Computing the Cost II

Now that you have the matrix with cosine similarity scores, which is the product of two matrices, we go ahead and compute the cost.

$\cos(v_1, v_2)$

v_2 \ v_1	_1	_2	_3	_4
_1	0.9	-0.8	-0.3	-0.5
_2	-0.8	0.5	-0.1	-0.2
_3	-0.3	-0.1	0.7	-0.8
_4	-0.5	-0.2	-0.8	1.0

mean_neg: speeds up training

closest_neg: helps penalize the cost more

We now introduce two concepts, the **mean_neg,** which is the mean negative of all the other off diagonals in the row, and the **closest_neg,** which corresponds to the highest number in the off diagonals.

$$\text{Cost} = \max(-\cos(A, P) + \cos(A, N) + \alpha, 0)$$

So we will have two costs now:

$$\text{Cost1} = \max(-\cos(A, P) + mean\_neg) + \alpha, 0)$$

$$\text{Cost2} = \max(-\cos(A, P) + closest\_neg + \alpha, 0)$$

The full cost is defined as: Cost 1 + Cost 2.