



گزارش تکلیف دوم

داریوش حسن پور آده

۹۳۰۸۱۶۴

۱ قسمت ۱

در کد ارسالی داده‌ها به صورت گفته شده بارگذاری شده‌اند.

۲ قسمت ۲

داده‌های آموزشی در دو متغیر x و marks بارگذاری شده‌اند که توسط رابطه‌ی ۱ وزن‌های بهینه بدست آمد. که وزن‌های بدست آمده به صورت زیر بدست آمده است.

$$w^* = [0.30 \quad 0.30 \quad 0.05 \quad 0.05 \quad 0.15 \quad 0.15]^T$$

با توجه به این که مجموع وزن‌ها برابر با ۱ می‌باشد می‌توان مقادیر هریک از وزن‌ها را به عنوان درصد اهمیت هریک از ویژگی‌ها که برای محاسبه‌ی نمره‌ی نهایی دانشجویان در نظر گرفته شده است، در نظر گرفت. در نتیجه میزان اهمیت ویژگی‌ها به صورت جدول ۱ می‌باشد.

$$w = (x^T x)^{-1} x^T \cdot \text{marks} \quad (۱)$$

Feature	Importance%
midterm	30%
finalterm	30%
tak1	5%
tak2	5%
research	15%
project	15%

جدول ۱: درصد اهمیت هریک از ویژگی‌ها برای محاسبه‌ی نمره‌ی نهایی

Student#	Grade
1	6.80
2	12.05
3	17.10
4	11.60

جدول ۲: جدول نمره‌های پیش‌بینی شده برای داده‌های آموزشی برای قسمت دوم

نمره‌های بدست آمده برای داده‌های تست با استفاده از وزن‌های بدست آمده به صورت جدول ۲ می‌باشد.

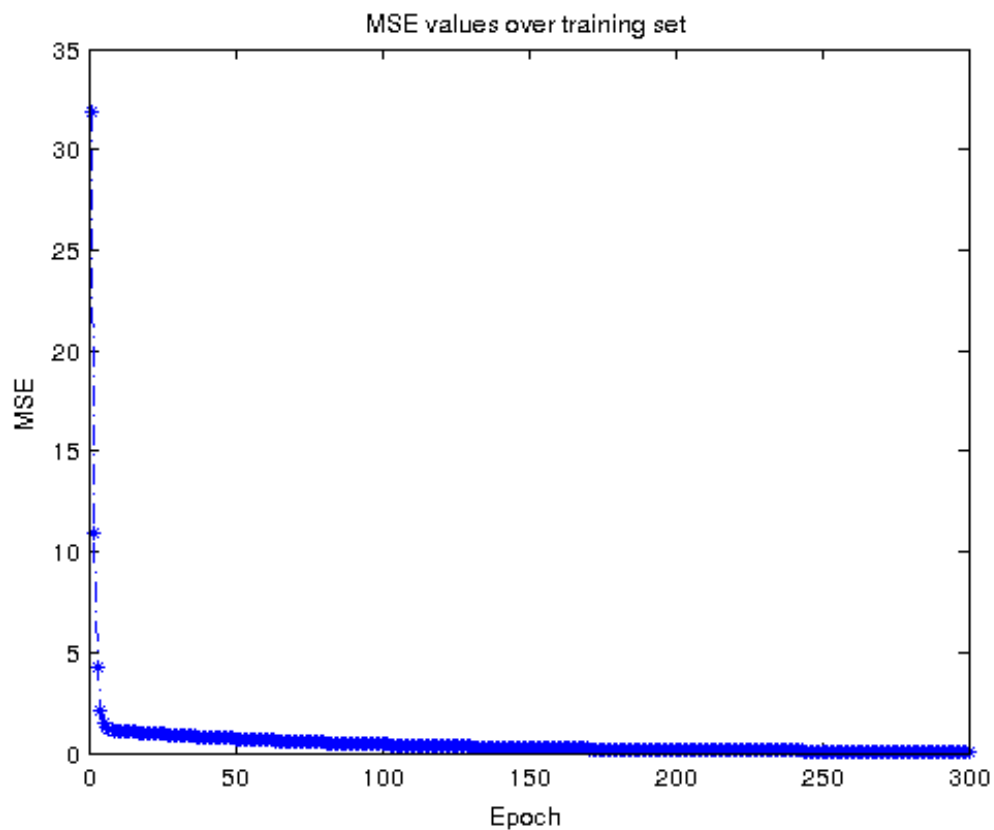
۳ قسمت ۳

وزن‌های بدست آمده در این قسمت به صورت زیر می‌باشد.

$$w^* = [0.27 \quad 0.34 \quad 0.09 \quad 0.07 \quad 0.16 \quad 0.08]^T$$

نمودار خطا برحسب تعداد دفعات تکرار را در شکل ۲ آمده است. در یادگیری ضریب یادگیری $2e-5$ و حداکثر تعداد دوره 300 در نظر گرفته شده است. نتایج پیش‌بینی بروی داده‌های تست در جدول ۳ آمده است همان طور که مشاهده می‌شود نمره‌های بدست آمده در این قسمت نزدیک به مقادیر بدست آمده در جدول ۲ می‌باشد که نشان می‌دهد الگوریتم نوشته شده درست کار می‌کند. همچنین وزن‌های بدست آمده در این روش با ضریب یادگیری ذکر شده نزدیک به وزن‌های بدست آمده در قسمت قبل می‌باشد. در صورتی که ضریب یادگیری را به مقدار $2e-4$ افزایش دهیم نمودار MSE و وزن‌های بدست آمده به صورت زیر می‌باشد. همان‌طور که مشاهده می‌شود به این هیچ عنوان خوب یادگرفته نشده است - چون از به دلیل طول گام بزرگ بهینه‌های محلی/جهانی را رد می‌کند.

