



SNS COLLEGE OF ENGINEERING



Kurumbapalayam(Po), Coimbatore – 641 107

Accredited by NAAC-UGC with 'A' Grade

Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

Department of AI &DS

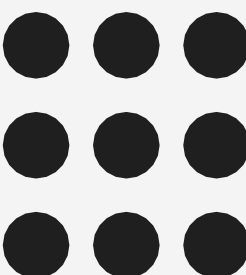
Course Name – 19AD602 DEEP LEARNING

III Year / VI Semester

Unit 3-DIMENSIONALITY REDUCTION

Topic: Autoencoders and dimensionality reduction in networks

GULSHAN BANU.A/ AP/AI AND DS /Autoencoders and dimensionality reduction in networks/SNSCE



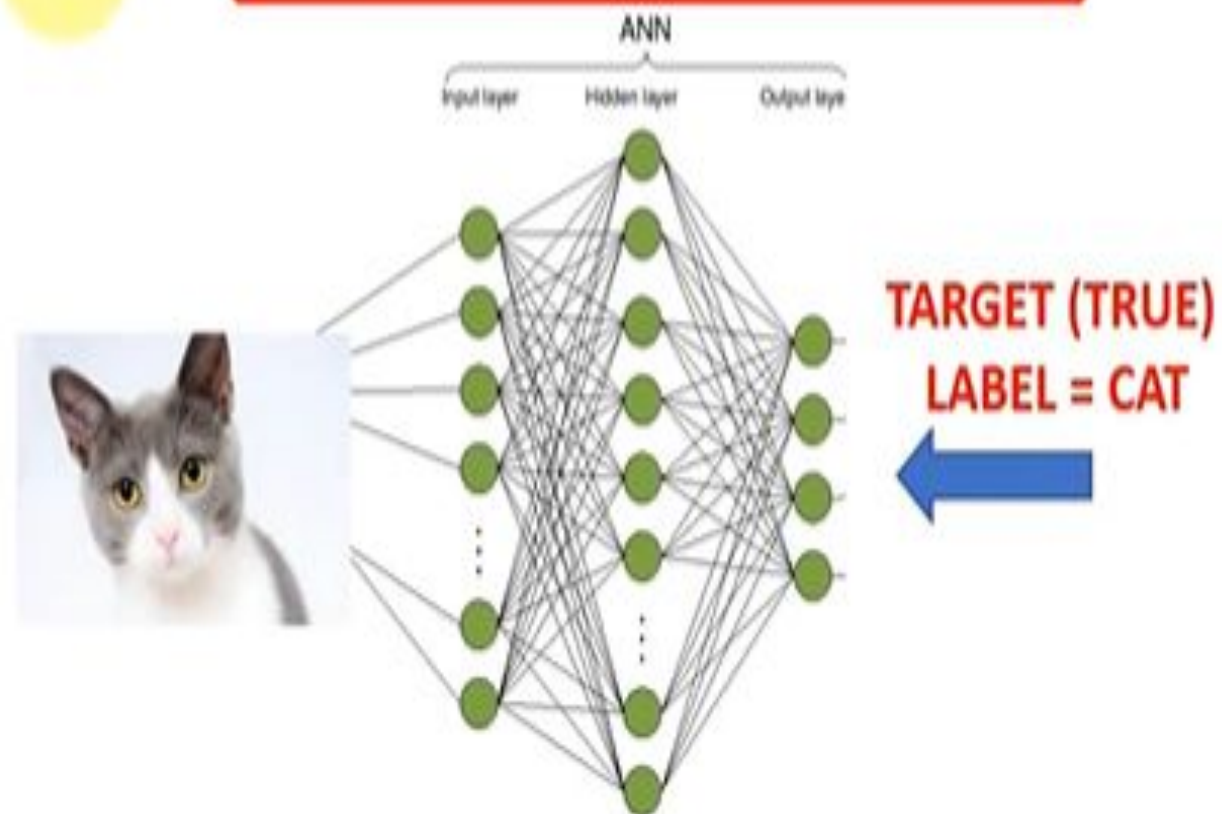


Case Study: Image Compression for Storage Optimization

