

NAMA = AZIZAH RAHMA ASRI  
NIM = 1103213025  
UTS DEEP LEARNING

## 1. Arsitektur CNN dengan X lapisan konvolusi menghasilkan akurasi training 98% tetapi akurasi validasi 62%.

- **Vanishing Gradient:** Fenomena ini terjadi ketika gradien (turunan error) yang mengalir mundur dari output ke input menjadi sangat kecil, terutama pada lapisan awal CNN. Hal ini menyebabkan pembaruan bobot menjadi tidak signifikan, sehingga pembelajaran terhambat.
  - **Mitigasi:**
    - Gunakan **aktivasi ReLU atau LeakyReLU** (yang sudah Anda gunakan) karena membantu menghindari gradien mendekati nol.
    - **Batch Normalization** di awal lapisan dapat membantu menstabilkan distribusi aktivasi dan mempertahankan gradien.
  - **Mengapa BN justru memperburuk generalisasi setelah lapisan ke-Y?**
    - Batch Normalization setelah lapisan tertentu dapat memperkenalkan noise statistik yang tidak stabil jika batch size kecil.
    - BN juga bisa mengganggu representasi jika digunakan terlalu banyak atau di layer yang tidak terlalu dalam.
  - **Alternatif stabilisasi pembelajaran:**
    - Gunakan **Group Normalization** atau **Layer Normalization** (terutama saat batch size kecil).
    - Gunakan **skip connections** seperti pada ResNet untuk membantu aliran gradien.
- 

## 2. Loss training stagnan di nilai tinggi setelah XXX epoch.

- **Tiga penyebab potensial:**
  1. **Learning Rate terlalu rendah:** Gradien sangat kecil sehingga pembelajaran tidak berkembang.
  2. **Inisialisasi bobot buruk:** Jika bobot awal terlalu kecil/besar, gradien bisa menghilang/meledak.
  3. **Model terlalu kompleks untuk data sederhana:** Overparameterisasi menyebabkan kesulitan optimasi awal.
- **Cyclic Learning Rate (CLR):**

NAMA = AZIZAH RAHMA ASRI  
NIM = 1103213025  
UTS DEEP LEARNING

- CLR mengatur LR naik-turun secara periodik untuk membantu keluar dari **local minima** atau **saddle points**.
  - **Momentum pada SGD:**
    - Menyimpan sebagian gradien sebelumnya agar arah langkah lebih stabil.
    - Meningkatkan kecepatan konvergensi dan membantu "melompati" lembah dangkal dalam loss landscape.
- 

### 3. Penggunaan ReLU tidak meningkatkan akurasi setelah 50 epoch.

- **Dying ReLU:**
    - ReLU meng-output nol jika input negatif. Setelah beberapa epoch, neuron bisa "mati" (selalu output nol), sehingga tidak berkontribusi pada pembelajaran.
  - **Dampak:**
    - Gradien menjadi nol untuk neuron yang mati, menyebabkan hilangnya kapabilitas representasi.
  - **Solusi:**
    - Ganti ReLU dengan **LeakyReLU** atau **ELU** untuk memastikan neuron tetap aktif sebagian meski input negatif.
- 

### 4. AUC Spesies X stagnan di 0.55, sementara kelas lain >0.85.

- **Class-weighted loss gagal karena:**
  - Bobot hanya memperkuat gradien, tapi tidak menangani distribusi/representasi fitur buruk.
- **Tiga penyebab potensial:**
  1. **Data imbalance ekstrem:** Jumlah gambar Spesies X terlalu sedikit.
  2. **Fitur visual Spesies X mirip dengan kelas lain:** Sulit dibedakan oleh CNN.
  3. **Augmentasi tidak relevan:** Augmentasi tidak cukup beragam untuk kelas minoritas.
- **Solusi alternatif:**
  - Gunakan **oversampling** atau **GAN-based augmentation**.

NAMA = AZIZAH RAHMA ASRI  
NIM = 1103213025  
UTS DEEP LEARNING

- Gunakan **focal loss** untuk memberi penalti lebih besar pada prediksi salah dari kelas minoritas.

---

**5. Peningkatan kompleksitas model menurunkan akurasi validasi dari 85% ke 65%.**

- **Overfitting:** Model belajar terlalu detail dari data training, termasuk noise, sehingga gagal generalisasi ke data baru.
- **Penambahan kapasitas model ≠ performa lebih baik**, karena:
  - Model kompleks butuh data lebih banyak untuk generalisasi.
  - Model besar lebih rentan overfitting pada dataset kecil.
- **Tiga kesalahan desain yang bisa memicu degradasi:**
  1. **Terlalu banyak layer konvolusi tanpa regularisasi memadai.**
  2. **Dropout atau pooling yang terlalu sedikit.**
  3. **Tanpa early stopping, atau tidak memantau val\_loss secara ketat.**