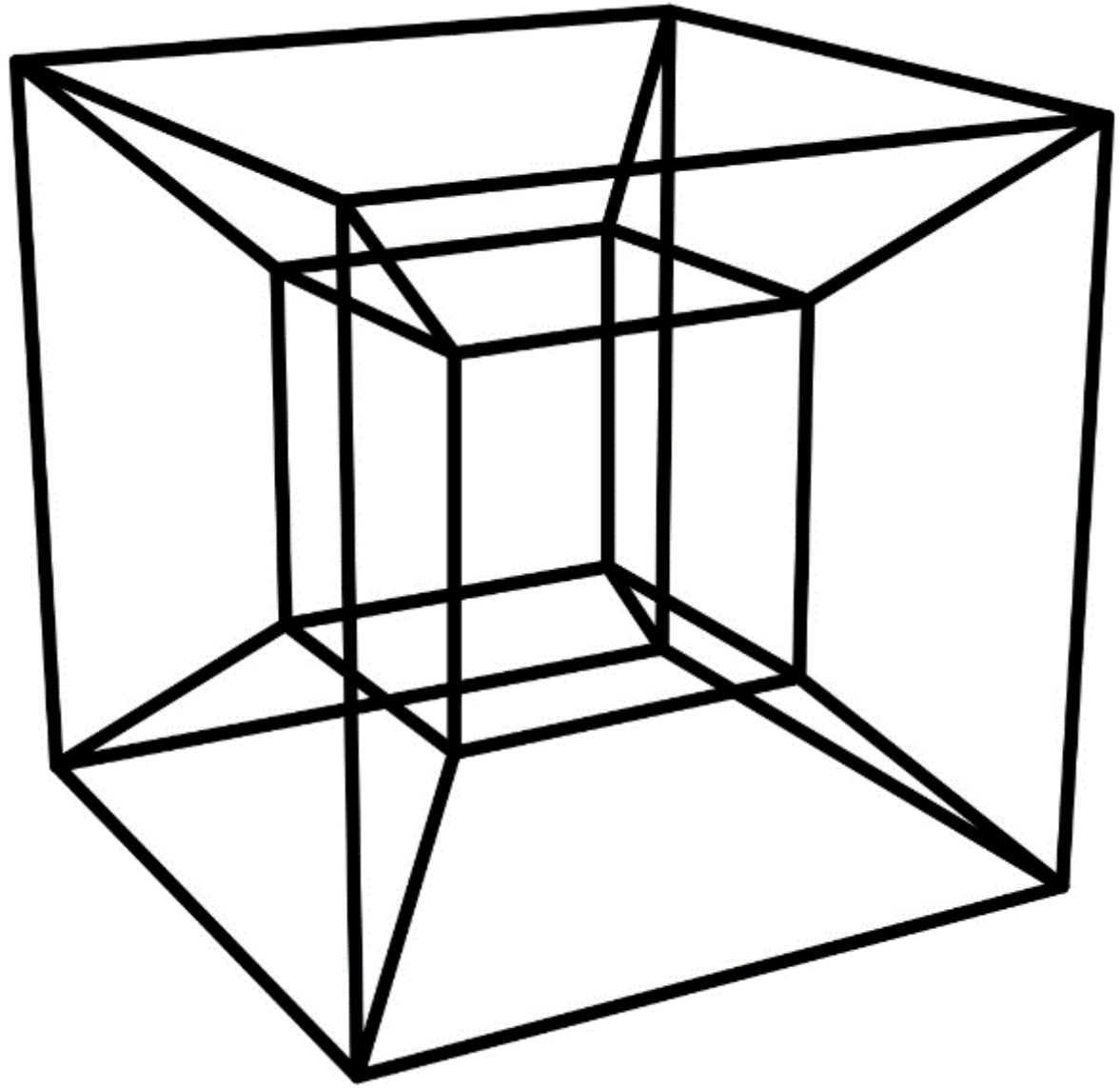
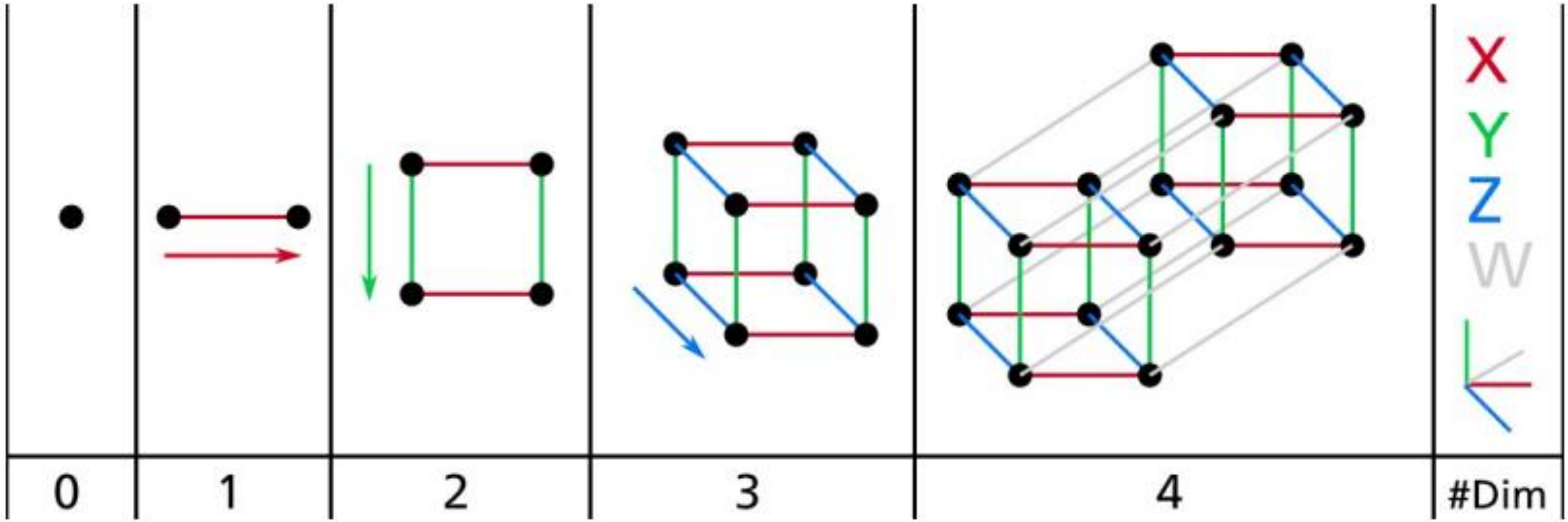
Hands-On Machine Learning with Scikit-Learn & TensorFlow

