

به نام خدا

تمرین سری ششم - درس مبانی بینایی کامپیوتر

سید محمد علی نخاری - شماره دانشجویی : 99521496

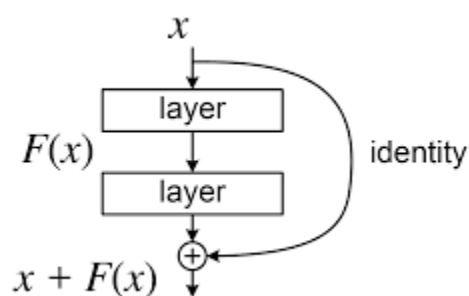
سوال اول (الف) ناپدید شدن گرادیان و انفجار گرادیان دو مشکل بوجود آمده در شبکه های عصبی عمیق هستند.

ناپدید شدن گرادیان مشکلی است که بیشتر در شبکه های عصبی عمیق که از روش های مبتنی بر گرادیان کاهشی و backpropagation برای آموزش داده ها و بروز رسانی وزن های استفاده شده در نورو ن های موجود در هر لایه استفاده می کنند، رخ میدهد. یعنی زمانی که مقدار خطا به ازای یک ورودی مشخص به مدل را محاسبه کردیم باید به صورت عقب گرد از لایه خروجی به سمت لایه ورودی حرکت کرده و با استفاده از روش گرادیان کاهشی و مینیمم کردن مقدار خطا، وزن های گره های هر لایه را بروز رسانی کنیم در واقع مقدار مشتق برای هر لایه باید به صورت زنجیره ای در هم ضرب شوند. در شبکه های عمیق با تعداد لایه های میانی زیاد، تعداد مشتق بیشتری در هم ضرب میشوند. حال اگر مقادیر این مشتق ها به قدری کم باشد و به صفر نزدیک باشد حاصل ضرب این مقادیر در هم عدد خیلی کوچکی شده به این ترتیب وزن گره های لایه های ابتدایی به خوبی بروز رسانی نمی شوند و در نتیجه ناپدید شدن گرادیان رخ میدهد. به عبارت دیگر برای بروز رسانی وزن گره ها از وزن قبلی و مقدار گرادیان به دست آمده استفاده کرده و وزن جدید را به دست می آوریم زمانی که مقدار گرادیان خیلی کم باشد این وزن بروز رسانی نمیشود. پس زمانی که تعداد لایه های شبکه زیاد باشد، لایه های اول نمیتوانند به خوبی آپدیت شوند و محو شدگی گرادیان رخ میدهد.

اگر بخواهیم در مورد نحوه رخداد آن توضیح دهیم میتوان گفت زمانی که در شبکه عصبی مورد استفاده از برخی از توابع فعال سازی مثل sigmoid استفاده میکنیم به ازای تغییرات بسیار زیاد یا کم در ورودی، مقدار آن در خروجی تفاوت چندانی نمی کند و مشتق به صفر میل میکند و در نتیجه مقدار گرادیان تابع ضرر به سمت صفر میل کرده و سبب ایجاد مشکل گفته شده میشود.

در مقابل، انفجار گرادیان در شبکه های عمیق به موقعیتی اشاره میکند که گرادیان های محاسبه شده در طول فرآیند به روزرسانی وزن ها به اندازه ای بزرگ میشوند که باعث میشود شبکه عصبی به سرعت ناپایدار و بهینه سازی به درستی انجام نشود. همانطور که میدانیم در یک شبکه عصبی با n لایه، به تعداد لایه ها مشتق به صورت زنجیره ای در هم ضرب میشوند. در صورتی که این مشتق ها مقادیر بزرگی داشته باشند گرادیان به صورت نمایی افزایش یافته و با حرکت به سمت لایه های ابتدایی این مقدار بسیار بزرگ شده و انفجار گرادیان رخ میدهد.

ب) در شبکه های معرفی شده قبل از ResNet با عمیق تر شدن لایه ها با مشکل ناپدید شدن گرادیان مواجه می شدیم. مشکل ناپدید شدن گرادیان از آنجا نشأت میگیرد که در فرآیند backpropagation گرادیان ها به لایه های ابتدایی نمیرسند و مدل نمیتواند به خوبی آموزش بیابد. شبکه ResNet برای حل این مشکل از Residual Block ها استفاده کرد. یعنی اگر ما لایه ای داشته باشیم که با گرفتن ورودی x قرار است خروجی Hx را یاد بگیرد در این معماری با گرفتن ورودی x خروجی $x + Fx$ را یاد میگیریم که مقدار x همان مقدار ورودی و مقدار Fx مقدار خروجی لایه های شبکه است. با این کار اگر شبکه به هر دلیلی نتوانست مقدار Fx را یاد بگیرد حداقل همان مقدار ورودی را به لایه های بعدی انتقال دهد.



به طور کلی در شبکه ResNet تعداد زیادی بلوک باقیمانده (Residual Block) وجود دارد که هرکدام دارای دو لایه کانولوشنی 3×3 هستند. در ابتدای شبکه هم یک لایه کانولوشنی برای استخراج ویژگی از تصویر قرار دارد. در نهایت در لایه آخر ابعاد داده ها با استفاده از average pooling کاهش یافته و از یک لایه کاملاً متصل برای دسته بندی استفاده می شود.