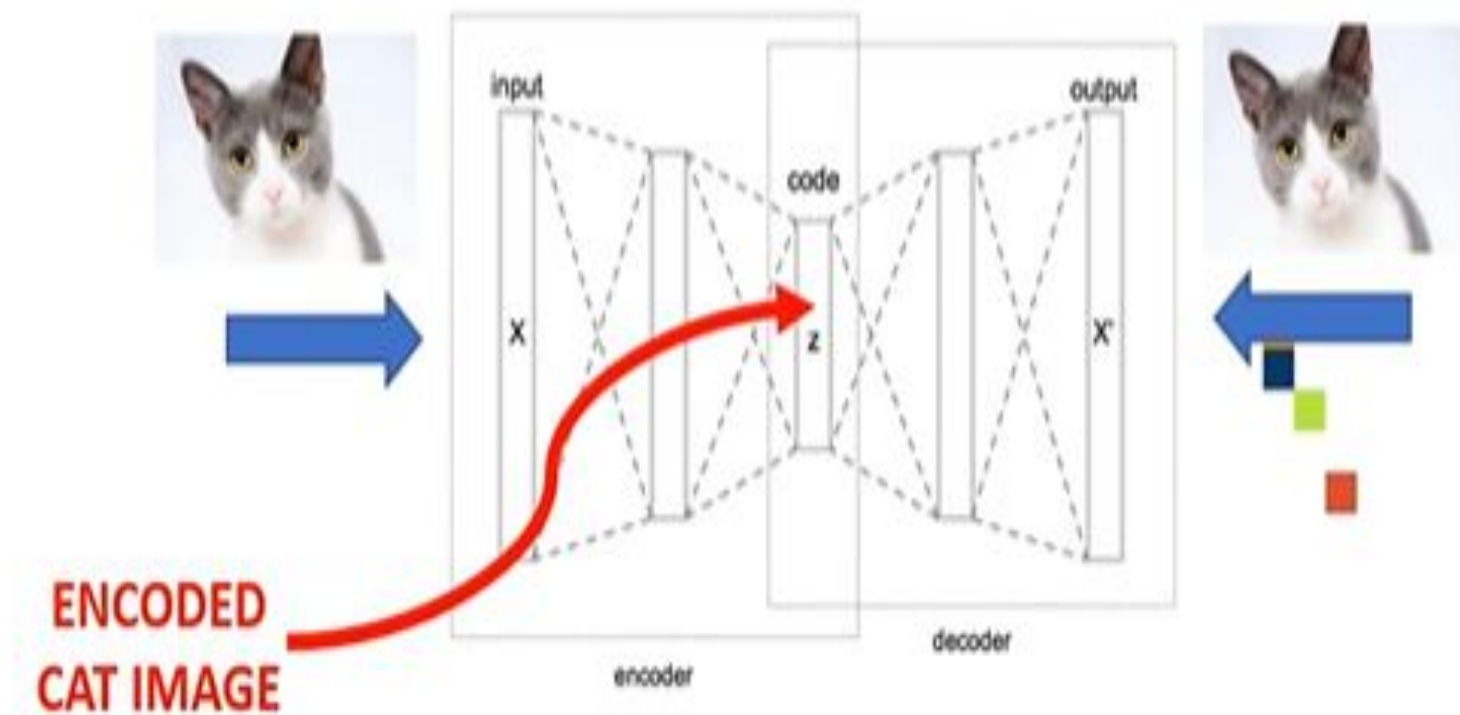
A tech company aims to store millions of grayscale images efficiently. Autoencoders are trained to compress image data into a compact latent space and then reconstruct the images with minimal loss. This reduces storage requirements while preserving key image features for retrieval.

- Auto encoders are a type of Artificial Neural Networks that are used to perform a task of data encoding (representation learning).
- Auto encoders use the same input data for the input and output, Sounds crazy right!?

