

شبکه عصبی و یادگیری عمیق	نیمسال دوم ۱۴۰۳-۱۴۰۴	تمرین سری اول
موعد تحویل: پنجشنبه ۱۴۰۴/۰۱/۲۲		

یک مدل خطی $y_o = mx + b$ با مقادیر دلخواه m و b را در نظر بگیرید. با توجه به نویزی که در محیط و فرایند نمونه برداری وجود دارد، نمونه‌هایی که از این خط در اختیار ما قرار می‌گیرند از رابطه کلی $y = y_o + n = mx + b + n$ تبعیت می‌کنند، که n نویز با توزیع نرمال در بازه مشخص $[-a, a]$ است. فرض کنید K نمونه از این مدل در اختیار داریم. هدف مساله، تخمین m و b با استفاده از روش gradient descent است.

۱- مقادیر دلخواهی برای m و b در نظر بگیرید و $K=50$ نمونه تصادفی در بازه $0 \leq x \leq 100$ با فرض $a = 0.1$ بسازید. با فرض مدل خطی و با استفاده از تابع خطای mean squared error و الگوریتم gradient descent مقادیر m و b را پیدا کنید.

یادآوری:

$$\theta = \begin{bmatrix} m \\ b \end{bmatrix}, \theta^* = \arg \min_{\theta} L(\theta)$$

$$L(\theta) = L(m, b) = \frac{1}{2} \sum_{k=1}^K (y_k - \hat{y}_k)^2 = \frac{1}{2} \sum_{k=1}^K (y_k - (mx_k + b))^2$$

۲- گام ۱ را با فرض $K=500$ ، $K=5000$ و $K=100000$ ، تعداد تکرار (iterations)، نرخ یادگیری و خطای قابل پذیرش (tolerance) های مختلف تکرار کنید. نمودار خطا را رسم نمایید.

۳- گام‌های فوق را برای $a = 0.3$ و $a = 0.5$ هم تکرار نمایید.

۴- به نظر شما کدام یک از موارد زیر نقش بیشتری در تخمین بهتر پارامترهای مدل دارند؟

a. تعداد نمونه

b. تعداد تکرار

c. نرخ یادگیری مناسب

d. خطای قابل پذیرش

* گزارش باید شامل روش تولید نمونه‌ها، روش انجام محاسبات، کد، نتایج و جمع بندی باشد.

** برنامه باید با زبان پایتون نوشته شود و سایر زبان‌ها مورد قبول نیستند. لطفاً کتابخانه‌هایی را که برای انجام تمرین استفاده می‌کنید معرفی نمایید.

*** تمرین فردی است و برای گزارشاتی که در هر یک از بخشها شبیه باشند نمره صفر منظور خواهد شد.

موفق باشید