

به نام خدا



دانشکده مهندسی کامپیوتر

درس یادگیری عمیق

دکتر مرضیه داودآبادی

تمرین سری سوم

دستیاران آموزشی:

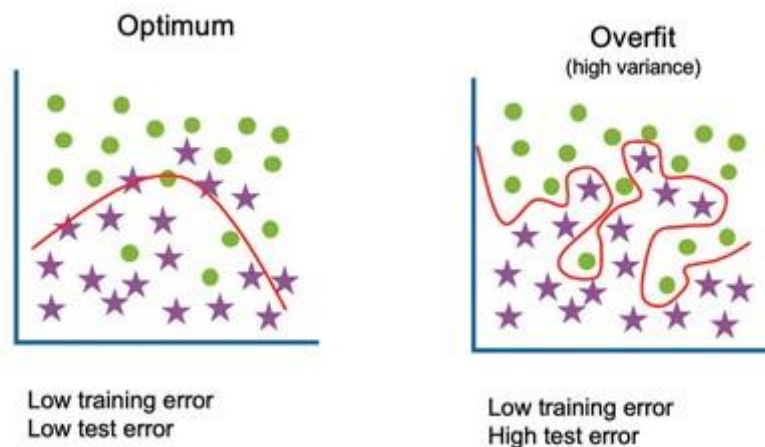
ارشیا اموری سرابی

مهدی خورشیا

نکات تکمیلی

۱- الف) مفهوم overfit در شبکه‌های عصبی را توضیح دهید؟ چه موقع گوییم شبکه‌ی عصبی overfit شده است؟ راه‌کارهای خود برای جلوگیری و حل overfit شدن شبکه‌های عصبی را نام ببرید. (۱۵ نمره)

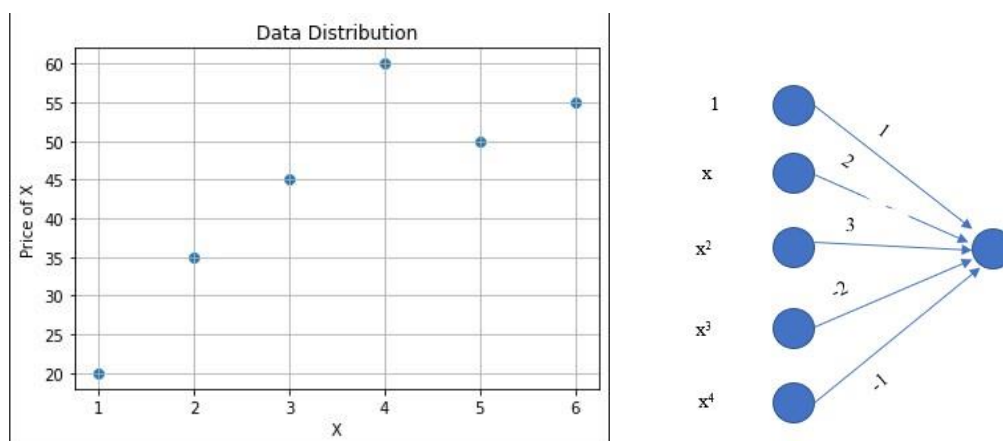
ب) شکل شماره‌ی ۱ نمایان‌گر مدل overfit شده بر روی مجموعه‌ی داده است و شکل شماره‌ی ۲ مدل normal را نشان می‌دهد. چرا در شکل شماره‌ی ۱ مدل برای داده‌هایی که تاکنون ندیده است، خطای بیشتری نسبت به مدل شکل شماره‌ ۲ تولید خواهد کرد؟ (۱۰ نمره)



