



مسئله‌ی ۱. یادآوری یادگیری ماشین

فرض کنید می‌خواهیم مسئله رگرشن خطی را با مدل

$$g(x^{(k)}) = \sum_{j=1}^D w_j x_j^{(k)} + b$$

تخمین بزنیم. که k نمایانگر شماره داده و j نمایانگر فیچرهای آن است.

برای جلوگیری از overfit شدن مدل، می‌خواهیم تابع regularizer

$$R(w) = \frac{\lambda}{2} \sum_{j=1}^D w_j^2$$

را به cost function اضافه کنیم. یعنی داشته باشیم:

$$E_S(g) = \frac{1}{N} \sum_{k=1}^N [t^{(k)} - g(x^{(k)})]^2 + \frac{\lambda}{2} \sum_{j=1}^D w_j^2$$

الف

توضیح دهید چرا اضافه کردن این عبارت باعث کاهش overfitting مدل می‌شود.

ب

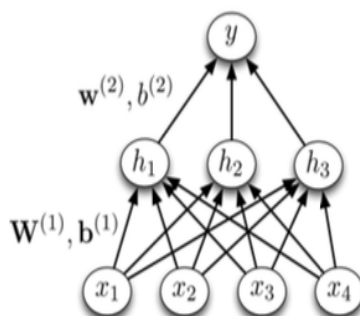
فرض کنید از الگوریتم gradient descent برای بهینه‌سازی مسئله استفاده کرده‌ایم. قوانین آپدیت را برای متغیرهای w_j و b محاسبه کنید. قانون آپدیت w_j را در این حالت اصطلاحاً weight decay می‌نامند. با توجه به رابطه بدست آمده توضیح دهید چرا همچنین نامی برای آن انتخاب شده است و رابطه آن با بخش الف سوال چیست.

مسئله ۲. شبکه تشخیص ترتیب

تابع فعالساز زیر را در نظر بگیرید:

$$f(x) = \begin{cases} 1 & x \leq 0 \\ 0 & o.w \end{cases}$$

با در نظر گرفتن این تابع به عنوان تابع فعالساز شبکه زیر، وزن‌ها و بایاس‌های شبکه زیر را به صورتی تعیین کند که اگر چهار ورودی حقیقی شبکه به صورت مرتب (یعنی $x_1 \leq x_2 \leq x_3 \leq x_4$) بودند، خروجی شبکه یک شود و در غیر اینصورت صفر.



موفق باشید (:)