



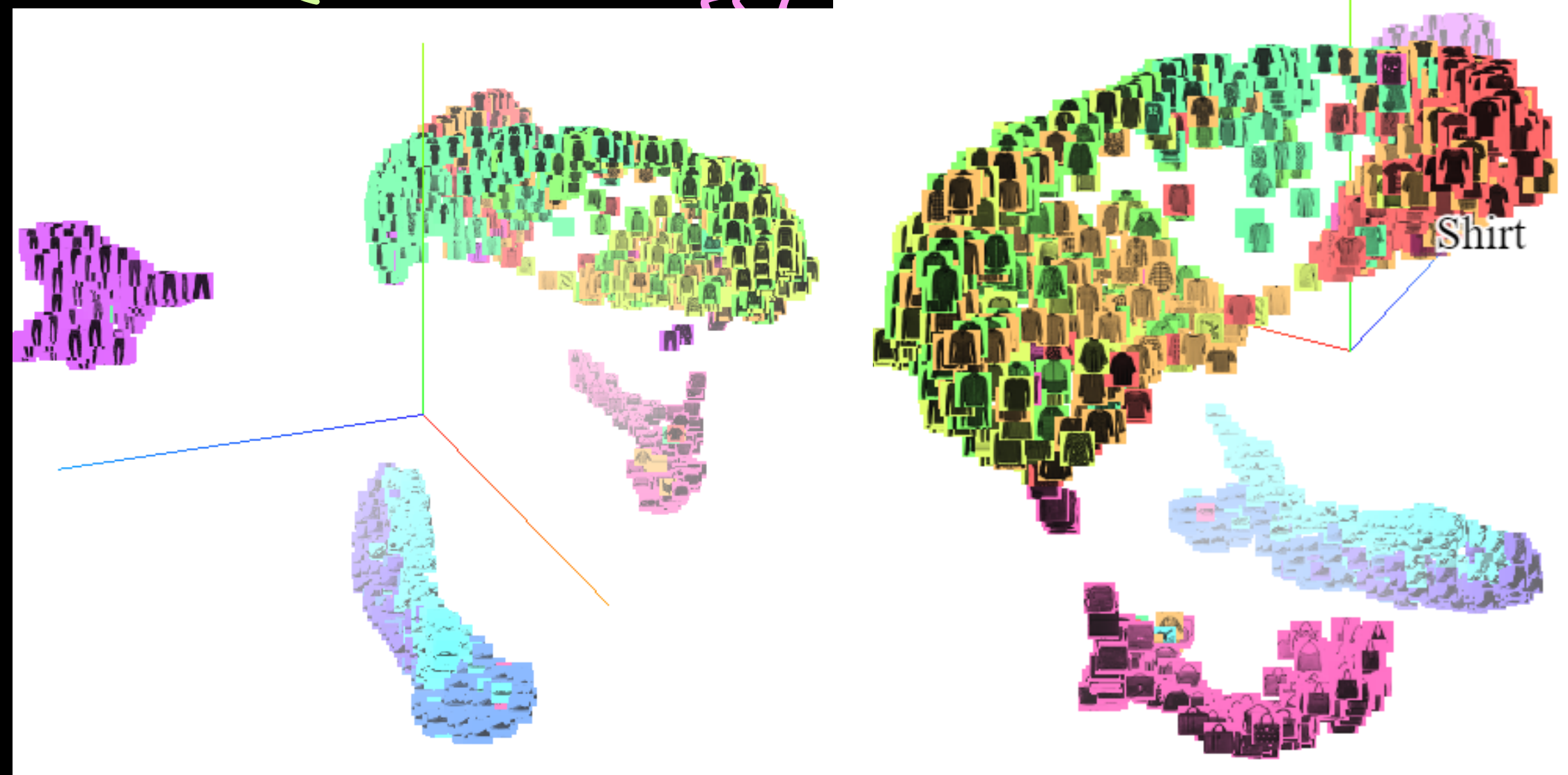
UMAP

Machile
learning

Team 6

What is UMAP

UMAP, or Uniform Manifold Approximation and Projection, is a dimensionality reduction technique used in machine learning. It helps us visualize high-dimensional data by preserving its underlying structure.



