

Exploring User Sentiments in Mobile Apps

Analisis Sentimen Aspek Ulasan Aplikasi Tinder Menggunakan Teknik Deep Learning

Presented by Galih Isa Yudhistira
(00000048134)

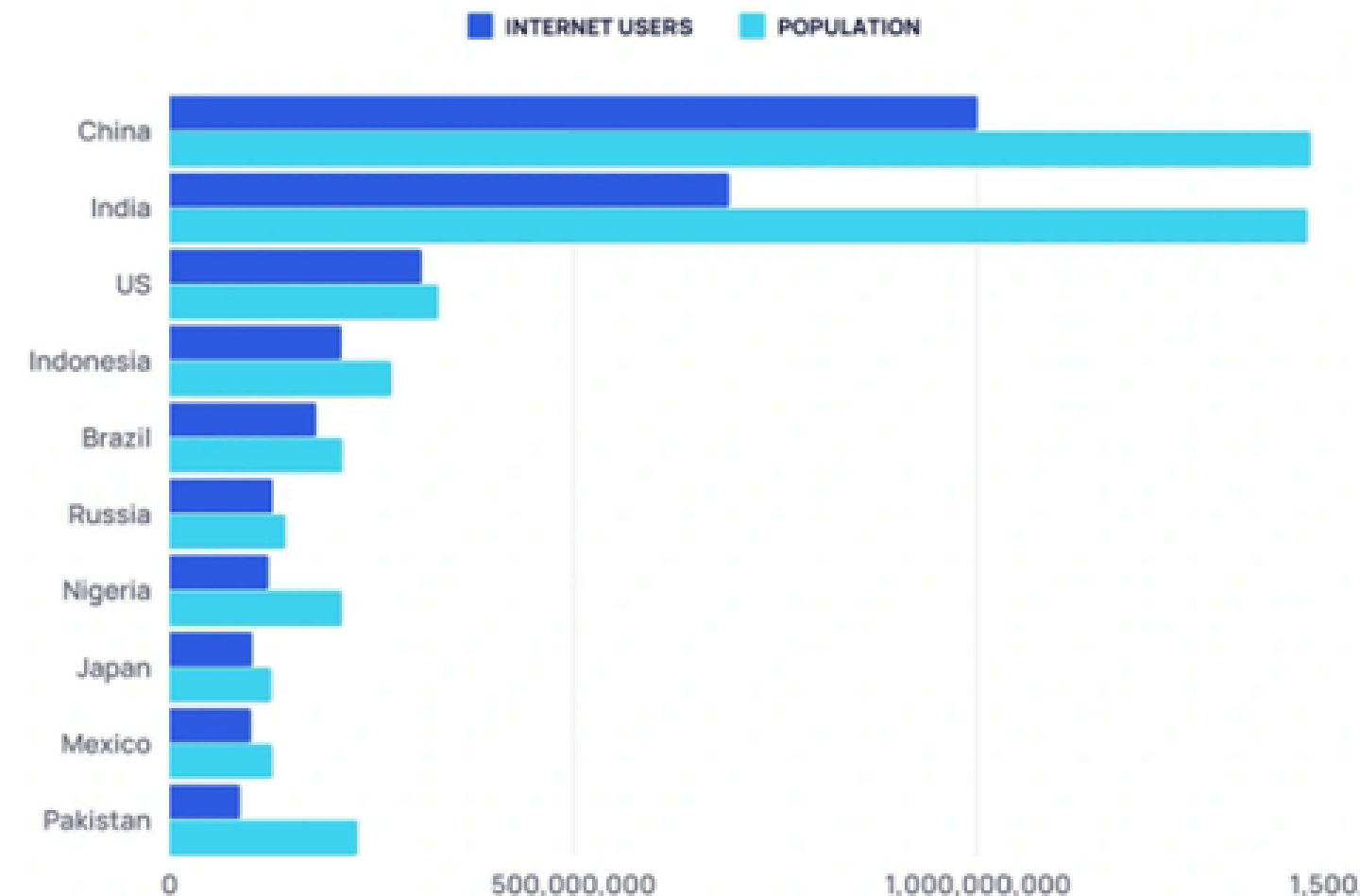
Dosen Pembimbing : Dr. Erick
Fernando, S.kom., M.S.I.



Research Overview and
Objectives

Presented by Galih Isa
Yudhistira

Countries with the most internet users



SOURCE: <https://explodingtopics.com/blog/countries-internet-users>

Perkembangan Dunia Teknologi dan Internet

Transformasi Kehidupan melalui Teknologi dan Internet

Teknologi dan internet telah mengubah banyak aspek kehidupan manusia dengan menghadirkan berbagai aplikasi yang memudahkan aktivitas sehari-hari, seperti belanja, transportasi, dan hiburan.

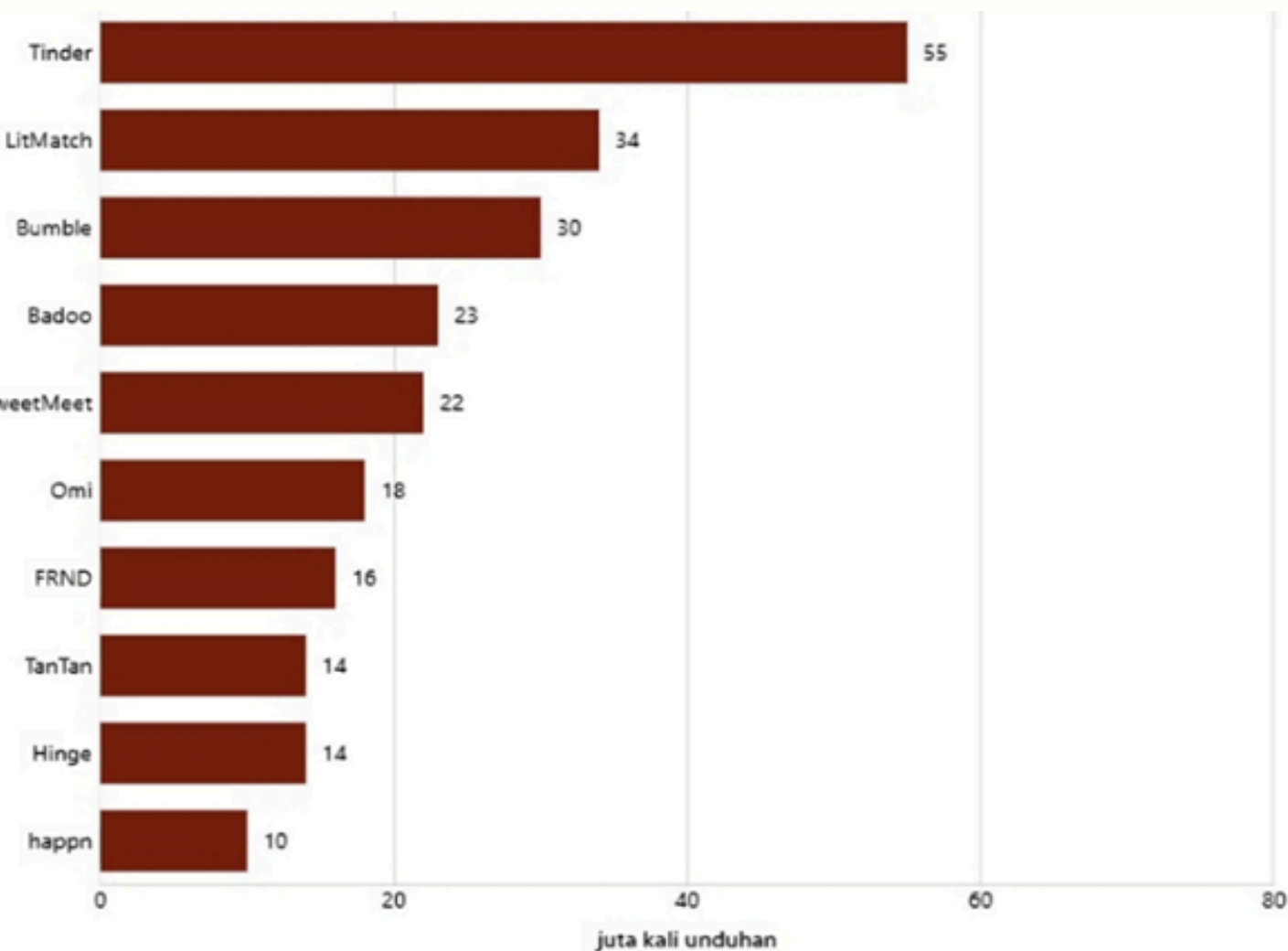
Perkembangan Pesat Teknologi dan Infrastruktur Komunikasi di Indonesia

