

# **DQM-DC**

# **Project Report**

Patomporn (Jab)

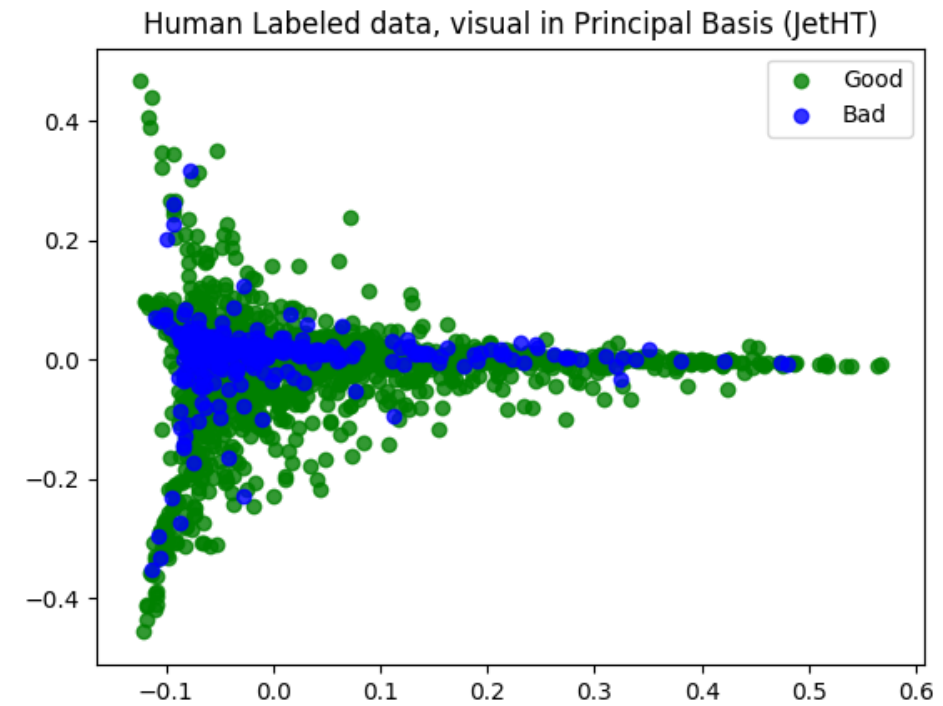
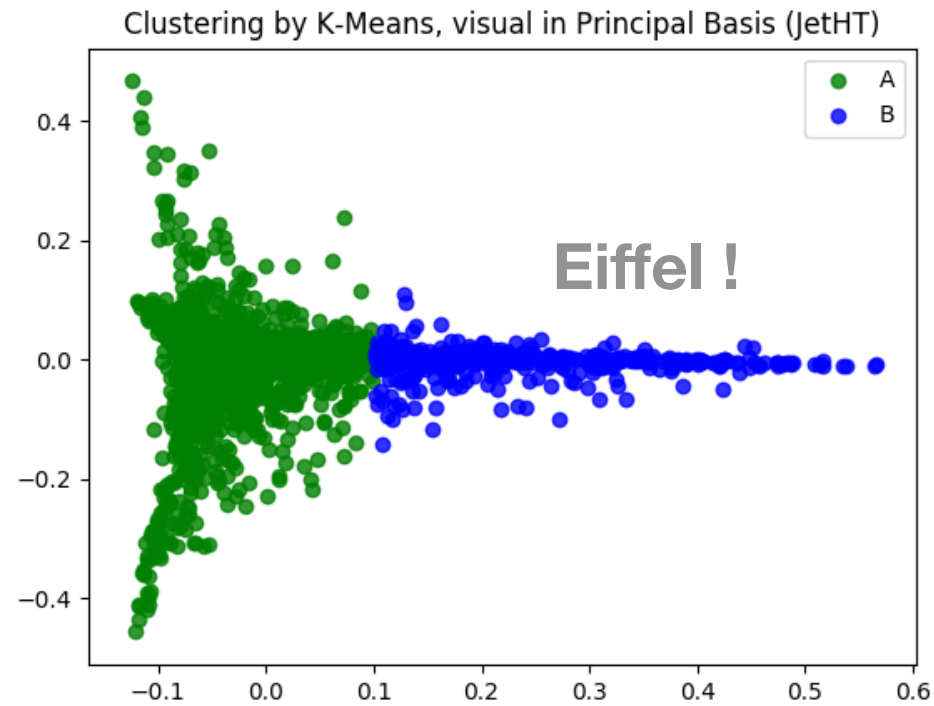
14 June 2019

