

Course Description Text

Deep Embedding Models

BAAI/bge-small-en-v1.5

avsolatorio/GIST-Embedding-v0

nvidia/NV-Embed-v2

Dimension Reducers

PCA

t-SNE

PaCMAP

4/7 dimensions &
70%/80%/90% explained variance

Cosine Similarity

Elementwise
Difference

Composite Distance Vectors

$$\Delta_c = \left(a_1 - b_1, \dots, a_k - b_k, \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} \right)$$

Traditional ML Classification

Linear Regression, Ridge Regression, Lasso,
KNN, SVM, Random Forest, LDA, QDA