



به نام خدا



دانشگاه تهران
دانشکده مهندسی برق و کامپیوتر
شبکه های عصبی و یادگیری عمیق
تمرین ۲ امتیازی

نام و نام خانوادگی	سجاد پاکدامن ساوجی
شماره دانشجویی	۸۱۰۱۹۵۵۱۷
تاریخ ارسال گزارش	

فهرست گزارش سوالات

3

سوال 1 Separation Index

14

نحوه اجرای کدها

در این تمرین شاخص جدایی را برای تمامی لایه های شبکه عصبی محاسبه می کنیم.
 (آ) در این قسمت از شاخص جدایی که در مقاله مربوطه آمده بود استفاده شده است و نمودار جدایی پذیری این شاخص در شکل ۲ رسم شده است.

$$SI(\{x_L^q\}_{q=1}^Q, \{l^q\}_{q=1}^Q) = \frac{1}{Q} \sum_{q=1}^Q \varphi(l^q - l_{near}^q) \quad \varphi(v) = \begin{cases} 1 & v = 0 \\ 0 & v \neq 0 \end{cases}$$

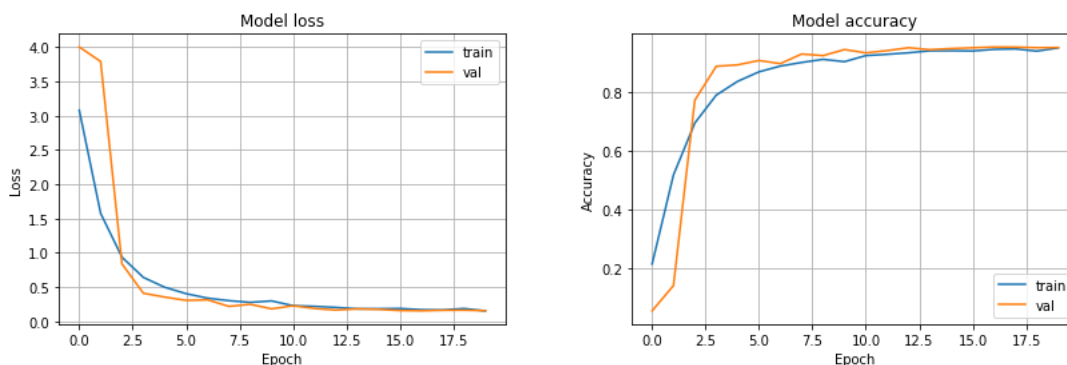
$$q_{near}^L = \arg \min_h \|x_L^q - x_L^h\| \quad h \in \{1, 2, \dots, Q\} \text{ and } h \neq q$$

این شاخص، جدایی پذیری را با شمارش نقاطی که نزدیکترین همسایه آن ها همکلاس آن ها است محاسبه می کند.

