# Lab 7

1. Explain the relationship between linear AE and principal component analysis (PCA).

|  |  |  |
| --- | --- | --- |
| **Aspect** | **Linear Autoencoder (AE)** | **Principal Component Analysis (PCA)** |
| **Objective** | Minimize reconstruction error (MSE) | Maximize variance, minimize information loss |
| **Method** | Neural network (linear encoder & decoder) | Linear algebra (eigen-decomposition) |
| **Type** | Can be linear or non-linear | Strictly linear |
| **Projection** | Learns to project onto latent space via gradient descent | Projects onto orthogonal principal components |
| **Non-linear Capability** | Yes, with non-linear activations (e.g., ReLU, Sigmoid) | No, limited to linear patterns |
| **Training** | Iterative, using gradient-based optimization | Closed-form solution (one-step calculation) |
| **Output** | Encoded features (latent space) + reconstruction | Principal components + data projection |

4. Observe the model performance improvements between the above two models and give reasons for the observed improvements.

|  |  |  |  |
| --- | --- | --- | --- |
| **Aspect** | **Autoencoder (Fully Connected)** | **Vanilla CNN Autoencoder** | **Improvement Reason** |
| Architecture | Fully connected layers (Dense layers) | Convolutional layers (Conv2D, Conv2DTranspose) | CNNs capture spatial structure (e.g., nearby pixels) better than fully connected layers. |
| Input Data Representation | Flattened 1D representation of image | Preserves 2D image structure (28x28x1) | Preserving the 2D structure allows CNNs to learn from the spatial relationships between pixels. |
| Reconstruction Quality | Poorer quality, especially for detailed or complex images | Better quality reconstructions, particularly for spatial features | CNNs are better suited for image data due to their ability to detect local features through convolutional filters. |
| Model Complexity | Simpler, fewer parameters | More complex with additional layers and parameters | CNN-based autoencoders generally require more parameters, leading to better learning for complex data. |
| Loss (Mean Squared Error) | Higher MSE | Lower MSE | CNN captures patterns better, thus reducing reconstruction error. |
| Training Time | Typically faster | Slower due to additional convolution operations | The added complexity improves performance but at the cost of training time. |
| Generalization to Test Data | Higher likelihood of underfitting or overfitting | Better generalization to unseen data | CNNs generalize better because they are designed to learn from hierarchical features in the data. |

6. Observe the model performance improvements between the Image De-noising AE and the Vanilla CNN AE.

|  |  |  |  |
| --- | --- | --- | --- |
| **Aspect** | **Image Denoising Autoencoder** | **Vanilla CNN Autoencoder** | **Observed Improvements** |
| Input Data | Noisy image data as input and original as target | Raw image data without added noise | Image Denoising AE is specifically tailored for noisy data, while CNN AE may not be designed to handle noise directly. |
| Architecture | Uses convolutional layers for both encoding and decoding | Also uses convolutional layers but no noise handling | Both use CNN, but Image Denoising AE handles noise effectively by reconstructing clean images from noisy input. |
| Reconstruction Quality | Excellent for noisy images (noise removed) | Good for regular image reconstruction | Image Denoising AE performs better in noise removal, while CNN AE may suffer when dealing with noisy input. |
| Loss (Mean Squared Error) | Lower MSE for noisy image reconstruction | Higher MSE compared to denoising AE when noise is introduced | The Denoising AE is explicitly trained to handle noisy inputs, leading to a significant improvement in MSE. |
| Training Purpose | Trained specifically to denoise images | Trained to reconstruct clean images | Denoising AE has better generalization when applied to noisy data. |
| Feature Extraction | Extracts noise-independent features | Extracts general spatial features | Noise removal is the key feature of Image Denoising AE, while CNN AE learns image features without accounting for noise. |
| Aspect | Image Denoising Autoencoder | Vanilla CNN Autoencoder | Observed Improvements |

7. Explain the differences between AE and Variational AE (VAE).

|  |  |  |
| --- | --- | --- |
| **Aspect** | **Autoencoder (AE)** | **Variational Autoencoder (VAE)** |
| Latent Space Representation | Deterministic latent space (single point encoding) | Probabilistic latent space (distribution encoding, typically Gaussian) |
| Loss Function | Minimizes reconstruction error (e.g., MSE) | Combines reconstruction error with a regularization term (KL divergence) |
| Objective | Focuses on encoding and reconstructing data | Focuses on encoding data as well as generating new data similar to training data (generative model) |
| Output | Deterministic output from input | Stochastic output; introduces variability in the latent space |
| Latent Space Interpretation | Harder to interpret latent space | Interpretable latent space with a continuous, structured representation |
| Sampling | No inherent sampling mechanism | Latent space involves sampling from distributions (enables generative models) |