# Data

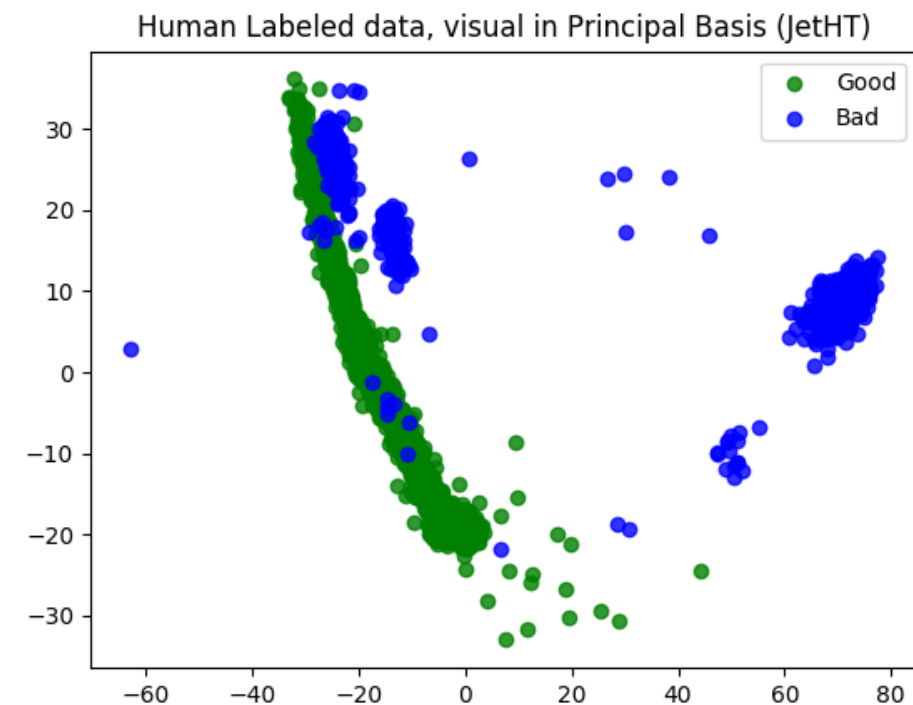
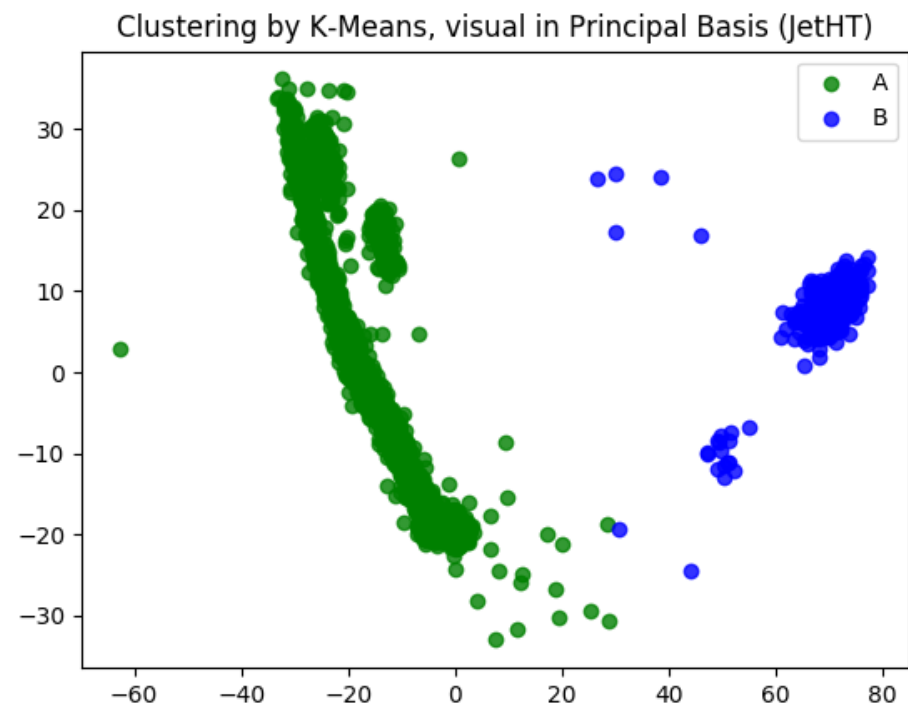
- 2016 Datasets ?
  - JetHT
    - 2806 Features
      - Has human label for “Good” or “Bad” (Might not be ground truth)
- Preprocess
  - **Normalize:  $x_i \in [0,1]$**

