The experiment is carried out on four datasets in CTU13. I split data for training and testing as the same in Nur Heywood’ paper I attached (although training data I quite large, up to 30,000 samples, I train the model on the whole training data). The network architecture, parameters are the same as the previous experiments on NSL-KDD, UNSW datasets. Overall, our VAE out-perform the state of the art, and Zur Heywood method.

Parameters: patience: 100, validate\_every: 1, monitor: 200, max\_epoch: 1000

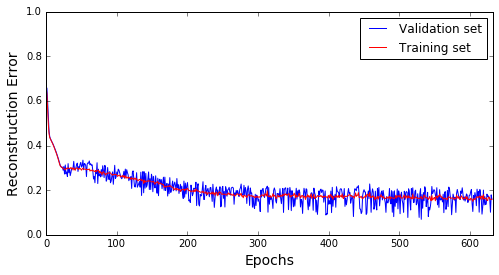
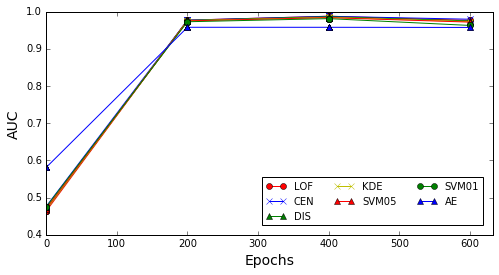
**+ Data: ['CTU13\_08', 'CTU13\_13', 'CTU13\_10', 'CTU13\_09']**

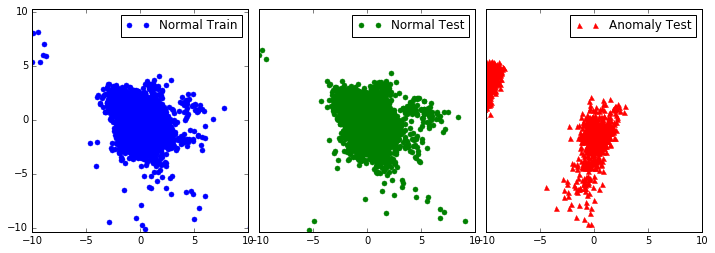
**+ Scaler: maxabs**

**1. CTU13\_08 ...**

+ Hidden Sizes: 40 [29, 18, 7] - Batch\_sizes: 1165

+ Data: 29128 (23303 train, 5825 vali) - 43694 normal, 3677 anomaly





[ 631. 0.159089 0.15971 ]

**+ Standard Deviation of Hidden data:**

[ 0.000243 0.000466 0.000228 0.000402 0.00027 0.000221 0.000565]

**+ AUC input, AUC hidden:**

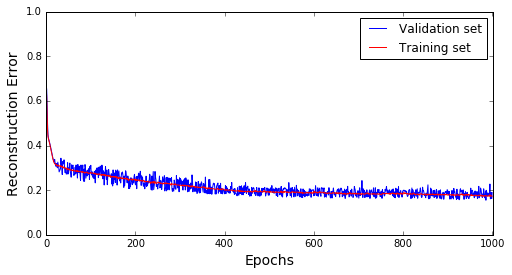
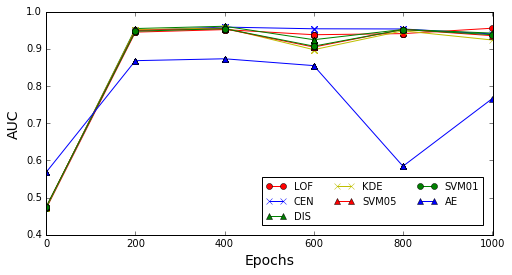
[[ 0.754 0.971 0.966 0.958 0.958 0.797 0. 0. ]

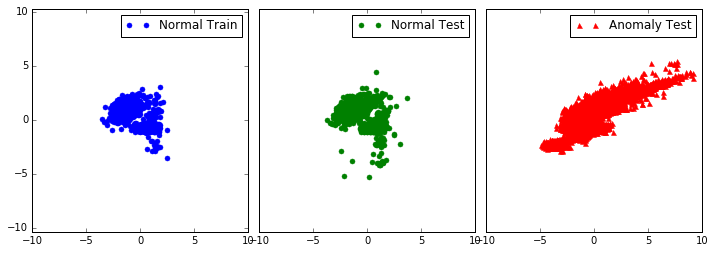
[ 0.978 0.981 0.976 0.973 0.972 0.965 0.958 15.971]

**2. CTU13\_13 ...**

+ Hidden Sizes: 40 [29, 18, 7] - Batch\_sizes: 511

+ Data: 12775 (10220 train, 2555 vali) - 19164 normal, 24002 anomaly





[ 1000. 0.188987 0.173096]

**+ Standard Deviation of Hidden data:**

[ 0.001669 0.004364 0.001326 0.004295 0.002362 0.003095 0.003252]