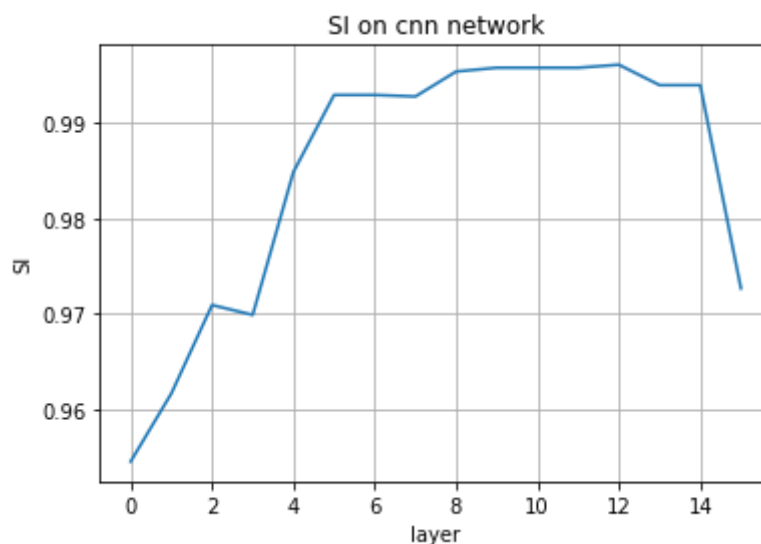
"Model: "sequential_1

# Layer (type)	Output Shape	Param
=====		
conv2d_1 (Conv2D)	(None, 28, 28, 32)	896
=====		
max_pooling2d_1 (MaxPooling2D)	(None, 14, 14, 32)	0
=====		
batch_normalization_1 (Batch Normalization)	(None, 14, 14, 32)	128
=====		
conv2d_2 (Conv2D)	(None, 12, 12, 64)	18496
=====		
max_pooling2d_2 (MaxPooling2D)	(None, 6, 6, 64)	0
=====		
dropout_1 (Dropout)	(None, 6, 6, 64)	0
=====		
batch_normalization_2 (Batch Normalization)	(None, 6, 6, 64)	256
=====		
conv2d_3 (Conv2D)	(None, 4, 4, 128)	73856
=====		
max_pooling2d_3 (MaxPooling2D)	(None, 2, 2, 128)	0
=====		
dropout_2 (Dropout)	(None, 2, 2, 128)	0
=====		
flatten_1 (Flatten)	(None, 512)	0
=====		
batch_normalization_3 (Batch Normalization)	(None, 512)	2048
=====		
dense_1 (Dense)	(None, 128)	65664
=====		
dropout_3 (Dropout)	(None, 128)	0
=====		
dense_2 (Dense)	(None, 43)	5547
=====		
Total params: 166,891		
Trainable params: 0		
Non-trainable params: 166,891		

در بالا مشخصات شبکه آموزش داده شده آمده است. این شبکه در کل ۱۵ لایه دارد. نمودار های آموزش این شبکه در شکل ۱ آمده است.



شکل ۱. نمودار های دقت و هزینه هنگام آموزش شبکه



شکل ۲. نمودار جدایی پذیری پس از هر لایه شبکه عصبی

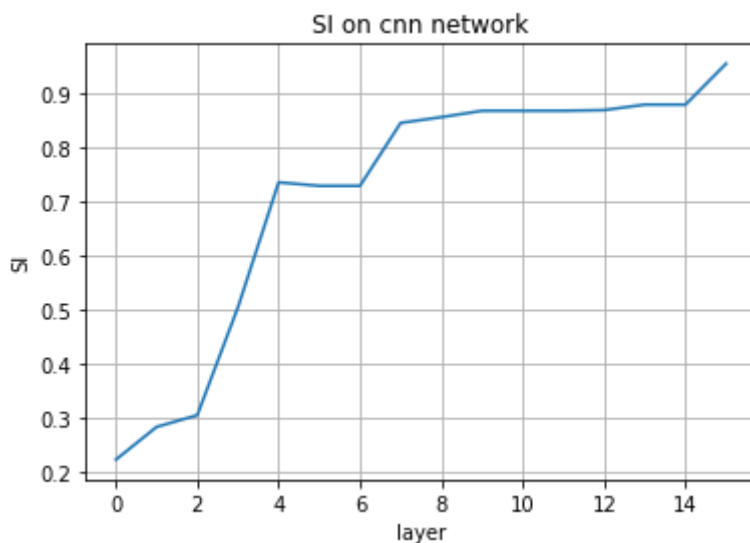
از آنجایی که برای محاسبه این معیار باید فاصله هر نقطه از تمامی نقاط دیگر محاسبه شود، اوردر زمانی $O(Q^2n)$ طول می کشد که بسیار هزینه بر است.

ب) محاسبه شاخص جدایی پذیری شبکه با استفاده از روش دوم

$$I(l) = \frac{\sum_{q=1}^Q \varphi\left(\begin{smallmatrix} l \\ k \end{smallmatrix} x_q\right)}{Q}$$

$$\varphi\left(\begin{smallmatrix} l \\ k \end{smallmatrix} x_q\right) = \begin{cases} 1 & \text{if } \operatorname{argmin}_{k_1} \|x_q^l - M_{k_1}\| = k \\ 0 & \text{else} \end{cases}$$

در این روش ابتدا مراکز دسته بدست آورده می‌شوند و تعداد نقاطی که به مرکز دسته خود از دیگر مراکز دسته نزدیکتر هستند شمارش می‌شود. این روش مرتبه زمانی $O(CQn)$ طول می‌کشد. نتیجه این روش در شکا ۳ آمده است.



شکل ۳. محاسبه شاخص جدایی برای تمامی لایه های شبکه

نحوه اجرای کدها

کد های تمرین به تفکیک سوال در نوتبوک های مربوطه پیوست شده است و نیازی به اجرای مجدد آن ها نیست. نتایج در نوتبوک ها وجود دارند.