T-shirt/top

Shirt

Pullover

Coat

Dress

Sandal

Sneaker

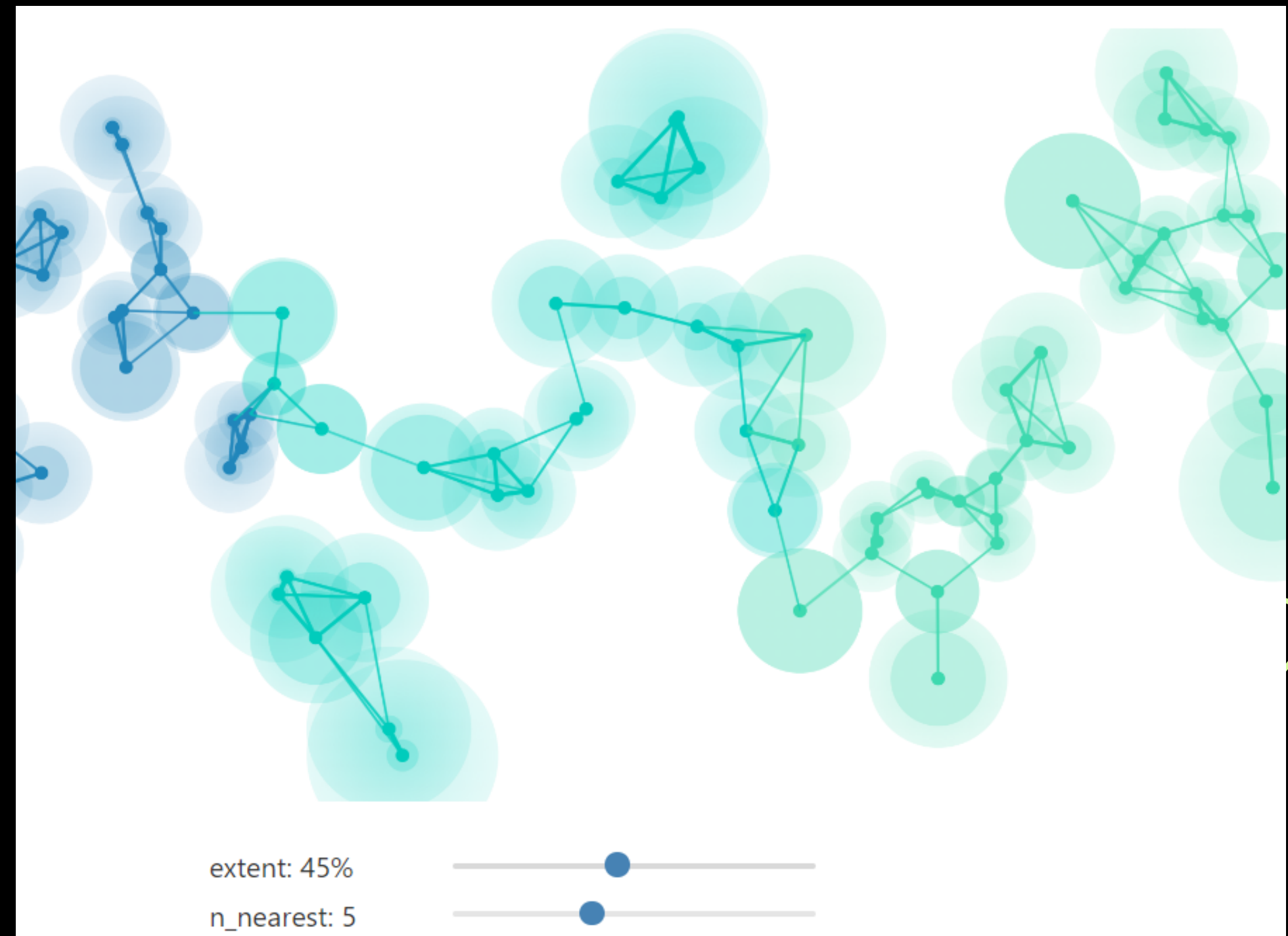
Ankle boot

Trouser

Bag

Basic principles of UMAP

UMAP operates based on fundamental principles such as neighborhoods and neighborhood graphs. It minimizes the graph's energy to find an optimal representation.



Steps to Implement UMAP



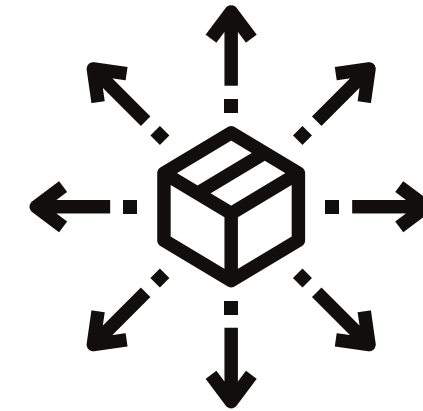
1

Creating a neighborhood graph



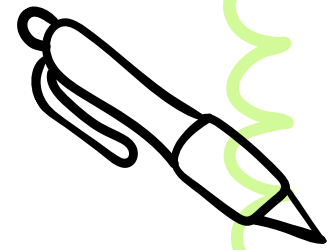
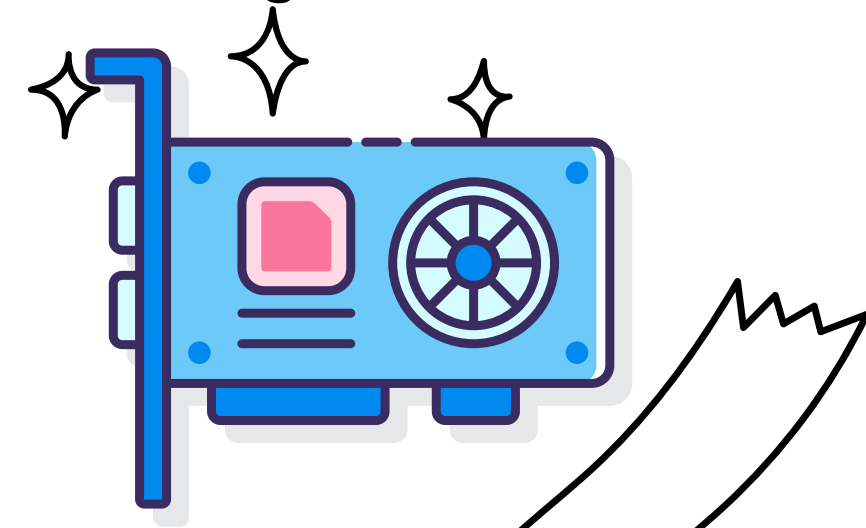
2

