

sentiment analysis

Presented by : Kelompok 3

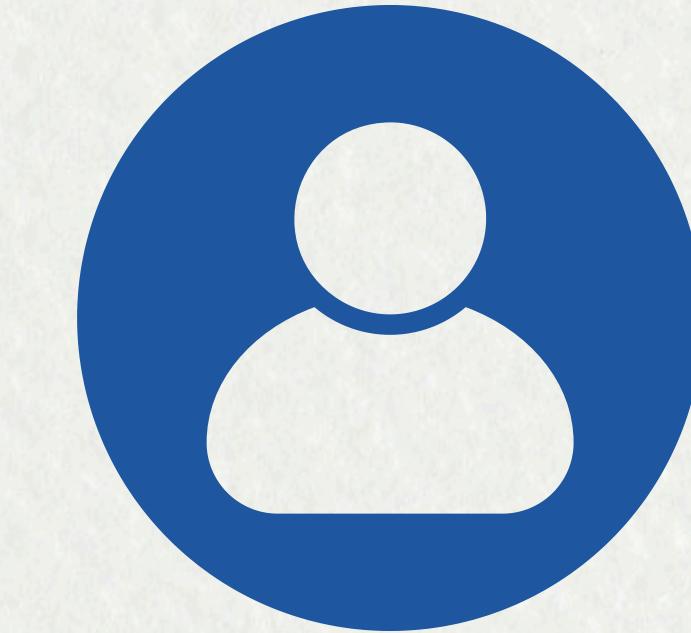
OUR TEAM



Andi Alisha Faiqihah
(H071221010)



Dewa Ayu Eka Natalia Pratiwi
(H071221021)



Laode Fahmi Hidayat
(H071221022)

BACKGROUND

Media sosial telah merevolusi komunikasi global, membawa manfaat sekaligus memunculkan fenomena negatif seperti komentar buruk yang merugikan secara psikiatrik lebih besar daripada bullying langsung. Komentar negatif ini tidak hanya merendahkan martabat individu dan menyebabkan stres, tetapi juga berkontribusi pada masalah kesehatan mental seperti depresi dan perilaku tidak etis, terutama pada anak-anak. Untuk memahami fenomena ini dan mengembangkan metode mitigasi, diperlukan dataset analisis sentimen yang representatif guna mengidentifikasi pola, intensitas, dan profil pengguna komentar negatif. Dataset ini juga mendukung pengembangan sistem deteksi otomatis untuk menghapus komentar tidak pantas secara efisien, sekaligus mendorong edukasi publik tentang pentingnya berperilaku positif di internet.