Di Indonesia, perkembangan teknologi dan internet berjalan sangat cepat dan didukung oleh infrastruktur komunikasi yang maju.

Posisi Indonesia sebagai Negara dengan Pengguna Internet Terbesar di Dunia

Hal ini menjadikan Indonesia sebagai salah satu negara dengan jumlah pengguna internet terbesar di dunia, menempati peringkat ke-4 pada tahun 2024, setelah China, India, dan Amerika Serikat. Data tersebut menunjukkan pertumbuhan pesat pengguna internet yang sejalan dengan kemajuan teknologi di Indonesia.

10 Aplikasi Kencan Online yang Paling Banyak Diunduh Secara Global (2023)



SOURCE:  databoks

10 Aplikasi Kencan Online yang Paling Banyak Diunduh Secara Global

Dominasi Tinder dalam Aplikasi Kencan Global

Tinder menjadi aplikasi kencan paling populer di dunia dengan jumlah unduhan mencapai 55 juta kali pada tahun 2023. Aplikasi ini berasal dari Amerika Serikat dan berada di bawah naungan Match Group yang juga mengelola beberapa aplikasi kencan lainnya seperti Hinge dan OkCupid.

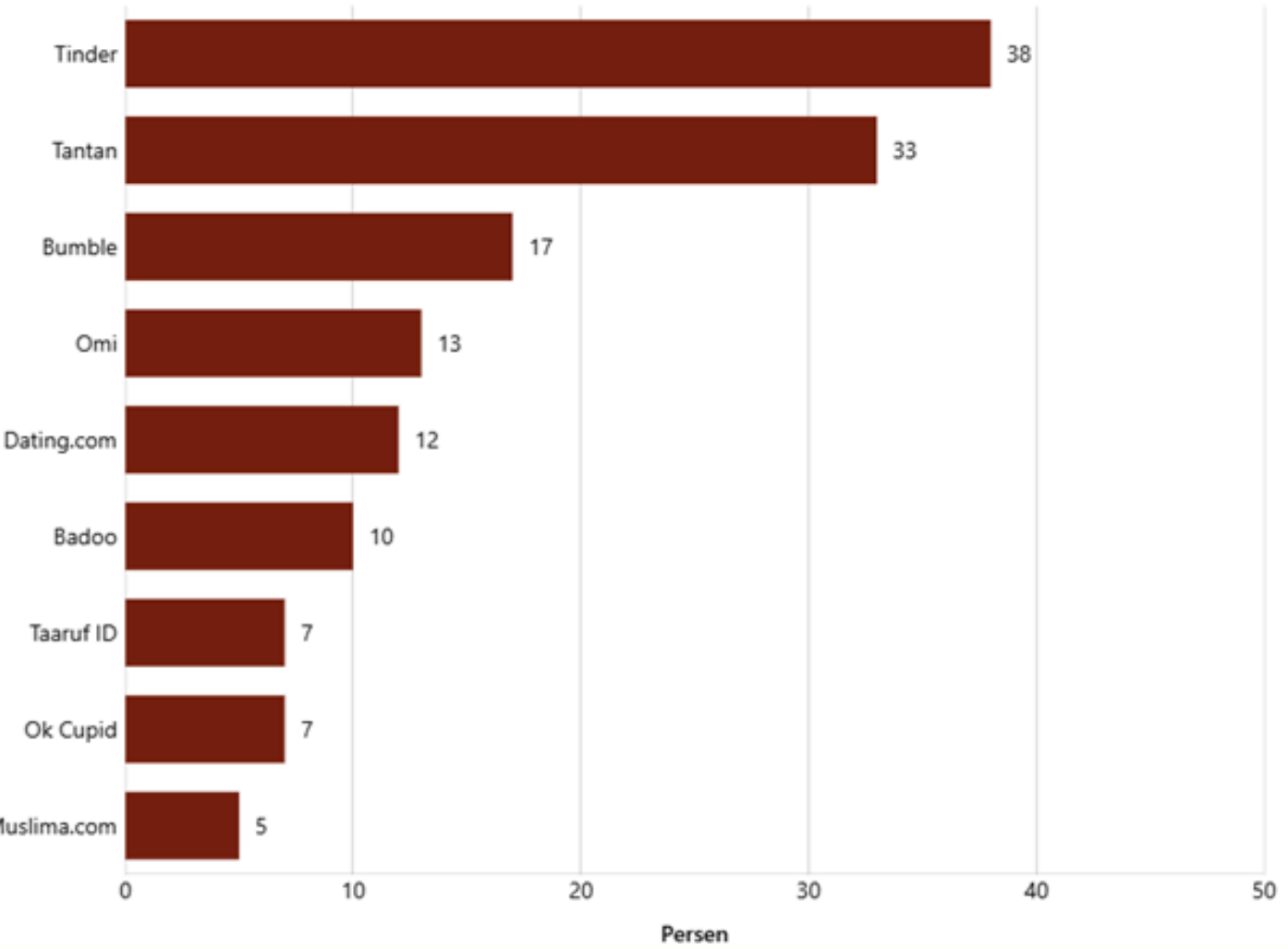
Pertumbuhan Pesat Aplikasi Kencan di Asia Tenggara

LitMatch menempati posisi kedua dengan 34 juta unduhan global di tahun 2023, menunjukkan peningkatan signifikan di wilayah Asia Tenggara meskipun tanpa rincian angka lengkap. Hal ini menandakan tren pertumbuhan aplikasi kencan yang kuat di kawasan ini.