Optimizing point distribution



3

Transforming the data



Implementing UMAP involves three main steps

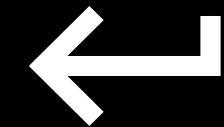




preservation of global structure

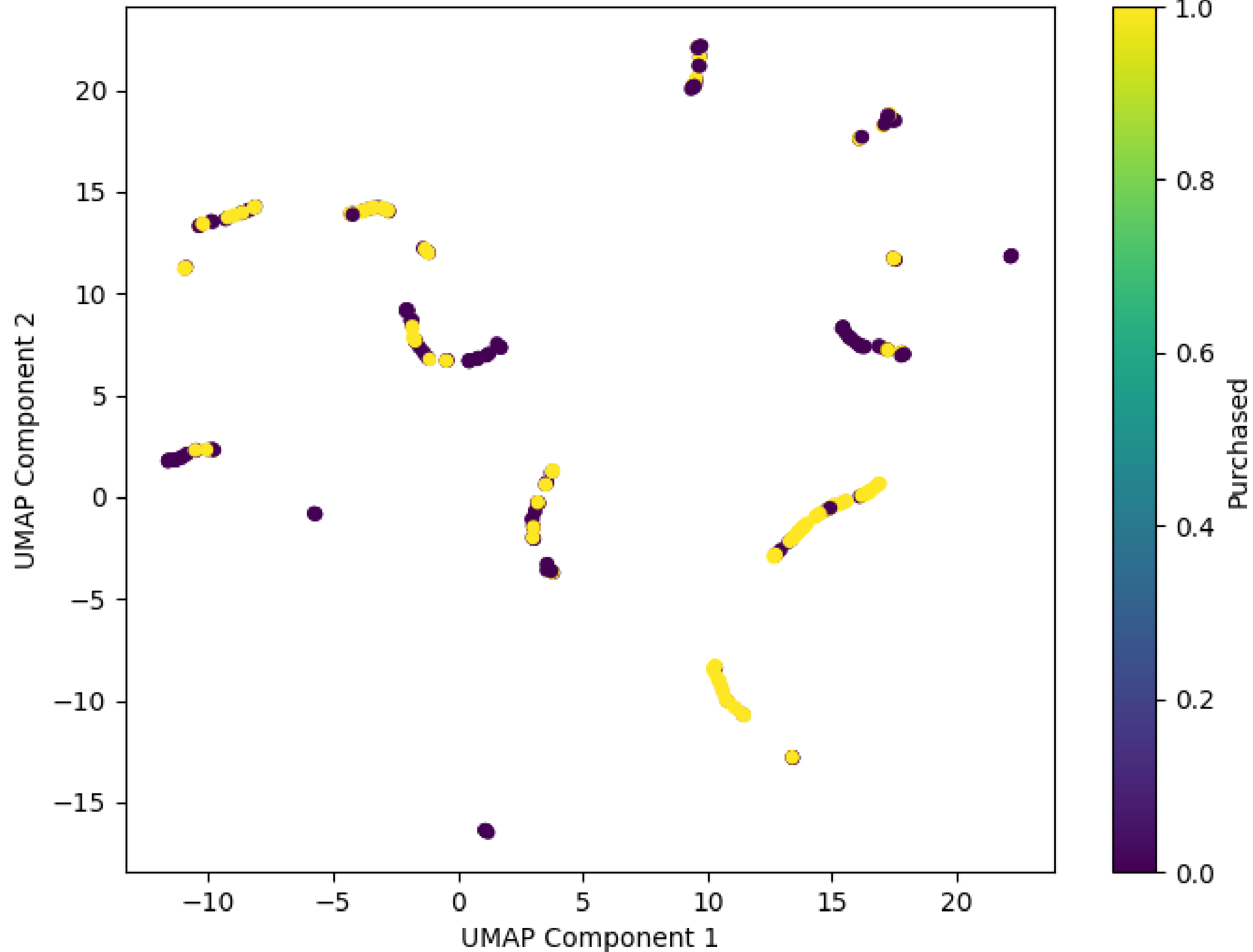
visualizing high-dimensional data preprocessing

efficient dimensionality reduction.