# Primary Analysis

## Normalize data



## Standardize data



# Autoencoder (Model)

- Body (All Fully Connected)
  - 128 PReLU
  - 64 PReLU
  - 16 PReLU
  - 64 PReLU
  - 128 Sigmoid
- **Variable initializer:** Truncated normal with sigma  $\sim 0.1$ ,
- **Batch\_size** 64

# Sparse Model

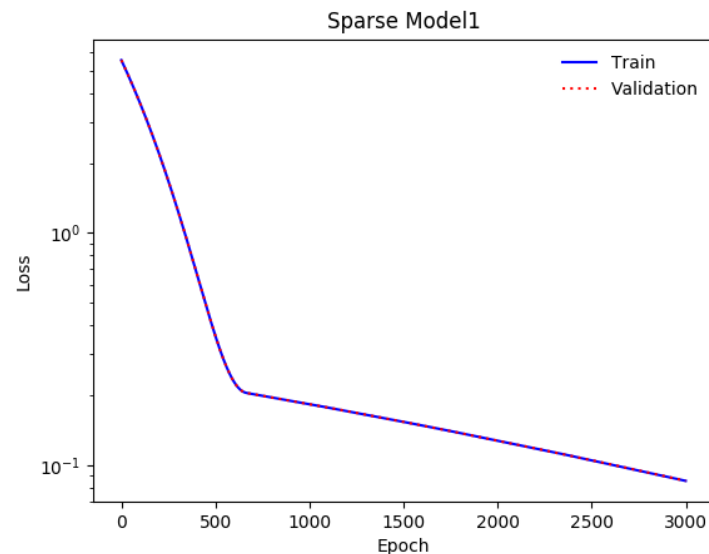
- Unsupervised
- Similar to Vanilla Autoencoder
- Tweak by **L1 Regularization ( Prevent overfitting )**

$$\mathcal{L}_{\text{tot}} \equiv \frac{1}{N} \sum_i^N |x_i - \tilde{x}_i|^2 + \lambda \sum_j ||w_j||$$

