

# **APLIKASI ANALISIS SENTIMEN MENGGUNAKAN MODEL ROBERTA+GRU-CNN DAN DETEKSI TOPIK MENGGUNAKAN MODEL ROBERTA+EFCM PADA TWEET MENGENAI CHATGPT**

**KELOMPOK 8**

# ANGGOTA KELompok



Safira  
Raissa  
2006568891



Rachel  
Octaviani  
2006463585



Regita  
Celia  
2006463540



Muhammad  
Daffa  
2006568626

# DAFTAR ISI

01

PENDAHULUAN

- A. LATAR BELAKANG
- B. RELATED WORKS
- C. DEFINISI MASALAH
- D. BATASAN MASALAH
- E. TUJUAN
- F. METODOLOGI PENELITIAN

02

METODE

- A. BERT DAN ROBERTA
- B. CNN, RNN, DAN GRU
- C. EFCM

03

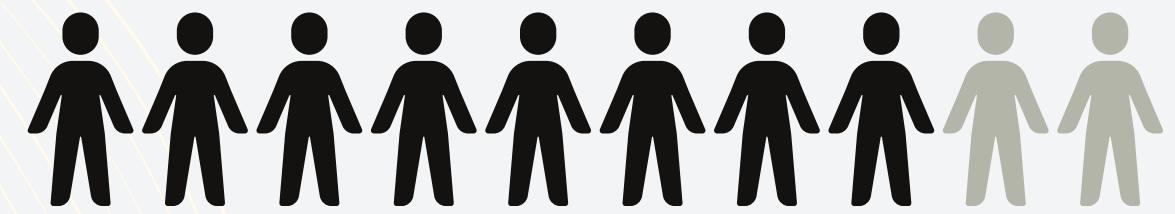
SIMULASI DAN ANALISIS HASIL

04

KESIMPULAN

# 1.

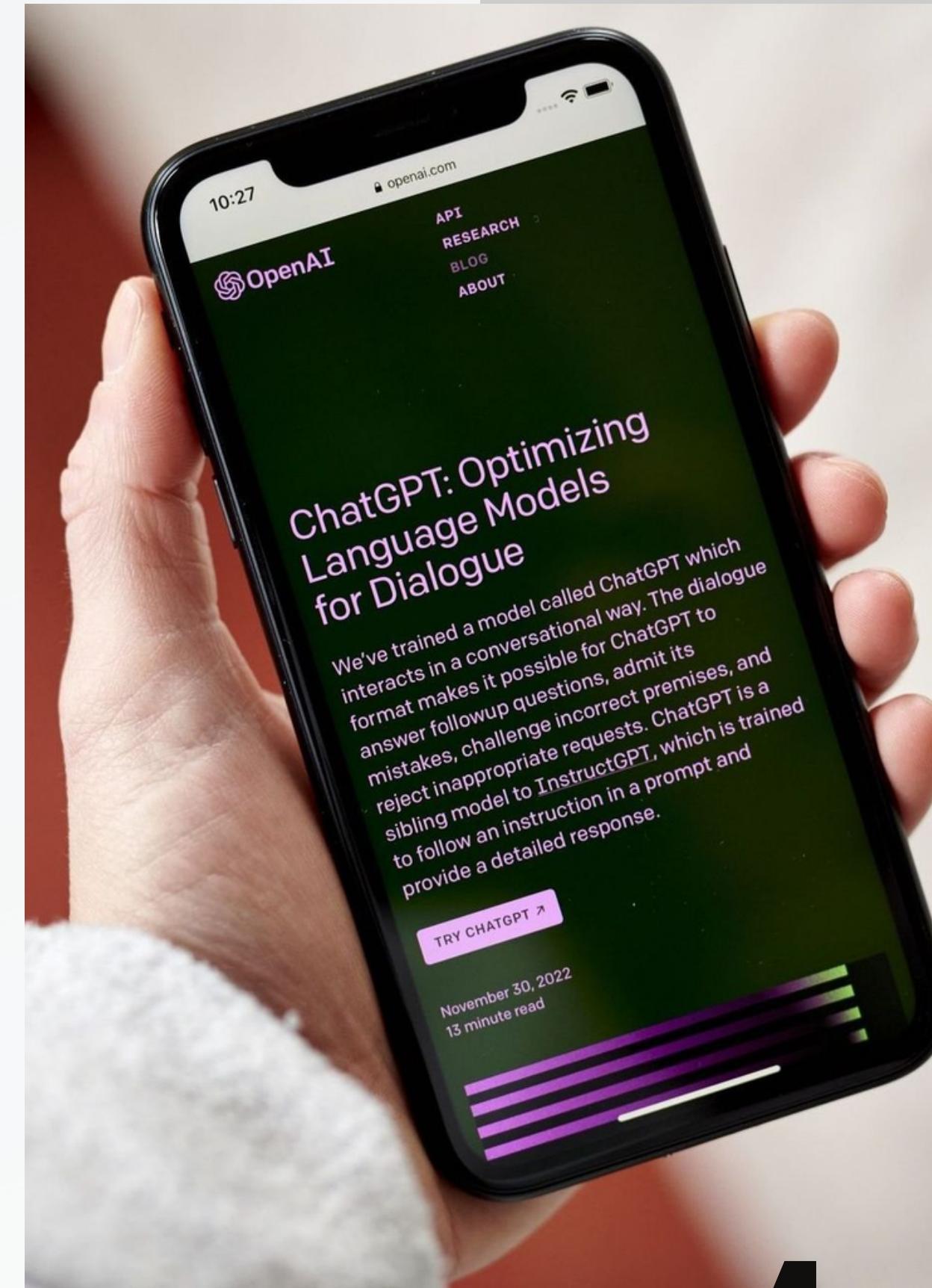
## PENDAHULUAN



# LATAR BELAKANG

Twitter merupakan media terbesar untuk publik berbagi pemikiran, cerita, dll. ChatGPT, bot percakapan AI yang sangat akurat, telah mendominasi Twitter dalam beberapa bulan terakhir, menghasilkan kegembiraan dan kekhawatiran di antara pengguna.

Sentimen ChatGPT dalam topik yang lebih spesifik perlu dianalisis untuk memberikan gambaran tentang bagaimana ChatGPT dirasakan di kalangan pengguna dan menunjukkan potensi diterimanya ChatGPT di masa depan.



# RELATED WORKS



MUBIN UL HAQUE, ISURU DHARMADASA, ZARRIN T. SWORNA, ROSHAN N. RAJAPAKSE, DAN HUSSAIN AHMAD, 2022

*"I think this is the most disruptive technology"* Exploring Sentiments of ChatGPT Early Adopters using Twitter Data

- Menganalisis sentimen dan mendekripsi topik pada data Twitter mengenai ChatGPT pada masa awal peluncuran.
- Melihat sentimen publik per topik terhadap ChatGPT.



KIAN LONG TAN, CHIN POO LEE, DAN KIAN MING LIM, 2023

RoBERTa-GRU: A Hybrid Deep Learning Model for Enhanced Sentiment Analysis

- Penggunaan model RoBERTa+GRU untuk analisis sentimen.
- Membandingkan model dan parameter yang digunakan dengan yang lainnya.



SYAMSURIANI, 2022

Analisis Kinerja BERT sebagai Metode Representasi Teks pada Model Hybrid Deep Learning untuk Analisis Sentimen Berbahasa Indonesia

- Analisis penggunaan BERT pada berbagai model, termasuk CNN-GRU, dalam analisis sentimen.

# RELATED WORKS



NICHOLAS RAMOS RICARDO, 2022

Analisis Performa EFCM dengan  
BERT sebagai Representasi Teks  
pada Pendekslsian Topik

- Menganalisis penggunaan model EFCM dengan BERT dalam mendekripsi topik.



LAURE THOMPSON DAN DAVID MIMNO

Topic Modeling with Contextualized  
Word Representation Clusters

- Penggunaan RoBERTa dalam mendekripsi topik.

