

یادگیری عمیق

تمرین سوم

نیکی نزاکتی

98522094

1.

الف) Overfitting یک خطای مدل سازی در شبکه های عصبی است که زمانی رخ می دهد که یک تابع خیلی نزدیک به مجموعه محدودی از نقاط داده تراز شده باشد. در نتیجه، این مدل فقط برای ارجاع به مجموعه داده های اولیه خود مفید است و نه به هیچ مجموعه داده دیگری. زمانی میگوییم شبکه عصبی overfit شده است که دقت خوبی بر داده به خصوص دارد و توانایی تعمیم و دقت بالا بر داده های دیگر را ندارد. راه کار های جلوگیری از overfit شدن و حل این مشکل برای شبکه شامل آموزش با داده های بیشتر، data augmentation، کم کردن اندازه شبکه و regularization است.

ب) زیرا مدل شکل ۱ روی این داده های به خصوص overfit شده است و الگوی کلی دسته بندی را یاد گرفته و قابلیت تعمیم ندارد. در نتیجه، داده هایی که تاکنون ندیده است را نمیتواند به خوبی دسته بندی کند و خطای بیشتری نسبت به مدل شکل ۲، که الگوی دسته بندی را بهتر یاد گرفته است، دارد.

2.

$$L_2 \text{ regularization: } \text{Error}(y, \hat{y}) + \underbrace{0.9}_{\lambda} \sum_{i=1}^N w_i^2 = \text{Loss} \quad \textcircled{2} \quad \text{الف)}$$

$$\text{Data: } (1, 20)$$

$$\text{Feed Forward: } 1 + \frac{2 \cdot 1}{2} + \frac{3 \cdot 1}{3} - \frac{2 \times 1}{2} - \frac{1 \times 1}{1} = 3$$

$$\text{MSE: } \frac{1}{2} (20 - 3)^2 = \frac{1}{2} \cdot 17^2 = 144.5$$

$$L_2: 144.5 + 0.9 (1 + 4 + 9 + 4 + 1) = 161.6$$

$$w = w - \lambda L_2 \Rightarrow \begin{cases} 1 - 16.16 = -15.16 \\ 2 - 16.16 = -14.16 \\ 3 - 16.16 = -13.16 \\ -2 - 16.16 = -18.16 \\ -1 - 16.16 = -17.16 \end{cases}$$

CS Scanned with CamScanner

(ب)

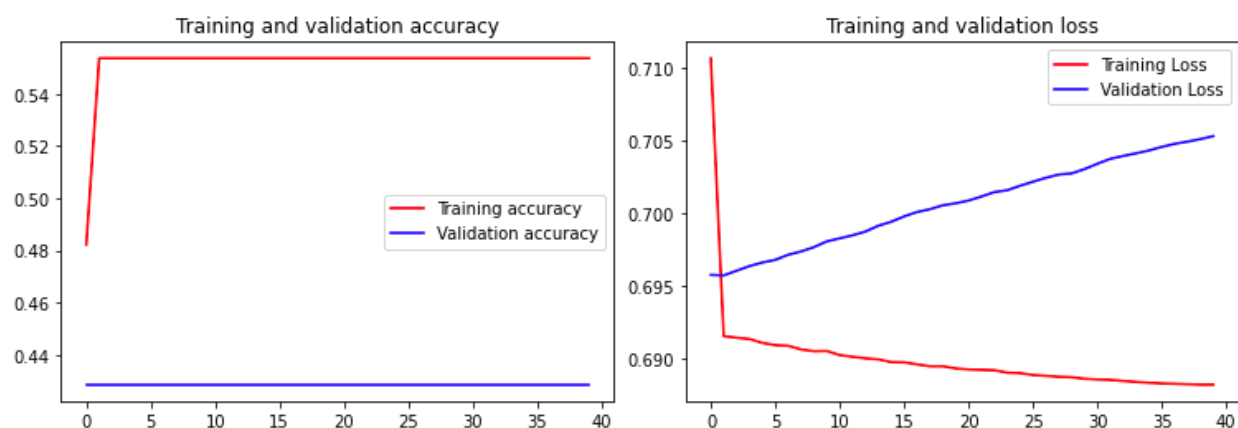
با کاهش مقدار λ مقدار جریمه وزن ها کاهش یافته و برعکس، به این ترتیب میتوان مقدار جریمه شدن و کاهش یافتن وزن ها را با λ کنترل کرد.

3.

