

## UTS Deep Learning

Analisa:

1. Jika menggunakan model MLP dengan 3 hidden layer (256-128-64) menghasilkan underfitting pada dataset ini, modifikasi apa yang akan dilakukan pada arsitektur? Jelaskan alasan setiap perubahan dengan mempertimbangkan bias-variance tradeoff!

➔ Modifikasi yang dapat dilakukan:

- **Meningkatkan jumlah neuron pada hidden layer:** Dengan lebih banyak neuron, model memiliki kapasitas yang lebih besar untuk menangkap pola
- **Menambah jumlah hidden layer:** Menambah hidden layer dapat membuat model mampu menangkap relasi yang lebih kompleks
- **Mengurangi regularisasi:** Mengurangi dropout atau L2 mampu membuat model model lebih fleksibel dalam penyesuaian terhadap data

Underfitting biasanya terjadi karena model terlalu sederhana untuk data.

Dengan beberapa cara tersebut, mampu meningkatkan kemampuan model untuk menangkap pola dalam data dan mengurangi bias

2. Selain MSE, loss function apa yang mungkin cocok untuk dataset ini? Bandingkan kelebihan dan kekurangannya, serta situasi spesifik di mana alternatif tersebut lebih unggul daripada MSE!

Loss Function	Kelebihan	Kekurangan	Lebih unggul dari MSE saat
MAE (Mean Absolute Error)	Mengukur absolute error yang lebih kuat terhadap outlier, karena tidak memberi bobot yang lebih besar dari kesalahan besar	Gradiennya konstan, sehingga lebih sulit untuk doptimalkan	Ingin menghindari efek dari outliers
Huber Loss	Gabungan dari MSE dan MAE, lebih kuat terhadap outliers	Perlu perhatian lebih terhadap tuning karena terdapat parameter delta yang perlu diatur	Dataset banyak outliers tapi ingin menangani kesalahan kecil lebih sensitif
Log-Cosh Loss	Mengurangi dampak besar dari outlier tapi	Berat komputasinya	Keseimbangan antara respons terhadap outlier

	mempertahankan sifat MSE		dan kesalahan kecil
--	--------------------------	--	---------------------

3. Jika salah satu fitur memiliki range nilai 0-1, sedangkan fitur lain 100-1000, bagaimana ini memengaruhi pelatihan MLP? Jelaskan mekanisme matematis (e.g., gradien, weight update) yang terdampak!

→ Mampu mempengaruhi pelatihan MLP

Mekanisme matematis yang terdampak:

- **Gradien:** Gradien yang dihitung untuk fitur akan lebih besar, sehingga menyebabkan model memberikan bobot lebih pada fitur dengan rentang nilai yang besar
  - **Weight Update:** Weight Update tidak seimbang antara fitur-fitur tersebut, sehingga model tidak optimal, menyebabkan pelatihan menjadi tidak efisien.
4. Tanpa mengetahui nama fitur, bagaimana Anda mengukur kontribusi relatif setiap fitur terhadap prediksi model? Jelaskan metode teknikal (e.g., permutation importance, weight analysis) dan keterbatasannya!

→ - **Permutation Importance:** Mengukur pengaruh fitur dengan menghitung perubahan skor kinerja model setelah fitur tertentu diacak. Kelemahannya bisa memakan waktu yang banyak, terutama ketika datasetnya besar

- **Weight Analysis:** Kontribusi relative dari setiap fitur dapat dilihat dari bobot yang dipelajari model. Kelemahannya hanya relevan untuk model sederhana dan tidak selalu akurat

5. Bagaimana Anda mendesain eksperimen untuk memilih learning rate dan batch size secara optimal? Sertakan analisis tradeoff antara komputasi dan stabilitas pelatihan!

→ Tradeoff:

Parameter	Efek terlalu besar	Efek terlalu kecil
Learning Rate	Overstepping pada minimum loss, model menjadi tidak stabil	Menyebabkan pelatihan menjadi sangat lama
Batch Size	Lebih cepat komputasi, tetapi menghasilkan gradien noise yang besar, model menjadi sulit untuk generalisasi	Menghasilkan model yang lebih akurat dan lebih sedikit memori, tetapi komputasi lama