Persaingan Ketat di Pasar Aplikasi Kencan

Bumble menempati posisi ketiga dengan 30 juta unduhan dan juga mengelola beberapa aplikasi kencan lain seperti Badoo yang memiliki total 23 juta unduhan pada tahun yang sama, menunjukkan pasar aplikasi kencan yang sangat kompetitif dengan beberapa pemain besar di dalamnya.

Aplikasi Kencan Online Terpopuler di Indonesia menurut Populix (Januari 2024)



SOURCE:  databoks

Tinder, Aplikasi Kencan Paling Banyak di Indonesia periode tahun 2024

Dominasi Tinder di Indonesia

Tinder adalah aplikasi kencan yang paling banyak digunakan di Indonesia tahun 2024, dengan 38% responden memilihnya sebagai aplikasi utama.

Aplikasi Kencan Populer Lainnya di Indonesia

Tantan dari Tiongkok juga cukup diminati dengan 33% pengguna, diikuti oleh Bumble, Omi, dan Datingcom yang memiliki pengguna lebih sedikit.

Keberagaman Pilihan Pengguna Aplikasi Kencan

Selain aplikasi utama itu, ada juga pengguna yang memakai berbagai aplikasi lain seperti Badoo, Taaruf, OkCupid, dan Muslima.com, menunjukkan variasi preferensi yang luas.

**Lalu??? Apa dampaknya?? baik
atau buruk?**

Korban Penipuan Aplikasi Kencan di Jakpus Tersebar di Vietnam-Thailand

Azhar Bagas Ramadhan - detikNews

Rabu, 29 Jan 2025 08:20 WIB



SOURCE : <https://news.detik.com/berita/d-7753771/korban-penipuan-aplikasi-kencan-di-jakpus-tersebar-di-vietnam-thailand>

Dampak Pengaruh Aplikasi Kencan Daring

Bahaya Love Scamming di Aplikasi Kencan Online

Aplikasi kencan sering disalahgunakan untuk penipuan seperti love scamming, di mana pelaku memanipulasi emosi korban untuk keuntungan finansial atau pencurian data pribadi.

Dampak Negatif bagi Korban

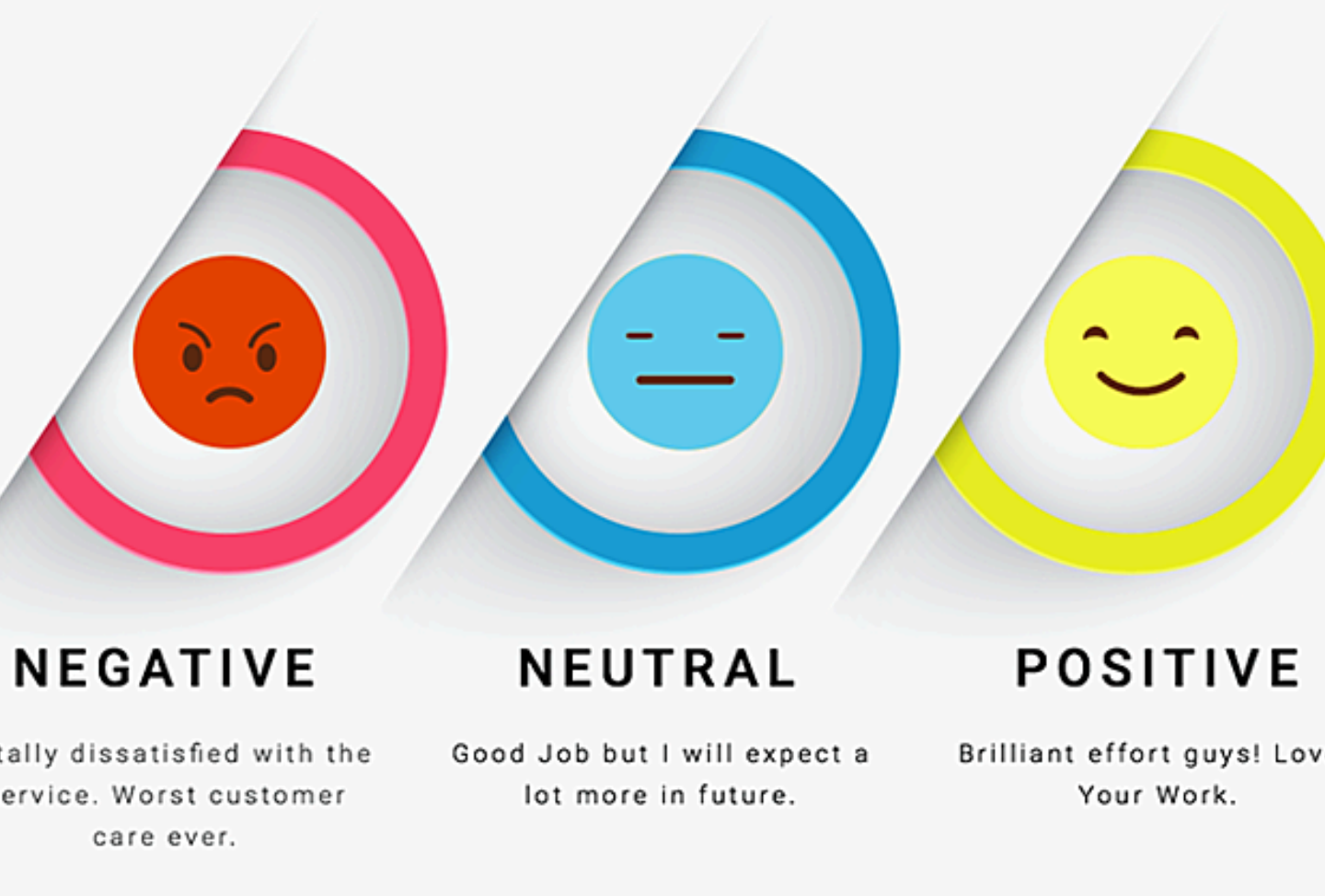
Korban love scamming mengalami kerugian materi sekaligus trauma emosional, hilangnya kepercayaan, gangguan sosial, bahkan pencurian identitas yang serius.

Risiko Psikologis dan Kebutuhan Evaluasi Aplikasi

Selain kejahatan digital, aplikasi kencan juga dapat memicu gangguan kesehatan mental seperti kecemasan dan stres, sehingga penting untuk mengevaluasi pengalaman pengguna demi perlindungan yang lebih baik.

Lalu??? Bagaimana Solusinya??

SENTIMENT ANALYSIS



Mengenal Analisa Sentimen

Apa Itu Analisis Sentimen?

Analisis sentimen mempermudah pemahaman cepat terhadap ribuan ulasan aplikasi dengan mengelompokkan sentimen menjadi positif, negatif, atau netral.

Mengapa Analisis Sentimen?

Ini adalah teknik komputer yang mengidentifikasi dan mengklasifikasikan opini dari teks untuk mengetahui perasaan yang terkandung di dalamnya.