الف) دقت و ضرر مدل ما با داده اولیه به صورت زیر است:

loss: 0.6882 - accuracy: 0.5536 - val_loss: 0.7053 - val_accuracy: 0.4286

Test = loss: 0.7047 - accuracy: 0.4333

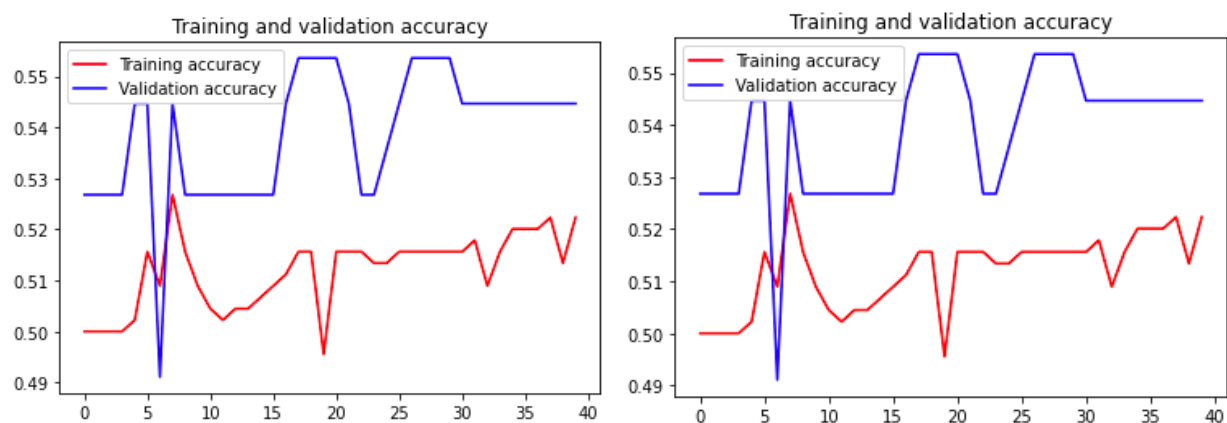


همانطور که مشاهده می شود دقت مدل بر داده train بهتر از validation است که یعنی مدل ما بر داده آموزش overfit شده است.

عملکرد مدل پس از data augmentation:

loss: 0.6868 - accuracy: 0.5223 - val_loss: 0.6829 - val_accuracy: 0.5446

Test = loss: 0.6958 - accuracy: 0.4875

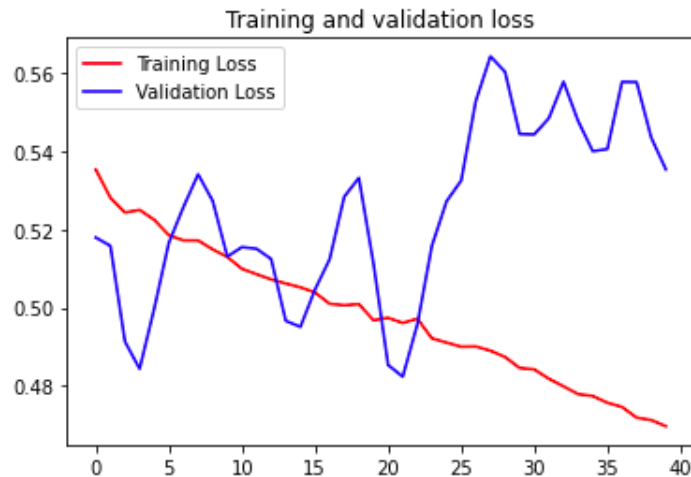


مشاهده می شود که پس از افزایش داده دقت مدل از آموزش بیشتر شده که این نشان می دهد مشکل overfit به کمک data augmentation حل شده است.

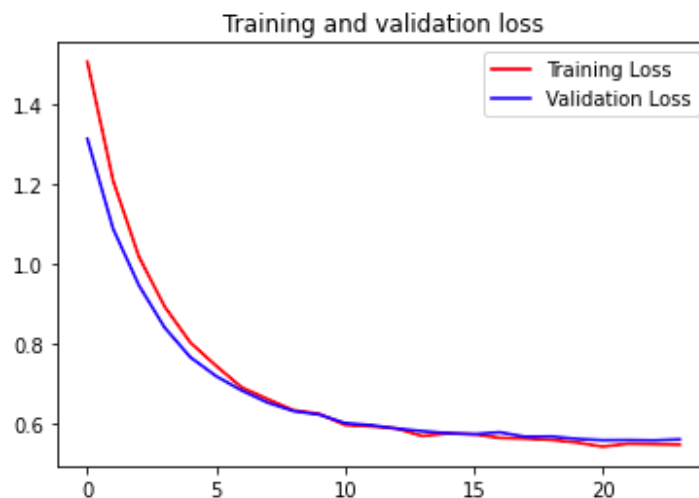
4.

عملکرد مدل در حالت اولیه به صورت زیر می باشد:

loss: 0.5335 - accuracy: 0.5152 - val_loss: 0.5600 - val_accuracy: 0.6667



از زیاد بودن ضرر validation در مقایسه با آموزش میبینیم که مدل overfit شده است. دلیل آن میتواند تعداد کم داده باشد که ۳۰ تا بود. برای حل این مشکل تعداد سَمپل ها را ۳۰۰ میکنیم. همچنین از L2 Regularization، Dropout، و EarlyStopping استفاده میکنیم که همه از overfitting جلوگیری میکنند. خروجی مدل بهبود یافته:



مشاهده می شود که مدل دیگر overfit نشده است.