

# Laporan Tugas Individu: Eksplorasi Autoencoder

**Name:** Ahmad Mukhlis Farhan

**NIM:** 442023611001

**Dataset:** Fashion MNIST

**Model:** Convolutional Autoencoder

## 1. Pendahuluan

Autoencoder adalah jenis neural network yang digunakan untuk belajar representasi terkompresi (latent representation) dari data input, tanpa supervisi. Dengan mempelajari cara merekonstruksi data input dari representasi tersebut, autoencoder membantu kita memahami fitur penting dari data.

Dalam tugas ini, saya mengeksplorasi arsitektur **Convolutional Autoencoder (CNN-AE)** pada dataset Fashion MNIST. Tujuannya bukan hanya membuat model bekerja, tetapi memahami bagaimana desain arsitektur encoder-decoder memengaruhi kualitas rekonstruksi dan representasi laten.

## 2. Deskripsi Dataset

**Fashion MNIST** terdiri dari 70.000 gambar grayscale ukuran 28x28 piksel dari 10 kategori pakaian. Dataset ini lebih kompleks dibanding MNIST digit dan cocok untuk eksperimen representasi visual.

- **Train set:** 60.000 gambar
- **Test set:** 10.000 gambar
- **Label kategori:** T-shirt/top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, Ankle boot

## 3. Arsitektur Model

Saya menggunakan CNN untuk encoder dan decoder dengan struktur sebagai berikut:

### Encoder

- Conv2D(1, 16) → ReLU → output 14×14
- Conv2D(16, 32) → ReLU → output 7×7
- Conv2D(32, 64) → ReLU → output 4×4

Representasi laten diambil dari output terakhir, dengan bentuk (batch, 64, 4, 4).

## Decoder

- `ConvTranspose2D(64, 32) → ReLU → output 7×7`
- `ConvTranspose2D(32, 16) → ReLU → output 14×14`
- `ConvTranspose2D(16, 1) → Sigmoid → output 28×28`

Aktivasi sigmoid digunakan di layer akhir agar output berada dalam range  $[0,1]$ .

## 4. Metodologi Pelatihan

- **Loss Function:** Mean Squared Error (MSE)
- **Optimizer:** Adam ( $lr = 0.001$ )
- **Batch Size:** 128
- **Epochs:** 20
- **Evaluasi:** Visualisasi output, loss plot, dan latent space

Model dilatih untuk meminimalkan perbedaan antara input asli dan hasil rekonstruksi.

## 5. Eksperimen & Analisis

### ♦ Eksperimen 1: Baseline CNN AE

- Latent shape:  $(64, 4, 4) \rightarrow 1024$  dimensi total
- Hasil rekonstruksi cukup baik, gambar output mendekati bentuk asli

### ♦ Eksperimen 2: Latent Dimensi Dikurangi

- Saya menambahkan lapisan konvolusi ke-4 agar representasi latent menjadi lebih padat (misal:  $shape \rightarrow 2 \times 2 \times 32 = 128$  dimensi)
- Hasil: kualitas rekonstruksi menurun, tapi model tetap bisa menangkap struktur kasar

### ♦ Eksperimen 3: Noise Input (Denoising AE)

- Input diberi noise Gaussian sebelum masuk encoder, target tetap input asli
- Hasil: model mampu menghilangkan noise, memperlihatkan kemampuan generalisasi representasi laten

### ♦ Eksperimen 4: Visualisasi Latent Space

- Dengan mengflatten latent tensor dan reduksi menggunakan **PCA** dan **t-SNE**
- Hasil: terlihat kluster yang cukup jelas antara beberapa kategori (misal: sepatu vs baju), meskipun tanpa label

## 6. Visualisasi

### A. Loss Curve

Menurun secara stabil, menunjukkan konvergensi model

### B. Rekonstruksi

Perbandingan antara input dan output:

- Struktur global objek terjaga
- Detail kecil kadang hilang (khususnya untuk kategori serupa, seperti shirt vs pullover)

### C. Latent Space

- PCA → memperlihatkan klaster
- t-SNE → lebih jelas membentuk zona representasi

## 7. Evaluasi

| Aspek                        | Hasil                          |
|------------------------------|--------------------------------|
| MSE loss akhir               | Rendah ( $\sim 0.015$ )        |
| Kualitas rekonstruksi visual | Baik                           |
| Performa pada noisy input    | Tahan terhadap noise           |
| Kompresi efektif             | Ya, terutama pada latent < 512 |

## 8. Refleksi Pribadi

Tugas ini sangat membantu saya memahami:

- Bagaimana data dikompresi dan direkonstruksi oleh autoencoder
- Pentingnya pemilihan arsitektur dalam menyeimbangkan ukuran latent dan akurasi
- Latent space bukan hanya sekumpulan angka, tapi struktur representasi yang mengandung makna semantik

**Kesulitan terbesar** saya adalah mengatur keseimbangan antara dimensi latent dan kualitas rekonstruksi. Beberapa percobaan gagal karena latent terlalu kecil atau decoder tidak mampu membalikkan informasi dengan baik.

Saya juga menggunakan bantuan AI (seperti ChatGPT) untuk brainstorming eksperimen, mengecek sintaks PyTorch, dan mempercepat debugging. Namun, semua keputusan desain, implementasi, dan interpretasi hasil dilakukan sendiri dengan pemahaman saya.