**+ AUC input, AUC hidden:**

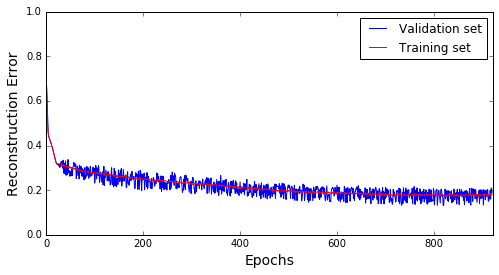
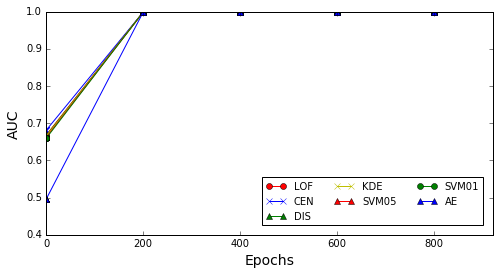
[[ 0.69 0.916 0.891 0.889 0.925 0.898 0. 0. ]

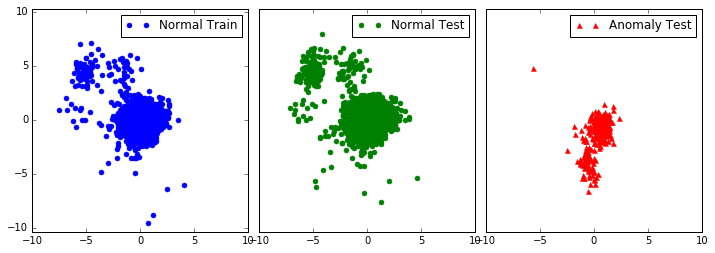
[ 0.956 0.94 0.943 0.925 0.937 0.94 0.765 17.31 ]]

**3. CTU13\_10 ...**

+ Hidden Sizes: 38 [28, 17, 7] - Batch\_sizes: 253

+ Data: 6338 (5071 train, 1267 vali) - 9509 normal, 63812 anomaly





[ 920. 0.186723 0.178054]

**+ Standard Deviation of Hidden data:**

[ 0.000404 0.000251 0.00028 0.005457 0.001205 0.000335 0.000397]

**+ AUC input, AUC hidden:**

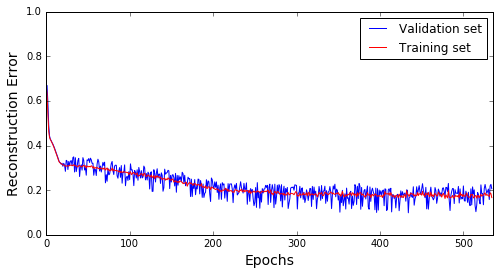
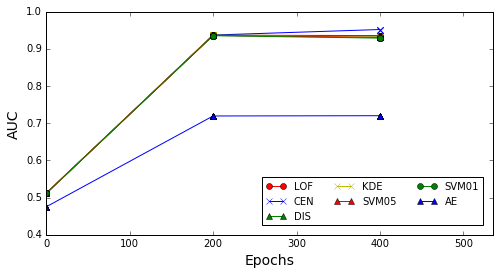
[[ 0.569 0.996 0.998 0.998 0.998 0.983 0. 0. ]

[ 0.999 0.999 0.999 0.999 0.999 0.999 0.998 17.805]]

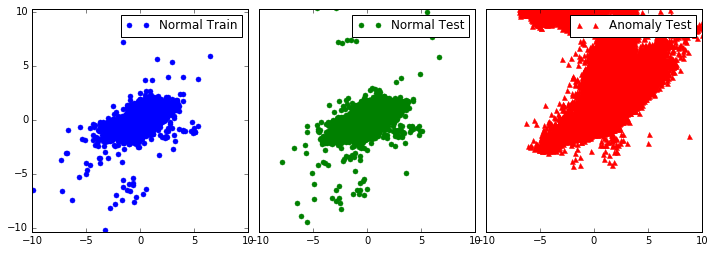
**4. CTU13\_09 ...**

+ Hidden Sizes: 41 [30, 18, 7] - Batch\_sizes: 479

+ Data: 11986 (9589 train, 2397 vali) - 17981 normal, 110993 anomaly



[ 534. 0.206538 0.166679]



**+ Standard Deviation of Hidden data:**

[ 0.000462 0.00087 0.0006 0.000771 0.000694 0.000379 0.000518]

**+ AUC input, AUC hidden:**

[[ 0.743 0.915 0.734 0.72 0.851 0.852 0. 0. ]

[ 0.937 0.952 0.937 0.934 0.933 0.934 0.72 16.668]]

**The different AUCs when training on hidden data of VAE and input data:**

[[ 0.224 0.01 0.01 0.015 0.014 0.168 0.958 15.971]

[ 0.266 0.024 0.052 0.036 0.011 0.042 0.765 17.31 ]

[ 0.43 0.004 0.001 0.001 0.002 0.016 0.998 17.805]

[ 0.193 0.037 0.203 0.214 0.083 0.082 0.72 16.668]]