# DEFINISI MASALAH

- Masalah yang diajukan pada proyek kali ini adalah **analisis sentimen** menggunakan model **RoBERTa+GRU-CNN** dan **deteksi topik** menggunakan model **RoBERTa+EFCM** pada **tweet mengenai ChatGPT**.
- Dari sudut pandang *machine learning*, masalah **deteksi topik** termasuk ke dalam masalah **clustering**, sedangkan **analisis sentimen** masuk ke dalam masalah **klasifikasi**.
- Tipe klasifikasi yang digunakan pada **analisis sentimen** dalam masalah yang diajukan adalah **klasifikasi biner**, yaitu **sentimen positif** dan **sentimen negatif**.

# BATASAN MASALAH

---

## DATASET

**Dataset** yang digunakan adalah dari dari Kaggle.com yang berisi tweet mengenai ChatGPT.

### Data Train:

Diberikan label sentimen *good* (positif) dan *bad* (negatif). Diambil 10.000 data dari 217.662 data secara random sample.

<https://www.kaggle.com/datasets/charunisa/chatgpt-sentiment-analysis>

### Data Prediksi:

Tidak ada label. Diambil 10.000 data dari 493.743 data secara random label.

Diambil 10.000 data dari 493.743 data secara random sample. sample.

<https://www.kaggle.com/datasets/khalidryder777/500k-chatgpt-tweets-jan-mar-2023>

## BAHASA

Bahasa yang digunakan pada data adalah **Bahasa Inggris**.

## MODEL

Model yang digunakan untuk **analisis sentimen** adalah **RoBERTa+GRU-CNN** dan untuk deteksi topik adalah **RoBERTa+EFCM**.

## KEMAMPUAN MODEL

Kemampuan model model **analisis sentimen** diukur dengan **accuracy**, sedangkan **deteksi topik** diukur dengan nilai **coherence**.

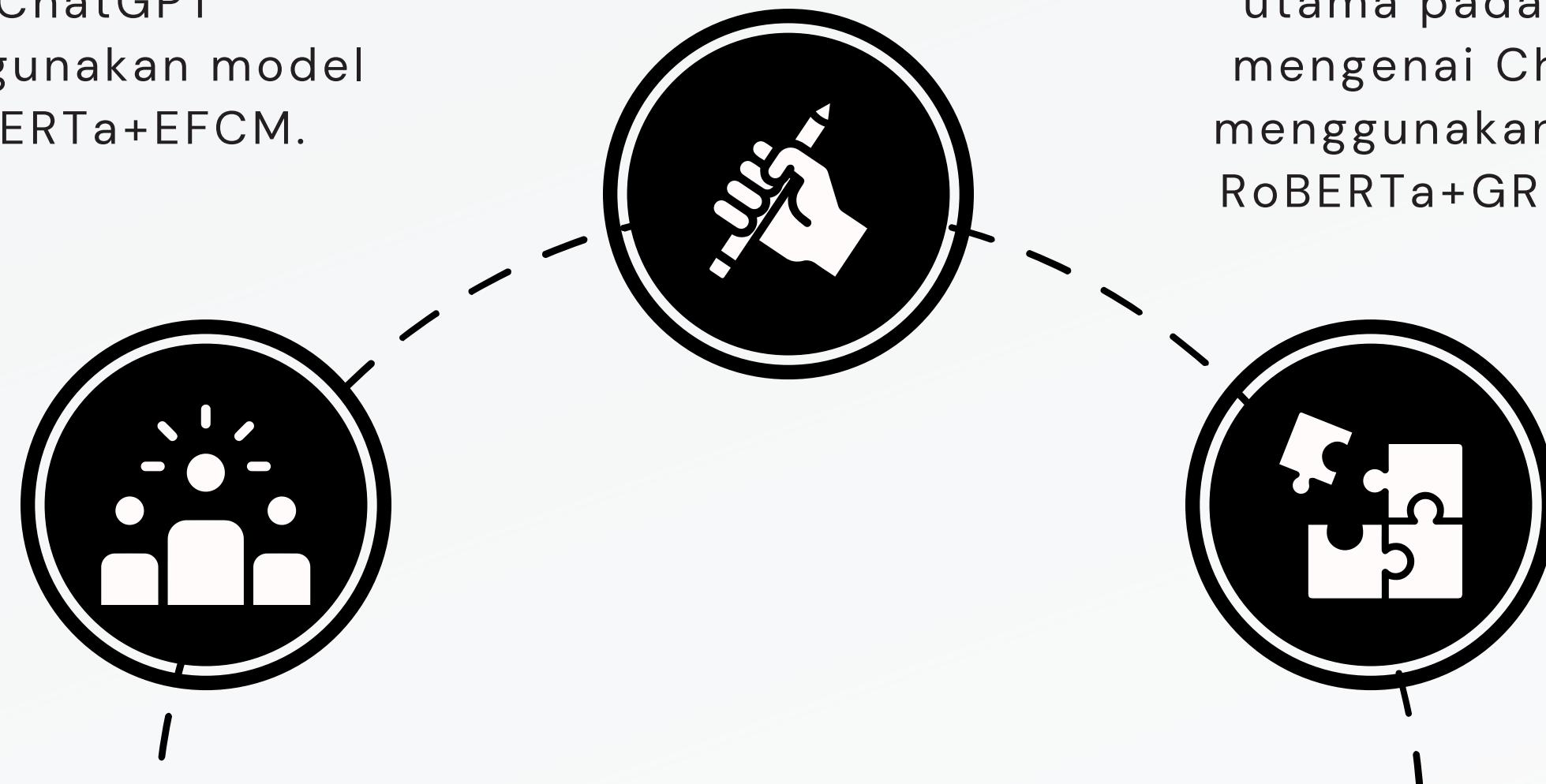
# TUJUAN

1.

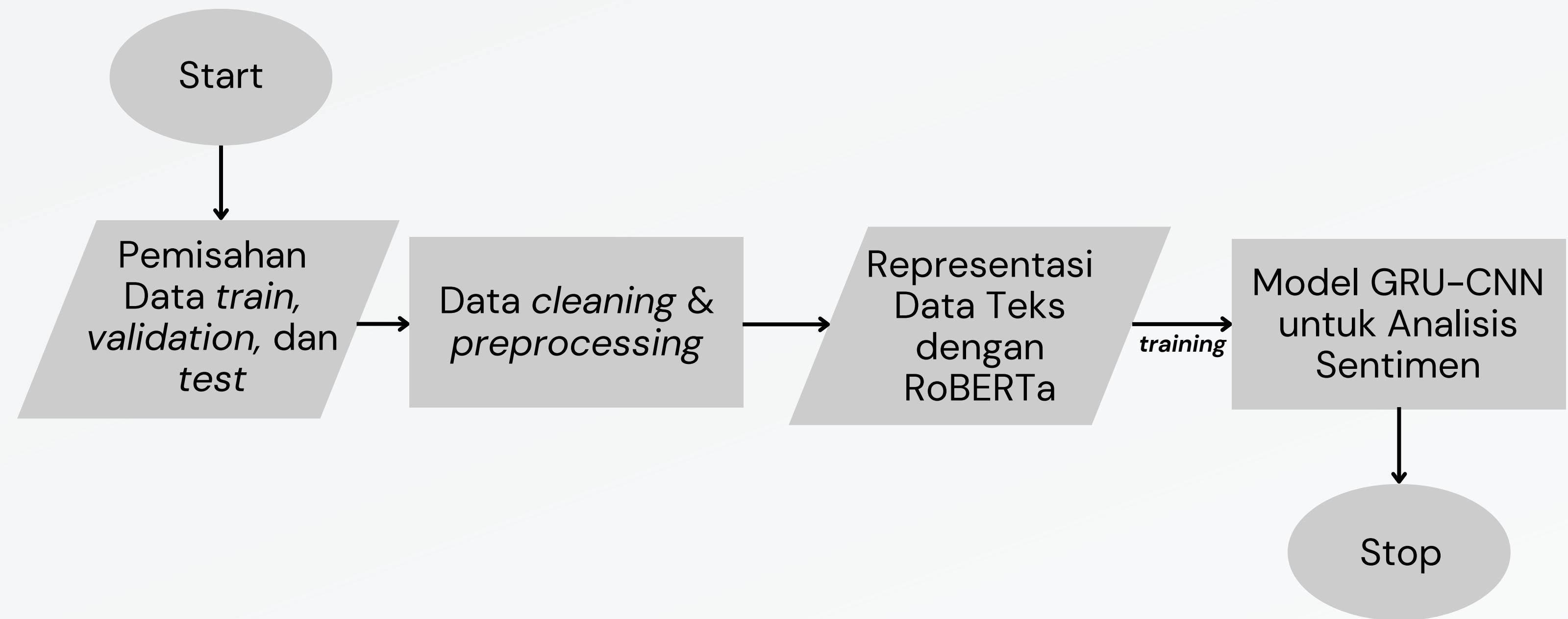
Mendeteksi topik-topik utama dari *tweet* mengenai ChatGPT menggunakan model RoBERTa+EFCM.