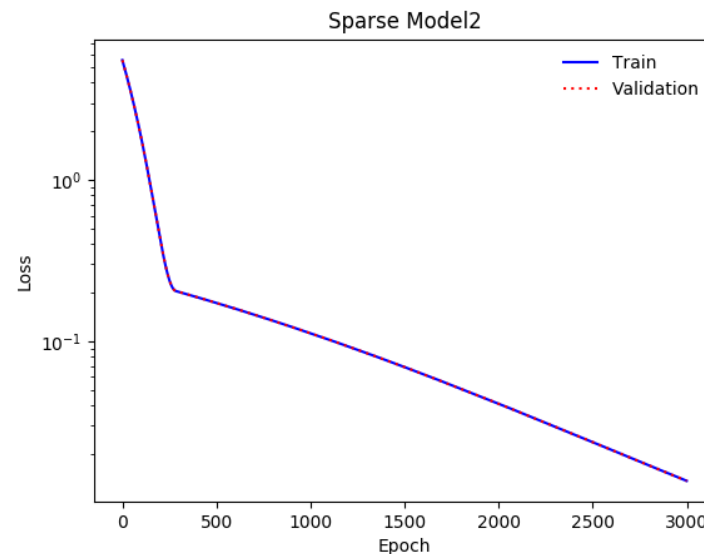
Set  $\lambda = 1e - 4$

# Sparse Training

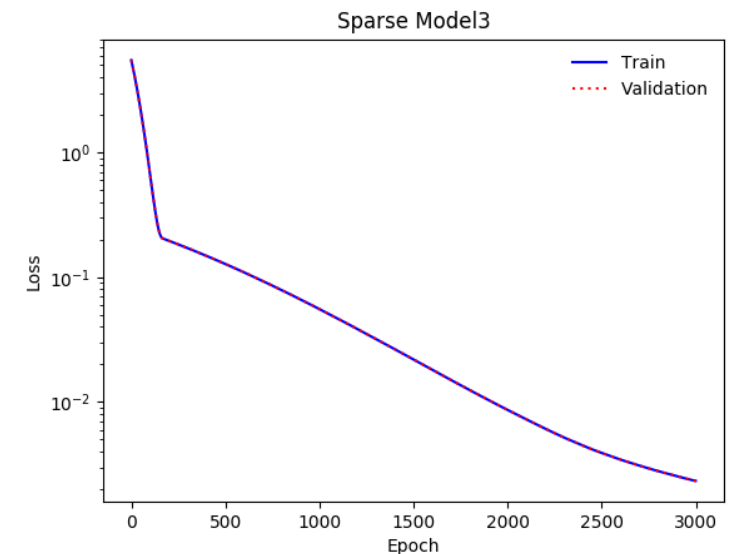
Fraction of data feeding 20%



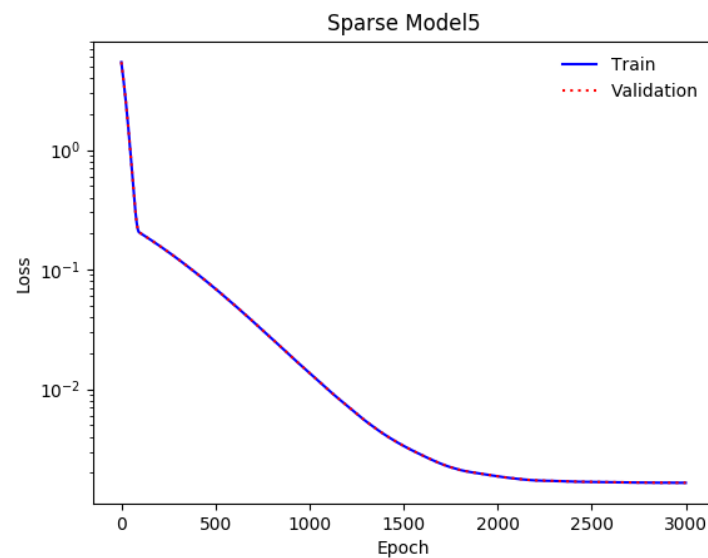
40%



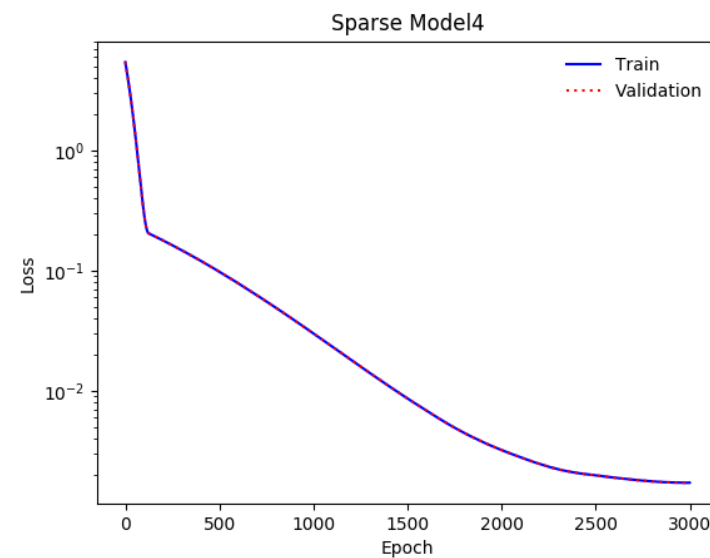
60%



80%



100%



Why it's almost same curve ??