DATASET

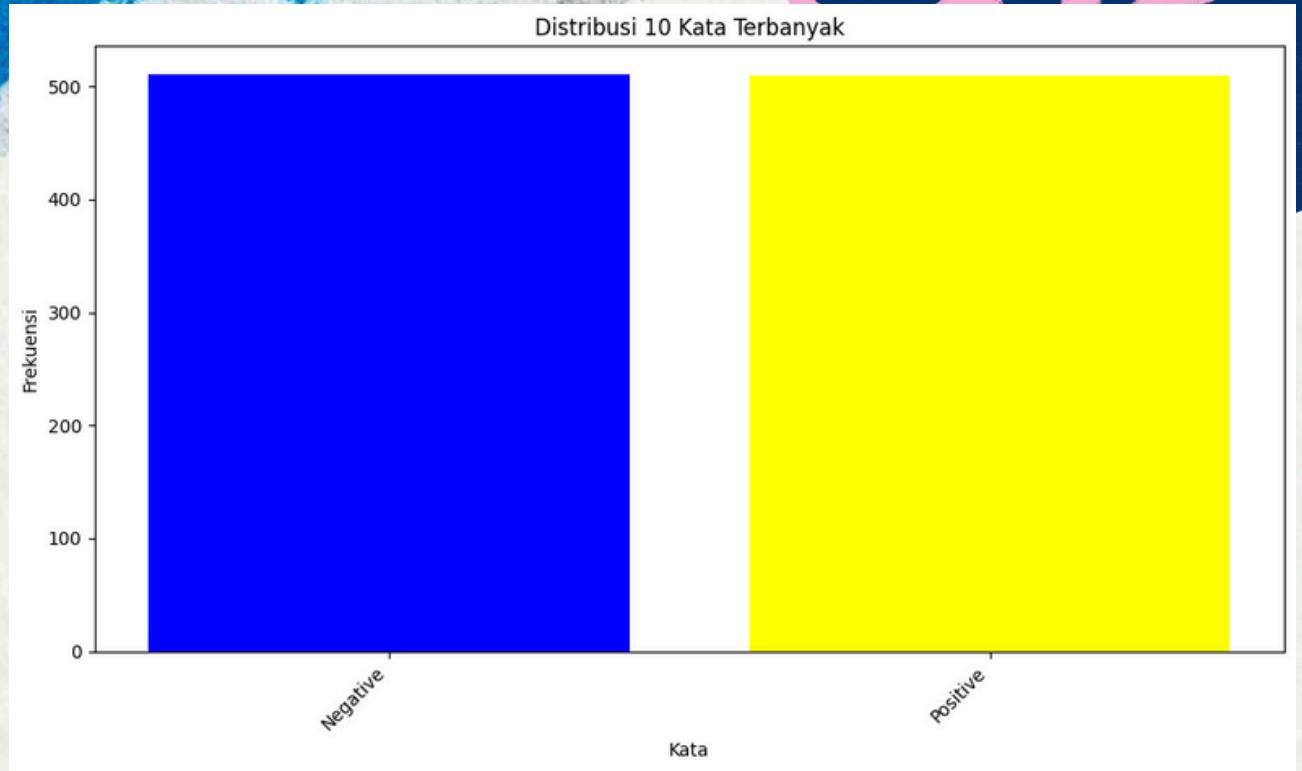
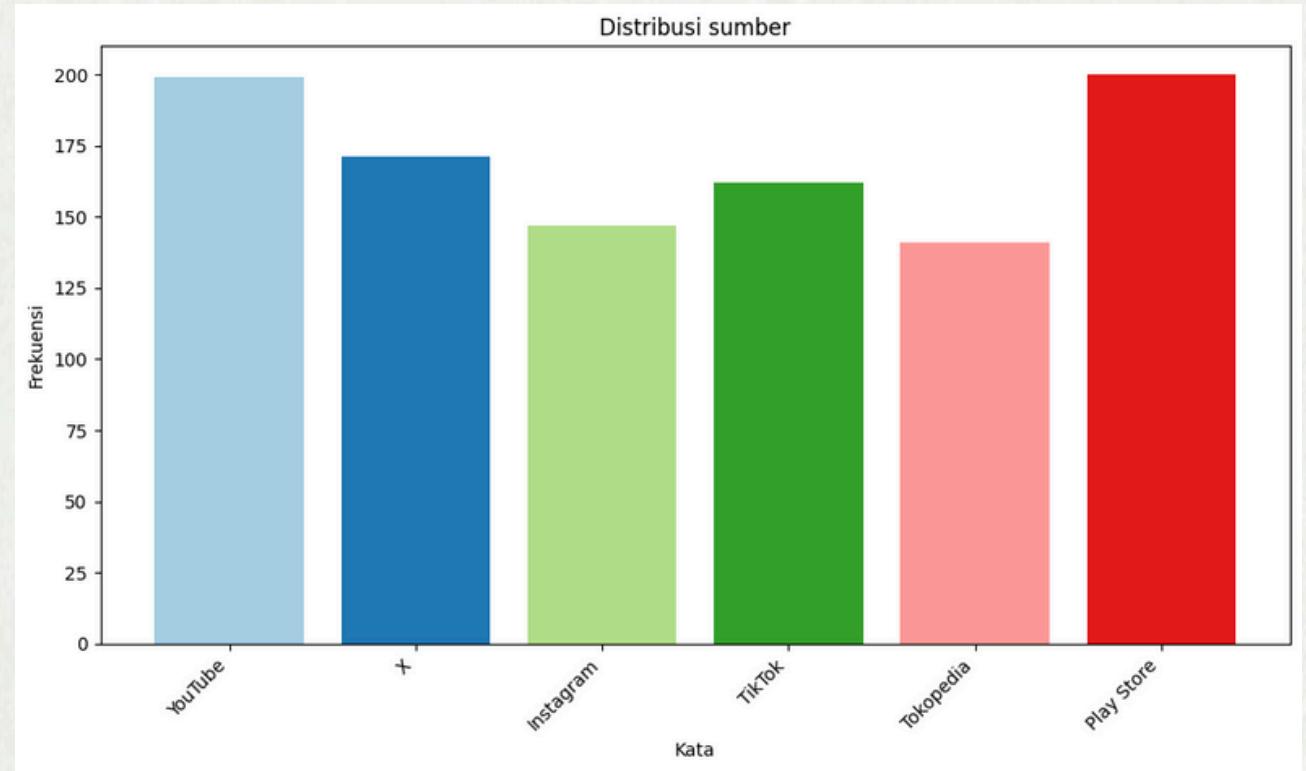
| Romentar | Timestap | Sumber | Label |
|----------|----------|--------|-------|
|----------|----------|--------|-------|

PROGRESS DATASET

TABEL JUMLAH DATA SETIAP PEKAN

| 25 Oktober 2024 | 1 November 2024 | 8 November 2024 | 15 November 2024 | 22 November 2024 | 28 November 2024 | Total Data |
|-----------------|-----------------|-----------------|------------------|------------------|------------------|------------|
| 81 | 22 | 71 | 341 | 88 | 417 | 1020 |

EDA



MODEL



LSTM



WORD2VEC-CBOW

ALASAN



Alasan LSTM digunakan:

- Dapat mengingat informasi dalam jangka waktu yang lebih lama dibandingkan RNN biasa.
- Cocok untuk masalah yang melibatkan konteks jangka panjang, seperti analisis teks dan prediksi urutan.