2.

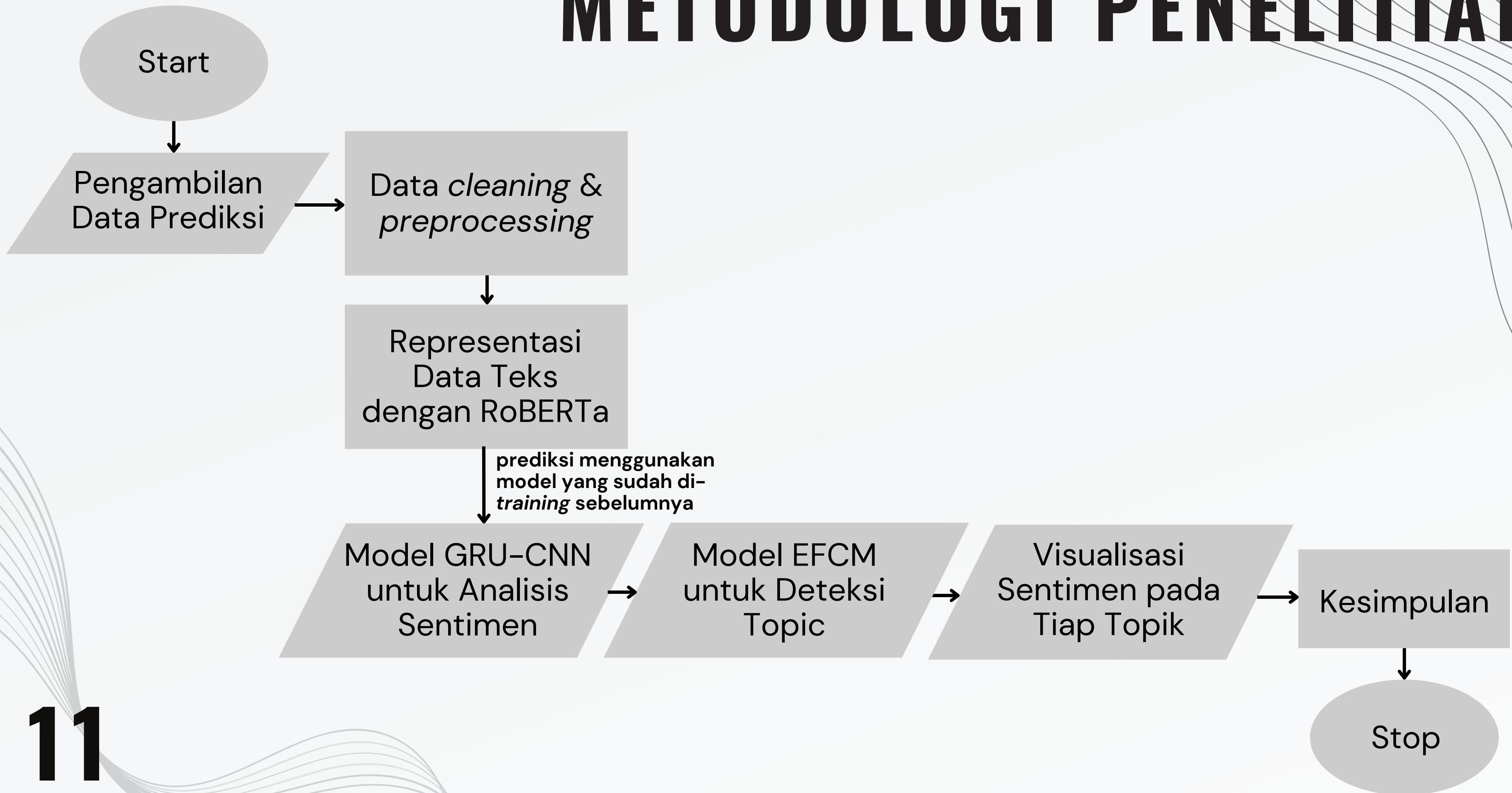
Mengklasifikasi sentimen untuk setiap topik-topik utama pada *tweet* mengenai ChatGPT menggunakan model RoBERTa+GRU-CNN.