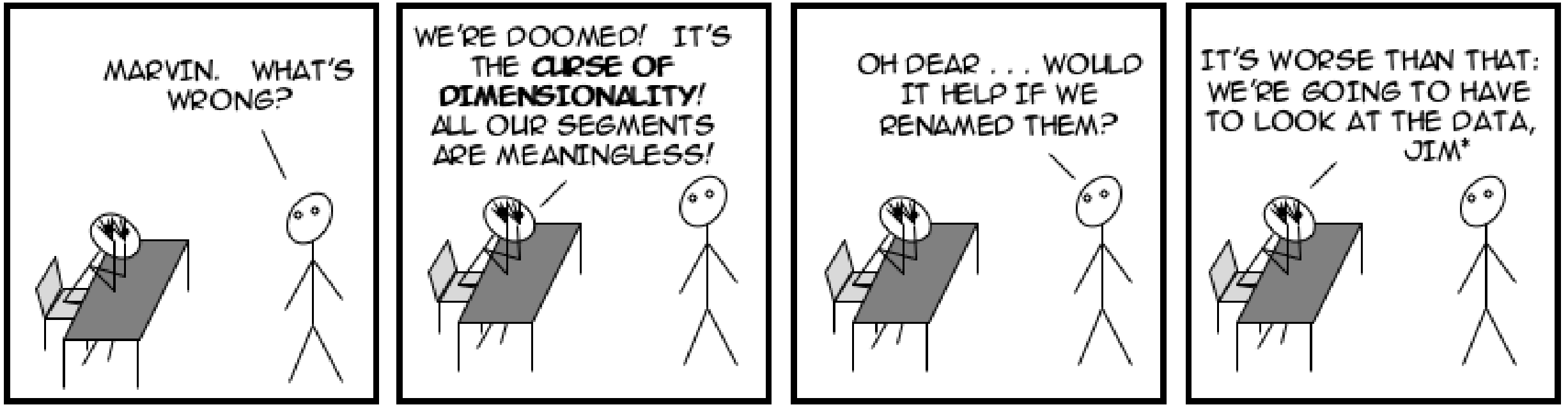
Ana Maria Sandoval Jimenez, Jannis Busch & Sabrina Steinert

Chapter 8.
Dimensionality Reduction




X
Y
Z
W
0
1
2
3
4
#Dim


Curse of
Dimensionality



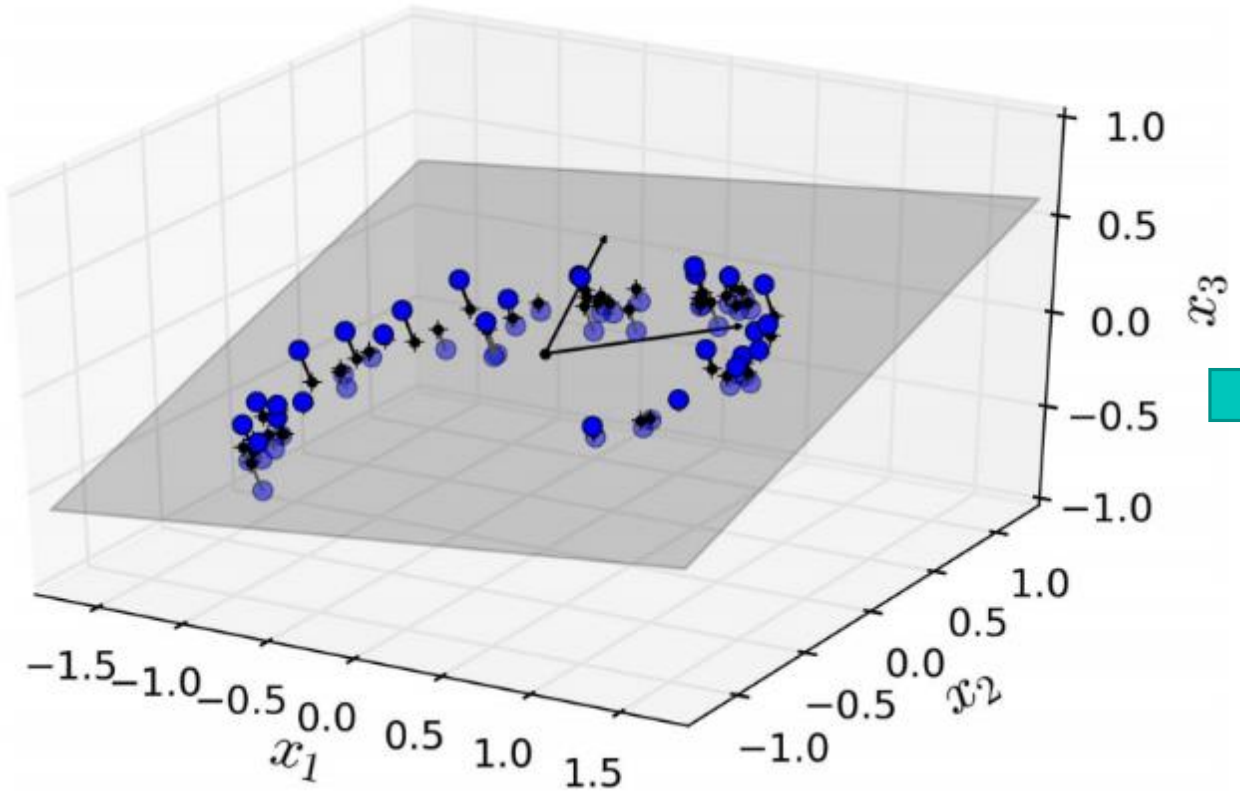
The problem caused by the exponential increase in volume associated with adding extra dimensions to a (mathematical) space.

