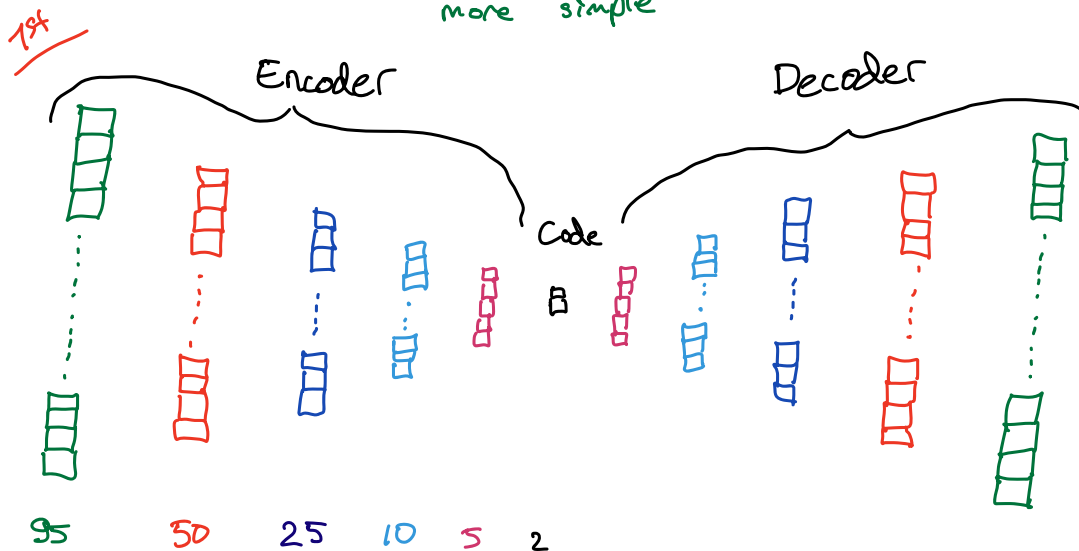


Risk Assessment Process

- ① Dataset 53503×30 Matrix $\xrightarrow{\text{Preprocessing}}$ 53503×35 Matrix
- Encoding for Categorical Data
- All columns are normalized to $[0,1]$

- ② Custom DEC Model ; \rightarrow performs better than DEC for this problem even though it is more simple



- 2nd
- Standard K-Means \rightarrow Cluster Centers \rightarrow Initial Clusters For The Next Step

- 3rd
- Clustering with KL Divergence
- includes
- Soft Assignment w/ Student's t -distribution
 - Target Distribution
 - KL (Kullback-Leibler) Divergence Minimization
 - Optimization (Iterations)

- 4th
- Hyperparameter Optimization with Model Metrics
- \rightarrow t -SNE, silhouette score, Davies-Bouldin Index, Calinski-Harabasz Index

- ③ Post-Processing \rightarrow 40% - 60% Range Values in Gaussian
- \rightarrow Mean of these data points for each cluster
- \rightarrow Cluster Risk Profiles