Manfaat untuk Aplikasi Kencan Online

Analisis ini berguna untuk menilai pengalaman pengguna dan kualitas aplikasi secara efisien tanpa harus membaca seluruh ulasan satu per satu.

Rumusan Masalah

- Bagaimana performa dan efektivitas model LSTM, BiLSTM, dan CNN dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Tinder, termasuk pengaruh tuning hyperparameter seperti unit LSTM, dropout rate, dimensi embedding, kernel size, dan filter CNN terhadap peningkatan akurasi dan performa model?
- Sejauh mana ketidakseimbangan kelas sentimen memengaruhi performa ketiga model tersebut, serta bagaimana efektivitas teknik Class Weights dalam mengatasi masalah ini? Selain itu, bagaimana hasil analisis sentimen ini dapat digunakan sebagai tolok ukur penilaian kualitas layanan aplikasi Tinder?

Tujuan Penelitian

- Menganalisis dan membandingkan performa algoritma LSTM, BiLSTM, dan CNN dalam klasifikasi sentimen ulasan pengguna aplikasi Tinder, termasuk evaluasi dampak tuning hyperparameter dan penerapan class weighting terhadap peningkatan akurasi model.
- Mengkaji pengaruh ketidakseimbangan kelas pada performa model serta efektivitas teknik class weighting, sekaligus menginterpretasi hasil klasifikasi sentimen sebagai indikator persepsi pengguna terhadap kualitas layanan aplikasi.

Manfaat Penelitian

Teoritis

Memberikan acuan pemilihan algoritma deep learning yang efektif untuk analisis sentimen dan wawasan teknis terkait penanganan data tidak seimbang melalui penerapan class weighting dan tuning hyperparameter, sekaligus berkontribusi pada pengembangan metode analisis sentimen khususnya pada ulasan aplikasi sosial di platform digital.

Praktis

Menyediakan gambaran umum dan interpretasi persepsi serta pengalaman pengguna aplikasi Tinder berdasarkan hasil klasifikasi sentimen, yang dapat digunakan sebagai bahan evaluasi dan pengambilan keputusan untuk peningkatan kualitas layanan aplikasi.

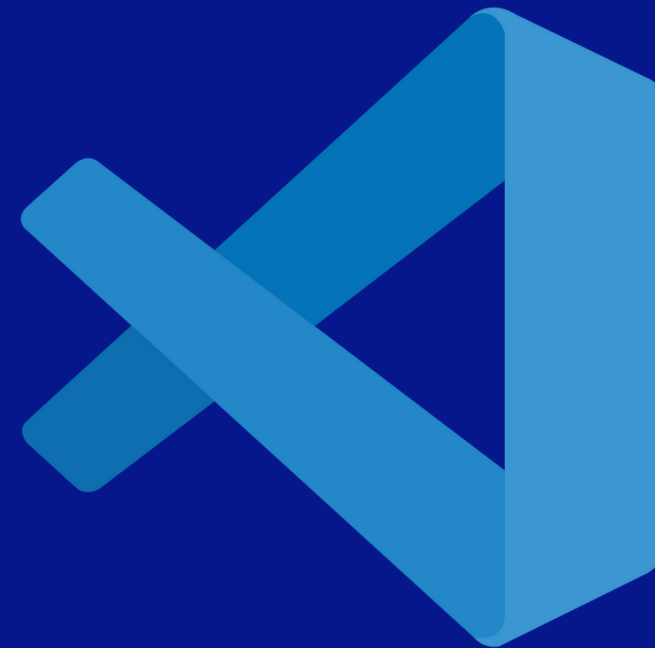
Batasan Penelitian

1. Dataset yang digunakan berupa 5000 ulasan pengguna aplikasi kencan Tinder di Google Play Store, dengan data terbatas pada ulasan terbaru dalam kurun waktu lima tahun terakhir hingga 12 Februari 2025.
2. Data yang dianalisis merupakan teks ulasan dengan klasifikasi sentimen dalam tiga kategori: positif, negatif, dan netral.
3. Penelitian menggunakan tiga algoritma utama: Long Short Term Memory (LSTM), Bidirectional LSTM (BiLSTM), dan Convolutional Neural Network (CNN), dengan fokus pada evaluasi performa setelah tuning hyperparameter dan penerapan teknik Class weights.
4. Studi ini hanya membahas pengaruh class weighting sebagai metode penanganan ketidakseimbangan kelas, tanpa membandingkan dengan teknik oversampling atau metode lain.
5. Evaluasi tuning hyperparameter meliputi jumlah unit LSTM/BiLSTM, dropout rate, dimensi embedding, serta parameter utama pada CNN, tanpa membahas hyperparameter lain di luar yang disebutkan.

Penelitian Terdahulu

- Dany Eka Saputra & Auliya Rahman Isnain (2024) Penelitian CNN pada analisis sentimen komentar YouTube bacapres 2024 mencapai skor F1 hingga 94%, dengan SMOTE dan tuning meningkatkan performa, namun ketidakseimbangan data masih berpengaruh (F1 makro 60-65%).
- Titis Arwindarti Esther Irawati Setiawan, Syaiful Imron Ahmad (2023) Model LSTM dengan Word2Vec efektif dalam analisis komentar Instagram, meraih akurasi 84% mampu menangkap makna semantik dan konteks sekuensial data.
- Fitroh, Fahmi Hudaya (2023) Model BiLSTM untuk klasifikasi sentimen menunjukkan performa sangat tinggi dengan akurasi 99,14% dan F1-score 99,12%, seimbang antara presisi dan recall, menandakan kinerja hampir sempurna.