Alasan kami memilih Word2Vec:

- Hasil uji coba dengan Word2Vec memberikan hasil yang lebih baik dalam memahami hubungan antar kata.
- Memungkinkan representasi kata dalam ruang vektor yang lebih efisien untuk model-model lainnya.

OPTIMASI



Optimasi dari LSTM yang digunakan:

- Melakukan Analisis Dan Penyesuaian Data
- menggunakan data bertahap



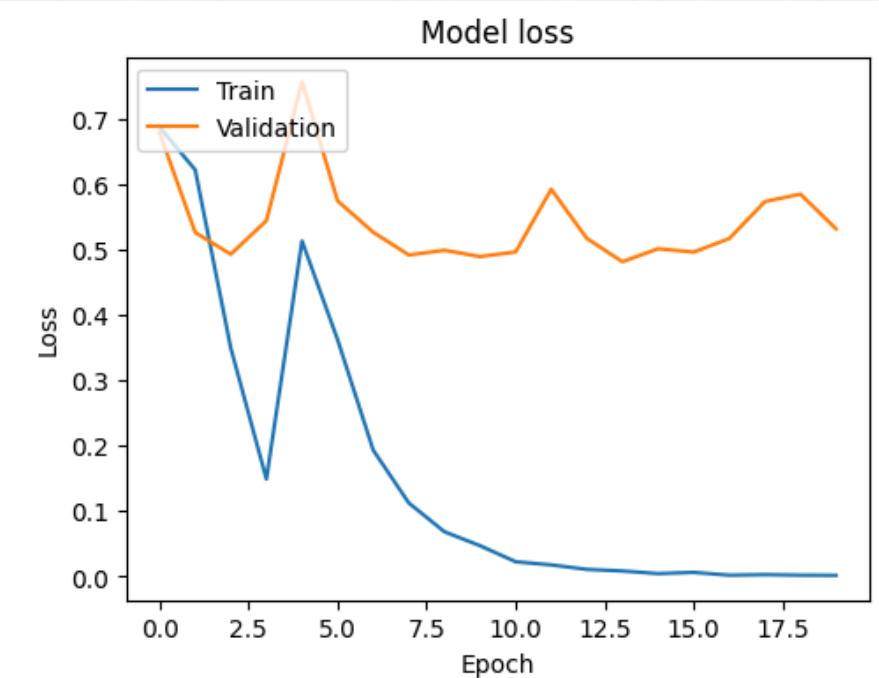
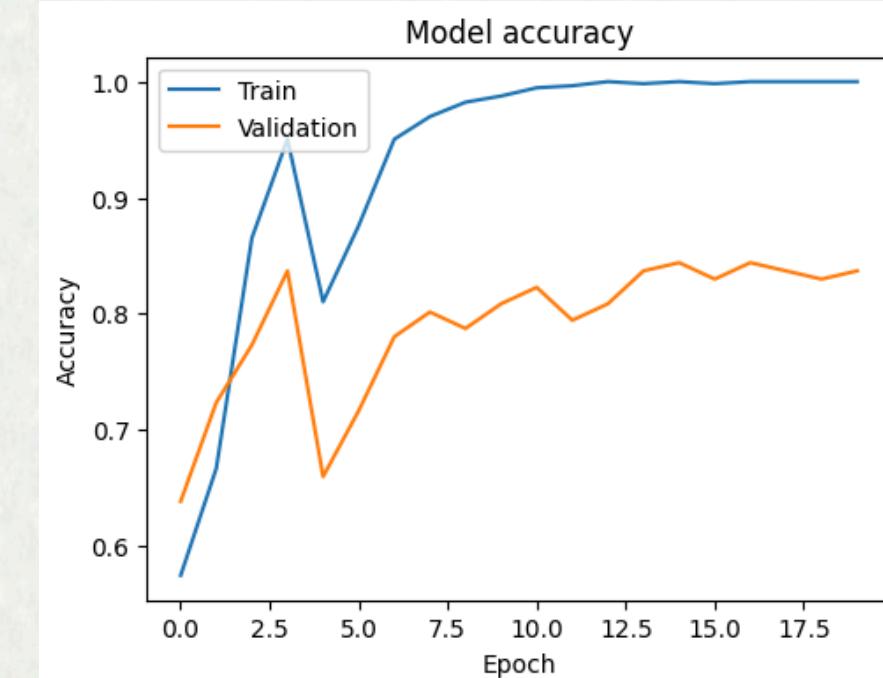
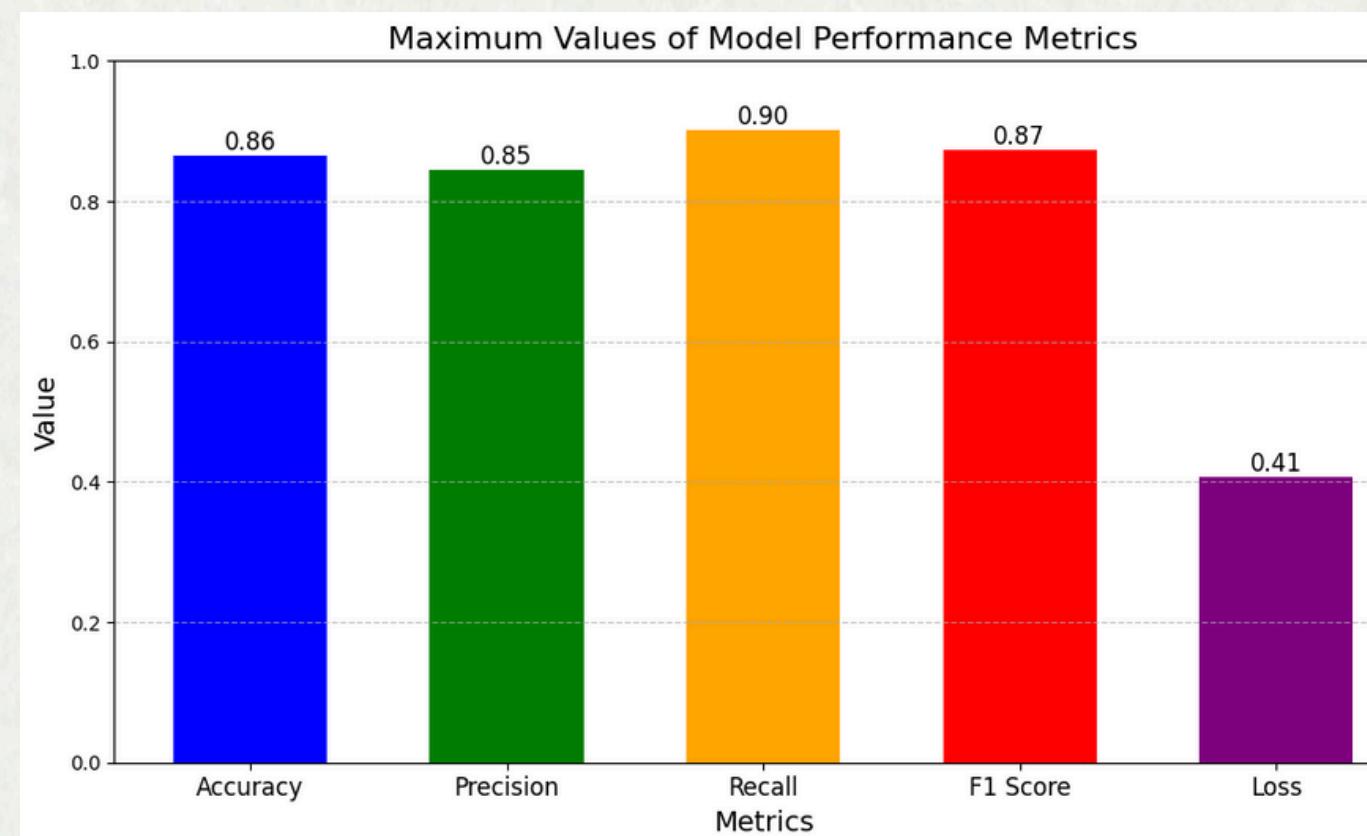
Optimasi dalam Word2Vec:

