

# ANALISA PERBANDINGAN HASIL EVALUASI MODEL RNN, LSTM, DAN GRU

Perbandingan Hasil Evaluasi Model RNN, LSTM, dan GRU

Model	Akurasi	Presisi	Recall	F1 Score	AUC
RNN (PyTorch)	0.8333	0.8290	0.8398	0.8344	0.9030
RNN (TensorFlow)	0.8582	0.8569	0.8602	0.8585	0.8582
LSTM (PyTorch)	0.8717	0.8556	0.8943	0.8745	0.9410
LSTM (TensorFlow)	0.8609	0.8426	0.8717	0.8569	0.8544
GRU (PyTorch)	0.8609	0.8880	0.8259	0.8558	0.8609
GRU (TensorFlow)	0.8609	0.8880	0.8259	0.8558	0.8609

## HASIL MASING MASING :

### 1. AKURASI

Model	Akurasi
RNN (PyTorch)	0.8333
RNN (TensorFlow)	0.8582
LSTM (PyTorch)	0.8717
LSTM (TensorFlow)	0.8609
GRU (PyTorch)	0.8609
GRU (TensorFlow)	0.8609

#### Interpretasi:

- LSTM (PyTorch) memberikan akurasi terbaik dengan 87.17%, yang mengindikasikan bahwa model ini memiliki performa yang lebih baik dalam mengklasifikasikan ulasan menjadi positif atau negatif dibandingkan dengan model lainnya.
- RNN (TensorFlow) juga menunjukkan hasil yang cukup baik, dengan akurasi 85.82%, lebih tinggi daripada RNN pada PyTorch.
- GRU (PyTorch, TensorFlow) dan LSTM (TensorFlow) memiliki akurasi yang serupa, tetapi LSTM sedikit lebih unggul, khususnya di TensorFlow.

---

### 2. Presisi (Precision):

Model	Presisi
RNN (PyTorch)	0.8290
RNN (TensorFlow)	0.8569
LSTM (PyTorch)	0.8556
LSTM (TensorFlow)	0.8426
GRU (PyTorch)	0.8880
GRU (TensorFlow)	0.8880

#### Interpretasi:

- GRU (PyTorch dan TensorFlow) menunjukkan presisi tertinggi dengan 88.80%, yang berarti model ini memiliki kemampuan terbaik untuk meminimalkan prediksi positif palsu.
  - LSTM (PyTorch) memiliki presisi yang sangat tinggi juga, tetapi sedikit lebih rendah dibandingkan dengan GRU, 85.56%.
  - RNN (TensorFlow) menunjukkan nilai presisi yang lebih tinggi (85.69%) daripada RNN (PyTorch), yang mencerminkan kemampuannya untuk meminimalkan prediksi positif palsu.
- 

#### 3. Recall:

Model	Recall
RNN (PyTorch)	0.8398
RNN (TensorFlow)	0.8602
LSTM (PyTorch)	0.8943
LSTM (TensorFlow)	0.8717
GRU (PyTorch)	0.8259
GRU (TensorFlow)	0.8259

#### Interpretasi:

- LSTM (PyTorch) memiliki recall terbaik 89.43%, yang menunjukkan bahwa model ini lebih baik dalam mendeteksi ulasan positif, mengurangi jumlah ulasan positif yang terlewatkan.
  - RNN (TensorFlow) juga memiliki recall yang sangat baik (86.02%), sedangkan GRU menunjukkan recall yang sedikit lebih rendah (82.59%), baik pada TensorFlow maupun PyTorch.
  - LSTM (TensorFlow) masih menunjukkan recall yang lebih baik (87.17%) dibandingkan GRU, namun tidak setinggi model LSTM (PyTorch).
-

#### 4. F1 Score:

Model	F1 Score
RNN (PyTorch)	0.8344
RNN (TensorFlow)	0.8585
LSTM (PyTorch)	0.8745
LSTM (TensorFlow)	0.8569
GRU (PyTorch)	0.8558
GRU (TensorFlow)	0.8558

#### Interpretasi:

- LSTM (PyTorch) memiliki F1 score tertinggi 87.45%, yang menunjukkan keseimbangan terbaik antara presisi dan recall, menghasilkan model yang dapat mendeteksi ulasan positif dan negatif dengan sangat baik.
- RNN (TensorFlow) menunjukkan F1 score yang cukup baik (85.85%), lebih tinggi dari RNN (PyTorch).
- GRU (PyTorch dan TensorFlow) memiliki F1 score yang serupa (85.58%), lebih rendah dibandingkan dengan LSTM (PyTorch) tetapi sangat baik untuk model yang lebih efisien secara komputasi.

---

#### 5. AUC (Area Under the Curve):

Model	AUC
RNN (PyTorch)	0.9030
RNN (TensorFlow)	0.8582
LSTM (PyTorch)	0.9410
LSTM (TensorFlow)	0.8544
GRU (PyTorch)	0.8609
GRU (TensorFlow)	0.8609

### Interpretasi:

- LSTM (PyTorch) memiliki AUC tertinggi 94.10%, yang menunjukkan bahwa model ini lebih baik dalam membedakan antara ulasan positif dan negatif. Model ini memiliki kemampuan diskriminasi yang sangat baik.
  - RNN (PyTorch) juga memiliki AUC yang sangat baik (90.30%), menandakan kemampuan yang baik dalam memisahkan ulasan positif dan negatif, meskipun sedikit lebih rendah dibandingkan dengan LSTM.
  - GRU (PyTorch dan TensorFlow) dan RNN (TensorFlow) memiliki AUC yang lebih rendah, namun tetap cukup baik untuk membedakan antara kelas positif dan negatif.
- 

## Kesimpulan

1. **Model LSTM (PyTorch)** menunjukkan kinerja terbaik dalam semua metrik evaluasi (Akurasi, Presisi, Recall, F1 Score, dan AUC). Dengan AUC yang paling tinggi (94.10%), model ini adalah pilihan terbaik jika akurasi dan kemampuan membedakan antara ulasan positif dan negatif adalah prioritas utama. LSTM (PyTorch) adalah model terbaik secara keseluruhan di semua metrik evaluasi, dengan kemampuan untuk menangani dependensi jangka panjang dalam data teks.
  2. **Model GRU (PyTorch dan TensorFlow)** sangat efisien dalam hal presisi dan F1 score meskipun memiliki kinerja sedikit lebih rendah dalam hal recall dan AUC dibandingkan dengan LSTM. GRU bisa menjadi pilihan yang baik jika kecepatan pelatihan dan efisiensi komputasi lebih diutamakan. GRU (PyTorch dan TensorFlow) memberikan efisiensi komputasi yang lebih baik dan presisi yang tinggi, tetapi dengan sedikit penurunan kinerja dalam recall dan AUC dibandingkan dengan LSTM.
  3. **RNN (TensorFlow)** memberikan kinerja yang sangat baik dengan presisi dan recall yang tinggi, namun LSTM dan GRU menunjukkan kinerja lebih baik dalam hal F1 score dan AUC. RNN (TensorFlow), meskipun lebih sederhana dan kurang efektif dibandingkan LSTM dan GRU, tetap memberikan kinerja yang cukup baik, terutama dalam presisi.
-