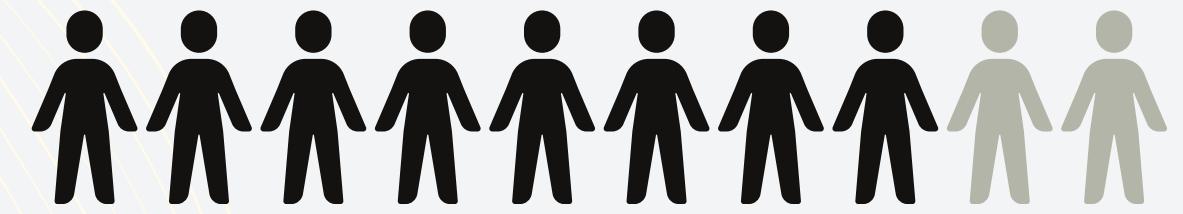
# METODOLOGI PENELITIAN



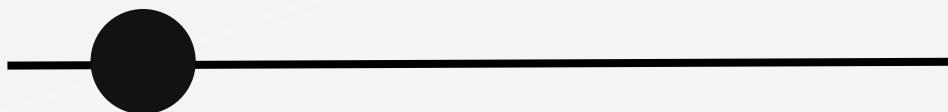
# METODOLOGI PENELITIAN



# 2. METODE



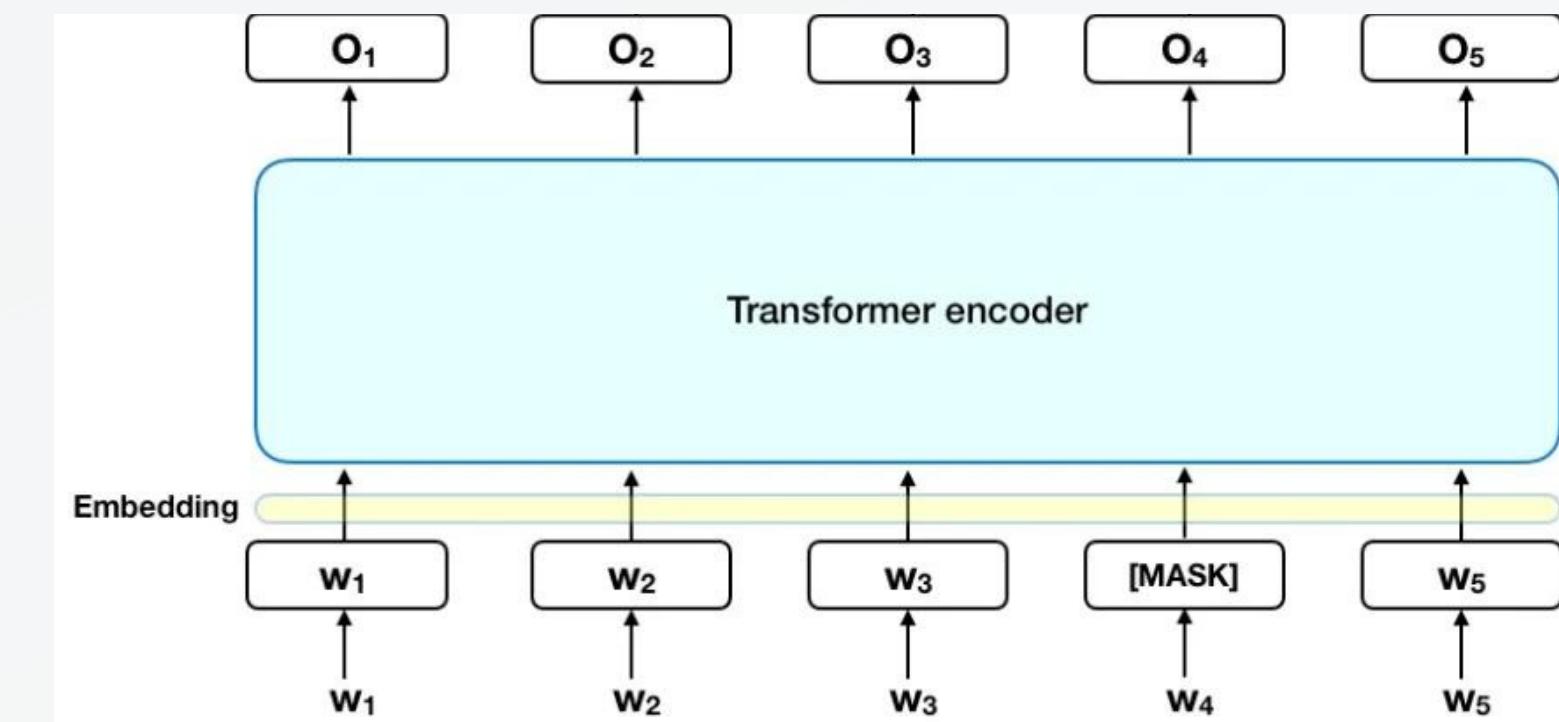
# BERT DAN ROBERTA



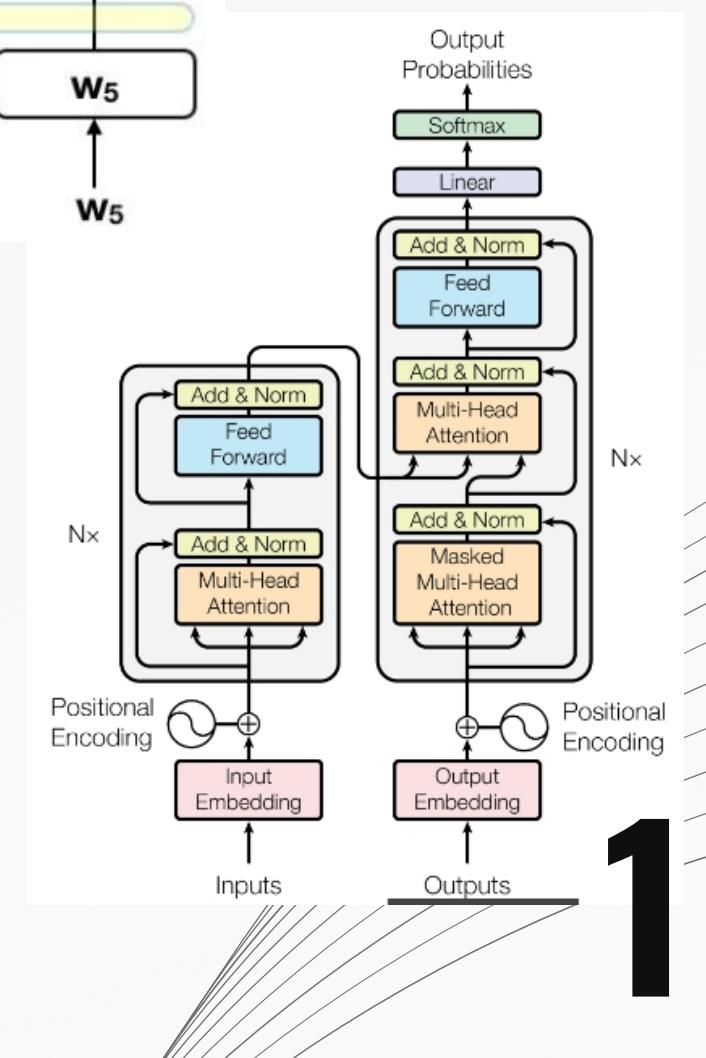
## BERT

**Bidirectional Encoder  
Representations from  
Transformers**

- Algoritma *deep learning* yang dirancang untuk mengolah *natural language processing* (Devlin et al., 2018).
- Teknik atau sistem berbasis *neural network* yang sangat berperan dalam *Natural Language Understanding* (NLU). NLU menginterpretasikan sebuah arti dari bahasa manusia dan mengklasifikasikan ke sesuatu yang dipahami oleh mesin.