TOOLS



Streamlit

KDD VS CRISP - DM

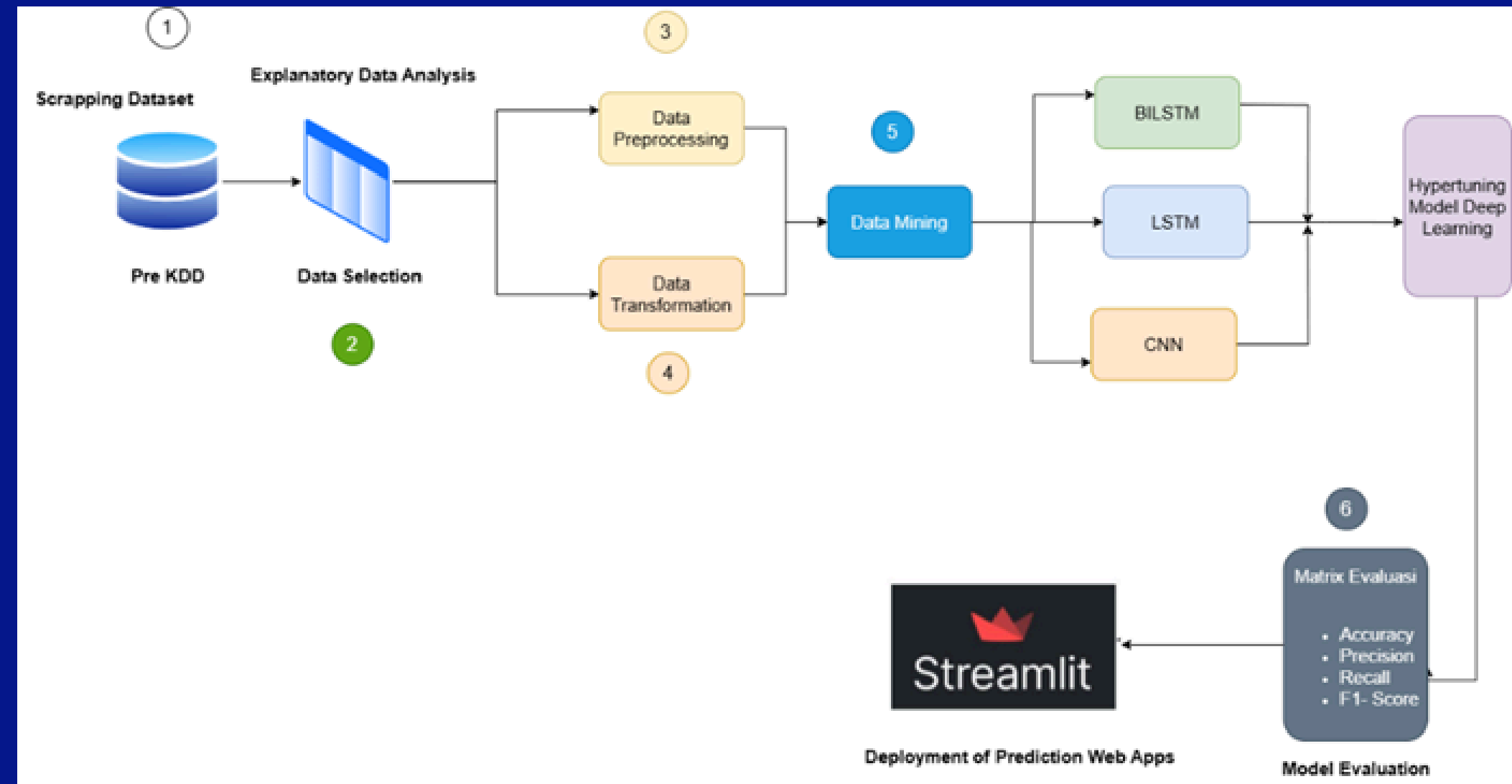
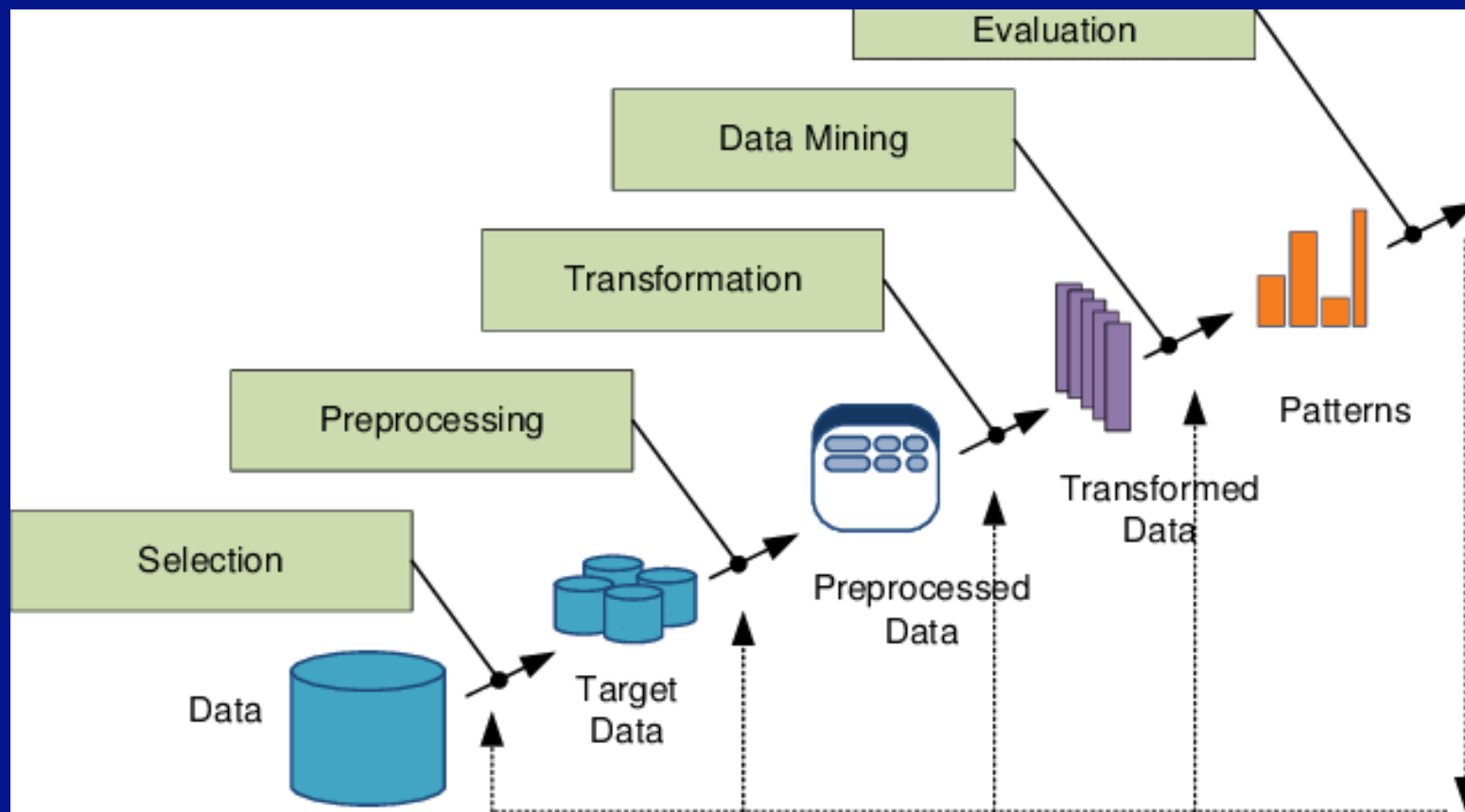
Aspek	Alasan Memilih KDD	Alasan Memilih CRISP-DM
Tahap Penyelesaian	Pre KDD, Data Selection, PreProcessing, Tranformation, Data mining, Evaluation	Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation
Keterlibatan Pemangku Kepentingan	Minim keterlibatan pemangku kepentingan di awal	Melibatkan tim produk dan pemasaran sejak awal
Sifat Data dan Kompleksitas	Efektif untuk data tidak terstruktur dan kompleks	Cocok untuk data yang terstruktur dan kebutuhan bisnis jelas
Pendekatan Proses	Peneliti terus gali insight baru.	Tim bisnis dan peneliti evaluasi bersama.
Tujuan Penggunaan Hasil	Mendapatkan wawasan mendalam untuk riset dan pengembangan	Hasil langsung relevan dan dapat diimplementasikan bisnis

Alasan Memilih KDD

Metode KDD tepat dipilih karena memudahkan eksplorasi data ulasan pengguna Tinder yang kompleks dan tidak terstruktur. KDD memberikan kebebasan bagi peneliti untuk menemukan pola baru tanpa harus bergantung pada keterlibatan awal pemangku kepentingan, sehingga proses bisa lebih fleksibel dan fokus pada riset mendalam untuk pengembangan aplikasi yang lebih baik.

