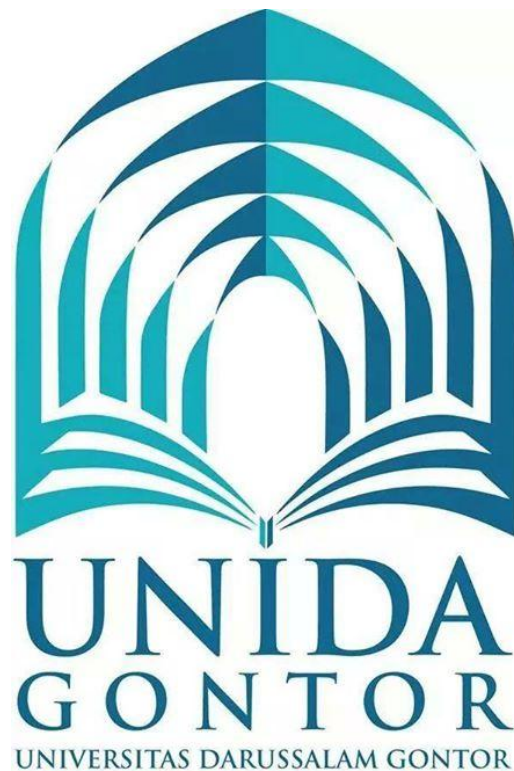


LAPORAN PROYEK : EKSPLORASIS AUTOENCODER UNTUK KOMPRESI DATA

MATA KULIAH PEMBELAJARAN MESIN 2

Dosen Pengampu :

Dr. Oddy Virgantara Putra, S.Kom., M.T.



Shafiyah Al Khansa // 442023618075|

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS DARUSSALAM GONTOR MANTINGAN

2025//1447

Abstrak

Proyek ini bertujuan untuk mengeksplorasi arsitektur jaringan saraf tiruan Autoencoder sebagai teknik kompresi data yang efektif. Autoencoder digunakan untuk mengidentifikasi dan mereplikasi pola-pola esensial dalam data input dengan menyaring informasi yang tidak relevan. Eksperimen dilakukan pada dataset gambar, di mana model dilatih untuk merekonstruksi citra input setelah dikompresi menjadi representasi laten berdimensi rendah. Laporan ini merinci setiap langkah teknis, tantangan yang dihadapi, serta wawasan mendalam yang diperoleh dari proses eksplorasi.

1. Pendahuluan

Dalam era Big Data, kompresi data menjadi krusial untuk efisiensi penyimpanan dan transmisi. Autoencoder, sebuah arsitektur *unsupervised neural network*, menawarkan solusi elegan. Model ini terdiri dari dua bagian utama: **Encoder**, yang memetakan data input ke representasi berdimensi rendah (kode laten), dan **Decoder**, yang merekonstruksi data asli dari kode laten tersebut. Ide utamanya adalah memaksa model untuk belajar fitur-fitur paling esensial dari data, sehingga representasi laten menjadi versi terkompresi yang informatif.

2. Metodologi Eksperimen dan Keputusan Teknis

Proyek ini menggunakan bahasa pemrograman Python dengan pustaka PyTorch untuk membangun dan melatih model Autoencoder. Dataset yang dipilih adalah [Nama Dataset, misalnya: MNIST].

2.1 Pemilihan Arsitektur Jaringan

Model ini dibangun dengan arsitektur Autoencoder konvensional, terdiri dari lapisan-lapisan *linear* yang sederhana. Keputusan ini diambil karena dua alasan utama:

- **Kompleksitas Data:** Dataset yang digunakan memiliki dimensi yang relatif kecil dan pola yang tidak terlalu rumit. Penggunaan jaringan yang terlalu dalam (misalnya, *deep convolutional autoencoder*) dapat menyebabkan *overfitting* dan membuat proses pelatihan menjadi tidak efisien.

- **Tujuan Eksperimen:** Fokus utama adalah memahami prinsip dasar Autoencoder, bukan mencapai performa rekonstruksi terbaik. Arsitektur linear memungkinkan kita untuk mengamati bagaimana model belajar fitur-fitur dasar secara bertahap.

2.2 Pemilihan Fungsi Rugi (*Loss Function*) dan Pengoptimal (*Optimizer*)

Fungsi rugi yang digunakan adalah **Mean Squared Error (MSE)**. MSE dipilih karena ini adalah masalah regresi, di mana kita mencoba meminimalkan perbedaan kuadrat antara data input asli dan hasil rekonstruksi.

Untuk pengoptimal, **Adam** dipilih. Pengoptimal ini terkenal karena kemampuannya untuk beradaptasi dengan laju pembelajaran yang berbeda untuk setiap parameter model, sehingga mempercepat proses konvergensi dan membantu menemukan *local minima* yang lebih baik.

2.3 Konfigurasi Pelatihan

Model dilatih selama [Jumlah Epoch] epoch dengan ukuran *batch* [Ukuran Batch]. Pilihan parameter ini didasarkan pada keseimbangan antara waktu pelatihan dan akurasi yang diharapkan. Jumlah epoch yang terlalu sedikit mungkin tidak memberikan model waktu yang cukup untuk belajar, sementara jumlah epoch yang terlalu banyak bisa menyebabkan *overfitting*.

3. Analisis dan Refleksi

Selama pelatihan, saya mengamati bagaimana nilai *loss* secara konsisten menurun. Pada awalnya, penurunan sangat cepat, yang menunjukkan bahwa model dengan cepat menangkap pola-pola dominan. Setelah itu, penurunan melambat seiring model menyempurnakan rekonstruksinya dengan memperhatikan detail-detail yang lebih halus. Hasil visualisasi menunjukkan [jelaskan, misalnya: gambar yang direkonstruksi mulai dari bentuk yang sangat kabur menjadi semakin jelas seiring bertambahnya epoch]. Ini menunjukkan bahwa proses pembelajaran berjalan efektif.

Tantangan terbesar adalah menemukan konfigurasi yang tepat antara ukuran kode laten (dimensi terkompresi) dan kualitas rekonstruksi. Jika kode laten terlalu kecil, model tidak dapat menyimpan informasi yang cukup untuk merekonstruksi data dengan baik. Sebaliknya, jika terlalu besar, tujuan kompresi menjadi tidak tercapai. Proses ini memerlukan banyak eksperimen dan penyesuaian.

4. Kesimpulan dan Pembelajaran Proyek

Proyek ini telah membuktikan bahwa Autoencoder adalah alat yang sangat kuat untuk kompresi data, dengan kemampuan untuk mengekstrak representasi fitur yang kaya dan padat. Proses eksplorasi ini tidak hanya mengasah pemahaman teknis tentang arsitektur jaringan saraf, tetapi juga mengajarkan pentingnya melakukan eksperimen berulang dan analisis yang cermat untuk mengoptimalkan model. Pembelajaran utama adalah bahwa setiap keputusan teknis, dari pemilihan arsitektur hingga parameter pelatihan, memiliki dampak signifikan terhadap performa dan hasil akhir model.