Arsitektur Transformer  
sumber: Vashwani et al., 2017



# BERT DAN ROBERTA

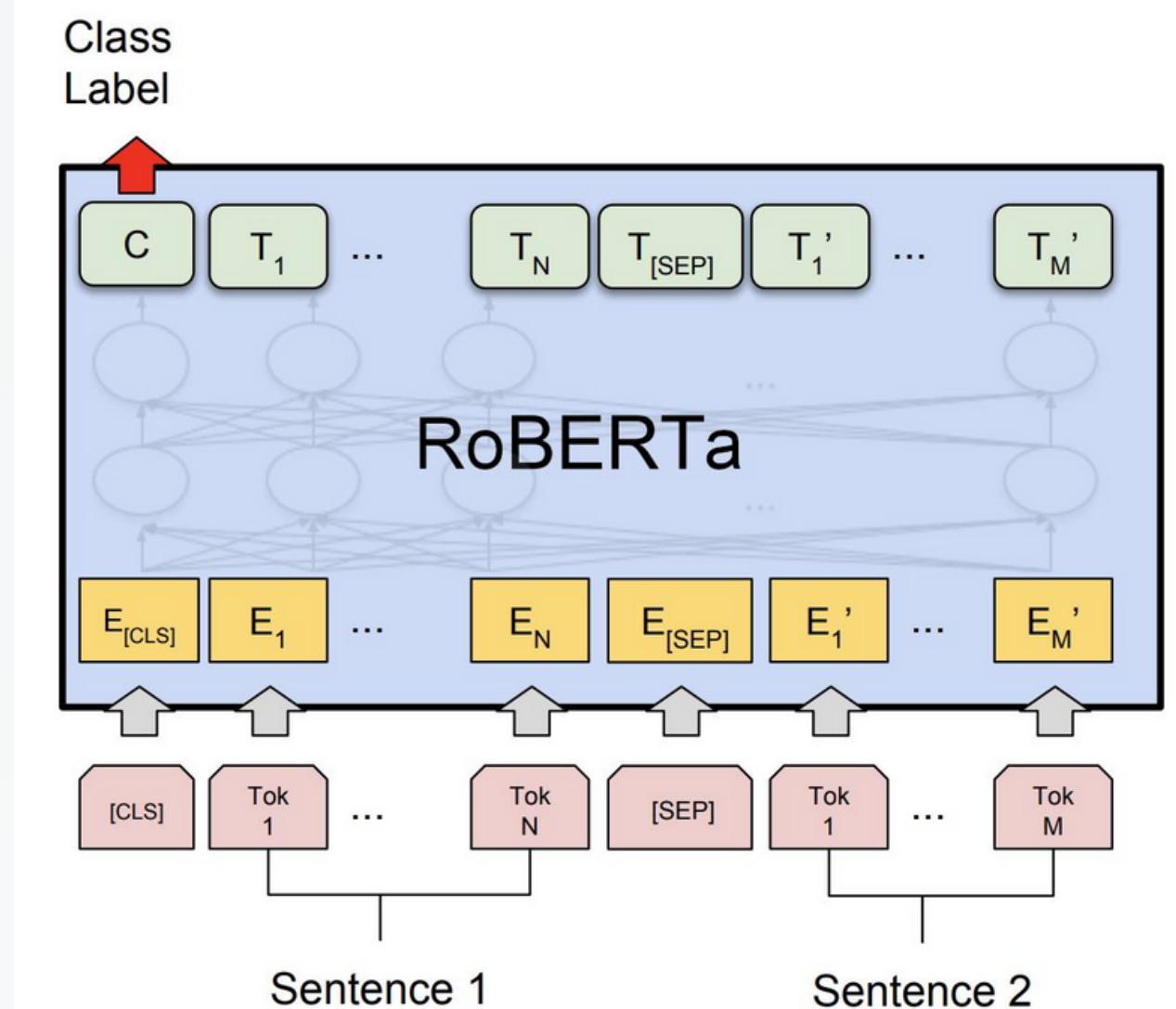


## RoBERTa

### Robustly Optimized BERT-Pretraining Approach

Perluasan dari arsitektur BERT yang telah untuk menjadi lebih optimal dan *robust*. Modifikasi tersebut antara lain:

1. Training yang lebih lama: RoBERTa dilatih pada corpus data yang lebih besar dan dalam durasi waktu yang lebih lama dibandingkan BERT.
2. Menghilangkan NSP (Next Sentence Prediction): RoBERTa hanya menggunakan task MLM (Masked Language Modelling) untuk melakukan *pre-training*.
3. Training dengan *dynamic masking*: Untuk mengubah masking secara dinamis pada setiap epoch sehingga memaksa model untuk lebih memerhatikan keseluruhan teks yang ada.

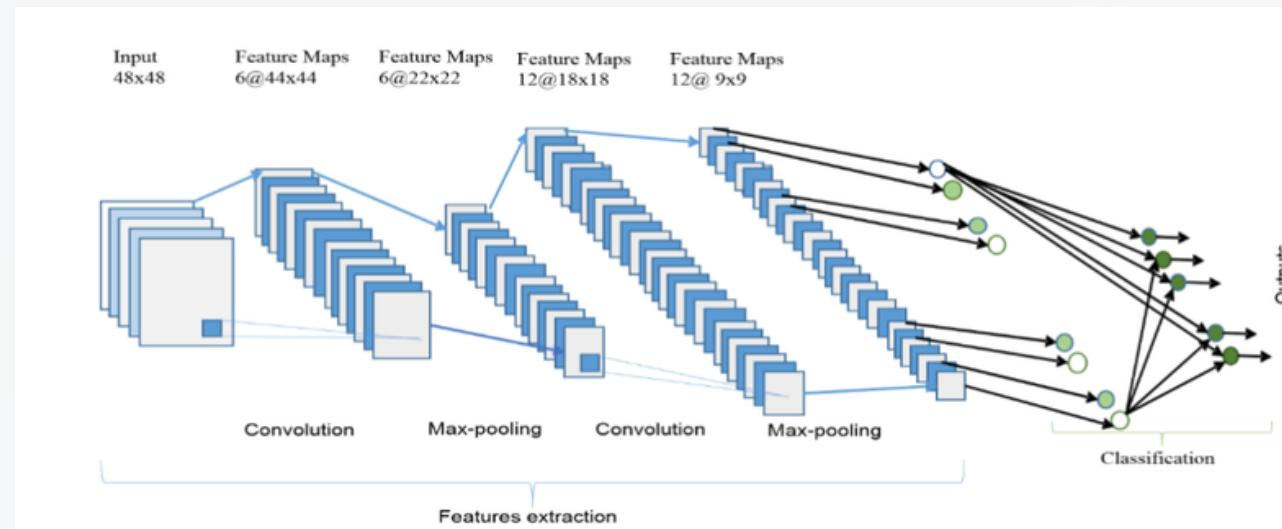


# CNN, RNN, DAN GRU

## CNN

### Convolutional Neural Network

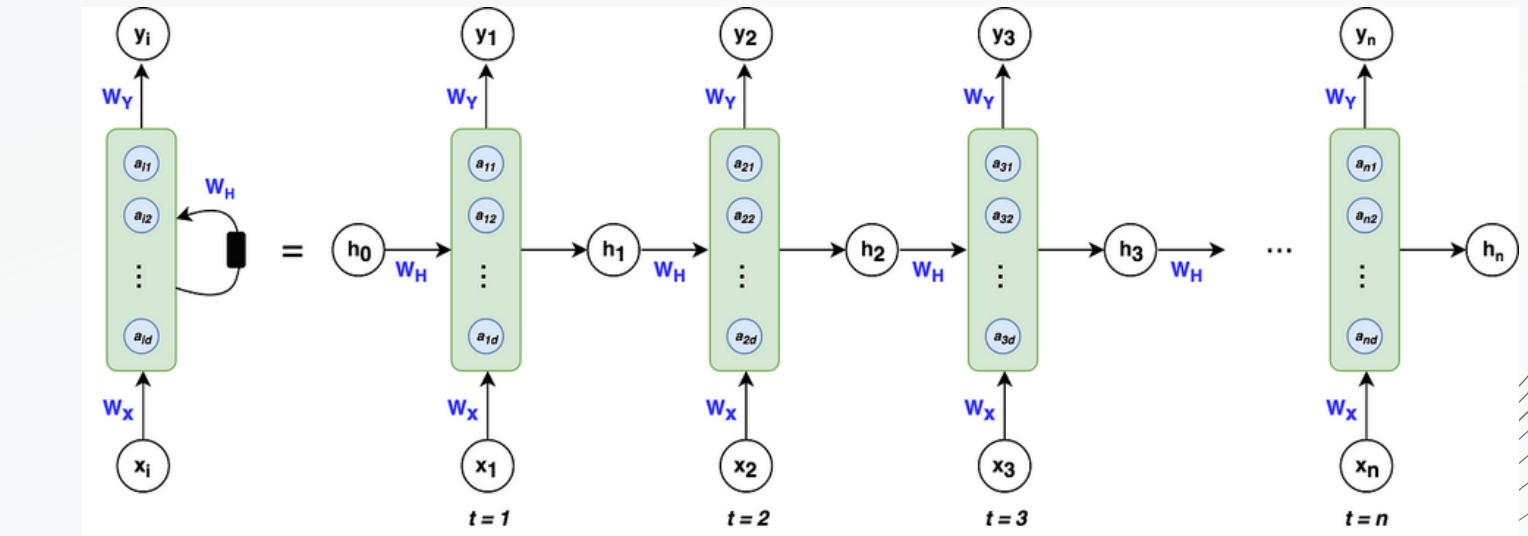
Feedforward neural network dengan lapisan convolution dan pooling yang digunakan untuk mempelajari representasi data dengan fitur berupa area dua dimensi.



## RNN

### Recurrent Neural Network

Salah satu arsitektur *Artificial Neural Network* yang dirancang khusus untuk memproses data yang berurutan/sequential seperti *time series*. RNN mampu menyimpan memori/ingatan (*feedback loop*) yang memungkinkan untuk mengenali pola data dengan baik, kemudian menggunakan untuk membuat prediksi yang akurat.