شکل شماره‌ی ۲

شکل شماره‌ی ۱

۲- الف) شبکه‌ی عصبی و مجموعه‌داده‌های زیر را در نظر بگیرید که قرار است در آن قیمت x ها را پیش‌بینی کنیم. نمی‌خواهیم این شبکه بر روی مجموعه‌داده‌ی داده شده overfit شود. می‌خواهیم با اعمال Regularization term، یادگیری داده‌ها را برای آن سخت کنیم. با انتخاب یک نمونه داده به‌دلخواه، ابتدا آن را در شبکه‌ی عصبی منتشر کنید (feed forward) و سپس هنگام بروز رسانی پارامترهای شبکه (back propagation)، به‌روزرسانی را با L2 regularization و $\lambda = 0.9$ انجام دهید. به خاطر داشته باشید که این کار باید بصورت دستی انجام گیرد و سعی کنید مراحل را کامل ذکر کنید. هم چنین تنها یک دور کافی است. (۱۵ نمره)



در این شبکه تابع فعال‌ساز خطی (linear) و نرخ یادگیری ۰.۱ است. تابع خطا را از نوع M.S.E. به فرمول زیر در نظر بگیرید:

$$loss = \frac{1}{2m} \sum_{i=1}^m (y - \hat{y})^2$$

ب) با تغییر مقدار λ چه تغییراتی در میزان خطای تولید شده و نحوه ی آموزش پارامترهای شبکه های عصبی بوجود خواهد آمد؟ (۱۰ نمره)

۳- الف) یکی از راه کارهای برای متعادل سازی پیچیدگی میان مدل و داده ها برای حالتی که مدل پیچیده تر از داده ها است، افزایش داده های آموزشی است. با این کار به مدل این فرصت داده می شود تا با دیدن حالات مختلف داده های آموزشی بتواند تعمیم پذیری بیشتری داشته باشد و در نتیجه بهتر آموزش ببیند. اما این راهکار دارای چالش هایی است. یکی از چالش های اساسی این است که توزیع داده ها به هم نخورد هم چنین، تعداد کمی داده برچسب دار در رابطه با بسیاری از مسائل وجود دارند که باعث می شود فرایند افزایش تعداد داده های آموزشی، فرایندی بسیار هزینه بر هم از نظر زمانی و هم از نظر مالی باشد. راهکارهای data augmentation برای حل این مشکل در افزایش داده ها با هزینه کمتر ارائه شدند. ما در DataAug.ipynb قرار است که برای مجموعه تصاویر داده شده، عملیات data augmentation را انجام دهیم و سپس نتیجه آن را ببینیم. برای این کار ابتدا یک مدل MLP داده شده است که باید بر روی داده های خام آموزش دهید. سپس مجموعه داده augmented را ایجاد کرده و همان مدل را بر روی مجموعه داده جدید دوباره از اول آموزش دهید. برای انجام این تمرین مراحل موجود در DataAug.ipynb را انجام دهید و آن را تکمیل کنید. تحلیل خود از نتایج به دست آمده برای دو آموزش مدل یکسان با مجموعه داده های متفاوت را بنویسید. درباره overfit شدن یا نشدن مدل ها صحبت کنید. دلیل خود از استدلال هایتان را نیز بیاورید. (۲۵ نمره)

ب) (امتیازی) با استفاده از کتابخانه keras مراحل مربوط به data augmentation را انجام دهید و مجموعه داده تولید شده خود را با مجموعه داده قبلی مقایسه کنید. (۱۵ نمره)

۴- در MLP_Overfit.ipynb یک مدل MLP و یک مجموعه داده به شما داده شده است. این مدل بر روی مجموعه داده داده شده، overfit شده است. با استفاده از ابزارهای موجود در کتابخانه keras سعی کنید از overfit شدن آن جلوگیری کنید. در نهایت نتایج به دست آمده خود را گزارش کرده و تحلیل کنید. (۲۵ نمره) (حتماً با دلیل ذکر کنید که چرا مدل overfit بود و چرا با تغییرات انجام داده شده مدل در راستای حل آن گام برداشته است)

(توجه: استفاده از regularization و dropout مورد انتظار اما به راهکارهای اضافه تر که بتوانند نتایج بهتری نیز ثبت کنند نمره امتیازی تعلق می گیرد.)

نکات تکمیلی

موفق باشید