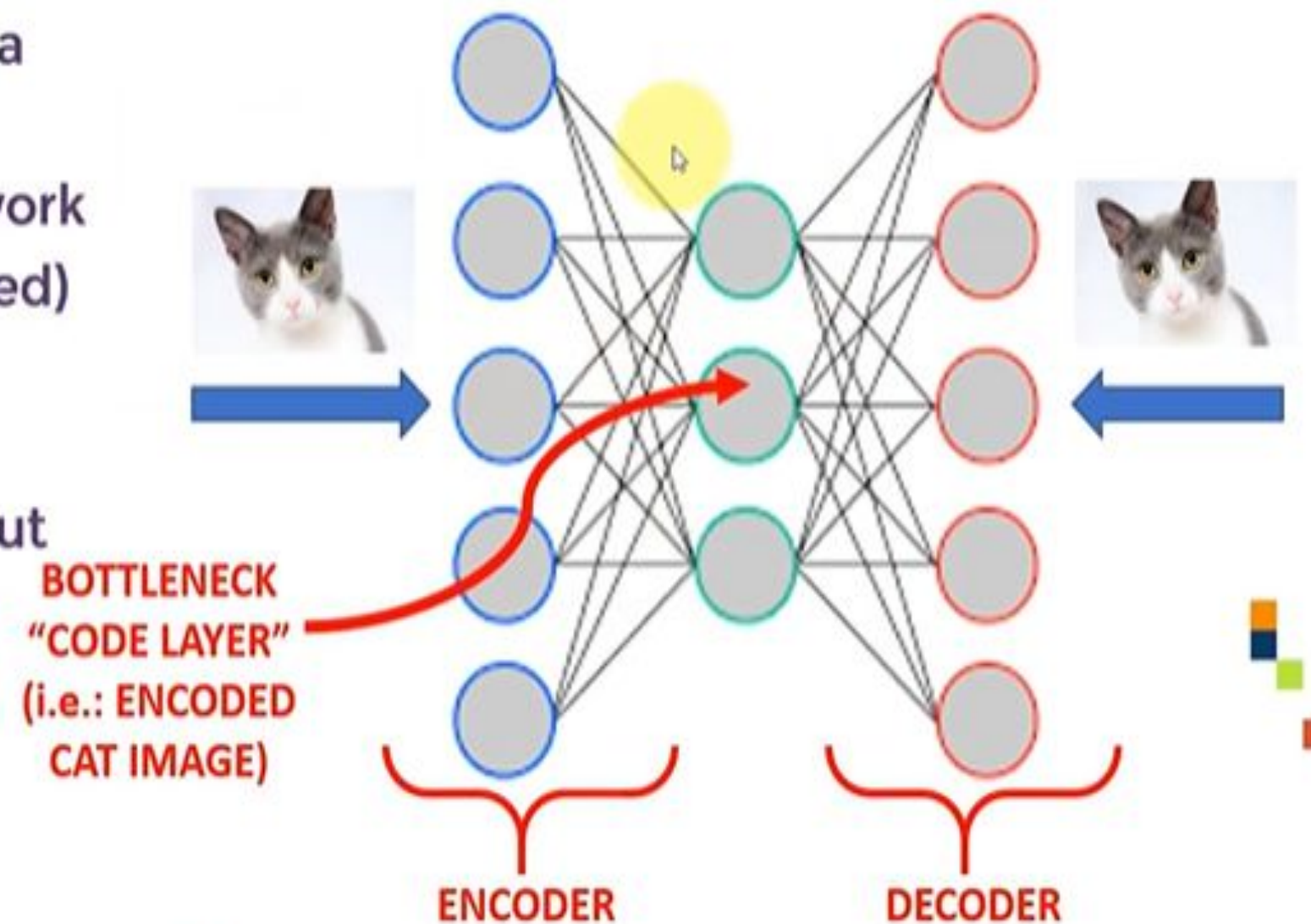
TRADITIONAL FEED-FORWARD ANN SUPERVISED TRAINING



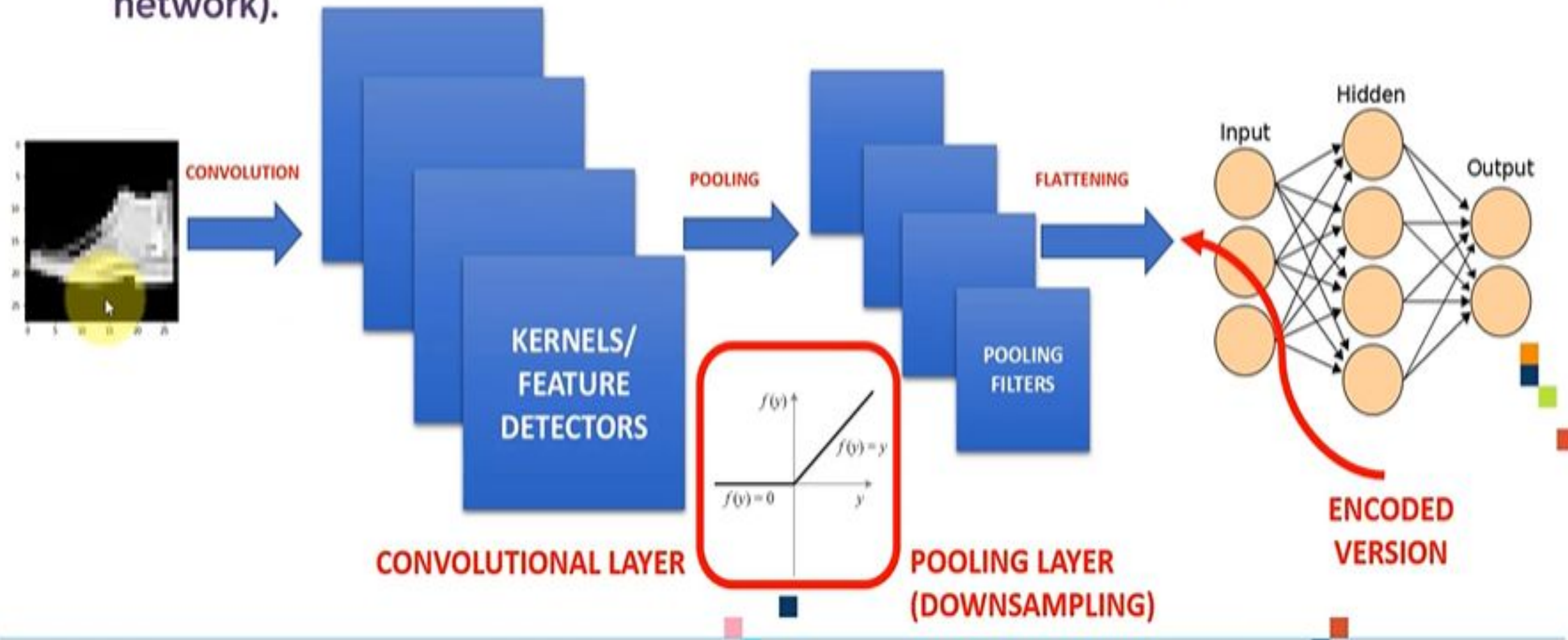
AUTOENCODERS UNSUPERVISED TRAINING



- Auto encoders work by adding a bottleneck in the network.
- This bottleneck forces the network to create a compressed (encoded) version of the original input
- Auto encoders work well if correlations exists between input data (performs poorly if the all input data is independent)
- *Great Reference: "Intro to Auto encoders by Jeremy Jordan"*



- CNNs performs encoding by taking the feature detectors and converting them into a compact single dimensional output (which is fed to the fully connected Dense network).



DIMENSIONALITY REDUCTION

Introduction

- The number of input features, variables, or columns present in a given dataset is known as *dimensionality*, and the process to reduce these features is called dimensionality reduction.
- A dataset contains a huge number of input features in various cases, which makes the predictive modeling task more complicated. Because it is very difficult to visualize or make predictions for the training dataset with a high number of features, for such cases, dimensionality reduction techniques are required to use.

