Activated Code a_c

0.10	0.03	0.08	0.15	0.16
0.11	0.40	0.15	0.87	0.14
0.14	0.31	0.68	0.10	0.12
0.08	0.62	0.19	0.22	0.17
0.08	0.03	0.06	0.07	0.06

Sparse Constraint:

KL-Divergence
It limits the average activated code to be close to 0.
Thus, the elements in activated code must mostly be near 0.

Drawbacks:
- **Can only apply Sigmoid or Tanh.**
- **Maybe all elements are too small.**

Activated Code a_c **Sparse Constraint: k-sparse (keep k largest elements)**

0	0	0	0	0
0	0	0	0.87	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

k = 1

0	0	0	0	0
0	0	0	0.87	0
0	0	0.68	0	0
0	0.62	0	0	0
0	0	0	0	0

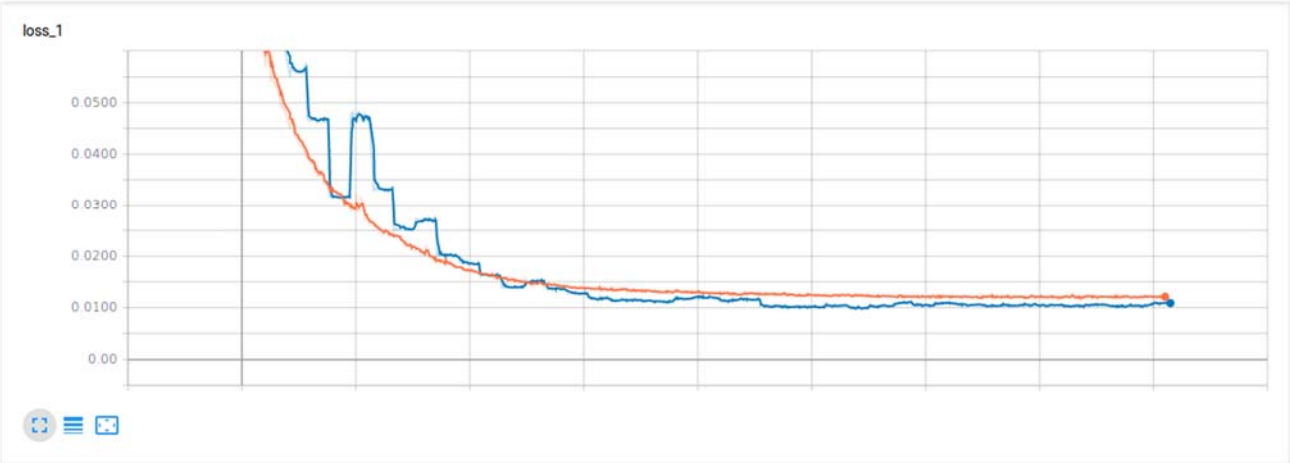
k = 3

0	0	0	0	0
0	0.40	0	0.87	0
0	0.31	0.68	0	0
0	0.62	0	0	0
0	0	0	0	0

k = 5

Good Points:
- **More flexible.**
- **Can keep features.**
- **Can apply ReLU or LeakyReLU.**
- **Able to get local features.**

Loss of Autoencoder



Problems:
- **Reconstruction slice is too blur.**
- **Classifier learns nothing from code.**

Potential Causes:
- **Model is too simple.**
- **Too many decode layers.**
- **Not enough weights.**
- **Needs a better sparse constraint.**