METODE PENELITIAN

Knowledge Discovery In Databases (KDD)



Pre KDD (Data Collection)

Tahap awal dari penelitian ini melibatkan pengumpulan data dari ulasan aplikasi Tinder di Google Play Store. Data tersebut diambil melalui proses pengambilan data dari platform Google Play Store dengan melakukan crawling data menggunakan google playstore web scrapper untuk kebutuhan penelitian.



```
from google_play_scraper import Sort, reviews

# Memperbarui untuk mengambil ulasan dari aplikasi tinder di Google Play Store
result, continuation_token = reviews(
    'com.tinder', # ID aplikasi tanpa parameter 'hl=id' di akhir
    lang='id', # Mengatur bahasa ke Bahasa Indonesia
    country='id', # Mengatur negara ke Indonesia
    sort=Sort.MOST_RELEVANT, # Urutkan berdasarkan relevansi
    count=1000, # Mengambil 1000 ulasan
    filter_score_with=None # Mengambil semua ulasan tanpa filter berdasarkan skor
)

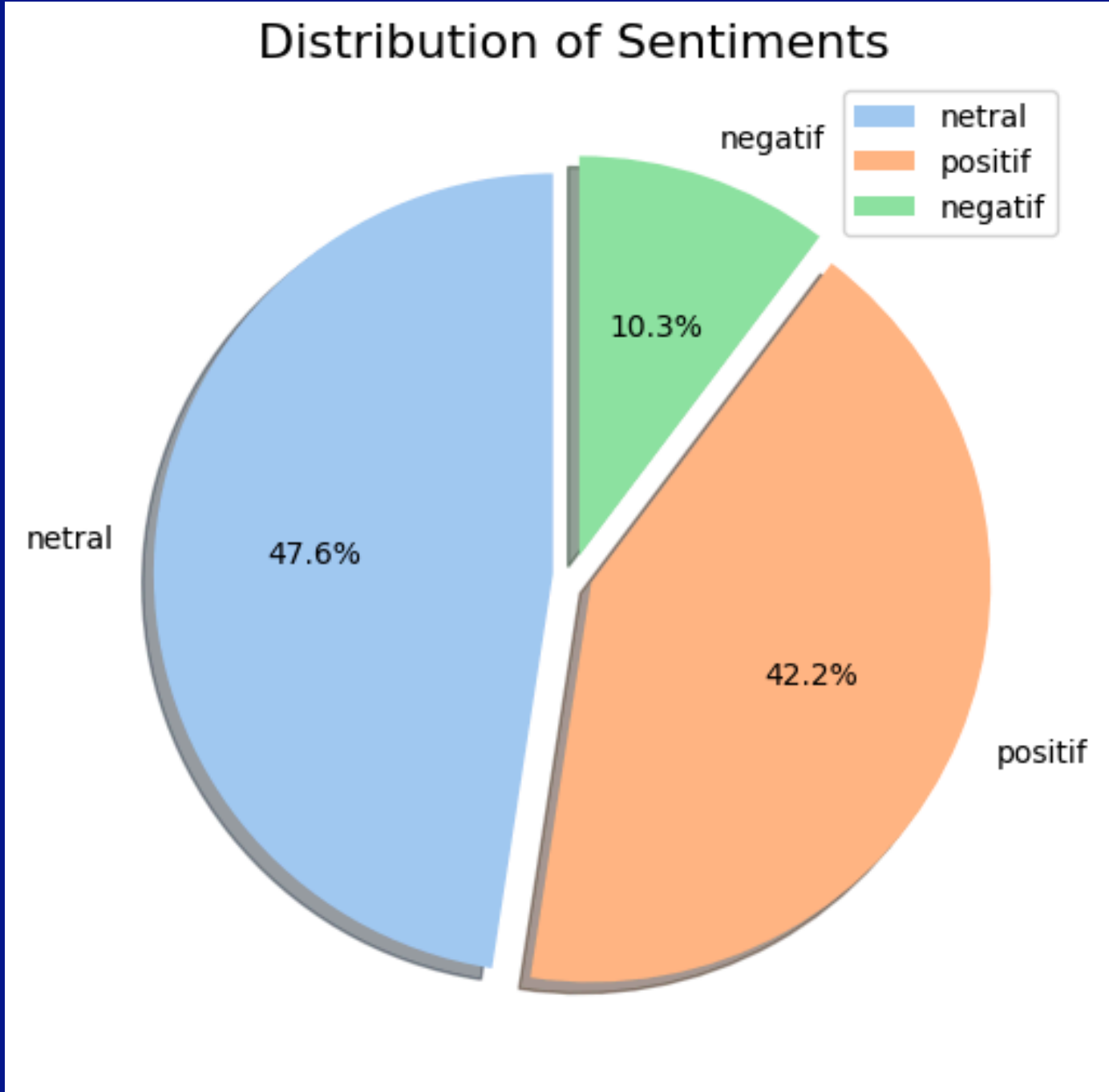
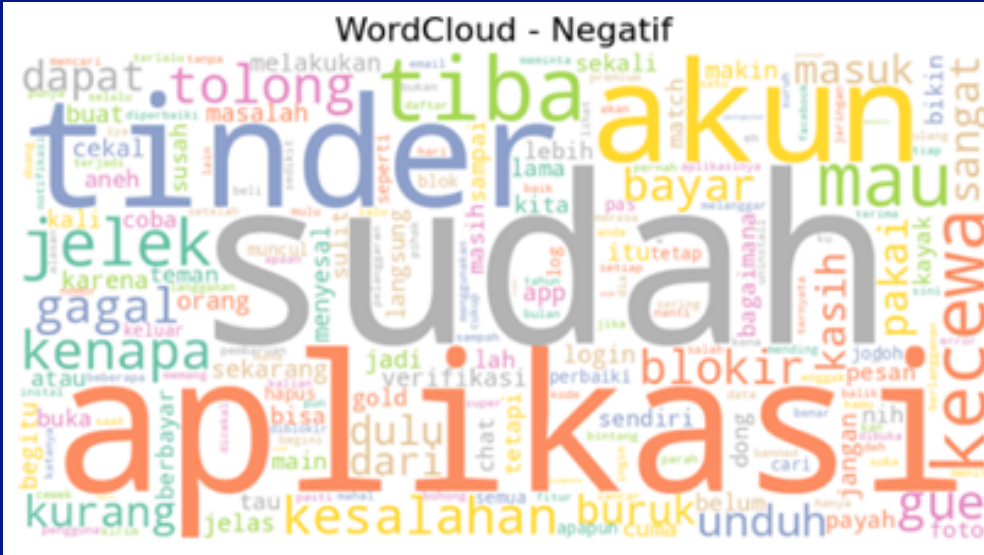
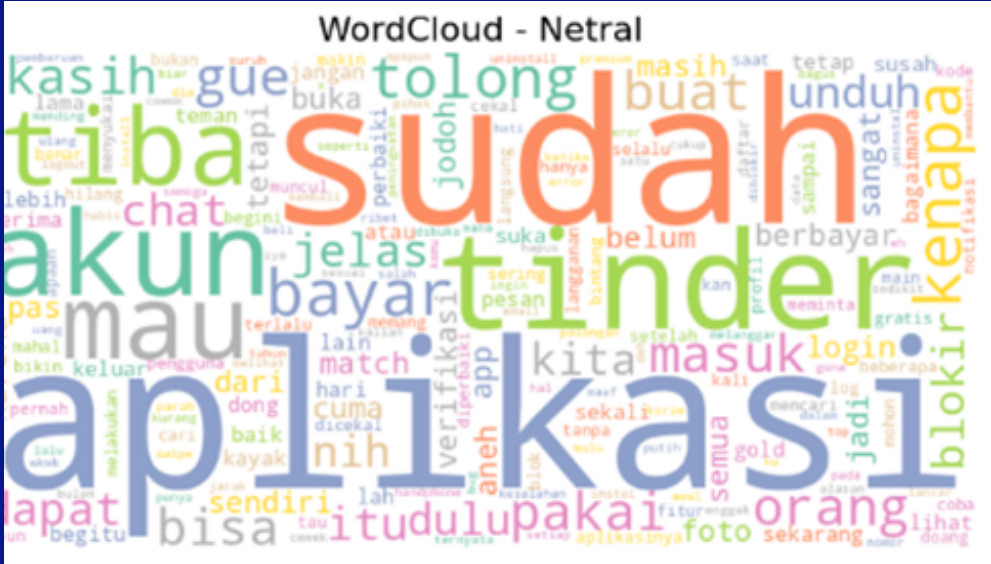
# Mengonversi hasil ke dalam DataFrame untuk memudahkan manipulasi data
import pandas as pd
import numpy as np
```