Advantages of UMAP

UMAP Projection of Social Network Ads Data



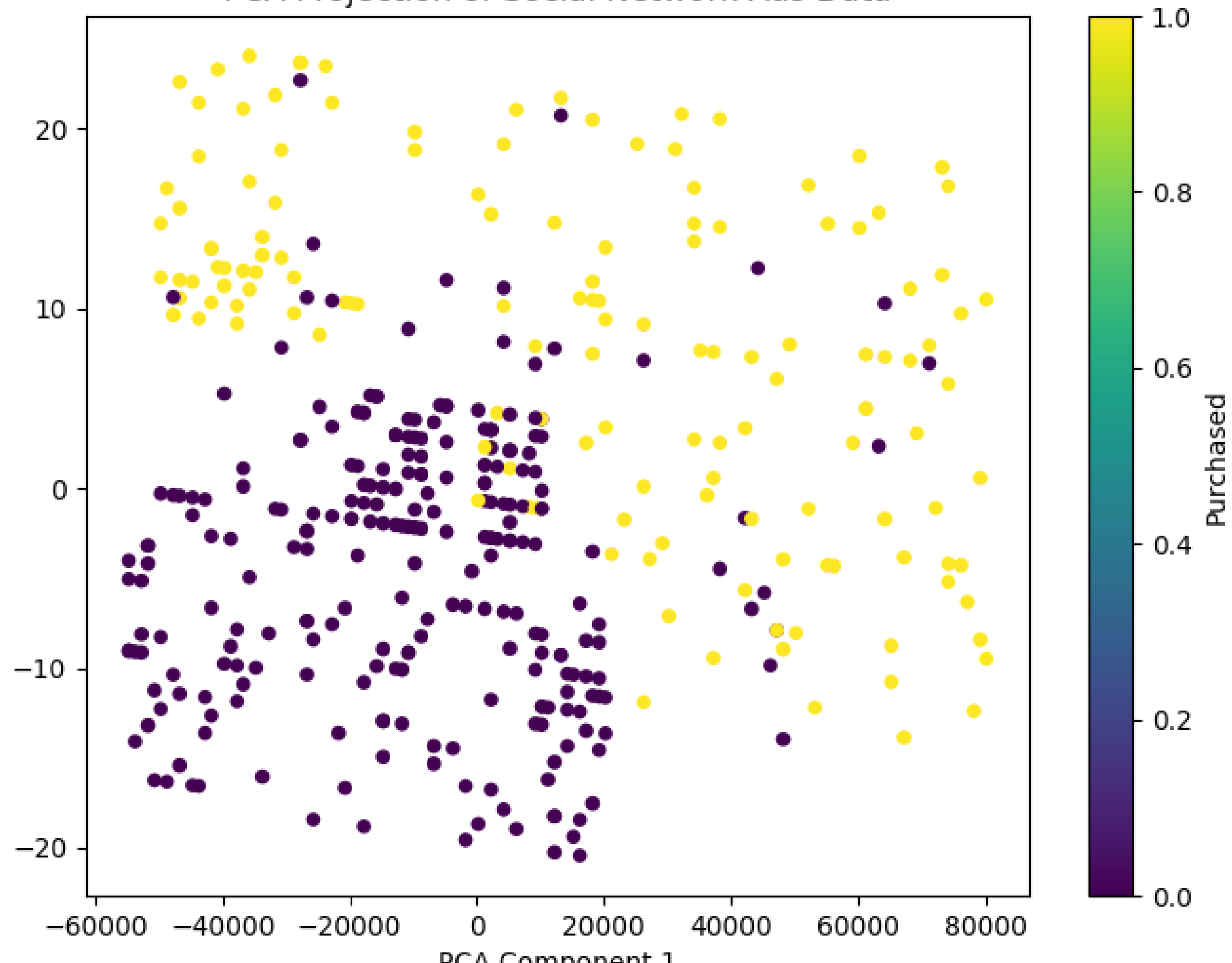
Output

Dimension 1

Dimension 2

"UMAP "1" is a dimension in the new space that captures a portion of the variability or structure in your data.

PCA Projection of Social Network Ads Data





Considerations and Limitations

1

parameter sensitivity

2

The challenge of interpreting the
reduced representation

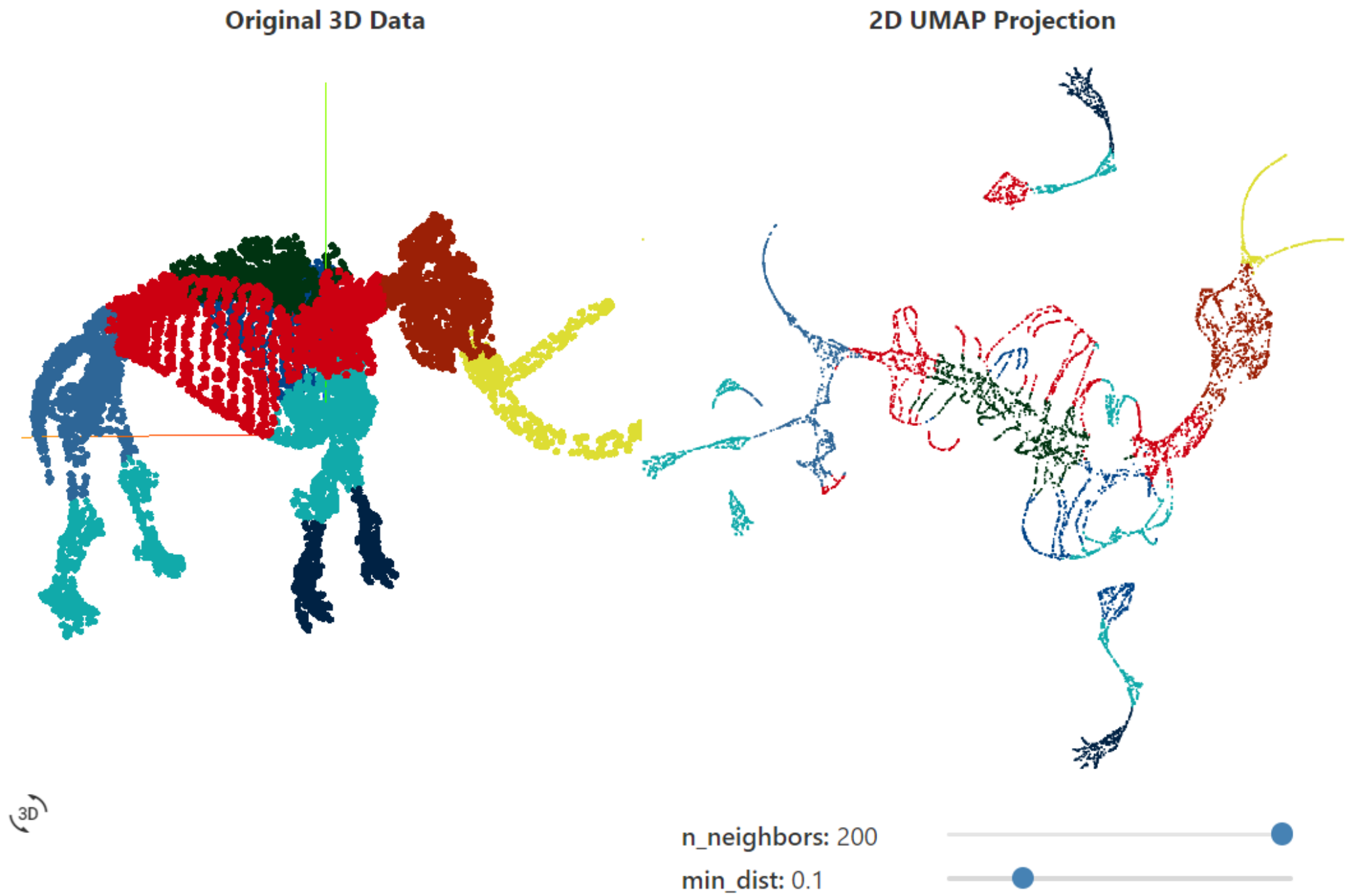
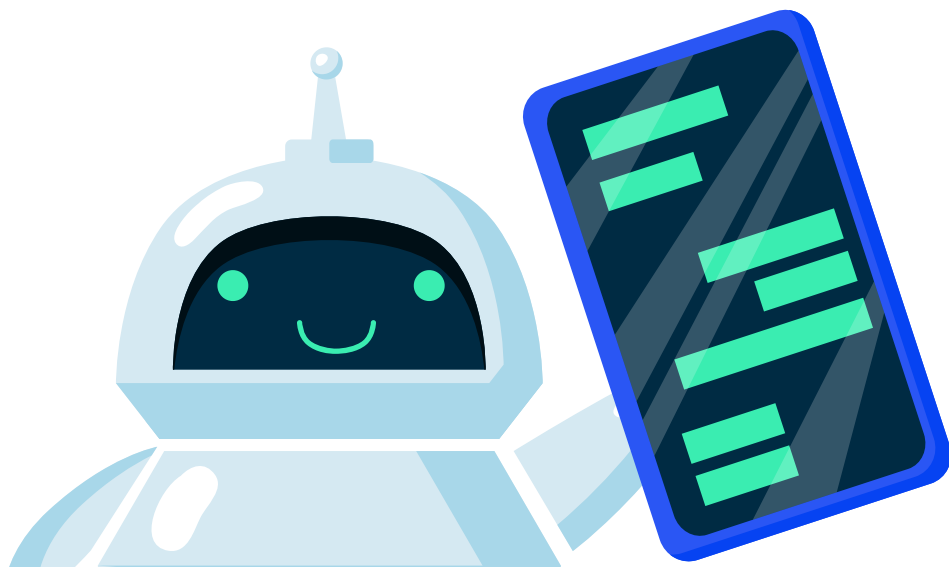
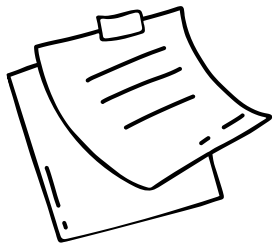
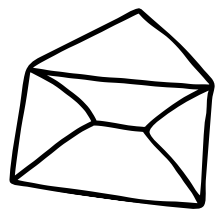
3