( Not exactly equal since I already take a look )

# Contractive Model

- Unsupervised
- Similar to Vanilla Autoencoder
- Tweak by **Jacobi Matrix ( Prevent variation in data)**

$$\mathcal{L} = \frac{1}{N} \sum_i^N |x_i - \tilde{x}_i|^2 + \boxed{\lambda ||J_h(x)||^2}$$

$$\text{Set } \lambda = 1e - 4$$

$$\text{Where } ||J_h(x)||^2 \equiv \sum_{ij} \left( \frac{\partial h_j}{\partial x_i} \right)^2$$

# Contractive Model

- Case PReLU activ. fn.

$$||J_h(x)||^2 \equiv \sum_j [\alpha_j H(-(w_{ji}x^i + b_j)) + H(w_{ji}x^i + b_j)] \sum_i (w_{ji})^2$$

- Case Sigmoid activ. fn.

$$||J_h(x)||^2 \equiv \sum_j [h_j * (I - h_j)] \sum_i (w_{ji})^2$$

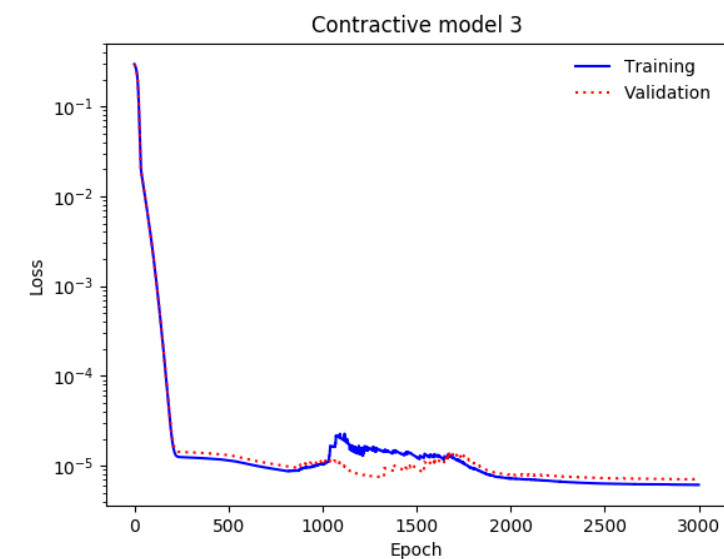
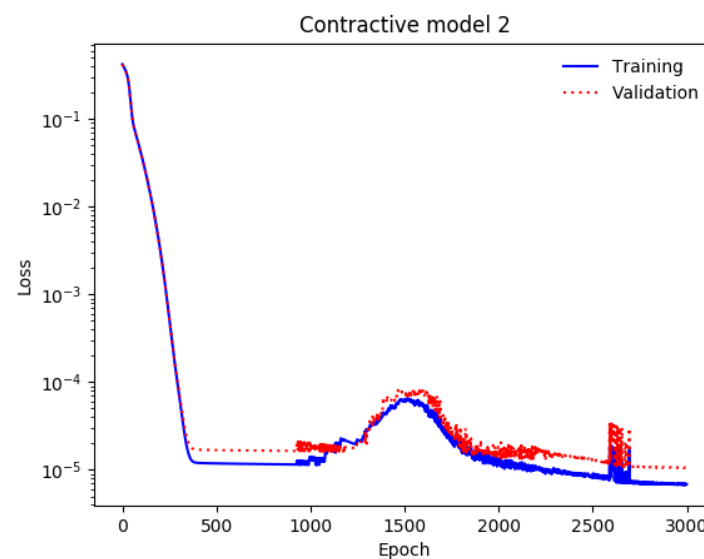
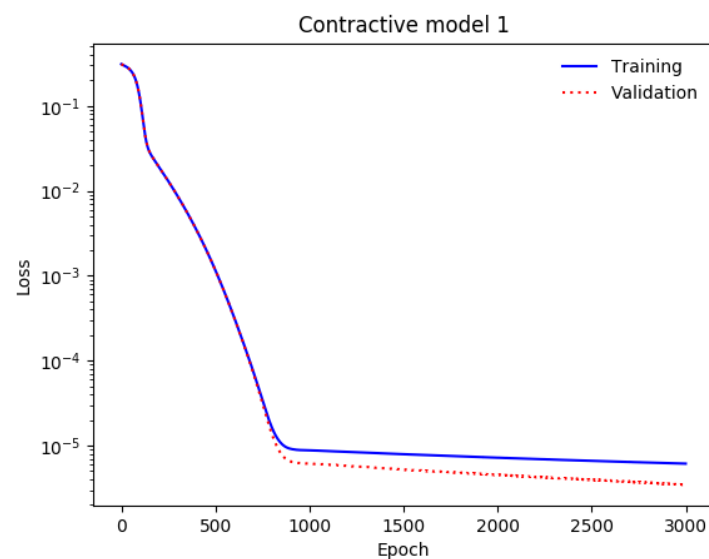


