

COVERT DATASET – WEEK 2

A. Tujuan

Cover Type adalah dataset yang berisi informasi geografis dan karakteristik tanah untuk mengklasifikasikan tipe vegetasi hutan. Dengan menggunakan teknik Deep Learning, khususnya Multilayer Perceptron (MLP), kita dapat membangun model prediktif yang mampu mengklasifikasikan tipe vegetasi berdasarkan fitur-fitur yang diberikan.

B. Dataset

Dataset yang digunakan berisi beberapa fitur geografis dan target Cover_Type, yang merupakan klasifikasi tipe hutan. Dataset ini memiliki fitur sebagai berikut:

- Elevation (ketinggian)
- Aspect (arah lereng)
- Slope (kemiringan)
- Horizontal dan Vertical Distance to Hydrology (jarak horizontal dan vertikal ke badan air)
- Soil Type (tipe tanah)
- Wilderness Area (area liar)
- Cover_Type (label target dengan 7 kelas)

C. Arsitektur MLP

Model yang digunakan adalah Multilayer Perceptron (MLP) dengan beberapa lapisan tersembunyi dan aktivasi ReLU. Model memiliki arsitektur sebagai berikut:

- Input Layer: Sesuai dengan jumlah fitur.
- Hidden Layers: 2 lapisan dengan 64 dan 32 neuron.
- Output Layer: Softmax activation untuk klasifikasi multi-kelas

D. Implementasi PyTorch

PyTorch digunakan untuk membangun model dengan pendekatan imperative programming yang fleksibel.

- Membangun model dengan nn.Module dan layer Linear + ReLU.
- Menggunakan CrossEntropyLoss dan optimizer Adam.
- Melatih model menggunakan backpropagation selama 50 epochs.
- Menghitung metrik evaluasi dengan torchmetrics.

E. Implementasi TensorFlow

Model dikembangkan dengan menggunakan Keras dari TensorFlow, dengan langkah-langkah sebagai berikut:

- Membangun arsitektur dengan Dense Layers dan aktivasi ReLU.
- Menggunakan Adam optimizer dan categorical cross-entropy loss.
- Melatih model selama 50 epochs.
- Mengevaluasi performa model dengan metrik yang disebutkan.

F. Hasil Training

Setelah melatih model dengan berbagai konfigurasi, diperoleh hasil evaluasi sebagai berikut:

Model	Akurasi	Recall	Presisi	F1-Score	AUC-ROC
PyTorch	0,63	0,66	0,63	0,6	0,78
TensorFlow	0,84	0,84	0,84	0,84	0,98

Hasil Eksperimen:

- Model berbasis TensorFlow menunjukkan performa yang lebih unggul dibandingkan model PyTorch dalam semua metrik evaluasi.

- Akurasi model TensorFlow mencapai 84%, jauh lebih tinggi dibandingkan dengan model PyTorch yang hanya 63%.
- AUC-ROC TensorFlow mencapai 0,98, yang menunjukkan model memiliki kemampuan sangat baik dalam membedakan kelas yang berbeda..

G. Kesimpulan

Berdasarkan hasil evaluasi, model TensorFlow lebih direkomendasikan untuk tugas klasifikasi Cover Type karena menunjukkan performa yang lebih baik di semua aspek. Namun, model PyTorch tetap bisa ditingkatkan dengan eksperimen tambahan, seperti:

- Penyesuaian learning rate dan optimizer
- Penambahan dropout layer untuk mengurangi overfitting
- Peningkatan jumlah epoch dengan early stopping.

H. Link Google Colab

PyTorch:

<https://colab.research.google.com/drive/1pv0Lf78clw7JZQKiwYWnhxNAKysNwNMV?usp=sharing>

Tensor Flow:

<https://colab.research.google.com/drive/1JTxJI6EM6Jn5PgiEuvKGtCoxLZsYmlOu?usp=sharing>