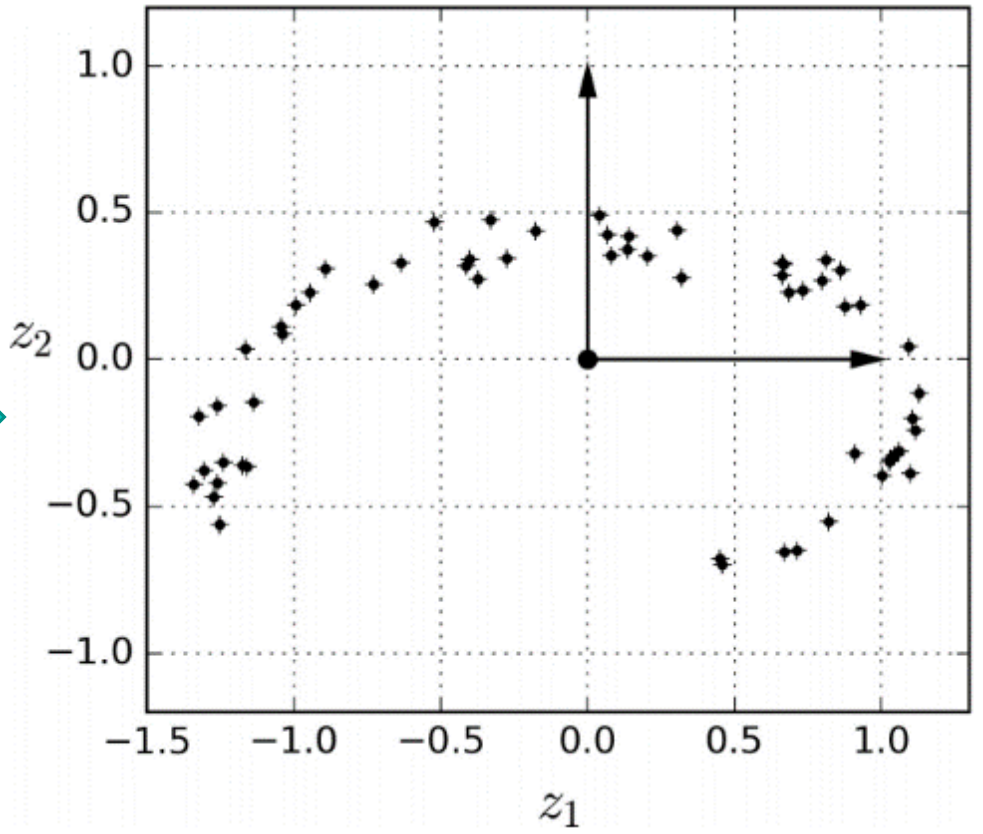
Dimensionality Reduction Approaches

Projection & Manifold Learning

PROJECTION

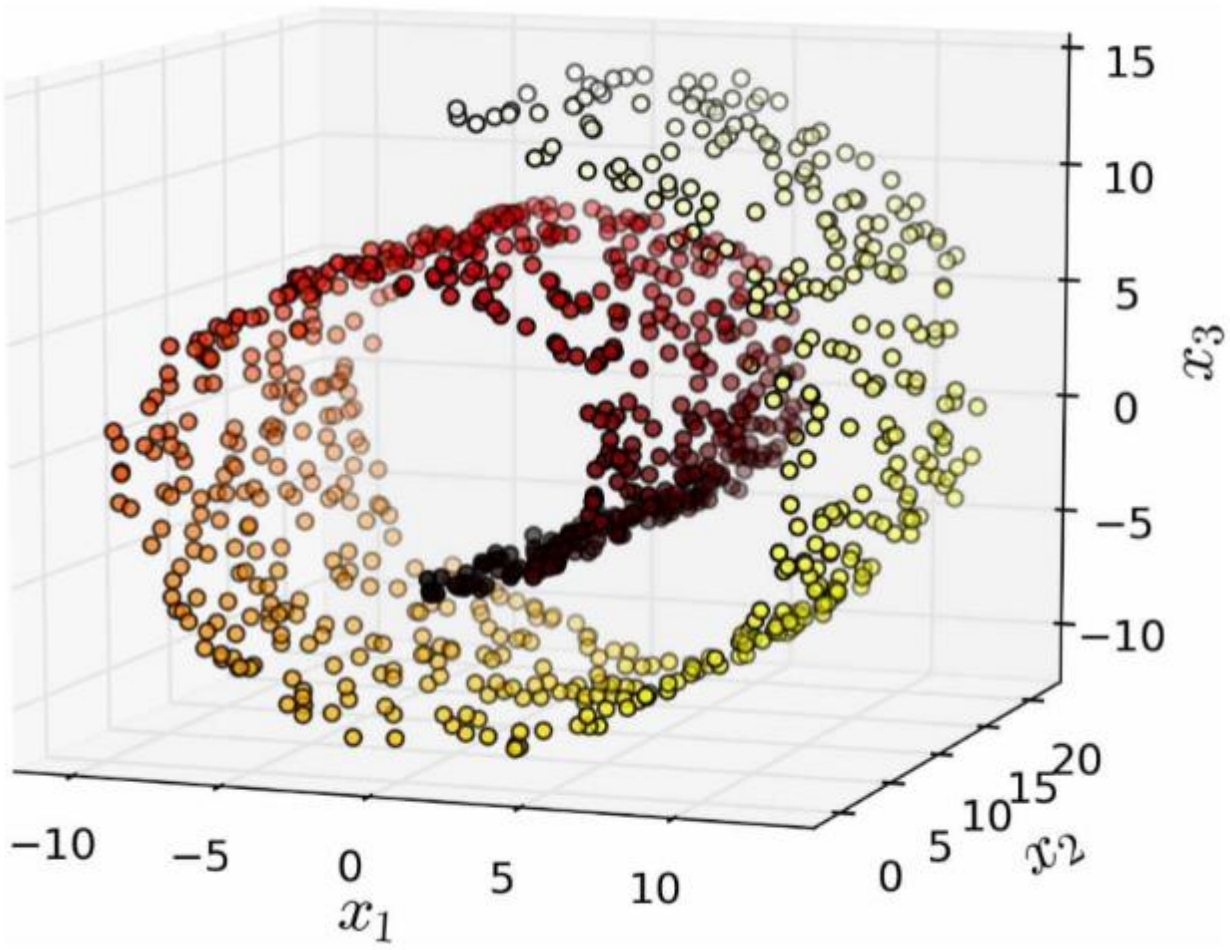
Reduce the dataset's dimensionality from 3D to 2D