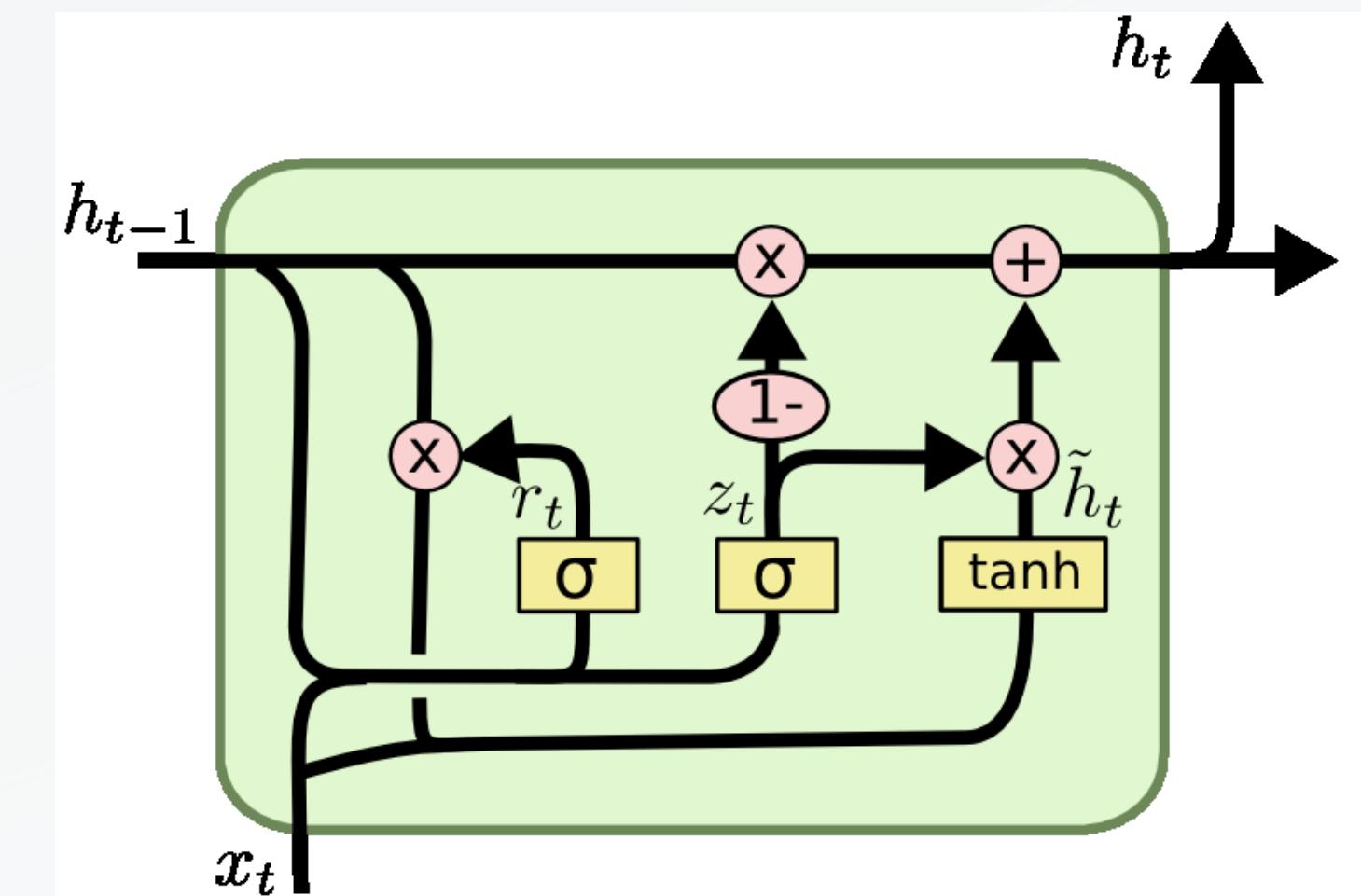
# CNN, RNN, DAN GRU



## GRU

### Gated Recurrent Unit

RNN hasil penyederhanaan dari arsitektur LSTM standar yang dikembangkan sebagai salah satu metode yang menggunakan lebih sedikit parameter tanpa mengurangi kinerja dari LSTM. GRU menggabungkan dua state dari LSTM, yaitu *cell state* dan *hidden state* menjadi satu *hidden state*. Akibanya, *output gate* dapat dibuang karena *output gate* hanya memutuskan beberapa banyak informasi pada *cell state* yang diteruskan ke *hidden state*.



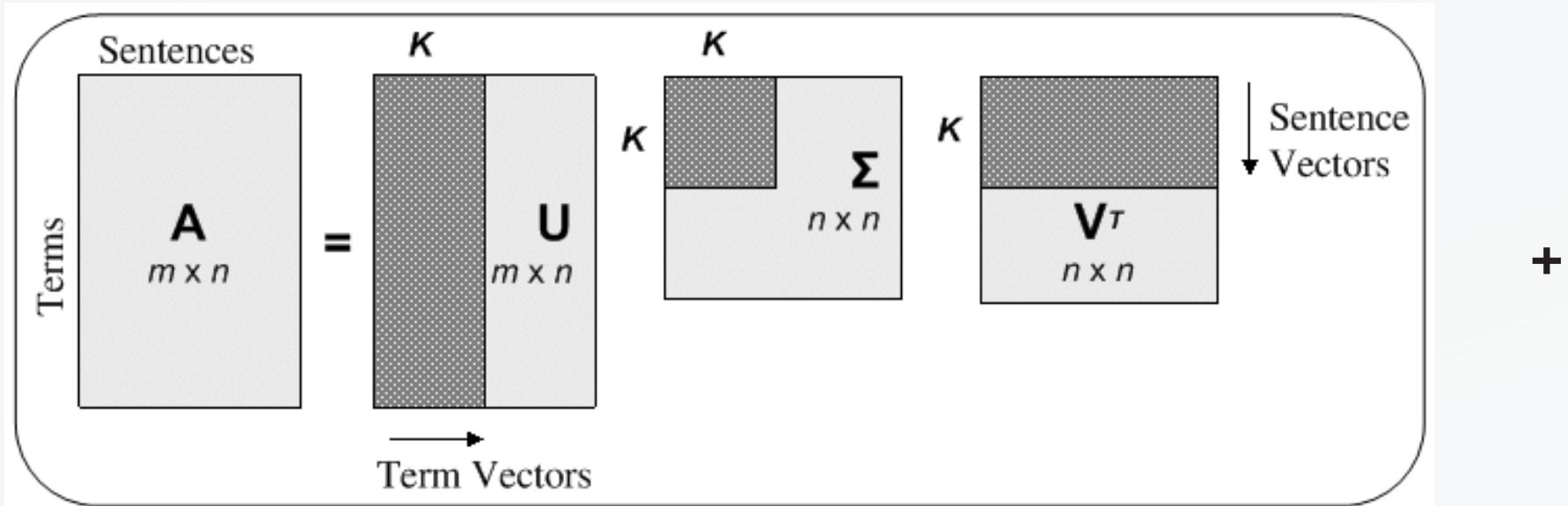
# EFCM



## EFCM

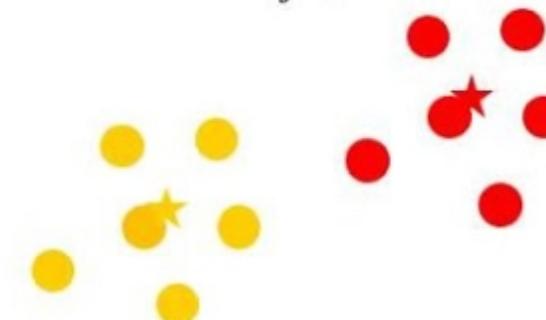
### Eigenspaced-Based Fuzzy C-Means

Metode *clustering* yang memanfaatkan metode reduksi dimensi *Truncated Singular Value Decomposition* (TSVD) dengan metode FCM (Murfi, 2018).



