

LAPORAN UJIAN AKHIR SEMESTER
MACHINE LEARNING

Dosen Pengampu: Izhan Fakhruzi, ST., MSc.



Disusun oleh:

Rangga Aditya Saputra	NIM. 221230043
Ferdian Putra Wijaksono	NIM. 221230045
Abbiyyu Farras	NIM. 221230029
Andika Putra Apriyatna	NIM. 221230071
Figo Firnanda	NIM. 221230039

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK & ILMU KOMPUTER
UNIVERSITAS MUHAMMADIYAH PONTIANAK
2026

Tujuan

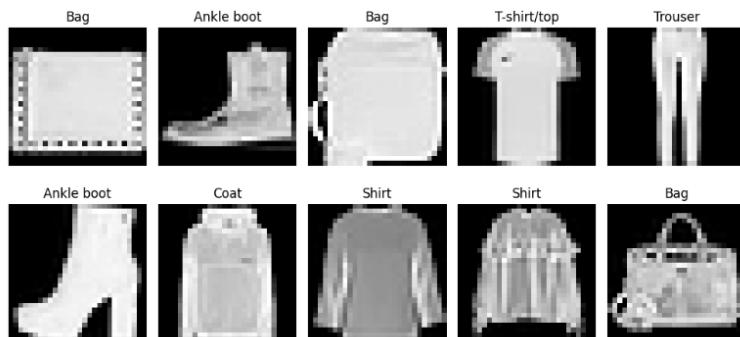
Tujuan dari *project* ini adalah menguji dan membandingkan beberapa model *machine learning* dan *deep learning* dalam mengklasifikasikan gambar Fashion-MNIST. Model yang digunakan meliputi SVM sebagai metode tradisional, MLP sebagai *neural network* berbasis fitur numerik, dan CNN sebagai model untuk pengolahan citra.

Deskripsi Data & Preprocessing

Dataset yang digunakan adalah Fashion-MNIST, yaitu kumpulan gambar pakaian berukuran 28×28 piksel grayscale dengan total 10 kelas. Data sudah dipisahkan menjadi data latih (60.000 gambar) dan data uji (10.000 gambar) dalam bentuk CSV, di mana kolom pertama berisi label kelas dan sisanya merupakan nilai piksel sebanyak 784 fitur. Pada tahap *preprocessing*, nilai piksel dinormalisasi dari rentang 0–255 menjadi 0–1 agar proses pelatihan lebih stabil, kemudian sebagian data latih dibagi lagi menjadi *train* dan *validation* untuk memantau performa model.

Hasil & Pembahasan

1. EDA

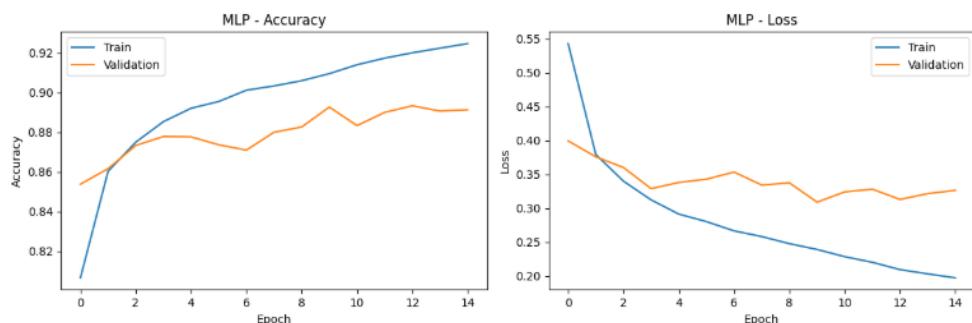


Gambar 1. Bentuk Data dari 10 Kelas

Pada bagian EDA, ditampilkan beberapa contoh gambar dari dataset Fashion-MNIST untuk melihat bentuk data secara langsung. Terlihat bahwa gambar berupa grayscale 28×28 piksel dengan berbagai kategori seperti bag, ankle boot, t-shirt/top, dan trouser, sehingga dataset ini cocok untuk tugas klasifikasi gambar.

