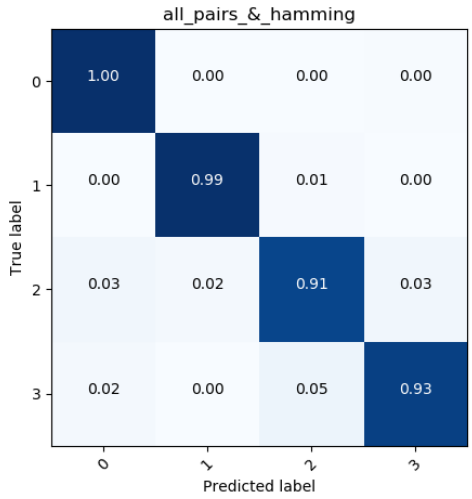
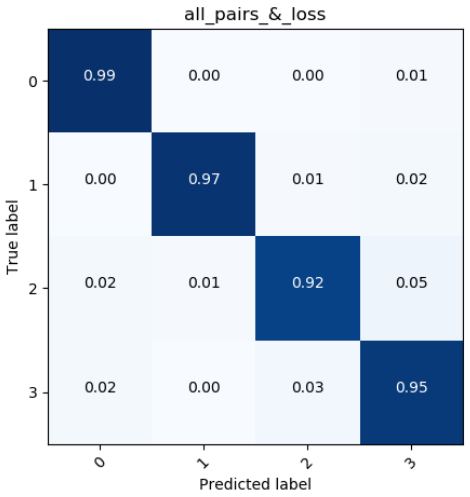
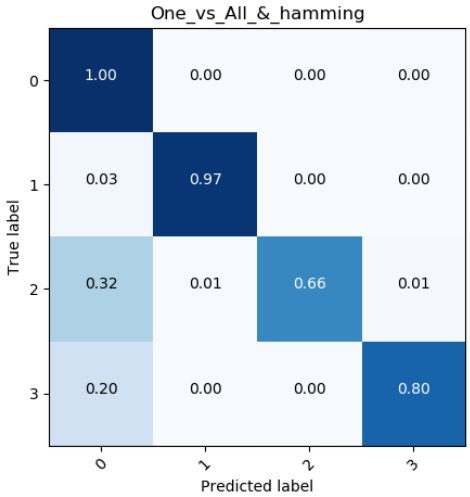
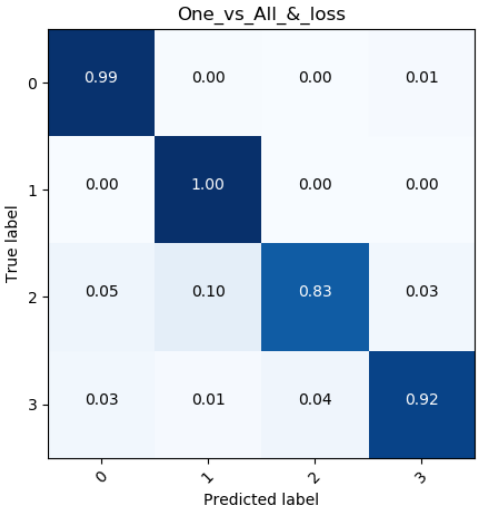
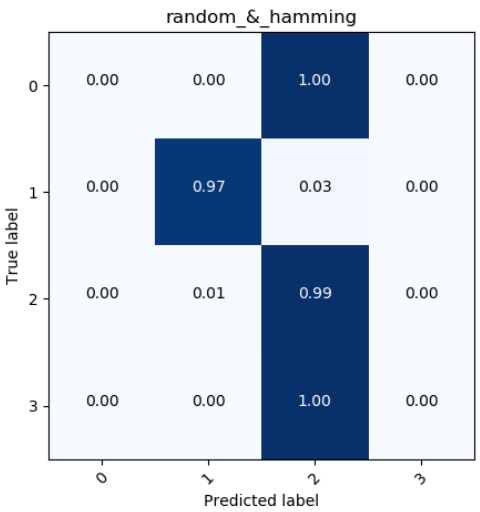
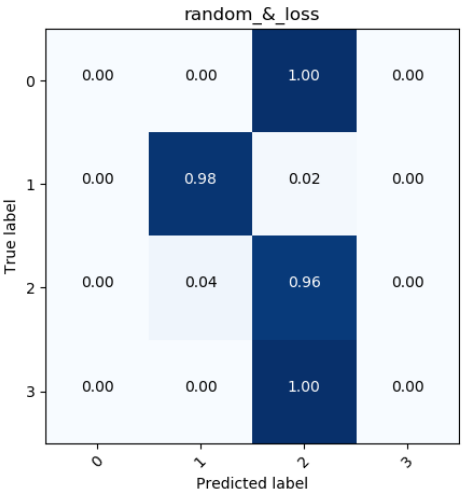
Ex1 – Conclusion

Student: Hodaya Koslowsky

In this exercise we ran 6 types of models, the normalized confusion matrix of the result is shown here.

	Hamming	Loss
All Pairs		
One vs All		
Random matrix		

Example of the accuracy of each model for a single run:

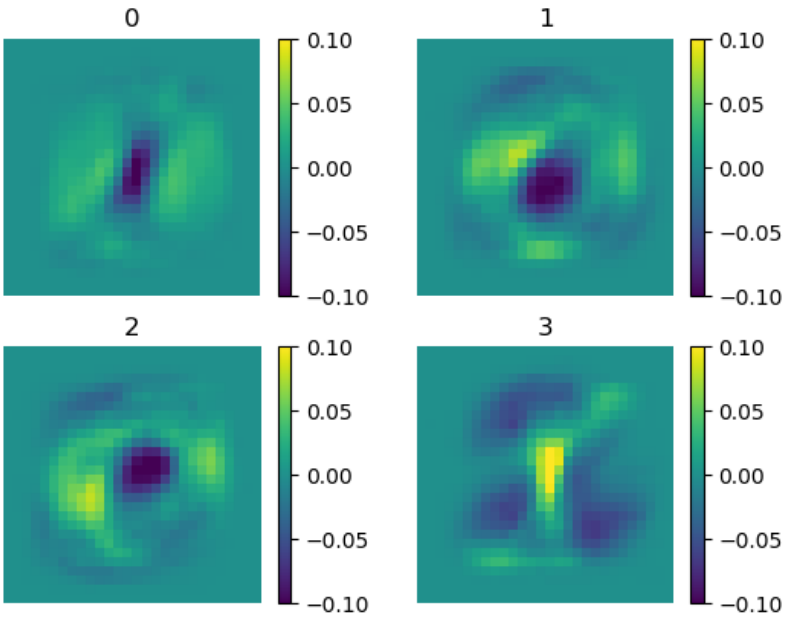
Accuracy	Loss	Hamming
All Pairs	0.964	0.96
One VS All	0.951	0.894
Random	0.512	0.472

Clearly the random matrix was the worse with learning. The All Pairs technique achieved the best results, although One VS All was also good.

The loss decoding was better in all the models, and the difference between the loss and the hamming decoding seems most significant in the One VS All model.

Bonus:

I was curious to see the matrix of W, so I took the model with the highest accuracy – all pairs with loss decoding, and this is the result:



I assume that with all the classes and all the train set the digits could have been seen more clearly, but it is possible to see the shape of the 0, and the fact that 0,1,2 all give negative weight to the middle while 3 has its most positive weights in the middle.