TSVD

$$J = \sum_{i=1}^2 \sum_{j=1}^N u_{ij}^m \|x_j - c_i\|^2 = \sum_{j=1}^N u_{1j}^m \|x_j - c_1\|^2 + \sum_{j=1}^N u_{2j}^m \|x_j - c_2\|^2$$



Fuzzy C-Means

# PARAMETER MODEL ANALISIS SENTIMEN

Parameter yang Digunakan Untuk Analisis Sentimen:

PARAMETER	DESKRIPSI
Filter	Jumlah filter yang digunakan pada convolutional layer.
Units	Dimensi vektor output pada layer GRU
Region Size	Ukuran region/kernel pada convolutional layer.
Learning Rate	Digunakan untuk mempercepat laju pembelajaran dalam menentukan bobot yang optimal
Regulasi L2	Teknik yang sering digunakan untuk regularisasi model. Parameter regularisasi L2 ini ada pada CNN layer, GRU layer, dan fully connected layer.

# OPTIMASI PARAMETER MODEL DETEKSI TOPIK

Parameter yang Digunakan untuk Deteksi Topik:

PARAMETER	DESKRIPSI
Number of Top Words	Jumlah kata per topik yang akan direpresentasikan
Number of Topics	Jumlah topik yang akan direpresentasikan
Number of n Components	Dimensi bentuk tereduksi
Value of m	Konstanta Fuzziness

# EVALUASI MODEL

## AKURASI

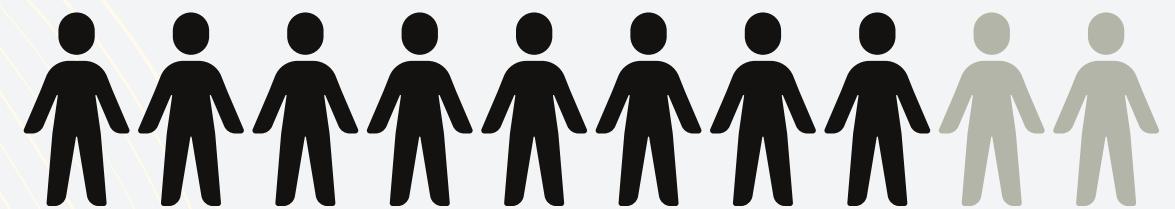
Akurasi =

$$\frac{\text{jumlah data yang sesuai}}{\text{jumlah seluruh data}} \times 100\%$$

## TOPIC COHERENCE WORD2VEC (TC-W2V)

$$TC - W2V = \frac{1}{\binom{N}{2}} \sum_{j=2}^N \sum_{i=1}^{j-1} sim(t_j, t_i)$$