2. Learning Curve (MLP & CNN)

- MLP

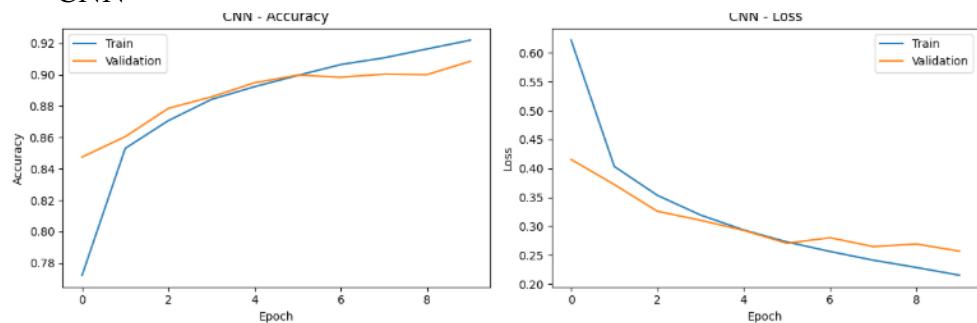


Gambar 2. Learning Curve MLP

Berdasarkan *learning curve* MLP di Gambar 2, terlihat bahwa akurasi training terus meningkat hingga sekitar 0,92, sedangkan akurasi validation berhenti di kisaran 0,89 dan cenderung stabil. Pada grafik *loss*, training loss terus menurun, namun *validation loss* tidak ikut turun secara signifikan dan cenderung datar. Pola ini menunjukkan adanya *overfitting* ringan, yaitu model semakin bagus di data latih tetapi peningkatannya tidak terlalu besar pada data validasi. Pada percobaan arsitektur, jumlah hidden layer dinaikkan dari 2 menjadi 4, namun peningkatan akurasi tidak signifikan, yaitu dari sekitar 0,88 (2 hidden layer) menjadi 0,89 (4 hidden layer). Karena selisihnya kecil, MLP dengan 2 hidden layer tetap dianggap cukup baik sebagai baseline yang lebih sederhana.

Pada bagian ini, AI digunakan untuk membantu menuliskan kode pembuatan grafik *learning curve* (*plot accuracy* dan *loss* untuk *train/validation*) serta merapikan parameter visualisasi seperti judul, label sumbu, dan legend agar hasil grafik mudah dibaca.

- CNN

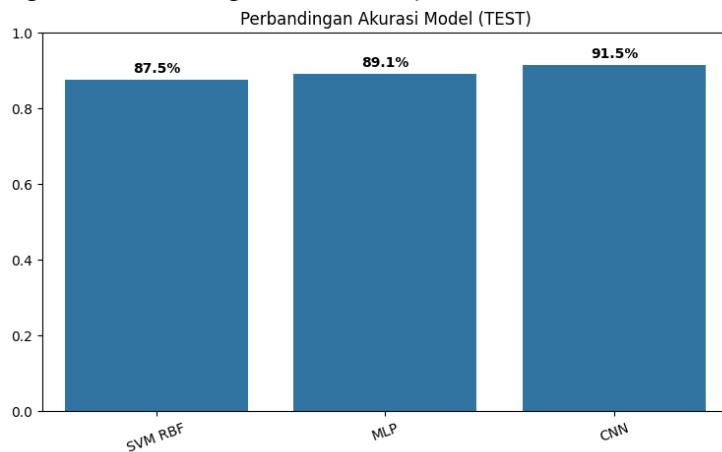


Gambar 3. *Learning Curve* CNN

Berdasarkan *learning curve* CNN Gambar 3, akurasi train dan validation meningkat dengan jarak yang kecil hingga berada di sekitar 0,91–0,92, sehingga performanya terlihat lebih stabil dan mampu melakukan generalisasi dengan baik. Grafik *loss* juga sama-sama menurun tanpa perbedaan yang jauh, yang menunjukkan *overfitting* tidak terlalu kuat.

Pada bagian ini, AI digunakan untuk membantu menuliskan kode pembuatan *learning curve* CNN (mengambil nilai dari `history_cnn.history`), menyusun interpretasi hasil grafik, serta membantu pengambilan keputusan bahwa arsitektur CNN dengan 2 *convolution layer* sudah cukup optimal karena saat dicoba menambah *layer*, akurasinya justru menurun.

3. Perbandingan hasil (tabel/grafik *accuracy*)



Gambar 4. Perbandingan Akurasi Model

Berdasarkan grafik Perbandingan Akurasi Model, model CNN memberikan hasil terbaik dengan akurasi sekitar 91,5%, diikuti oleh MLP sebesar 89,1%, dan SVM RBF sebesar 87,5%. Hasil ini menunjukkan bahwa CNN lebih unggul karena mampu mempelajari pola gambar secara spasial (bentuk dan tekstur), sedangkan MLP dan SVM menggunakan data dalam bentuk fitur datar sehingga informasi struktur gambar tidak ditangkap sebaik CNN. Pada bagian ini, AI digunakan untuk membantu membuat kode visualisasi perbandingan akurasi dalam bentuk *bar chart* serta membantu merangkum interpretasi hasil.

