

۱- شبکه عصبی کانولوشن را در نظر بگیرید که توسط لایه های ستون سمت چپ جدول زیر تعریف شده است. ابعاد خروجی و تعداد پارامترها را در هر لایه پر کنید. می توانید ابعاد فعال سازی را در قالب  $(H, W, C)$  بنویسید، که در آن  $H, W, C$  به ترتیب ارتفاع، عرض و ابعاد کانال هستند. به جز در مواردی که مشخص شده است،  $padding$  را ۱ فرض کنید، در صورت لزوم  $stride$  را ۱ بگیرید.

**Notation:**

CONV $x$ - $N$  denotes a convolutional layer with  $N$  filters with height and width equal to  $x$ .

POOL- $n$  denotes a  $n \times n$  max-pooling layer with stride of  $n$  and 0 padding.

FLATTEN flattens its inputs, identical to `torch.nn.flatten` / `tf.layers.flatten`

FC- $N$  denotes a fully-connected layer with  $N$  neurons

Layer	Activation Volume Dimensions	Number of parameters
Input	$32 \times 32 \times 3$	0
CONV3-8		
Leaky ReLU		
POOL-2		
BATCHNORM		
CONV3-16		
Leaky ReLU		
POOL-2		
FLATTEN		
FC-10		

۲- در یک شبکه عصبی عمیق، لایه  $U$  (با اندازه  $m$ ) خروجی های خود را به لایه  $V$  (با اندازه  $n$ ) با وزن های  $W$  ارسال می کند. فرض کنید  $u_i$  خروجی نورون  $i$ ام لایه  $U$  است،  $v_j$  خروجی نورون  $j$ ام لایه  $V$ ، و  $w_{i,j}$  وزن اتصال از  $u_i$  به  $v_j$  است. لایه  $U$  از تابع فعال سازی تانژانت هایپربولیک ( $\tanh$ ) استفاده می کند، در حالی که لایه  $V$  از تابع فعال سازی سیگموئید استفاده می کند. عبارت مناسب برای مشتق جزئی زیر چیست:

$$\frac{\partial v_j}{\partial u_i} = ?$$

- (a)  $\sum_{i=1}^m u_i w_{i,j} v_j (1 - v_j^2)$
- (b)  $\sum_{i=1}^m u_i w_{i,j} v_j (1 - v_j)$
- (c)  $u_i w_{i,j} v_j (1 - v_j)$
- (d)  $(1 - u_i^2) w_{i,j} v_j (1 - v_j)$
- (e)  $\sum_{i=1}^m (1 - u_i^2) w_{i,j} v_j (1 - v_j)$
- (f)  $\sum_{i=1}^m w_{i,j} v_j (1 - v_j)$
- (g)  $w_{i,j} v_j (1 - v_j)$