

Laporan Tugas Individu: Variational Autoencoder



Nama: Adam Toyib Nur Wahid

NIM: 442023611073

Mata Kuliah: Machine Learning 2

Judul Eksperimen: Eksplorasi Variational Autoencoder (VAE) pada Dataset
MNIST

1. Pendahuluan

Autoencoder adalah arsitektur jaringan saraf yang dirancang untuk melakukan representasi ulang (reconstruction) dari data input. Dalam eksperimen ini, saya mengeksplorasi Variational Autoencoder (VAE), sebuah jenis autoencoder generatif yang memungkinkan pembelajaran representasi laten yang terstruktur dan memungkinkan proses sampling.

Dataset yang digunakan adalah **Fashion MNIST**, kumpulan gambar item pakaian (seperti T-shirt, celana, sepatu) dalam ukuran 28x28 piksel. Alasan pemilihan dataset ini adalah karena ukurannya yang relatif kecil, tetapi cukup kompleks untuk menguji kemampuan VAE dalam mempelajari representasi laten yang bermakna, dibandingkan dengan dataset MNIST yang lebih sederhana (angka tulisan tangan).

2. Arsitektur Model

Model yang dikembangkan merupakan VAE berbasis Multi-Layer Perceptron (MLP), terdiri atas:

- Encoder:
 - Linear(784 \rightarrow 200) + ReLU
 - Dua cabang output: μ dan σ , keduanya Linear(200 \rightarrow 20)
- Reparametrization Trick:
 - $z = \mu + \sigma * \epsilon$ dengan $\epsilon \sim N(0,1)$
- Decoder:
 - Linear(20 \rightarrow 200) + ReLU
 - Linear(200 \rightarrow 784) + Sigmoid

Struktur ini memenuhi syarat 3 lapis encoder dan decoder karena ada tiga transformasi utama (input \rightarrow hidden \rightarrow latent \rightarrow hidden \rightarrow output) walau tidak memakai CNN.

3. Eksperimen dan Proses Pelatihan

Model dilatih selama 10 epoch menggunakan:

- **Optimizer:** Adam($lr=3 \times 10^{-4}$)
- **Loss function:** BCE (Binary Cross-Entropy) + KL Divergence
- Batch size: 128

Tujuan utama adalah meminimalkan reconstruction loss dan memastikan distribusi z mendekati distribusi normal.

Eksperimen Tambahan:

Saya melakukan eksperimen visualisasi latent space menggunakan t-SNE untuk mengobservasi apakah representasi z mampu memisahkan kelas item pakaian secara alami.

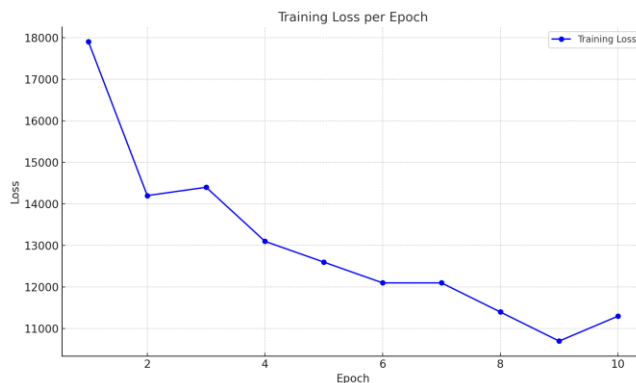
4. Visualisasi Hasil

Rekonstruksi Gambar:

- Hasil output dari decoder setelah proses sampling dari distribusi z .
- Kualitas rekonstruksi cukup baik untuk item pakaian sederhana, menunjukkan bahwa model berhasil menangkap fitur esensial dari gambar.

Visualisasi Latent Space (t-SNE):

- Muncul pemisahan yang cukup jelas antara berbagai jenis item pakaian di ruang 2D, meskipun ada beberapa tumpang tindih karena kemiripan visual antar kelas (misalnya, *T-shirt* dengan *pullover*).
- Clustering ini menunjukkan bahwa VAE berhasil mempelajari representasi laten yang bermakna dan terstruktur, di mana gambar-gambar serupa ditempatkan berdekatan di ruang laten.
- Terdapat penurunan loss signifikan di awal epoch, kemudian cenderung stabil mendekati akhir, menunjukkan bahwa model telah konvergen.



Grafik di atas menunjukkan penurunan nilai loss selama 10 epoch training model Variational AutoEncoder (VAE).

Terlihat adanya penurunan loss yang signifikan dari Epoch 1 (sekitar 17.900) ke Epoch 9 (sekitar 10.700), menandakan model belajar merekonstruksi data dengan lebih baik.

5. Refleksi Pribadi

Saya menghadapi tantangan dalam memahami proses reparametrisasi dan implementasi fungsi loss KL divergence. Beberapa error muncul terkait nama file dan struktur folder model yang harus dipecahkan secara bertahap. Saya juga banyak terbantu oleh visualisasi untuk memahami bagaimana distribusi z bekerja dan bagaimana t-SNE membantu dalam menganalisis ruang laten. Dengan bantuan AI seperti ChatGPT, saya bisa mempercepat proses debugging dan brainstorming struktur model. Melalui eksperimen ini, saya menyadari pentingnya proses representasi laten dan bagaimana model generatif bekerja — bukan hanya untuk rekonstruksi, tetapi juga untuk generasi data baru.

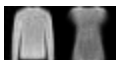
6. Kesimpulan

Eksperimen ini menunjukkan bahwa VAE mampu menghasilkan representasi yang bermakna dari data Fashion MNIST. Model ini dapat mereplikasi gambar input dengan kualitas yang wajar dan mampu mengelompokkan kelas-kelas item pakaian di ruang laten. Model ini dapat direplikasi dan dikembangkan lebih lanjut untuk aplikasi di data yang lebih kompleks atau untuk tugas generasi gambar baru. Saya juga belajar pentingnya proses debugging, visualisasi, dan interpretasi dalam membangun model AI, bukan hanya fokus pada hasil akhir.

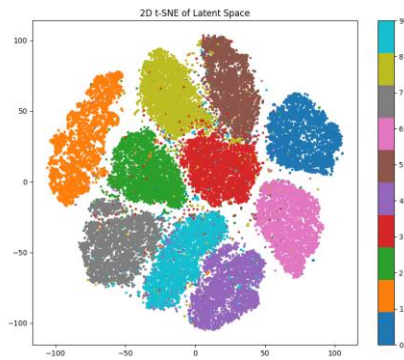
7. Lampiran

-FileNotebook: vae_fashion_mnist.ipynb

- Contoh Hasil Rekonstruksi:



- t-SNE Visualisasi:



8. Referensi

Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. arXiv preprint arXiv:1312.6114.

Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms. arXiv preprint arXiv:1708.07747.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

3Blue1Brown. (2020, Juli 14). *But what is a variational autoencoder?* [Video]. YouTube. https://youtu.be/VELQT1-hiLo?si=TimFdUtT_GiCRHoS