- Melakukan Training secara berulang
- Melakukan penyusunan data acak

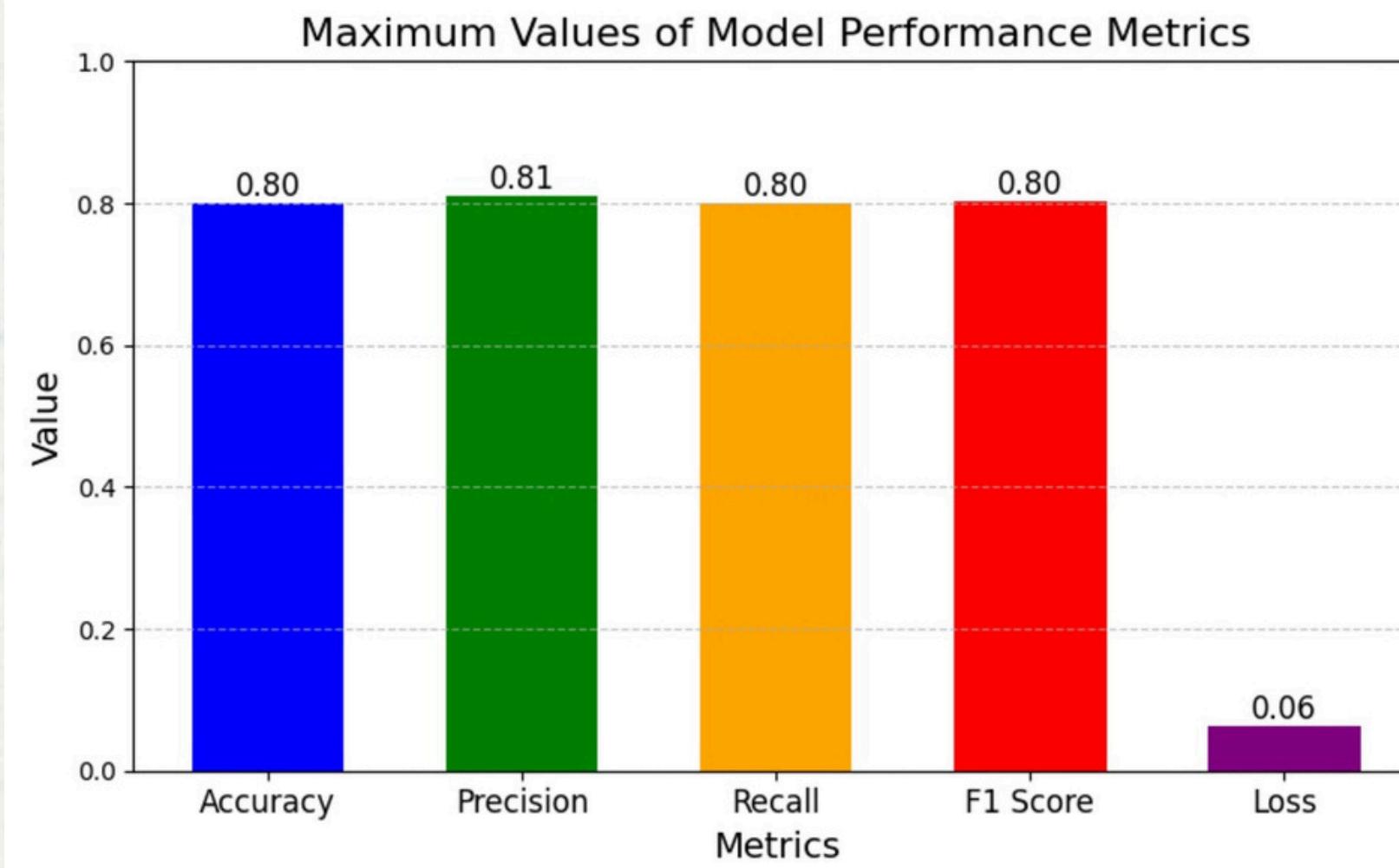
Evaluasi

- Kami Melakukan Analisis Dan melihat bahwa tanda baca dan emoji terkadang mempengaruhi emosi yang terkandung dalam sebuah kalimat sehingga saya memutuskan untuk tidak menghilang tanda titik dan emoji dalam proses training terutama untuk model Word2vec
- Model W2V ternyata membutuhkan epoch training yang tinggi untuk membuat model dapat mencapai hasil yang baik
- Evaluasi, perbaikan , dan variasi data menjadi solusi yang sangat mendukung kami