Parameter Tuning
Required

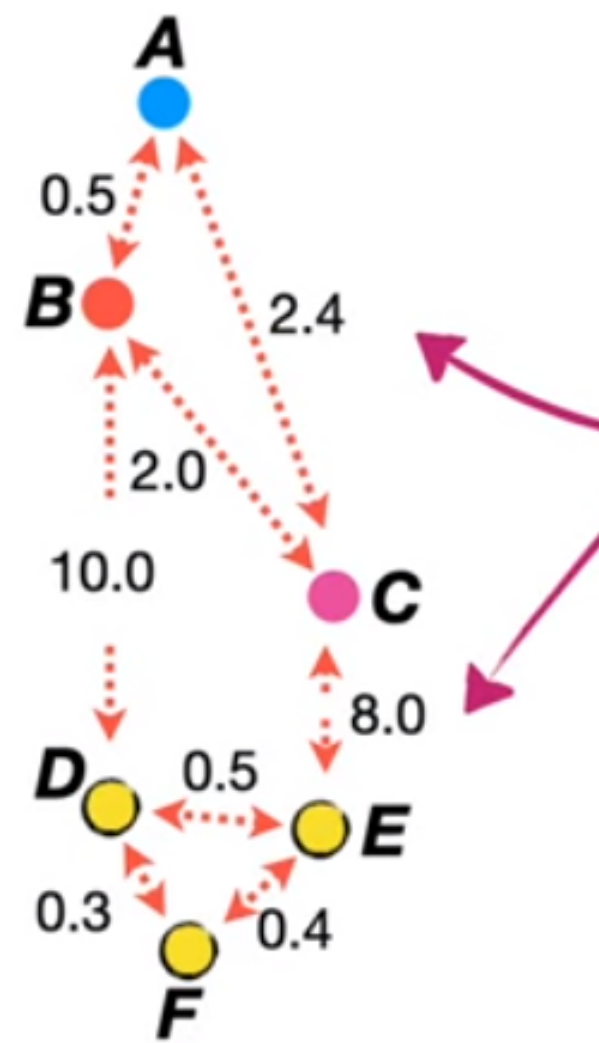
1



The visualization is an exploration of the impact of UMAP parameters on a 2D projection of 3D data.

