

1. Hidden Layers

- **Pengamatan:** Kombinasi jumlah dan ukuran neuron dalam hidden layer memengaruhi kemampuan model untuk menangkap pola kompleks.
- **Hasil:**
 - 1) Konfigurasi dengan lebih banyak layer cenderung menghasilkan akurasi lebih tinggi, tetapi dapat menyebabkan overfitting jika jumlah neuron terlalu besar.
 - 2) Layer kecil dengan lebih sedikit neuron (misalnya [4] atau [8]) cocok untuk masalah sederhana tetapi kurang optimal untuk data kompleks.

2. Activation Functions

- **Pengamatan:** Aktivasi menentukan bagaimana sinyal non-linear diperkenalkan ke model.
- **Hasil:**
 - 1) ReLU: Performa terbaik secara umum, cocok untuk deep networks karena mengatasi vanishing gradient.
 - 2) Sigmoid: Akurasi cukup baik tetapi cenderung lambat karena vanishing gradient.
 - 3) Tanh: Cocok untuk data dengan distribusi yang lebih terpusat tetapi kurang efektif dibandingkan ReLU.
 - 4) Identity: Performa terendah karena linearitasnya.

3. Epochs

- **Pengamatan:** Jumlah epoch menentukan seberapa banyak model dilatih pada dataset.
- **Hasil:**
 - 1) Epoch rendah (misalnya 1, 10) seringkali tidak cukup untuk konvergensi.
 - 2) Epoch tinggi (misalnya 100, 250) memberikan akurasi lebih baik tetapi meningkatkan risiko overfitting.

4. Learning Rate

- **Pengamatan:** Learning rate memengaruhi kecepatan dan stabilitas pelatihan.

- **Hasil:**

- 1) Learning rate tinggi (10, 1) menghasilkan pelatihan yang tidak stabil atau gagal konvergen.
- 2) Learning rate rendah (0.001, 0.0001) memberikan hasil yang stabil tetapi pelatihan lebih lambat.

5. Batch Size

- **Pengamatan:** Batch size memengaruhi stabilitas dan efisiensi pelatihan.

- **Hasil:**

- 1) Batch size kecil (16, 32) lebih lambat tetapi memberikan pembaruan gradien yang lebih akurat.
- 2) Batch size besar (256, 512) mempercepat pelatihan tetapi rentan terhadap generalisasi buruk.