Columns/Features= 10K

| Age | BMI | Glucose | Gender | Insulin | | Target |
|-----|-----|---------|--------|---------|-------|--------|
| - | - | - | - | - | - | - |
| - | - | - | - | - | - | - |
| - | - | -- | - | - | - | - |
| - | - | -- | - | - | -- | - |
| - | - | - | -- | - | - | - |
| - | - | - | -- | - | - | - |
| - | -- | - | --- | - | - | - |

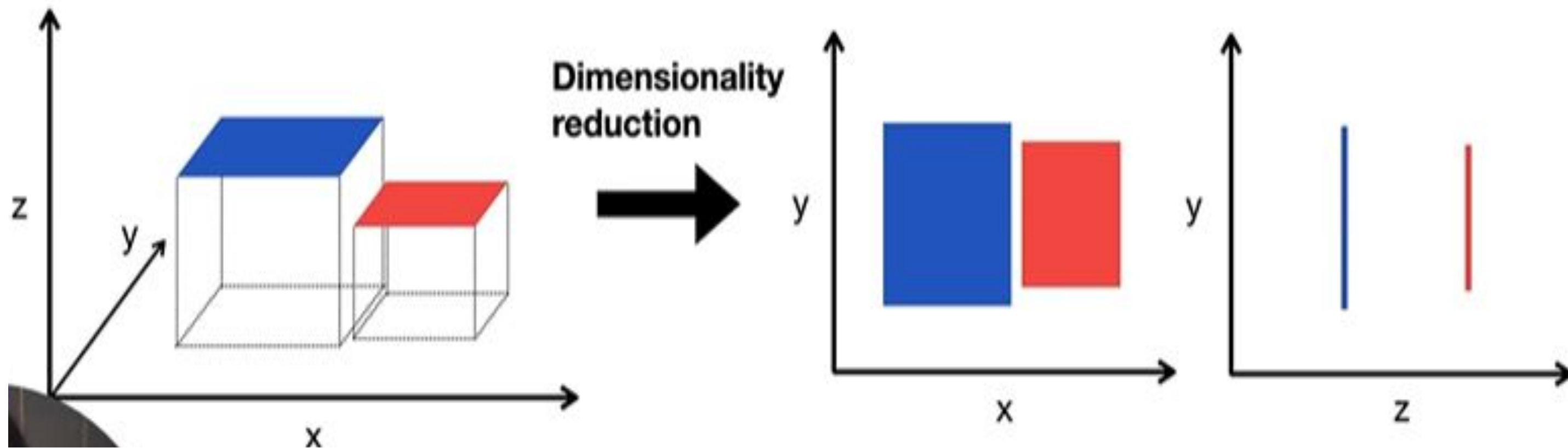
Dimensionality: number of columns/features =10k

Dimensionality Reduction

"It is a way of converting the higher dimensions dataset into lesser dimensions dataset ensuring that it provides similar information."

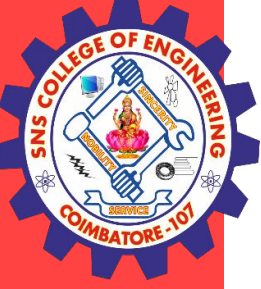
*It can be done using a technique called '**Feature Selection**' where we select only important features and discard irrelevant features.*

► Transforming 3D dataset into 2D dataset.



Common techniques of Dimensionality Reduction

- Principal Component Analysis (PCA)
- *Linear Discriminant Analysis (LDA)*
- *and many more*



Activities:

1. **Autoencoder Training:** Use an image dataset (e.g., MNIST) to train an autoencoder, visualize original and reconstructed images.
2. **Dimensionality Reduction:** Analyze the latent space (encoded representation) of the autoencoder for clustering or visualization using PCA or t-SNE.
3. **Performance Evaluation:** Measure reconstruction error and compare compressed size versus the original data size.



Autoencoders and dimensionality reduction in networks