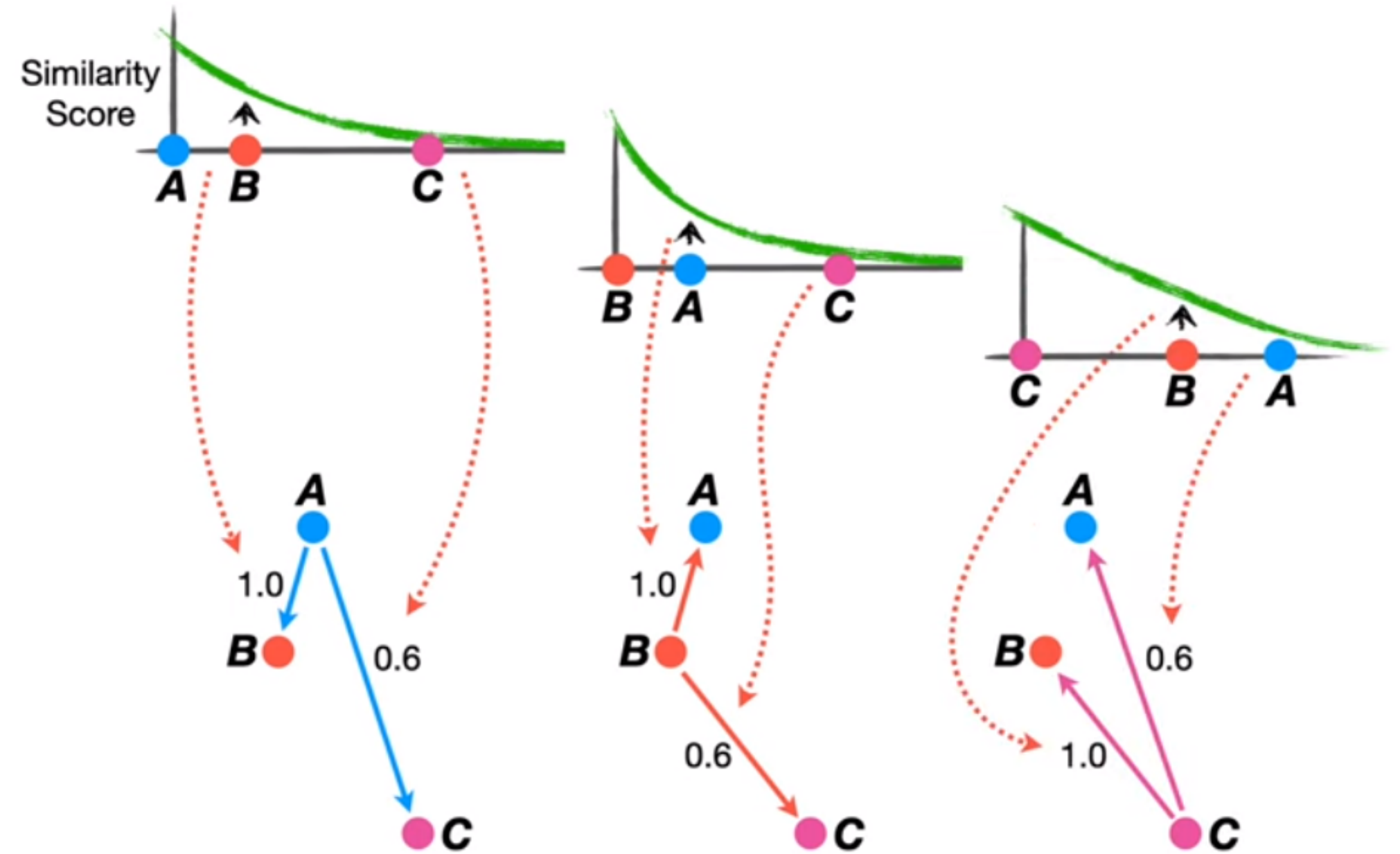
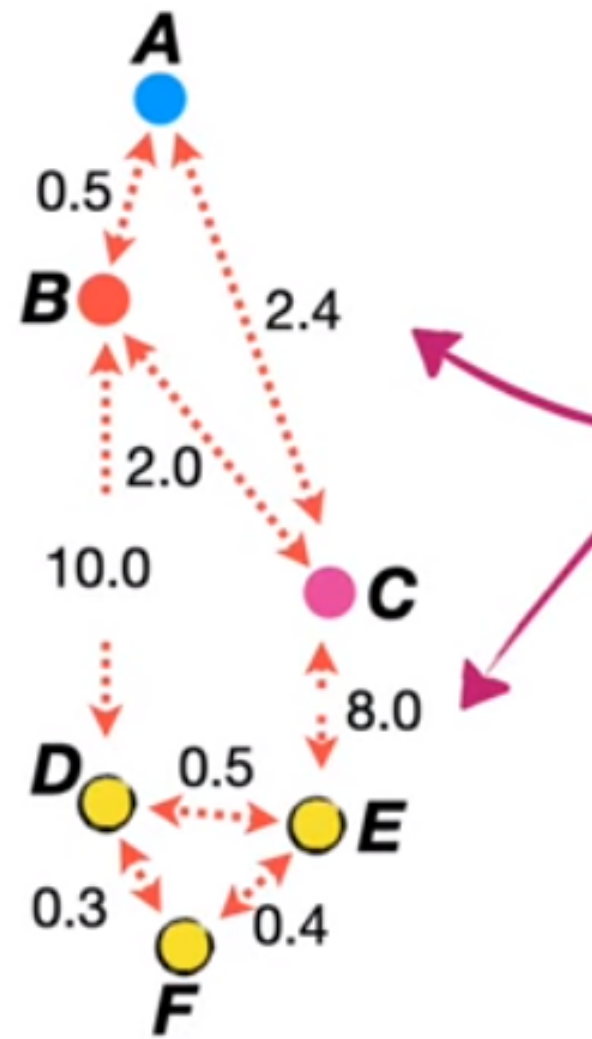


