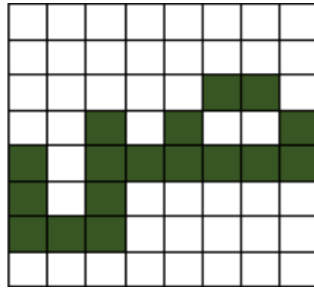
منابع :

<https://www.tutorials.com/explain-deep-neural-network-and-shallow-neural-networks/>

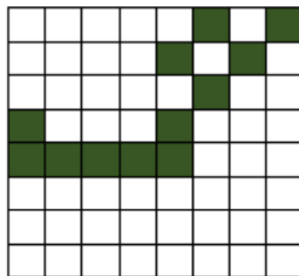
[What is the difference between deep and shallow neural networks? - Quora](#)

[Shallow Neural Networks. In this post, I have explained what... | by Rochak Agrawal | Towards Data Science](#)

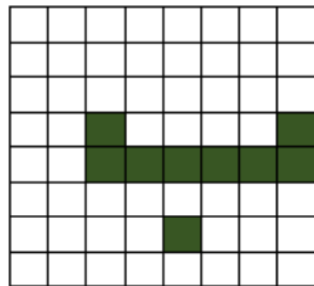
Problem 3



[0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,1,0,1,0,0,1,1,0,1,1,1,1,1,1,0,1,0,0,0,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0]



[0,0,0,0,0,1,0,1,0,0,0,0,1,0,1,0,0,0,0,0,0,1,0,0,1,0,0,0,1,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]



[0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,1,1,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0]

چنانچه نویز ورودی نداشته باشیم ، می توان پیکسل های با وضعیت مشابه را در سه عکس حذف کرد که در این صورت تنها ۲۴ فیچر برای مقایسه باقی خواهد ماند (لازم به ذکر است چنانچه نویز ورودی داشته باشیم ، نمیتوانیم فیچر های مشترک را حذف کنیم) .

در مرحله بعد هم پرسپترون بدون لایه میانی و با ۲۴ نورون وردی (در صورت وجود نويز ورودی با ۶۴ نورون ورودی) و ۳ نورون خروجی میتوان طراحی کرد که از عهده پیچیدگی مساله بر بیاید .