

درس یادگیری ماشین

# گزارش تکلیف Neural Networks

استاد درس:

دکتر افتخاری

نگارش:

امیرحسین ابوالحسنی

شماره دانشجویی: ۴۰۰۴۰۵۰۰۳

پاییز ۱۴۰۳

## فهرست مطالب

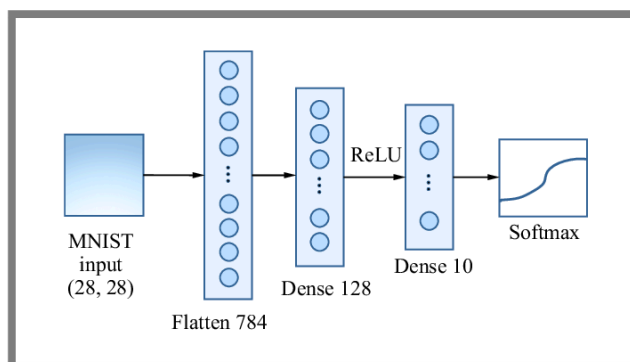
۱	مقدمه	۲
۲	معماری شبکه	۲
۳	بهینه‌ساز و تابع هزینه	۲
۴	آموزش مدل	۳
۵	ارزیابی مدل	۳
۶	خروجی مدل	۴

## ۱ مقدمه

یکی از بهترین کتابخانه‌ها برای پیاده‌سازی شبکه‌های عصبی، کتابخانه Pytorch می‌باشد. در این تکلیف به پیاده‌سازی شبکه عصبی ساده برای تسک طبقه بندی روی دیتاست MNIST پرداخته می‌شود.

## ۲ معماری شبکه

برای معماری شبکه عصبی استفاده شده، از دو لایه Perceptron به همراه تابع فعال ساز ReLU استفاده شده است.



شکل ۱: معماری شبکه عصبی

## ۳ بهینه‌ساز و تابع هزینه

برای آموزش مدل از بهینه‌ساز Adam استفاده شده است.

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$$

$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

$$\theta_t = \theta_{t-1} - \frac{\alpha \cdot \hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

Where:

- $m_t$ : First moment estimate (mean of gradients)
- $v_t$ : Second moment estimate (uncentered variance of gradients)
- $g_t$ : Gradient at time step  $t$
- $\beta_1, \beta_2$ : Exponential decay rates

- $\alpha$ : Learning rate
- $\epsilon$ : Small constant to prevent division by zero
- $\theta_t$ : Updated parameter

تابع هزینه‌ای که برای این مدل در نظر گرفته شده، به علت نوع تسک که طبقه بندی است، Cross Entropy می‌باشد.

## ۴ آموزش مدل

مدل با پارامترهای جدول ۱ آموزش دیده است.

نام هایپر پارامتر	مقدار
Epochs	۵
Learning Rate	۰.۰۰۱
Batch Size	۱۲۸

جدول ۱: مقادیر هایپر پارامترهای مدل

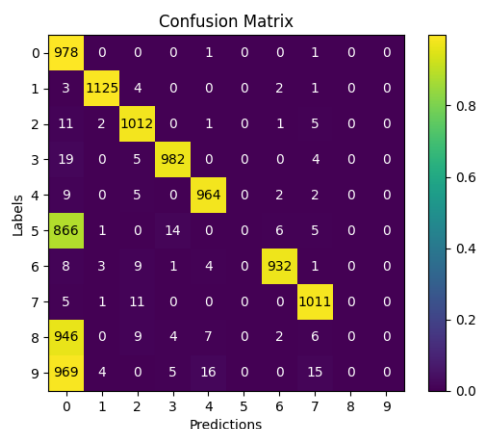
## ۵ ارزیابی مدل

پس از آموزش مدل نوبت به ارزیابی آن می‌رسد. دو متریک دقت و F1 Score برای ارزیابی عملکرد مدل استفاده شده است.

معیار	مقدار
دقت	۰.۶۹
F1 Score	۰.۶۹۸

جدول ۲: نمرات ارزیابی مدل ارائه شده

همچنین در شکل ۲، ماتریس سردرگمی نشان داده شده است.



شکل ۲: ماتریس سردرگمی

## ۶ خروجی مدل

نمونه ای از خروجی مدل در شکل ۳ قابل مشاهده است.

True: 7 Pred: 7	True: 2 Pred: 2	True: 1 Pred: 1	True: 0 Pred: 0	True: 4 Pred: 4	True: 1 Pred: 1	True: 4 Pred: 4	True: 9 Pred: 0
True: 5 Pred: 0	True: 9 Pred: 0	True: 0 Pred: 0	True: 6 Pred: 6	True: 9 Pred: 0	True: 0 Pred: 0	True: 1 Pred: 1	True: 5 Pred: 0
True: 9 Pred: 0	True: 7 Pred: 7	True: 3 Pred: 3	True: 4 Pred: 4	True: 9 Pred: 0	True: 6 Pred: 6	True: 6 Pred: 6	True: 5 Pred: 0
True: 4 Pred: 4	True: 0 Pred: 0	True: 7 Pred: 7	True: 4 Pred: 4	True: 0 Pred: 0	True: 1 Pred: 1	True: 3 Pred: 3	True: 1 Pred: 1
True: 3 Pred: 3	True: 4 Pred: 4	True: 7 Pred: 7	True: 2 Pred: 2	True: 7 Pred: 7	True: 1 Pred: 1	True: 2 Pred: 2	True: 1 Pred: 1
True: 1 Pred: 1	True: 7 Pred: 7	True: 4 Pred: 4	True: 2 Pred: 2	True: 3 Pred: 3	True: 5 Pred: 0	True: 1 Pred: 1	True: 2 Pred: 2
True: 4 Pred: 4	True: 4 Pred: 4	True: 6 Pred: 6	True: 3 Pred: 3	True: 5 Pred: 0	True: 5 Pred: 0	True: 6 Pred: 6	True: 0 Pred: 0
True: 4 Pred: 4	True: 1 Pred: 1	True: 9 Pred: 0	True: 5 Pred: 0	True: 7 Pred: 7	True: 8 Pred: 0	True: 9 Pred: 0	True: 3 Pred: 3

شکل ۳: خروجی مدل برای یک دسته از داده تست