Raw Distances

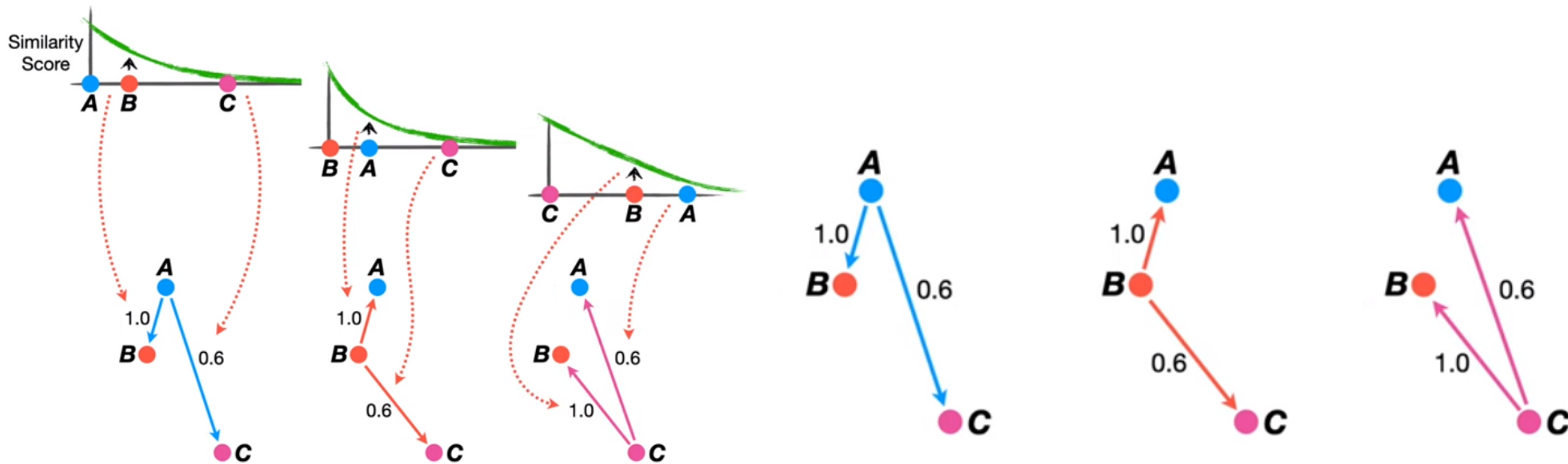


$$\text{Similarity Score} = e^{-(\text{raw dist.} - \text{dist. to nearest neighbor})/\sigma}$$

Raw Distances

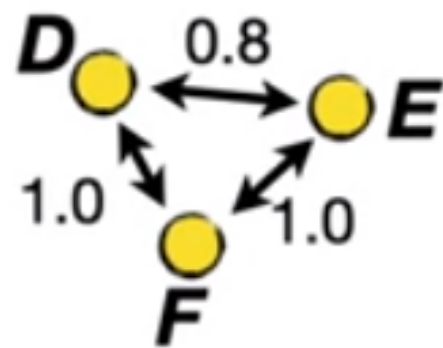
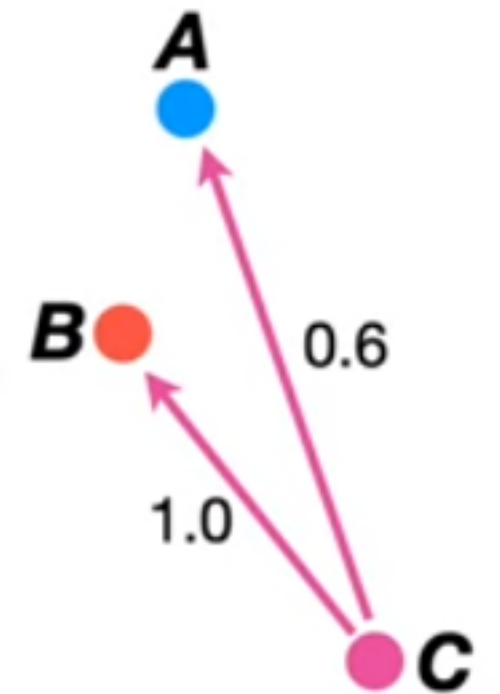
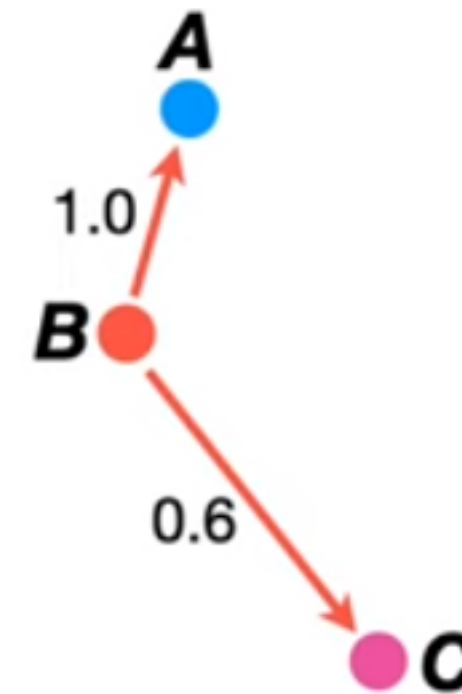
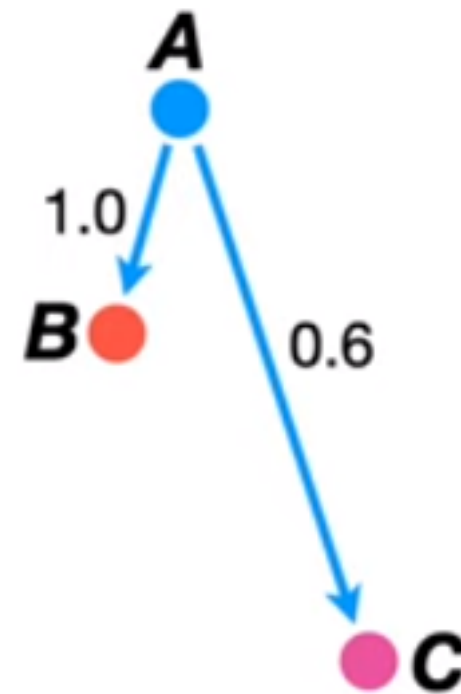
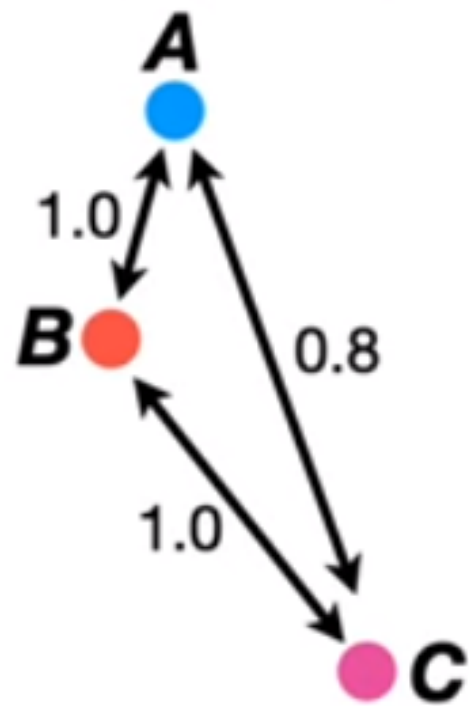