Data Selection

Hasil pengumpulan data kemudian akan disaring berdasarkan klasifikasi sentimen yang akan dibagi menjadi tiga kelas, yaitu sentimen positif sentimen negatif dan netral.

Variable	Jenis Variable	Keterangan
User name	Categorical variable	Variabel ini mewakili nama pengguna yang memberikan ulasan
Score	Numerical variable	Variabel ini mewakili peringkat atau score yang diberikan pada
At	Numerical variable	Variabel ini mewakili keterangan tanggal, bulan, tahun
Content	Categorical variable	Variabel ini mewakili isi komentar atau ulasan pengguna mengenai

Data Preprocessing Result



Data Mining & Evaluation Result

BILSTM

Penelitian Ini				
Algoritma	Akurasi	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
BILSTM	94%	Positif : 94% Negatif : 100% Netral : 92%	Positif : 99% Negatif : 89% Netral : 98%	Positif : 97% Negatif : 94% Netral : 99%
BILSTM with Class Weights	97%	Positif : 96% Negatif : 98% Netral : 97%	Positif : 100% Negatif : 87% Netral : 96%	Positif : 96% Negatif : 82% Netral : 93%
BILSTM with Class Weights Hyperparameter	98%	Positif : 99% Negatif : 98% Netral : 97%	Positif : 99% Negatif : 98% Netral : 97%	Positif : 99% Negatif : 98% Netral : 97%

Data Mining & Evaluation Result

LSTM

Penelitian Ini				
Algoritma	Akurasi	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
LSTM	96%	Positif : 98% Negatif : 99% Netral : 97%	Positif : 99% Negatif : 89% Netral : 98%	Positif : 99% Negatif : 94% Netral : 97%
LSTM with Class Weights	96%	Positif : 97% Negatif : 97% Netral : 97%	Positif : 99% Negatif : 91% Netral : 97%	Positif : 98% Negatif : 94% Netral : 97%
LSTM with Class Weights Hyperparameter	99%	Positif : 99% Negatif : 99% Netral : 99%	Positif : 99% Negatif : 99% Netral : 99%	Positif : 99% Negatif : 99% Netral : 99%

