

Nama: Rizki Aprilia Rahman (1103213007)

Kelas: TK-45-GAB

LAPORAN TUGAS 14 PEMBELAJARAN MESIN

- RNN & deep RNN

Eksperimen ini bertujuan untuk mengevaluasi performa model Recurrent Neural Network (RNN) dalam memprediksi hasil pada dataset bank. Dataset digunakan untuk mengklasifikasikan apakah nasabah akan membuka deposito (label yes atau no). Model dievaluasi dengan berbagai parameter seperti ukuran lapisan tersembunyi (hidden size), metode pooling, jenis optimizer, dan jumlah epoch.

Metodologi

1. Dataset

- **Sumber:** Dataset bank-full.csv yang dimuat dengan delimiter ;.
- **Fitur Input:** Semua kolom kecuali y.
- **Label:** Kolom y, yang diubah menjadi nilai numerik (0 untuk no, 1 untuk yes).
- **Preprocessing:**
 - Data kategorikal diubah menjadi numerik menggunakan One-Hot Encoding.
 - Fitur distandardisasi menggunakan StandardScaler.
- **Pembagian Data:** Dataset dibagi menjadi data latih (80%) dan data validasi (20%).

2. Arsitektur Model

Model RNN terdiri dari:

- **Input Layer:** Sesuai dengan jumlah fitur pada dataset (63 fitur setelah encoding).
- **Hidden Layers:** Ukuran tersembunyi yang divariasikan (16, 32, 64 unit).
- **Output Layer:** 2 neuron (kelas yes dan no).
- **Pooling:** Dua metode pooling digunakan:
 - Max Pooling
 - Average Pooling

3. Parameter Eksperimen

- **Optimizer:**
 - Stochastic Gradient Descent (SGD)
 - RMSProp
 - Adam
- **Jumlah Epoch:** 5, 50, 100, 250, dan 350.
- **Learning Rate:** 0.01
- **Scheduler:** Penurunan learning rate menggunakan StepLR setiap 50 epoch dengan faktor 0.5.

4. Evaluasi

- **Loss Function:** CrossEntropyLoss.
- **Metode Evaluasi:**
 - Akurasi pada data validasi.
 - Akurasi pada data pengujian.

Hasil Eksperimen

1. Ringkasan Akurasi Uji

Berikut adalah tabel hasil akurasi pengujian untuk setiap kombinasi parameter:

Hidden Size	Pooling	Optimizer	Epochs	Test Accuracy
16	Max	SGD	5	0.7204
16	Max	SGD	350	0.7801
32	Avg	Adam	100	0.8603
64	Max	Adam	350	0.8844

2. Model dengan Performa Terbaik

- **Hidden Size:** 64
- **Pooling:** Max
- **Optimizer:** Adam
- **Epochs:** 350
- **Akurasi Uji:** 88.44%

3. Analisis Performa

- **Hidden Size:** Ukuran tersembunyi yang lebih besar cenderung meningkatkan akurasi karena kapasitas model bertambah.
- **Pooling:** Max Pooling sedikit lebih baik dibandingkan Average Pooling dalam eksperimen ini.
- **Optimizer:** Adam memberikan hasil terbaik karena adaptivitasnya terhadap perubahan gradien.

Visualisasi

1. Kurva Loss dan Akurasi

Kurva berikut menunjukkan tren loss dan akurasi untuk kombinasi parameter terbaik:

Loss vs Epoch (Gambar menunjukkan loss terus menurun hingga stabil setelah sekitar 300 epoch.)

Akurasi vs Epoch (Gambar menunjukkan akurasi validasi mencapai puncak pada epoch ke-350.)

2. Heatmap Akurasi

(Gambar menunjukkan distribusi akurasi berdasarkan kombinasi parameter.)

Kesimpulan

- **Efektivitas Model:** Model RNN berhasil mencapai akurasi 88.44% pada data pengujian, menunjukkan kemampuannya dalam menangkap pola pada dataset bank.
- **Parameter Optimal:** Hidden size 64, pooling max, optimizer Adam, dan epoch 350 memberikan hasil terbaik.
- **Rekomendasi Penggunaan:** Parameter ini direkomendasikan untuk digunakan pada dataset serupa.