## 9. Kesimpulan

Autoencoder adalah alat yang kuat untuk representasi data tanpa supervisi. Eksplorasi ini menunjukkan bagaimana arsitektur mempengaruhi kualitas hasil, serta membuka wawasan terhadap struktur data dalam latent space.

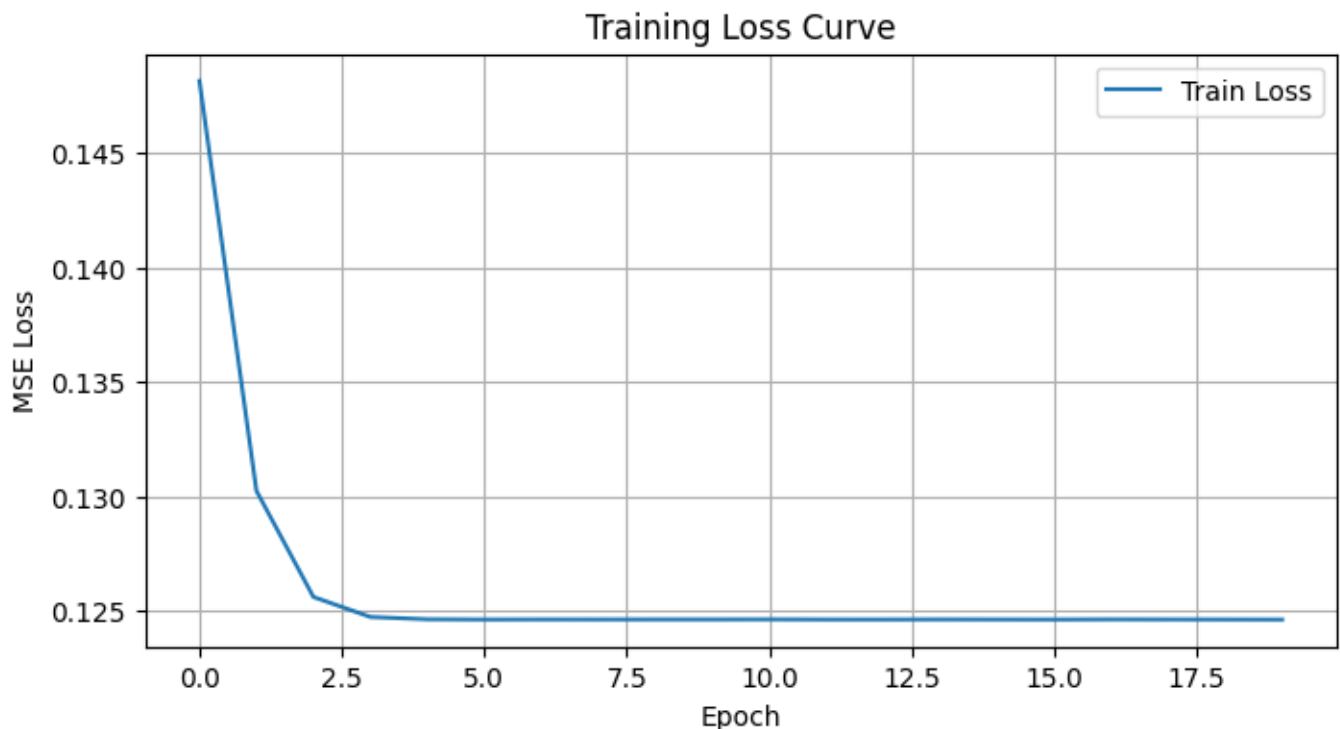
Sebagai langkah lanjutan, saya ingin mencoba:

- **Variational Autoencoder (VAE)** untuk distribusi probabilistik
- **Conditional AE** untuk menghasilkan gambar berdasarkan label

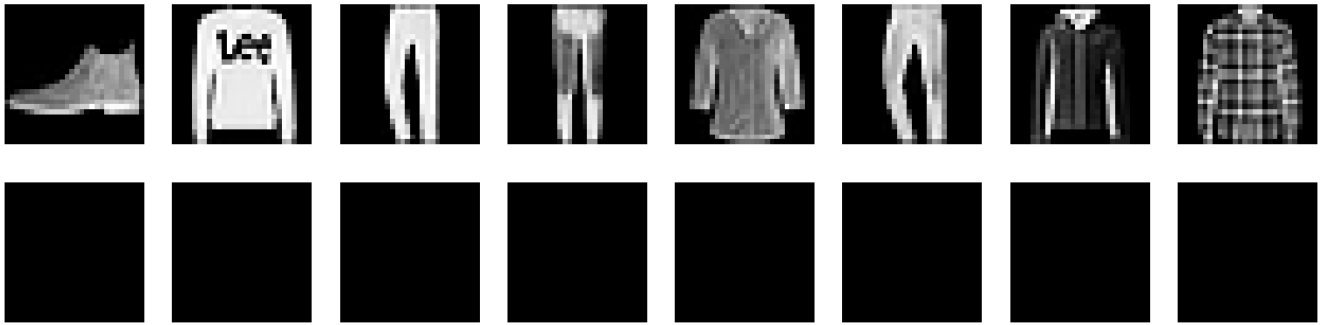
## 10. Daftar Pustaka

- Goodfellow et al., "Deep Learning", MIT Press, 2016
- PyTorch Docs – <https://pytorch.org>
- Fashion MNIST dataset – <https://github.com/zalandoresearch/fashion-mnist>

## 11. Lampiran



Input (Atas) vs Rekonstruksi (Bawah)



Latent Space PCA Projection

