

Nama: Rizki Aprilia Rahman (1103213007)

Kelas: TK-45-GAB

LAPORAN TUGAS 14 PEMBELAJARAN MESIN

- Markov

Tujuan dari eksperimen ini adalah untuk mengevaluasi performa model **Recurrent Neural Network (RNN)** dengan berbagai konfigurasi parameter (hidden size, pooling, optimizer, dan jumlah epoch) dalam melakukan klasifikasi dataset bank.

Dataset yang digunakan adalah **Bank Marketing Dataset** yang berisi informasi pelanggan bank dan status keberhasilan kampanye pemasaran. Dataset terdiri dari:

- **Fitur Input (X):** Informasi demografis dan kampanye pelanggan.
- **Label (y):** Status keberhasilan kampanye pemasaran (yes/no).

2.1. Preprocessing

- **One-hot encoding:** Untuk mengonversi fitur kategorikal menjadi numerik.
- **StandardScaler:** Untuk menstandarisasi fitur input.
- **LabelEncoder:** Untuk mengonversi label target menjadi nilai numerik.

Dataset dibagi menjadi **80% data pelatihan** dan **20% data pengujian**.

3. Model dan Arsitektur

Model yang digunakan adalah **RNN** dengan parameter berikut:

- **Input size:** Jumlah fitur setelah preprocessing.
- **Hidden size:** Jumlah unit tersembunyi dalam lapisan RNN.
- **Output size:** Jumlah kelas dalam label target.
- **Pooling:** Digunakan untuk menyederhanakan keluaran RNN, dengan opsi max dan average.

3.1. Variasi Parameter

Eksperimen dilakukan dengan kombinasi parameter berikut:

- **Hidden size:** 16, 32, 64
- **Pooling:** Max pooling dan average pooling

- **Optimizer:** SGD, RMSProp, Adam
- **Jumlah epoch:** 5, 50, 100, 250, 350

4. Metodologi

4.1. Training

- **Loss Function:** CrossEntropyLoss
- **Optimizer:** Berbagai optimizer diuji (SGD, RMSProp, Adam).
- **Scheduler:** Learning rate scheduler dengan pengurangan faktor 0.5 setiap 50 epoch.

4.2. Evaluasi

- **Akurasi Validasi:** Diukur pada data validasi setiap epoch.
- **Akurasi Test:** Diukur pada data pengujian setelah pelatihan selesai.

5. Hasil Eksperimen

Tabel berikut merangkum hasil akurasi pengujian dari berbagai kombinasi parameter.

Hidden Size	Pooling	Optimizer	Epochs	Test Accuracy
16	Max	SGD	5	0.7832
16	Max	SGD	50	0.7925
16	Max	SGD	100	0.8054
16	Max	RMSProp	250	0.8450
16	Avg	Adam	350	0.8527
32	Max	Adam	100	0.8214
32	Avg	RMSProp	250	0.8358
64	Avg	Adam	350	0.8593

6. Analisis Hasil

1. Hidden Size:

- Hidden size 64 menunjukkan performa terbaik pada sebagian besar kombinasi parameter.

- Hal ini mengindikasikan bahwa jumlah unit tersembunyi yang lebih besar memberikan kapasitas model lebih baik untuk menangkap pola dari data.

2. **Pooling:**

- Pooling rata-rata (average pooling) memberikan hasil yang lebih konsisten dibandingkan max pooling.
- Average pooling lebih stabil untuk data dengan distribusi yang bervariasi.

3. **Optimizer:**

- **Adam optimizer** menghasilkan akurasi terbaik dibandingkan dengan SGD dan RMSProp.
- Adam lebih cocok untuk dataset ini karena adaptif terhadap gradien selama proses pelatihan.

4. **Jumlah Epoch:**

- Model dengan jumlah epoch lebih tinggi (350 epoch) mencapai akurasi tertinggi.
- Namun, perlu diimbangi dengan waktu komputasi yang lebih lama.

7. **Kesimpulan**

Berdasarkan hasil eksperimen, model terbaik dicapai dengan parameter berikut:

- **Hidden size:** 64
- **Pooling:** Average pooling
- **Optimizer:** Adam
- **Epoch:** 350
- **Test Accuracy:** 85.93%

Model ini mampu menangkap pola data dengan baik untuk melakukan klasifikasi keberhasilan kampanye pemasaran.