HASIL LSTM SEBELUM LENGRAP

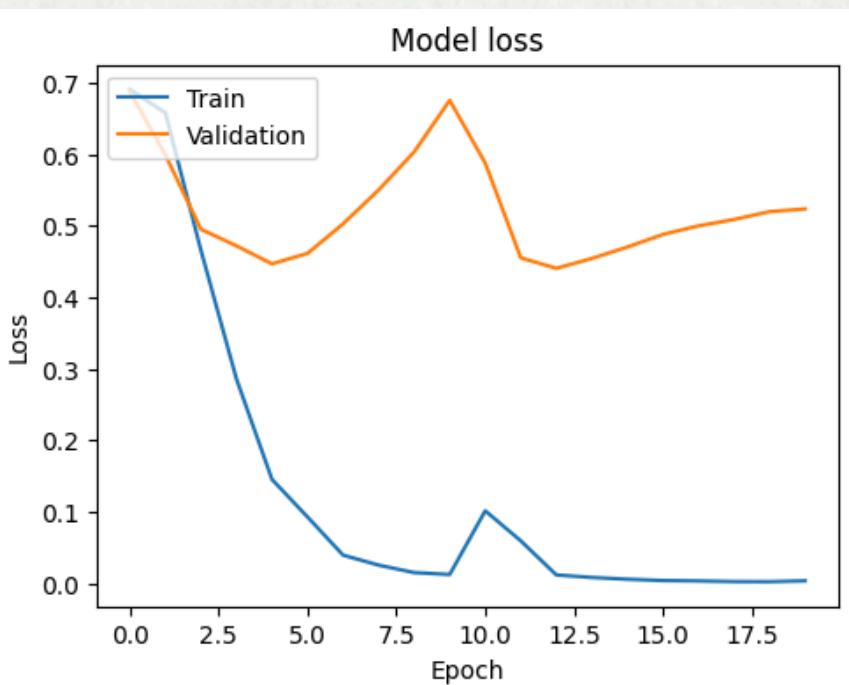
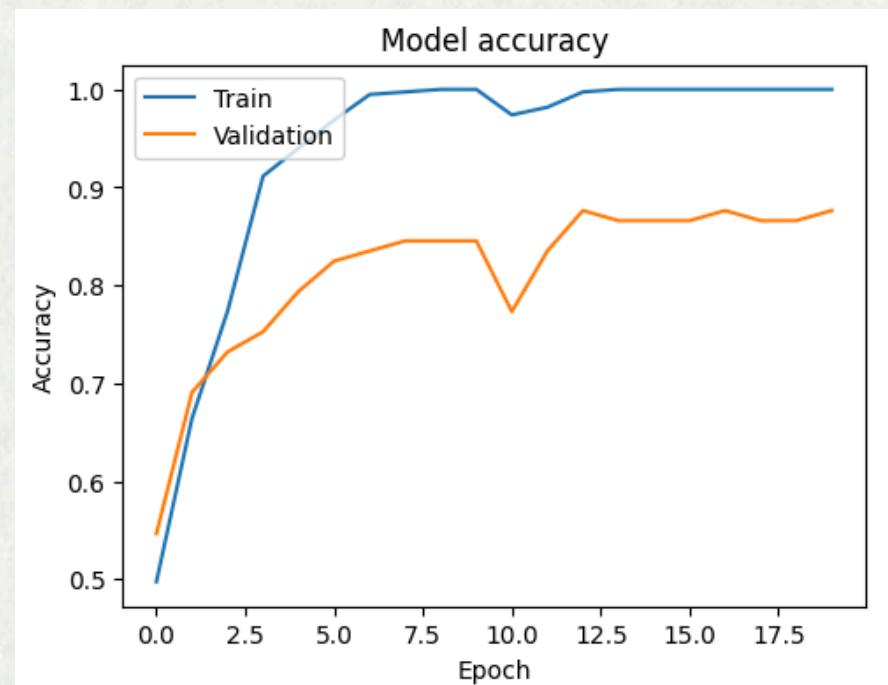
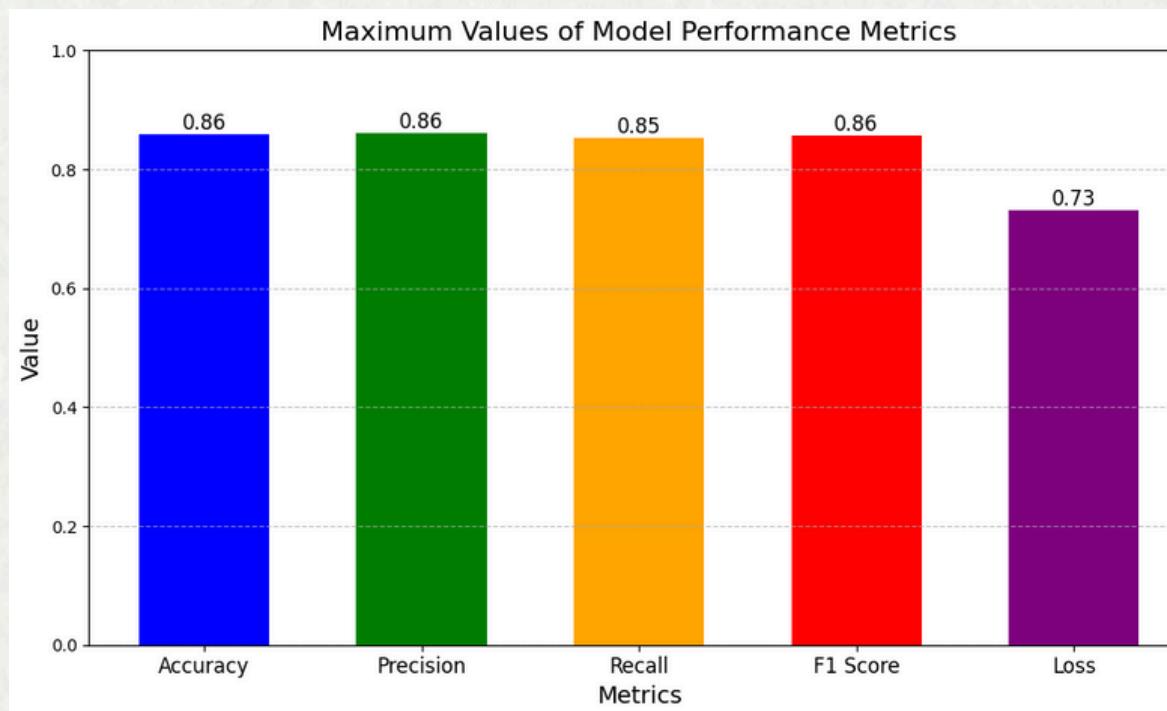