4. *Confusion matrix* dan *Classification Report* (model terbaik)

- *Confusion Matrix*

Confusion Matrix - CNN (Acc: 91.65%)											
Label Aktual	T-shirt/top	Label Prediksi									
		T-shirt/top	Trouser	Pullover	Dress	Coat	Sandal	Shirt	Sneaker	Bag	
T-shirt/top	879	3	18	9	1	1	85	1	3	0	
Trouser	1	985	1	9	0	0	3	0	1	0	
Pullover	11	1	885	11	37	0	52	0	3	0	
Dress	18	6	4	935	23	1	11	0	2	0	
Coat	0	3	44	17	863	1	71	0	1	0	
Sandal	0	0	0	0	0	962	0	32	1	5	
Shirt	128	4	59	21	49	0	734	0	5	0	
Sneaker	0	0	0	0	0	3	0	968	0	29	
Bag	3	1	3	0	2	1	2	3	983	2	
Ankle boot	0	0	0	0	0	0	0	28	0	971	

Gambar 6. *Confusion Matrix* dari Model Terbaik

Confusion matrix ini menunjukkan hasil prediksi CNN pada data uji, di mana sumbu vertikal adalah label asli dan sumbu horizontal adalah label prediksi. Nilai yang paling besar berada pada garis diagonal, artinya sebagian besar gambar berhasil diklasifikasikan dengan benar, sehingga akurasi model mencapai sekitar 91,65%.

Namun masih ada beberapa kesalahan prediksi yang terlihat pada kelas-kelas yang bentuknya mirip, misalnya *T-shirt/top* sering tertukar dengan *Shirt*, dan *Pullover/Coat/Shirt* juga beberapa kali saling tertukar.

- *Classification Report*

Classification Report (Model Terbaik):				
	precision	recall	f1-score	support
T-shirt/top	0.85	0.88	0.86	1000
Trouser	0.98	0.98	0.98	1000
Pullover	0.87	0.89	0.88	1000
Dress	0.93	0.94	0.93	1000
Coat	0.89	0.86	0.87	1000
Sandal	0.99	0.96	0.98	1000
Shirt	0.77	0.73	0.75	1000
Sneaker	0.94	0.97	0.95	1000
Bag	0.98	0.98	0.98	1000
Ankle boot	0.96	0.97	0.97	1000
accuracy			0.92	10000
macro avg	0.92	0.92	0.92	10000
weighted avg	0.92	0.92	0.92	10000

Gambar 7. *Classification Matrix* dari Model Terbaik

Classification report ini menampilkan evaluasi model per kelas, sehingga hasilnya lebih detail dibanding hanya melihat akurasi. Model terbaik memiliki akurasi sekitar 0,92, dengan kelas seperti Trouser, Sandal, Bag, dan Ankle boot memiliki nilai *precision/recall* tinggi, sedangkan kelas Shirt paling rendah karena sering tertukar dengan kelas yang bentuknya mirip.

Pada bagian ini, AI digunakan untuk membantu menambahkan kode *classification_report()* agar evaluasi mencakup *precision*, *recall*, dan *F1-score*, serta membantu menjelaskan kelas mana yang performanya bagus dan mana yang masih lemah.

Kesimpulan

Kesimpulannya, penelitian ini berhasil membandingkan performa model SVM RBF, MLP, dan CNN dalam klasifikasi gambar *Fashion-MNIST*. Dari hasil pengujian, CNN menunjukkan performa terbaik dengan akurasi sekitar 91%, diikuti oleh MLP sekitar 89% dan SVM sekitar 87%, sehingga CNN lebih efektif karena mampu mempelajari pola gambar secara spasial.

Selain itu, eksperimen arsitektur menunjukkan bahwa penambahan kompleksitas model (seperti menambah *hidden layer* pada MLP atau menambah *conv layer* pada CNN) tidak selalu meningkatkan akurasi secara signifikan. Evaluasi melalui *confusion matrix* dan *classification report* juga menunjukkan sebagian besar prediksi sudah tepat, namun beberapa kelas yang memiliki bentuk mirip seperti *T-shirt/top* dan *Shirt* masih sering tertukar.