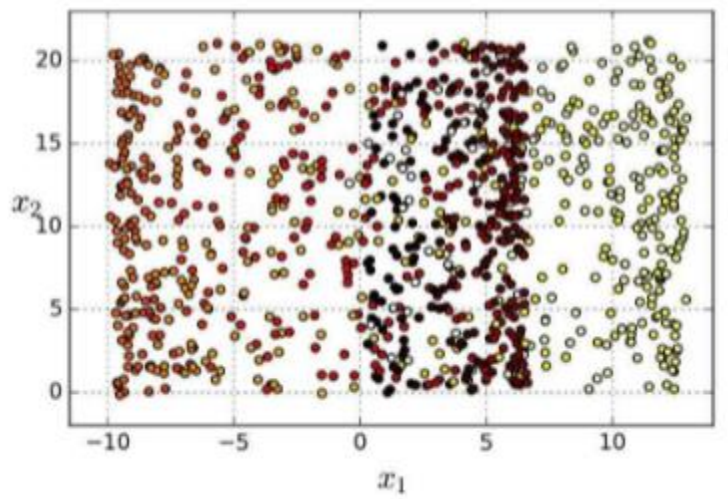


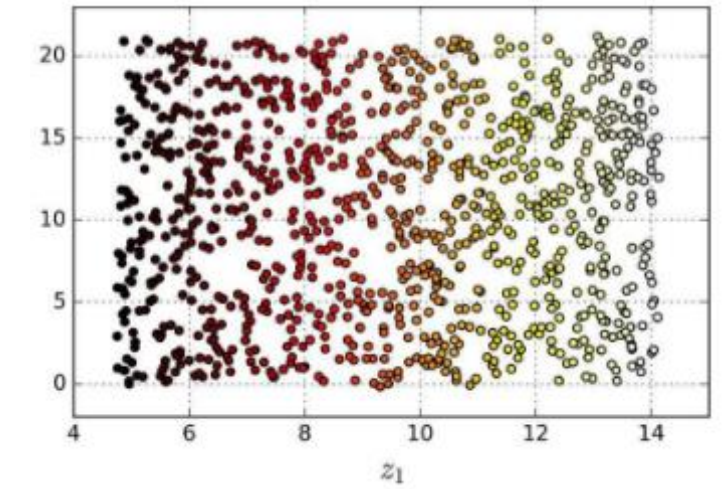


MANIFOLD

MANIFOLD LEARNING







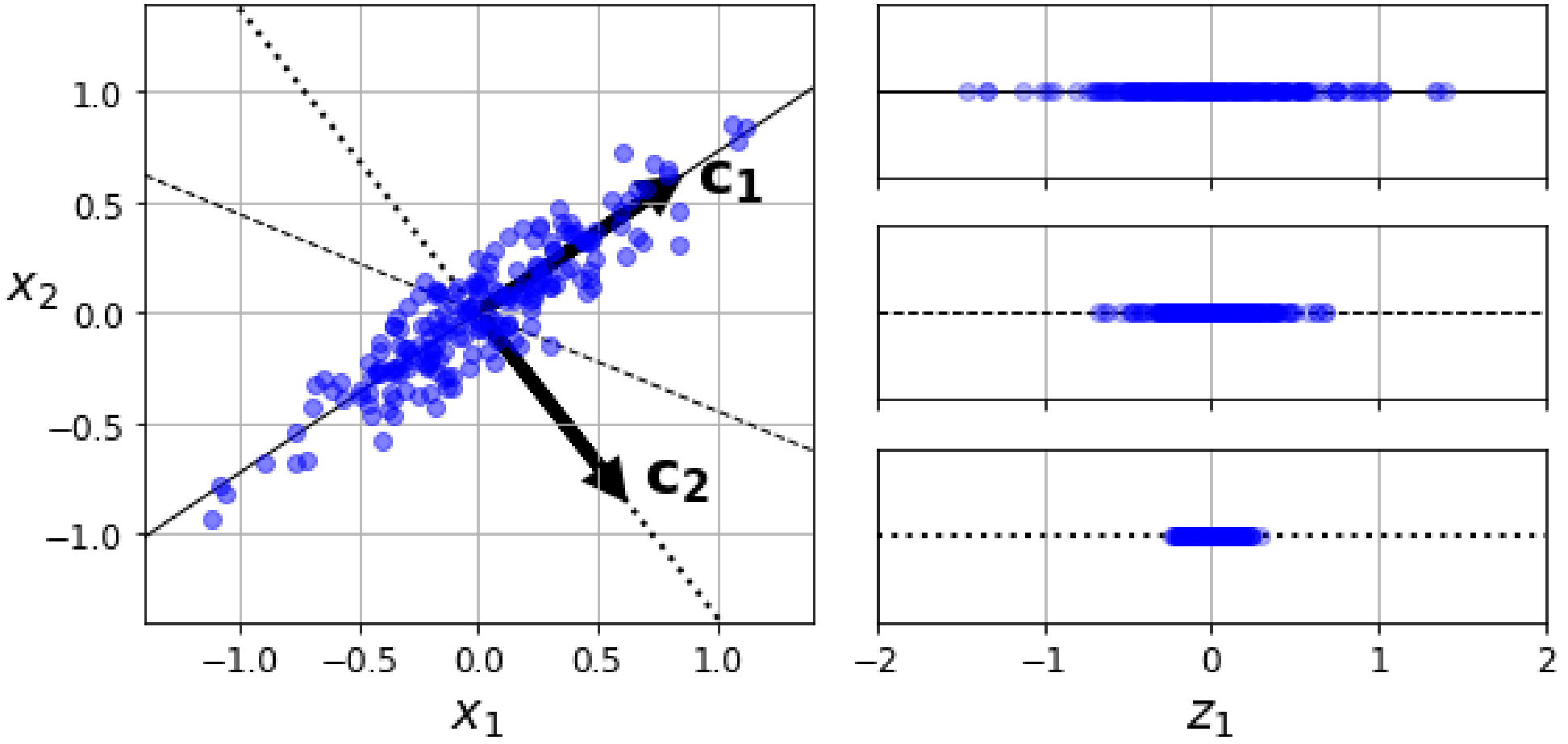
Dimensionality Reduction Technique

Projection

Principal **C**omponent **A**nalysis