HASIL W2V SEBELUM LENGRAP



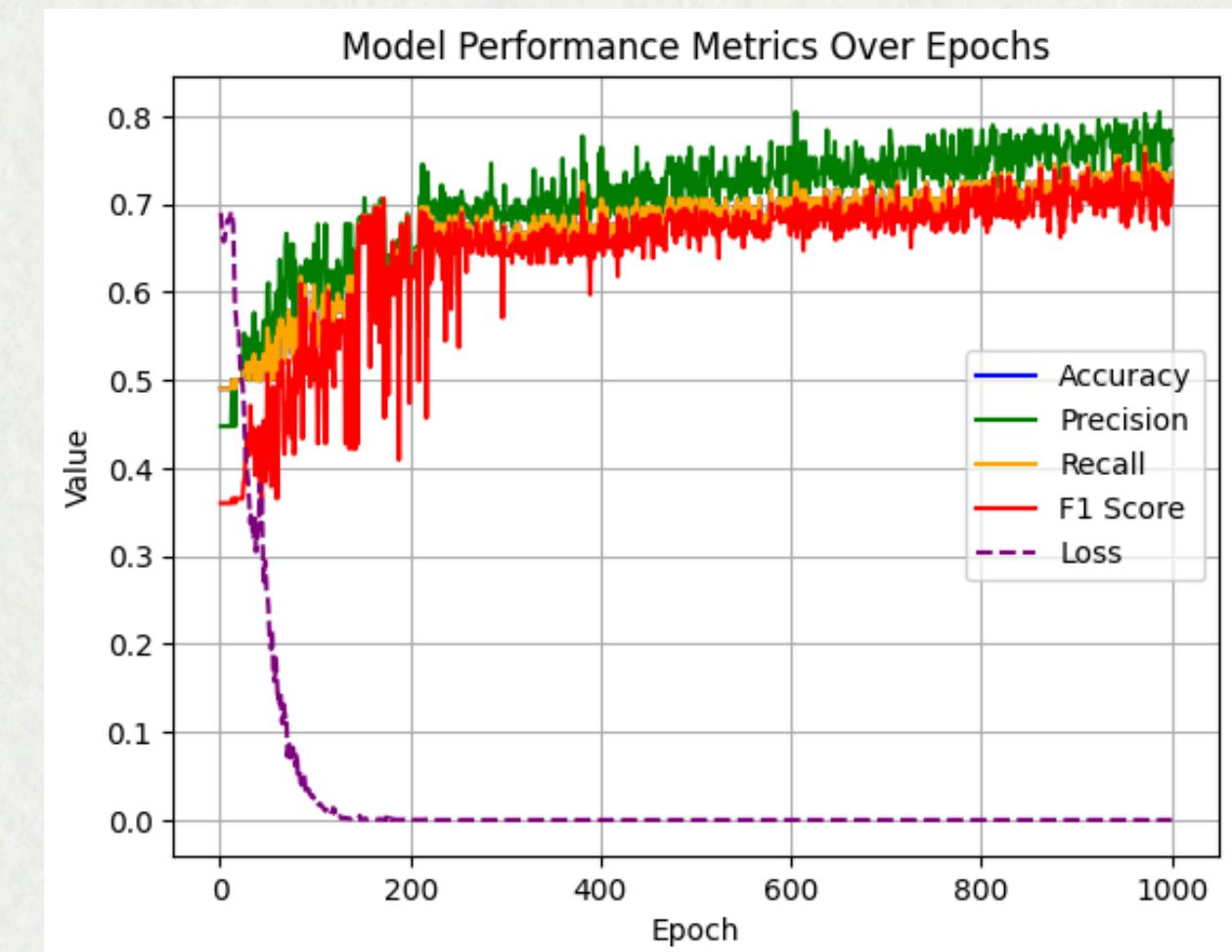
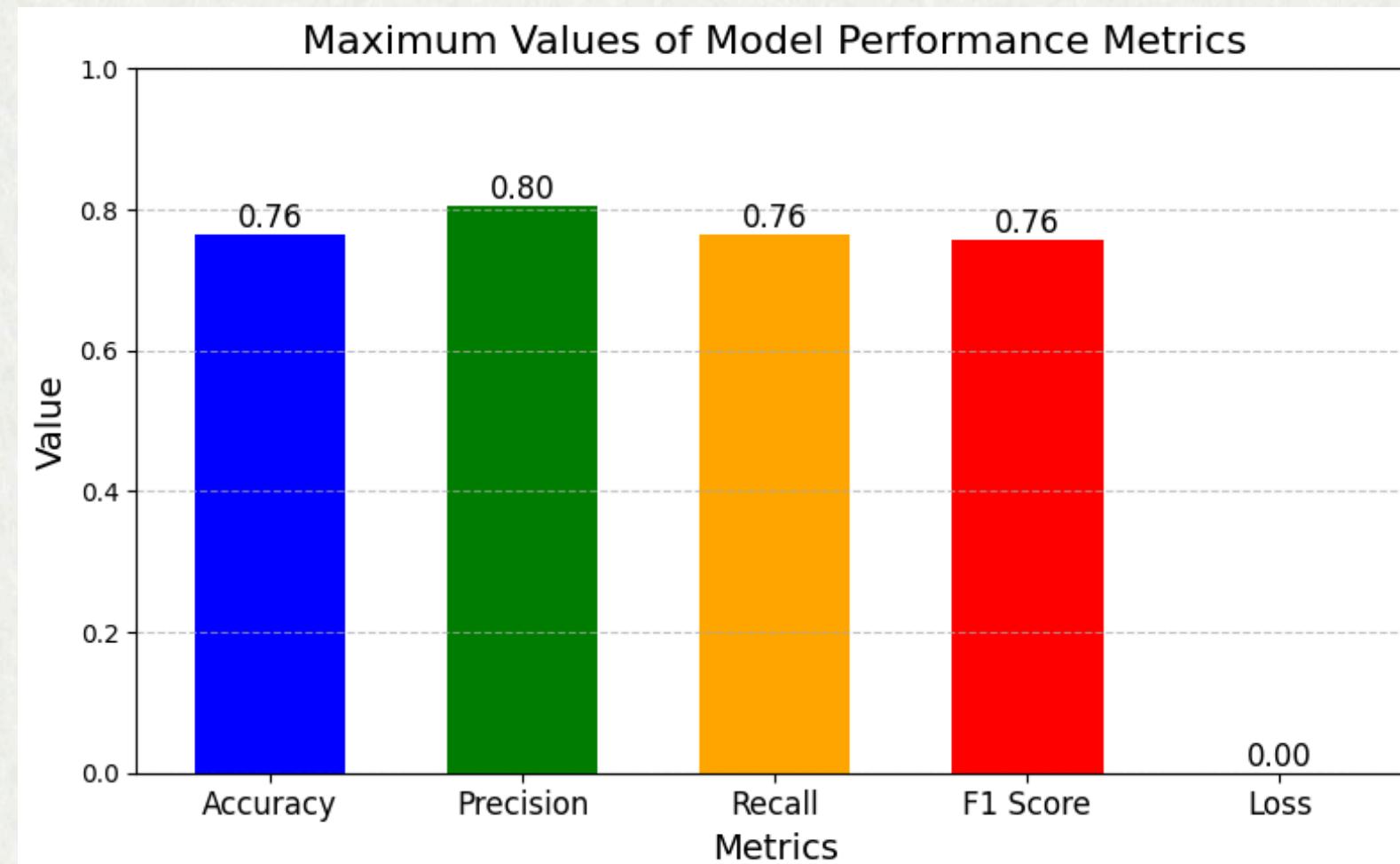
HASIL AKHIR

LSTM



HASIL AKHIR

W2V



CONCLUSION

Untuk penelitian di masa depan, disarankan untuk mengeksplorasi pengaruh metode preprocessing data terhadap performa Word2Vec-CBOW, serta menguji arsitektur LSTM yang lebih ringan untuk mengurangi waktu pelatihan. Selain itu, penelitian dapat diperluas ke dataset lain dengan karakteristik berbeda untuk mengonfirmasi generalisasi hasil ini. Kombinasi Word2Vec-CBOW dan LSTM dalam model hybrid juga dapat menjadi arah yang menjanjikan untuk mencapai hasil yang lebih optimal, Sehingga Bisa dikatakan bahwa Word2Vec-CBOW lebih stabil dalam proses pembelajaran dan LSTM lebih cepat dalam mendapat akurasi yang tinggi

THANK
YOU

Presented by : Kelompok 3