$$\text{Similarity Score} = e^{-(\text{raw dist.} - \text{dist. to nearest neighbor})/\sigma}$$



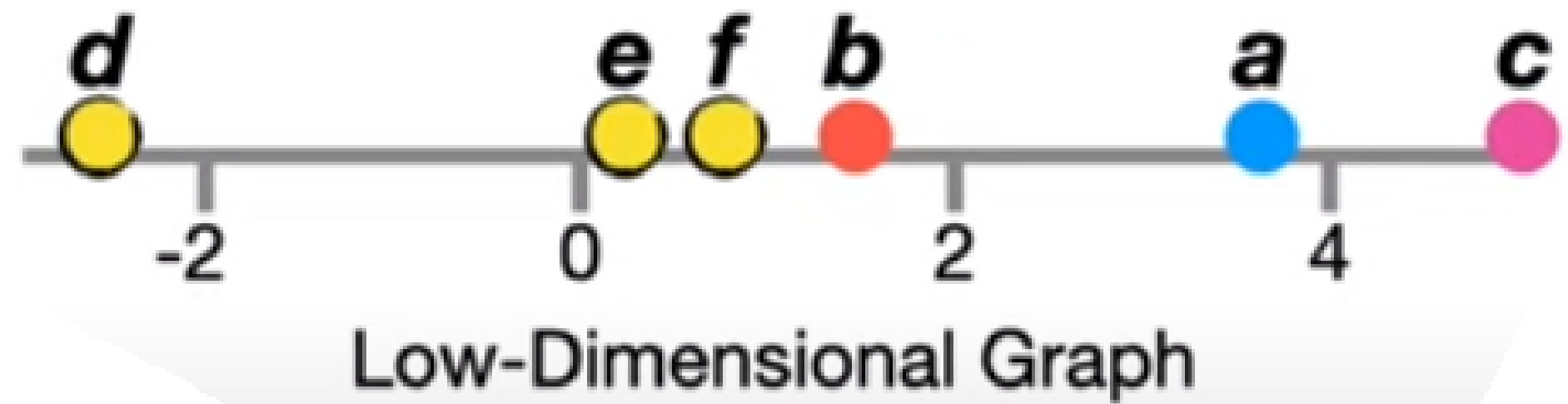
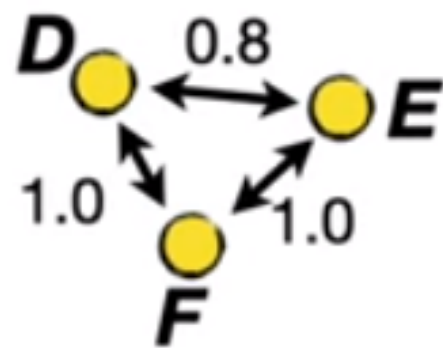
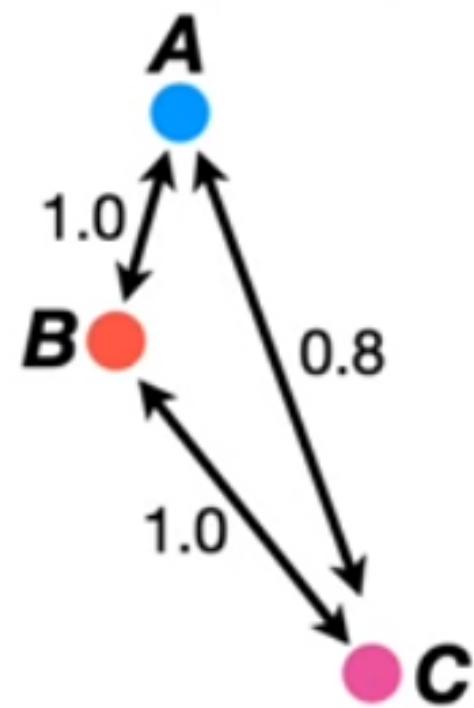
$$\text{Symmetrical Score} = (S_1 + S_2) - S_1 S_2$$

Symmetric
Similarity Scores



$$\text{Symmetrical Score} = (S_1 + S_2) - S_1 S_2$$

Symmetric
Similarity Scores



$$\text{Symmetrical Score} = (S_1 + S_2) - S_1 S_2$$

Symmetric
Similarity Scores