PCA

PRESERVING THE VARIANCE



PCA
PRINCIPAL COMPONENTS

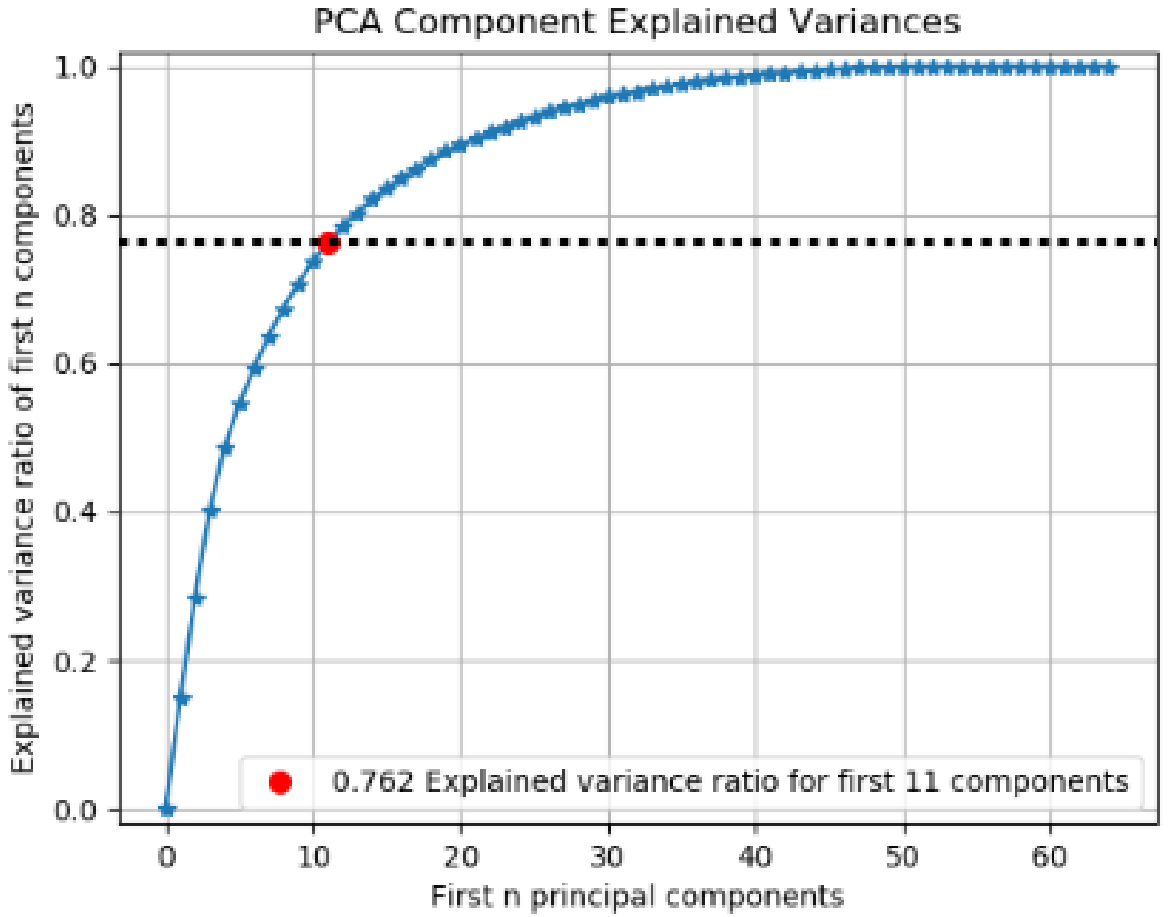
jupyter 8_DimensionalityReduction

Standard matrix factorization technique:

SINGULAR VALUE DECOMPOSITION (SVD)

CENTER THE DATA!!

