# Modeling Happiness Using One-Class Autoencoder

Yun-Gyung Cheong, Young-Rok Song, Byung-Chull Bae
Dept of Software, Sungkyunkwan University, South Korea
Dept. of Game, Hongik University, South Korea
aimecca@skku.edu, id4thoma@skku.edu, byuc@hongik.ac.kr

Our Approach

- Text Preprocessing: tokenization, vectorization, removal of stop words
- Machine Learning based Classification (SVM, KNN, Xgboost): 0.91 for social, 0.73-0.75 for agency class prediction
- **Labeling the unlabeled dataset: Autoencoder**

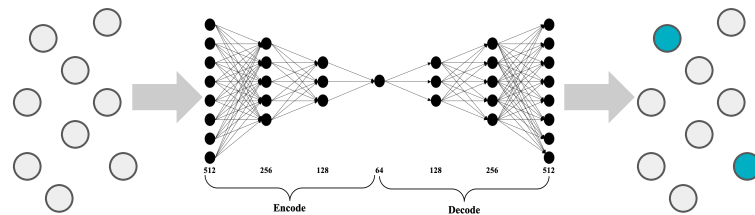One-hot encoding of a moment → [512 256 128 64 128 256 512 | Encode | Decode] → Recovered output
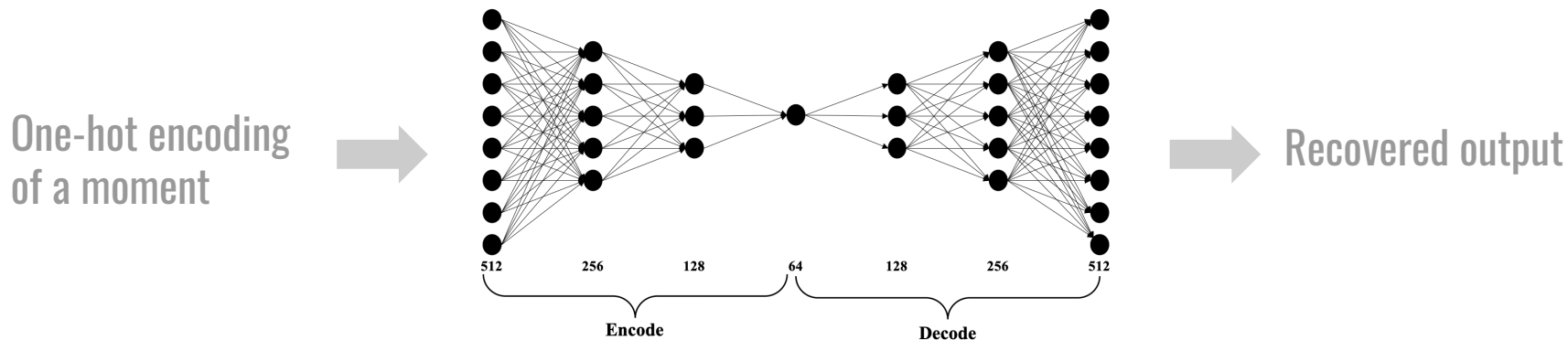
Train with one class only

Any inputs that display small difference between the input and output belongs to the class with which the model is trained

One AE for modeling one class for labeling unlabeled set
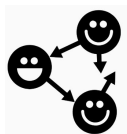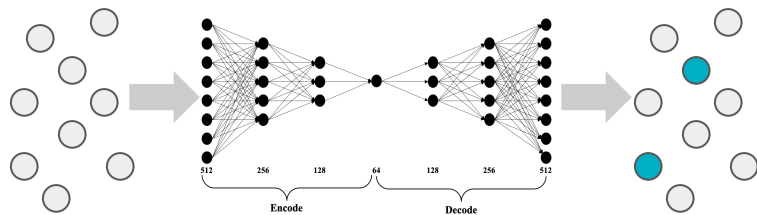
social

non-social

- Build 4 autoencoder models trained with social, non-social, agency, and non-agency
- For any moment, compute the difference between the input and the recovered output
- If the output is similar to the input, we determine that the moment belongs to the class the model is trained with
- As a result, only less than 3% of the unlabeled data were labeled (precision 0.9 for social, 1.0 for non-social)