Data Mining & Evaluation Result

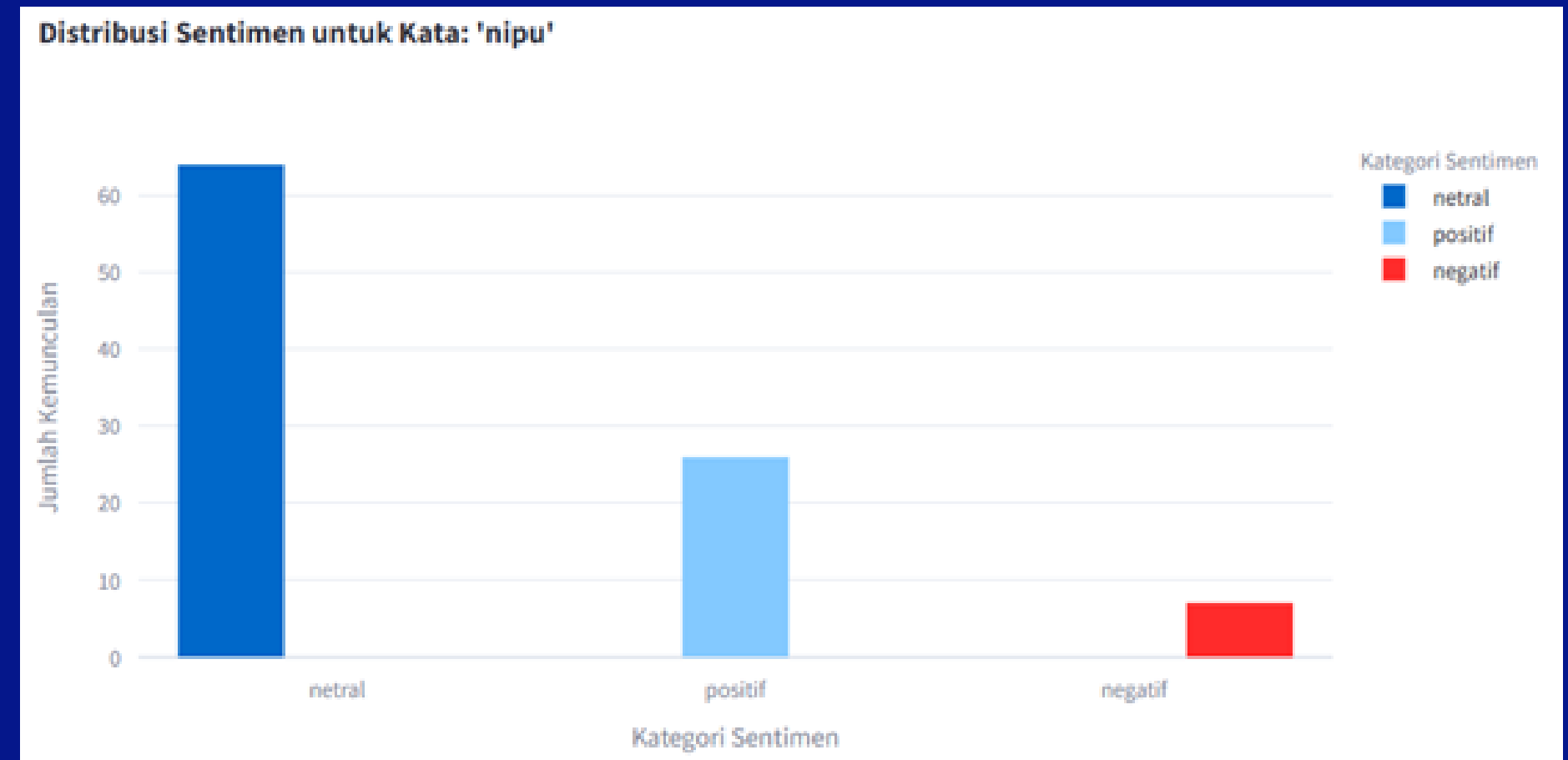
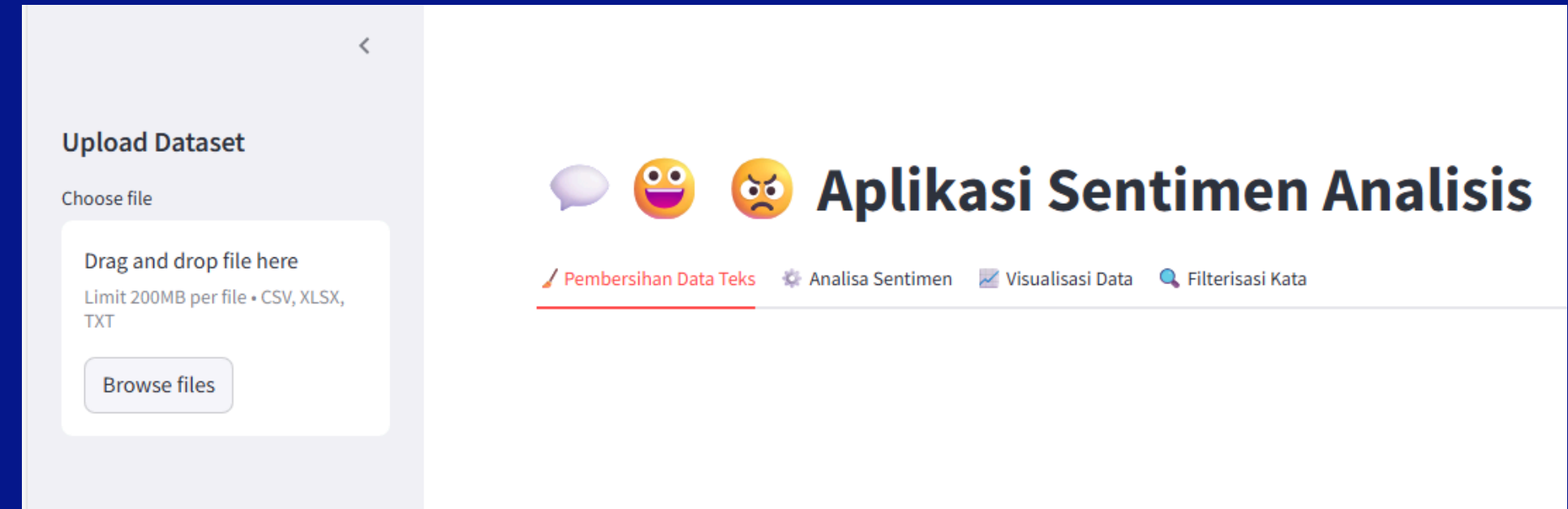
CNN

Penelitian Ini				
Algoritma	Akurasi	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
CNN	98%	Positif : 99% Negatif : 97% Netral : 97%	Positif : 97% Negatif : 96% Netral : 99%	Positif : 98% Negatif : 97% Netral : 98%
CNN with Class Weights	98%	Positif : 99% Negatif : 96% Netral : 97%	Positif : 98% Negatif : 98% Netral : 94%	Positif : 98% Negatif : 97% Netral : 98%
CNN with Class Weights Hyperparameter	97%	Positif : 99% Negatif : 98% Netral : 96%	Positif : 96% Negatif : 98% Netral : 99%	Positif : 99% Negatif : 98% Netral : 96%

Implementasi Model Aplikasi

Aplikasi prediksi analisis sentimen berbasis web ini dibuat menggunakan library Streamlit untuk memudahkan pengguna dalam melakukan prediksi melalui antarmuka yang simpel dan interaktif.

Proses pengembangannya mengikuti prinsip Knowledge Discovery In Databases (KDD) yang mencakup tahapan persiapan, seleksi, pembersihan, transformasi, penambahan data, dan evaluasi model.



Kesimpulan

- Penelitian ini membuktikan performa tinggi model BiLSTM, LSTM, dan CNN dalam klasifikasi sentimen multi-kelas pada ulasan Tinder. BiLSTM unggul dengan akurasi 98,7%, didukung class weighting dan tuning. Ketiga model menunjukkan precision, recall, dan F1-score antara 0,98-0,99. Aplikasi dikembangkan terpadu dengan pra-pemrosesan, model deep learning, klasifikasi, dan visualisasi interaktif seperti word cloud. Visualisasi memudahkan pengguna mengeksplorasi pola dan kalimat dalam data.
- Keterbatasan penelitian adalah data hanya dari Tinder, dan proses pelatihan membutuhkan sumber daya besar serta waktu lama. Pengembangan selanjutnya dapat memakai transfer learning atau model pra-latih seperti BERT untuk peningkatan performa. Integrasi data real-time dan analisis aspek juga menjadi fokus masa depan. Langkah ini diharapkan menjadikan aplikasi lebih adaptif dan efisien. Penelitian ini membuka peluang pengembangan model analisis sentimen yang lebih luas dan dinamis.

Saran

- Disarankan menggunakan data dari berbagai platform seperti Twitter, Instagram, atau marketplace digital untuk menguji kemampuan generalisasi model dengan ragam gaya bahasa dan konteks sosial yang beragam.
- Perlu dilakukan evaluasi model deep learning seperti GRU, Transformer, serta model pretrained seperti BERT dan RoBERTa. Bandingkan juga dengan model klasik seperti Random Forest, XGBoost, dan LightGBM untuk melihat keunggulan tiap metode.
- DANN dapat membantu mengatasi perbedaan distribusi antara data latih dan uji dari sumber berbeda, sehingga meningkatkan generalisasi model dalam skenario domain adaptasi.
- Analisis aspek memungkinkan pengenalan opini terhadap fitur spesifik dalam ulasan panjang dan kompleks, sehingga menghasilkan informasi sentimen yang lebih detail dan bernilai.

Thank

You!