RV College of Engineering®

UNIT 2
Autoencoder

By: Prof. Rajesh R M
Assistant Professor
Dept. of AI and ML
RVCE, Bangalore

Go, change the world®

Outline

- Learn how the architectural design of autoencoders makes them perfectly suited to generative modeling.
- Build and train an autoencoder from scratch using Keras.
- Use autoencoders to generate new images, but understand the limitations of this approach.
- Learn about the architecture of the variational autoencoder and how it solves many of the problems associated with standard autoencoders.
- Build a variational autoencoder from scratch using Keras.
- Use variational autoencoders to generate new images.
- Use variational autoencoders to manipulate generated images using latent space arithmetic.

The Autoencoder Architecture

Autoencoders are a type of neural network designed to learn efficient data representations, primarily for the purpose of **dimensionality reduction** or **feature learning.**

- An ***encoder*** network that compresses high-dimensional input data such as an image into a lower-dimensional embedding vector
- A ***decoder*** network that decompresses a given embedding vector back to the original domain (e.g., back to an image)

The Autoencoder Architecture

Dimensionality reduction

Dimensionality reduction is a technique used to transform high-dimensional data into a lower-dimensional representation while preserving the most important information.

Principal Component Analysis

Based on

Eigen Values

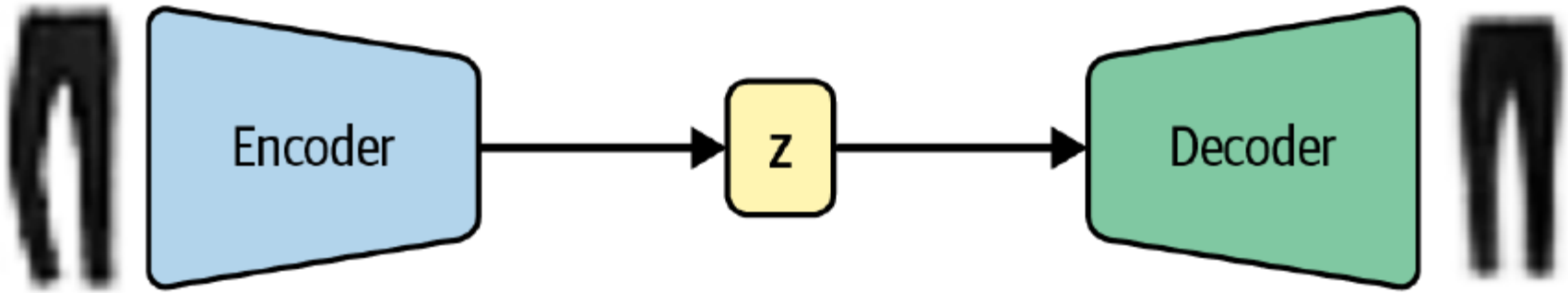
Eigen Vectors

The Autoencoder Architecture

Feature learning

Feature learning, also known as representation learning, is a machine learning technique where a system automatically discovers the most relevant features from raw data to improve its performance on tasks like classification or detection