$$\hat{w} = [0.30 \quad 0.30 \quad 0.05 \quad 0.05 \quad 0.15 \quad 0.15]$$

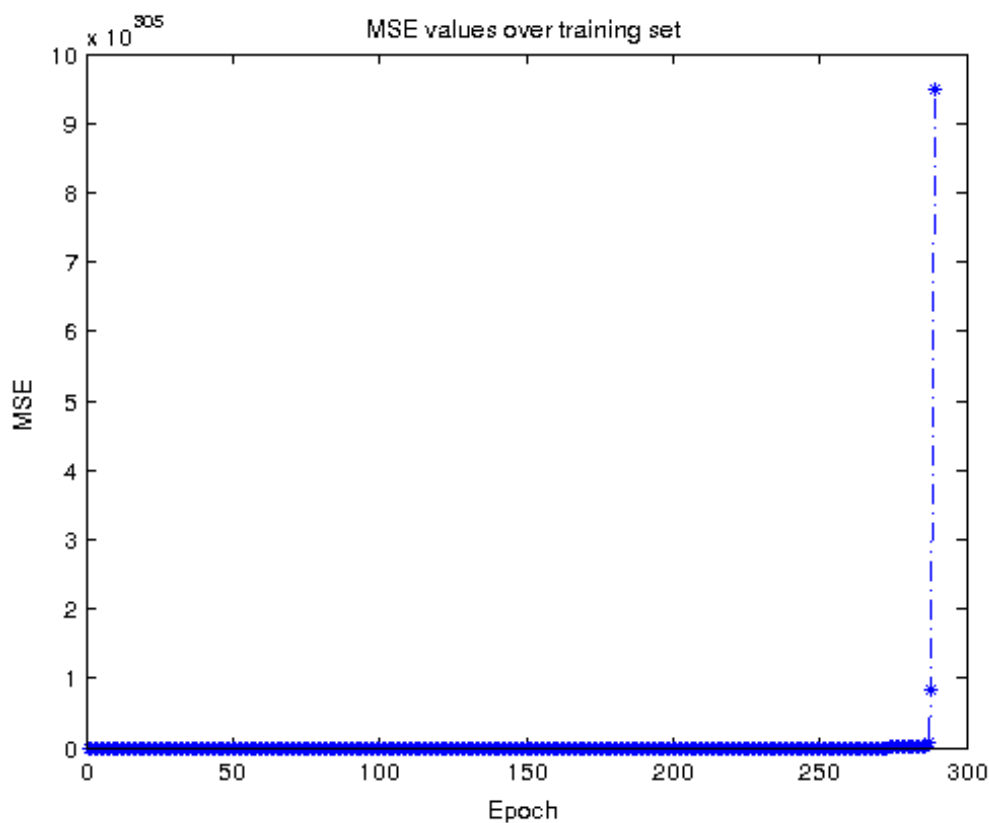


شکل ۱: نمودار خطا برحسب تعداد دفعات تکرار برای قسمت دوم با ضریب یادگیری $2e-5$

در صورتی که با ضریب یادگیری $2e-5$ با تعداد دوره‌ی ۱۰,۰۰۰ بار برنامه را اجرا کنیم ظرایب بدست آمده دقیقاً معادل با ظرایب بدست آمده در قسمت ۲ می‌باشد، که نشان می‌دهد با فرصت کافی و ضریب یادگیری مناسب برنامه می‌تواند به بهینه‌ی جهانی همگرا شود.

Student#	Grade
1	5.94
2	10.60
3	17.59
4	11.21

جدول ۳: جدول نمره‌های پیش‌بینی شده برای داده‌های آموزشی برای قسمت ۳



شکل ۲: نمودار خطا برحسب تعداد دفعات تکرار برای قسمت دوم با ضریب یادگیری $2e-4$

۴ قسمت ۴

در این قسمت به ازای دانشجویانی که نمره‌ی بالای ۱۲ را گرفته‌اند مقدار ۱ و برای کمتر از ۱۲ را مقدار ۰ به عنوان مقادیر هدف در نظر گرفته شده است. بعد از آموزش وزن‌ها و نتایج حاصل بروی داده‌های تست به صورت زیر می‌باشد.

$$w_* = \begin{bmatrix} 0.02 & 0.05 & -0.02 & 0.01 & 0.00 & 0.00 \end{bmatrix}$$

همان‌طور که مشاهده می‌شود نمره دانشجوی شماره‌ی ۴ که در قسمت ۱ زیر ۱۲ پیش‌بینی شده است و مردود

Student#	Output	Passed(threshold = 0.5)?
1	0.3224	false
2	0.7345	true
3	0.9509	true
4	0.5721	true

جدول ۴: جدول نمره‌های پیش‌بینی شده برای داده‌های آموزشی برای قسمت ۴

به حساب می‌آید در نتایج این قسمت قبول شده است - که نشان دهنده خطای مدل دسته‌بند بدست آمده است.