PCA
EXPLAINED VARIANCE RATIO



PCA

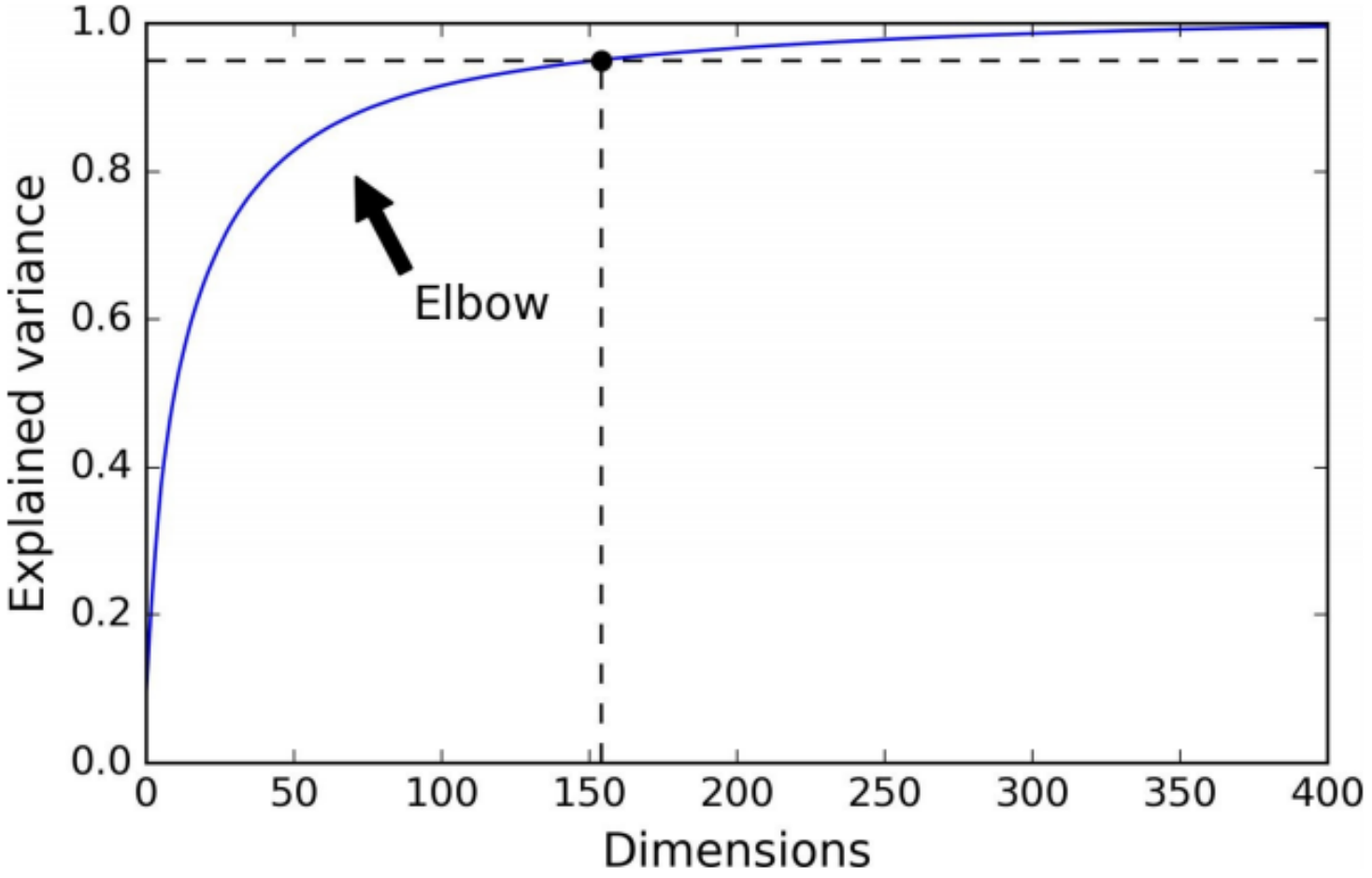
THE RIGHT NUMBER OF COMPONENTS

jupyter 8_DimensionalityReduction

Which ratio of variance do you wish to preserve?