# Contractive Training

Fraction of data feeding 20%

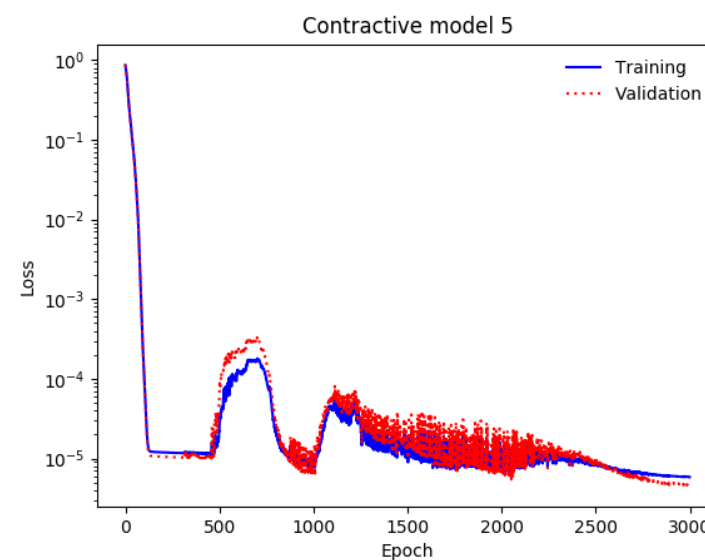
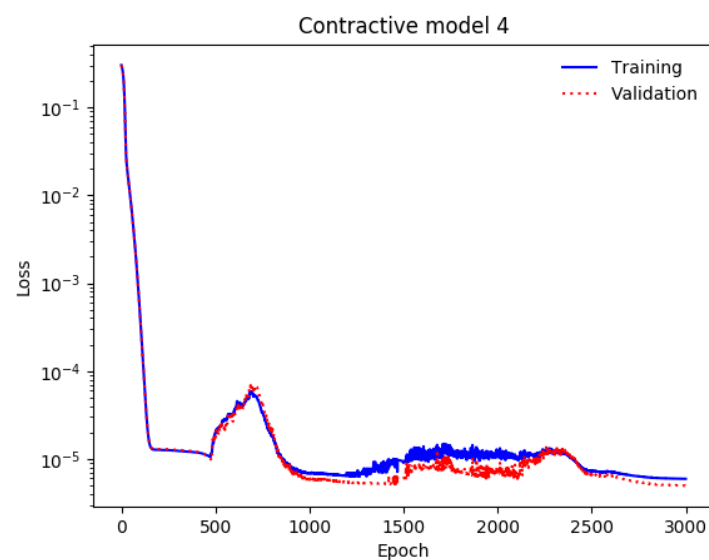
40%

60%



80%

100%



- First graph: validation even work better than training
- Fluctuation ?

# Future work

- Adjust model
- More on Autoencoder
  - Variational
- Model evaluation with more robust technique (ROC, AUC)