The Autoencoder Architecture

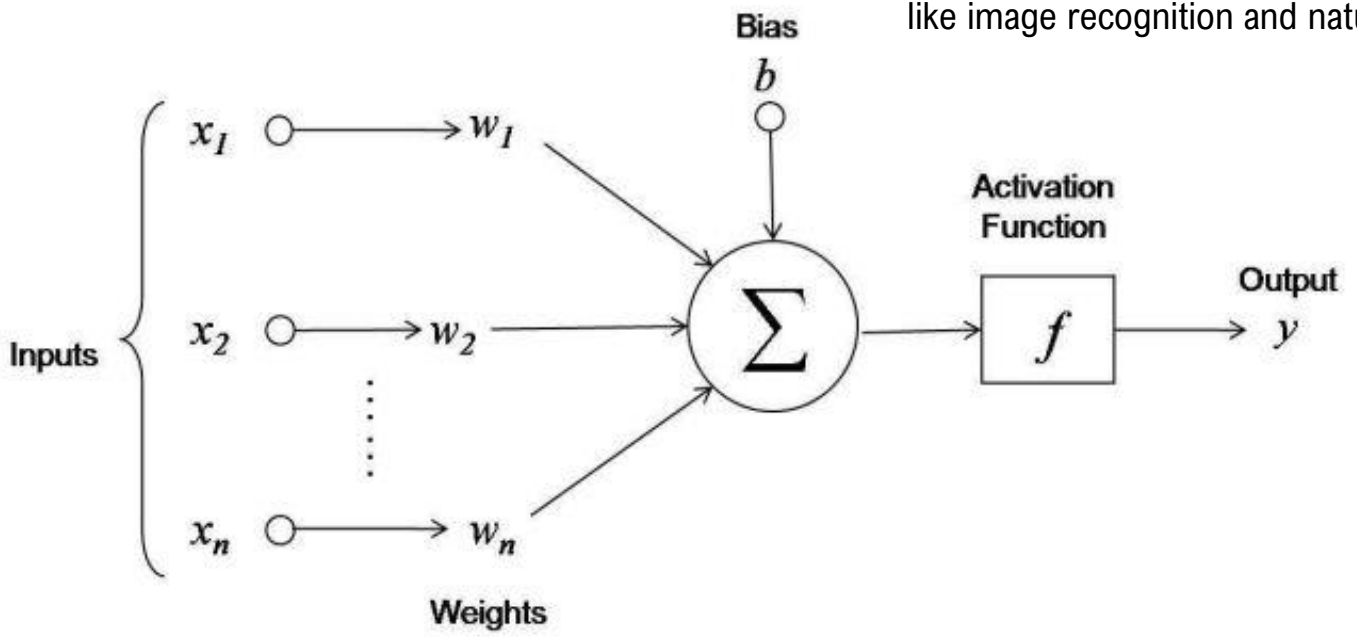


A diagram of the network architecture is shown in Figure An input image is encoded to a latent embedding vector *z*, which is then decoded back to the original pixel space.

The Autoencoder Architecture

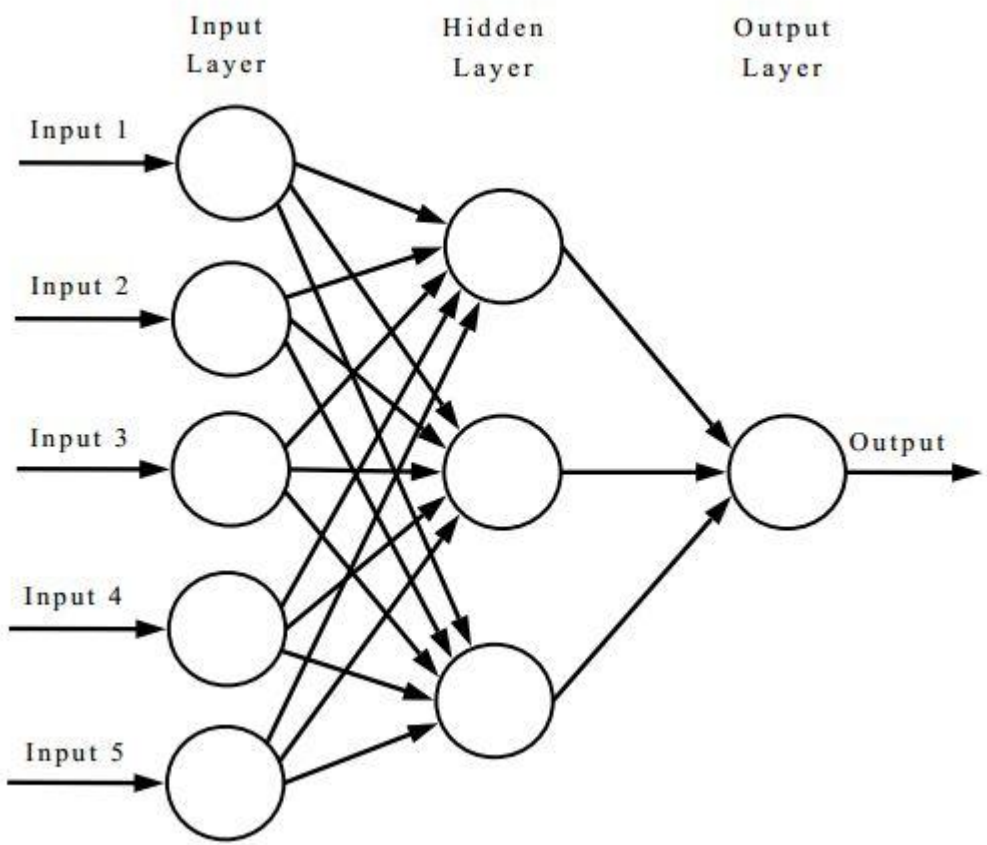
Artificial Neural Network

Artificial Neural Networks (ANNs), or simply neural networks, are computational models inspired by the human brain, consisting of interconnected nodes (neurons) that process information and learn from data, used in various applications like image recognition and natural language processing



The Autoencoder Architecture

Artificial Neural Network



The Autoencoder Architecture

Artificial Neural Network