PCA

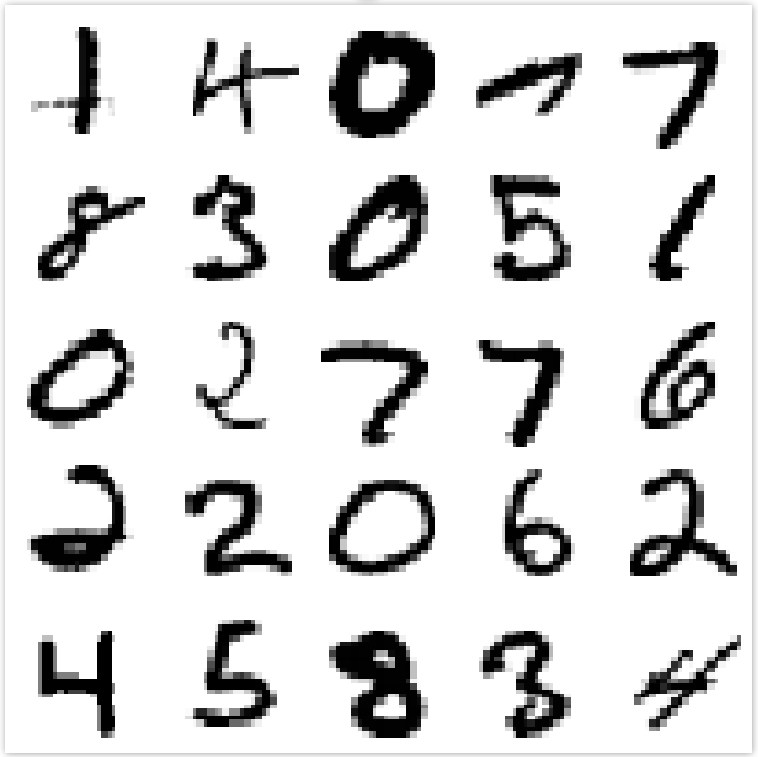
THE RIGHT NUMBER OF COMPONENTS



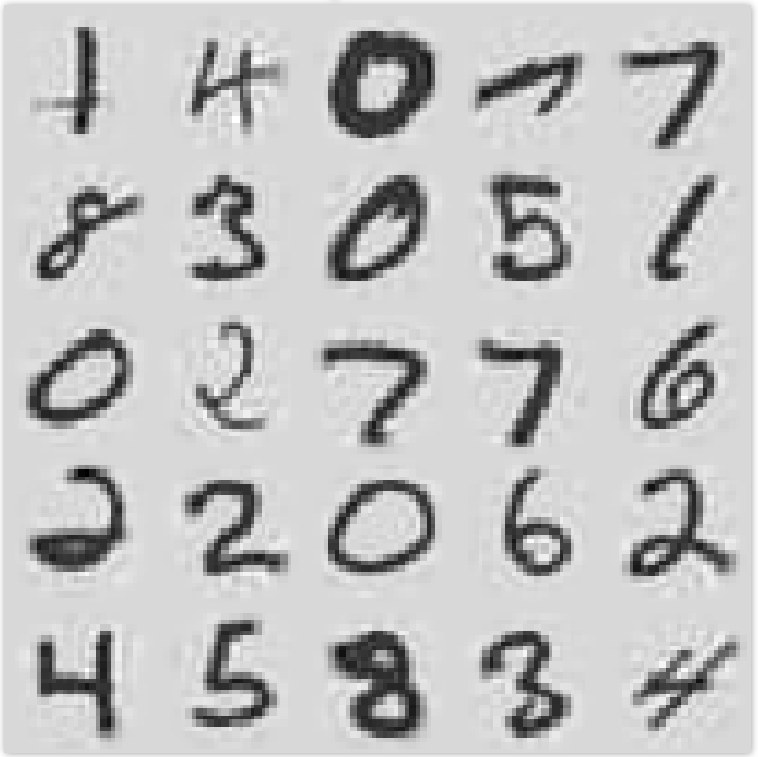
PCA
COMPRESSION

jupyter 8_DimensionalityReduction

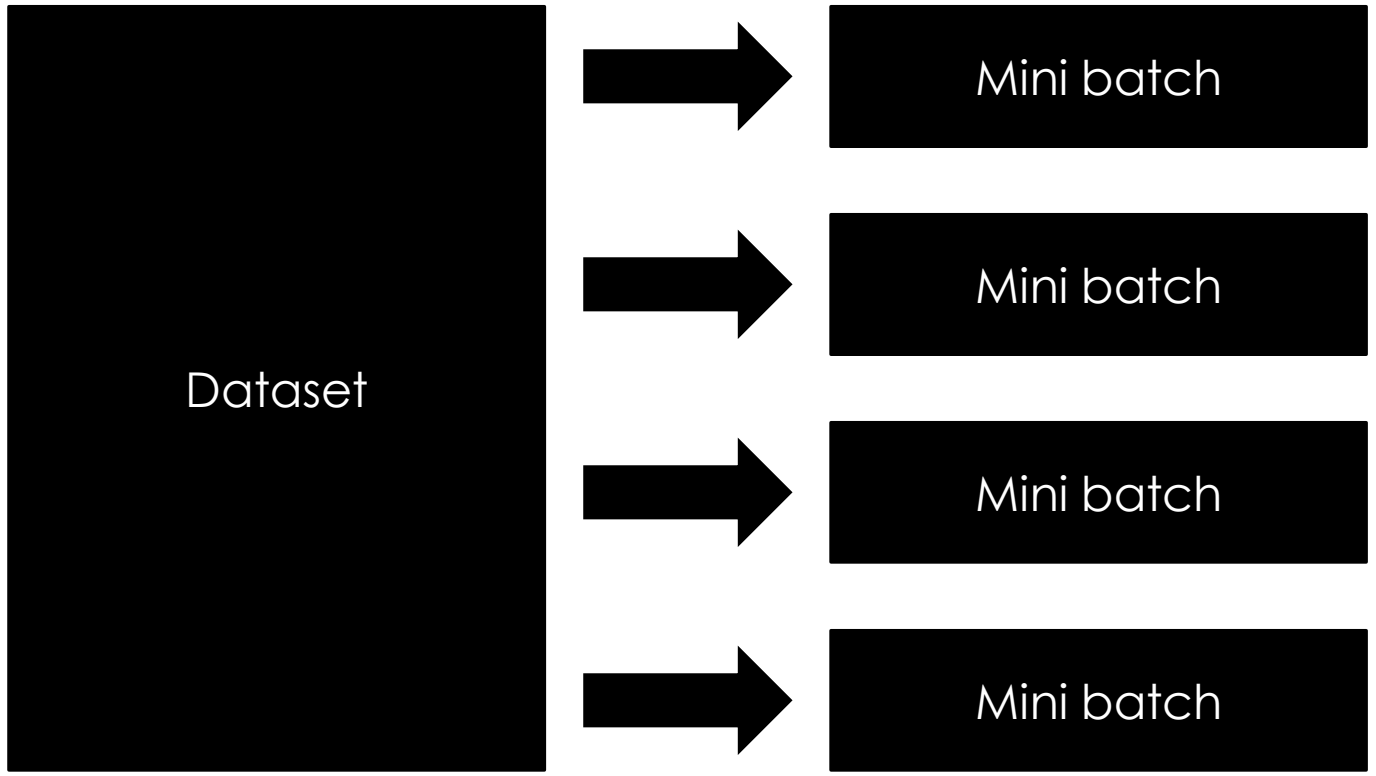
Original



Compressed



PCA
INCREMENTAL PCA

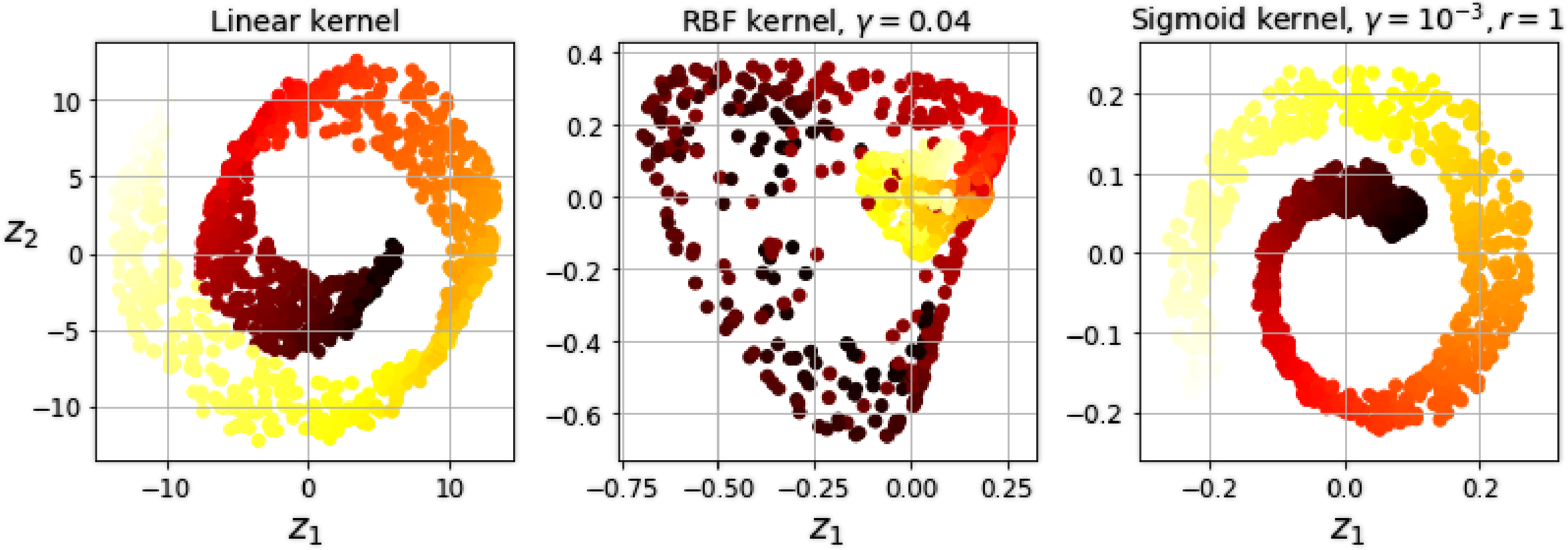


MEMORY-FRIENDLY

ONLINE

PCA
KERNEL PCA

Kernel Trick

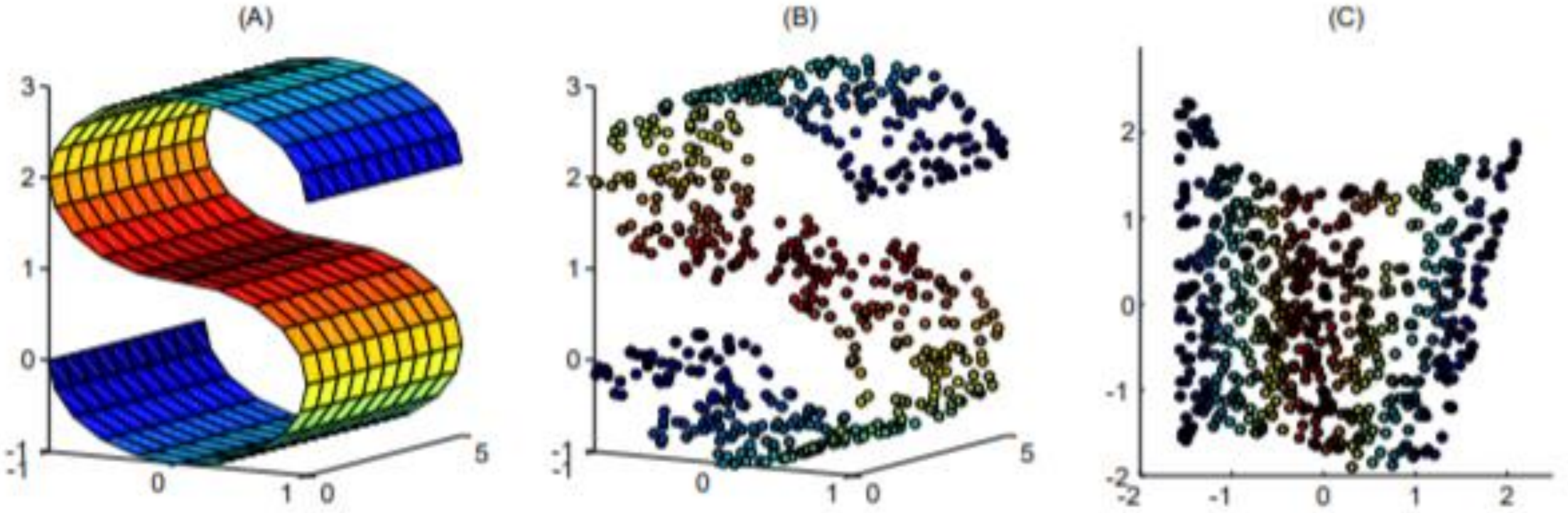


Dimensionality Reduction Technique

Manifold Learning

Locally
Linear
Embedding

LLE



Dimensionality Reduction Technique

Other Techniques

- Multidimensional Scaling
- Isomap
- t-Distributed Stochastic Neighbor Embedding
- Linear Discriminant Analysis

DIMENSIONALITY REDUCTION
CHAPTER 8

Github Repository

- Ipython notebook **Chapter 8 Tutorial**
- Ipython notebook **Iris Data**
- PDF presentation **Chapter 8 Slides**

Ana Maria Sandoval Jimenez, Jannis Busch & Sabrina Steinert

Kernel Trick (SVM)