# 3. SIMULASI DAN ANALISIS HASIL

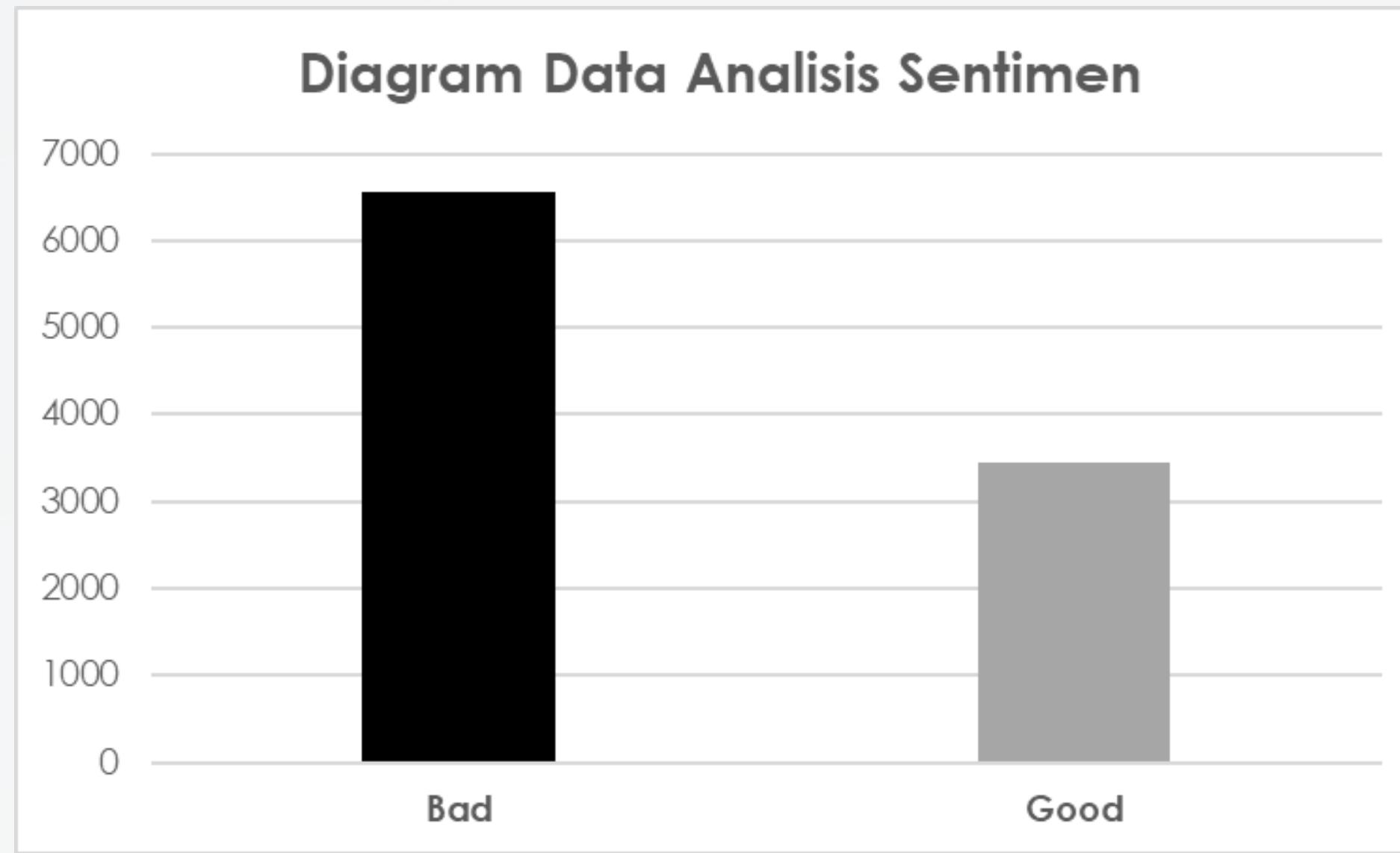




**LINK PROGRAM DAPAT DIAKSES PADA LINK  
BERIKUT:**

**[HTTPS://BIT.LY/TUGASAKHIRK2KELOMPOK8](https://bit.ly/tugasakhirk2kelompok8)**

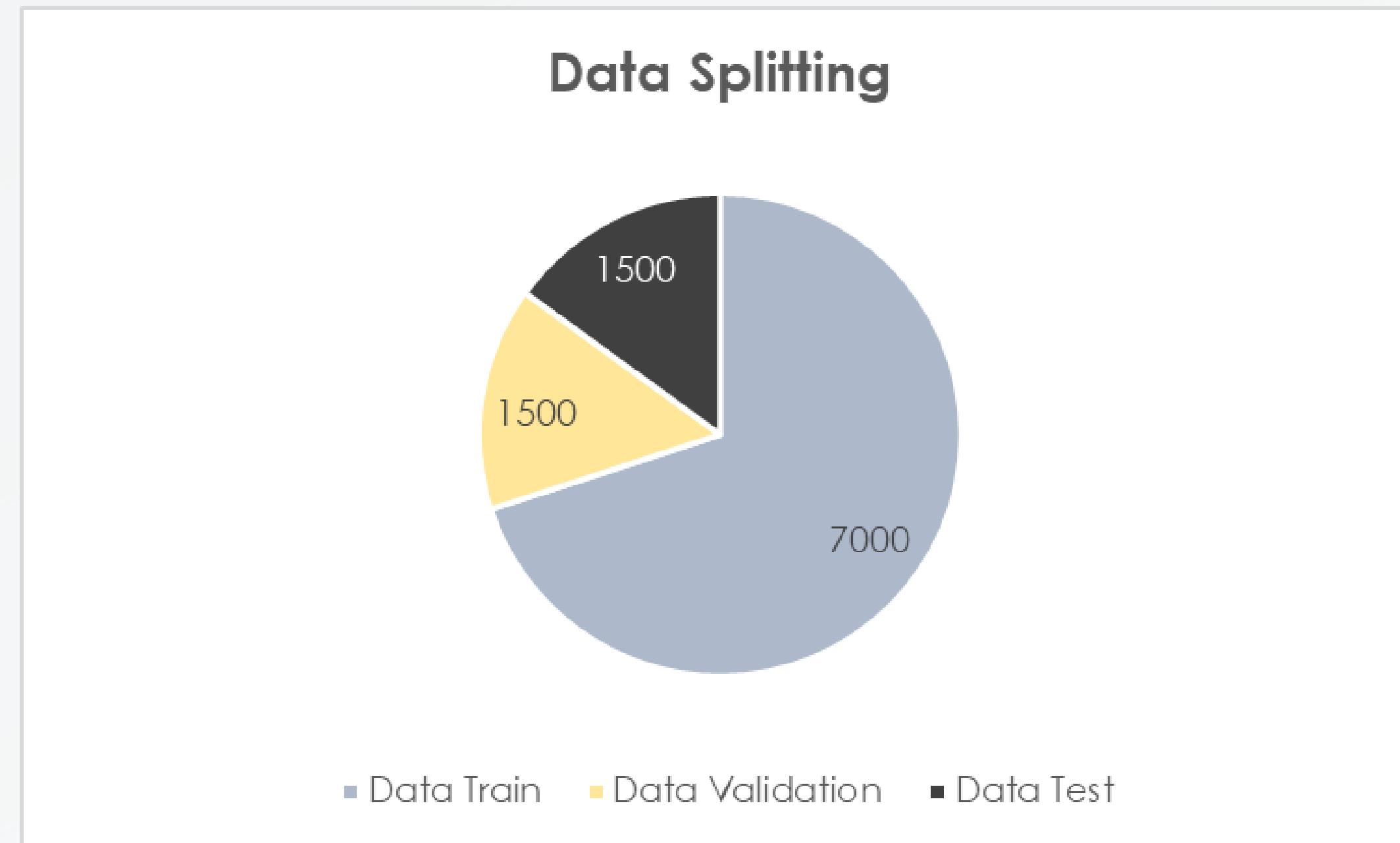
# PENGAMBILAN DATA TRAIN



Dataset yang berisi *tweet* mengenai ChatGPT berasal dari Kaggle.com. Total jumlah data sentimen sebanyak 10.000 (diambil menggunakan random sample dari total keseluruhan data asli) data, yang terdiri dari:

- Jumlah data sentimen negatif sebanyak 6555 data.
- Jumlah data sentimen positif sebanyak 3445 data.

# PEMISAHAN DATA TRAIN



Dari total 10.000 data dibagi menjadi data *train*, *validation*, dan *test* dengan rasio 70:15:15.

# DATA CLEANING DAN PREPROCESSING

## Pembersihan Data

Tahapan *preprocessing* data untuk pembersihan data:

1. Mengubah semua huruf besar menjadi huruf kecil
2. Menghapus www. dan https://
3. Menghapus @ (username) dan # (hashtag)
4. Menghapus tanda \, ', ",
5. Menghapus emoji
6. Menghapus spasi berlebih
7. Menghapus kata-kata yang tidak signifikan menggunakan stopwords.

# DATA CLEANING DAN PREPROCESSING

SEBELUM

ChatGPT is out in research preview! \n\nIt is optimized for dialogue and is ready to have a conversation. Don't be shy--give it a try and help us make it better!\n\n<https://t.co/GviygS>  
SLRo

SESUDAH

chatgpt is out in research preview! it is optimized for dialogue and is ready to have a conversation. don't be shy—give it a try and help us make it better!

# DATA CLEANING DAN PREPROCESSING

- **One-hot-encoding** adalah langkah untuk mengubah variabel kategorik menjadi variabel numerik, dengan cara membuat array binary untuk sentimen negatif dan sentimen positif yang ada di data training sehingga dapat diproses dalam model.
- **Sentimen negatif** memiliki array [1,0] dan **sentimen positif** memiliki array [0,1]
- Array yang dibentuk berdimensi 2 agar dapat di input ke dalam model BERT
- Dimensi pertama pada array untuk sifat negatif, dan dimensi kedua untuk sifat positif.
- Jika ulasan bersifat negatif, dimensi pertama dari array ulasan tersebut bernilai 1, dan lainnya 0. Begitu pula sebaliknya untuk ulasan bersifat positif.

ULASAN	LABEL	ULASAN	ARRAY
Ulasan 1	0	Ulasan 1	[1 , 0]
Ulasan 2	1	Ulasan 2	[0 , 1]

# DATA CLEANING DAN PREPROCESSING

## Representasi Data

Representasi data menggunakan metode RoBERTa

1. Tokenisasi *input* menggunakan RoBERTa *tokenizer*
2. Membatasi banyak token pada tiap satu data
3. Menyamakan panjang dokumen dengan *padding* dan *truncation*
4. Mengubah token menjadi bilangan bulat sehingga *input* dapat dibaca oleh model RoBERTa dengan tahapan numerikalisasi (*numericalization*).

# OPTIMASI PARAMETER MODEL ANALISIS SENTIMEN

Parameter Tuning pada Model Analisis Sentimen

LAYER	PARAMETER	NILAI
GRU	Unit	100; 150; 200
	L2 Kernel	0.001; 0.01
	L2 Recurrent	0.001; 0.01
CNN	Jumlah filter	200; 250; 300
	Region size	3; 4; 5
	L2 CNN	0.001; 0.01
Fully Connected	L2 Dense	0.001 ; 0.01

# OPTIMASI PARAMETER MODEL ANALISIS SENTIMEN

Teknik optimasi pembaruan bobot	:	Adam
Teknik hyperparameter tuning	:	Bayesian
Learning rate	:	0.001
Trial	:	5
Epoch	:	100
Batch size	:	32
Fungsi aktivasi		
• CNN Layer	:	ReLU
• Fully connected layer	:	Softmax
Early stopping		
• Kriteria	:	validation accuracy
• Patience	:	5

# OPTIMASI HYPERPARAMETER MODEL ANALISIS SENTIMEN

RUNNING KE-	HYPERPARAMETER	AKURASI
1	units: 150 kernel regularizer: 0.001 rec regularizer: 0.01 filters: 300 kernel size: 3 kernel cnn: 0.001 kernel dense: 0.01	85.91644167900080%
2	units: 100 kernel regularizer: 0.001 rec regularizer: 0.001 filters: 200	88.30645084381100%

RUNNING KE-	HYPERPARAMETER	AKURASI
2	kernel size: 3 kernel cnn: 0.001 kernel dense: 0.01	88.30645084381100%
3	units: 150 kernel regularizer: 0.001 rec regularizer: 0.001 filters: 300 kernel size: 3 kernel cnn: 0.001 kernel dense: 0.001	88.04566860198970%
4	units: 150 kernel regularizer: 0.001 rec regularizer: 0.001 filters: 300 kernel size: 3 kernel cnn: 0.001 kernel dense: 0.001	87.105441093444800%

RUNNING KE-	HYPERPARAMETER	AKURASI
5	units: 150 kernel regularizer: 0.001 rec regularizer: 0.001 filters: 300 kernel size: 4 kernel cnn: 0.001 kernel dense: 0.01	88.24714422225950%
6	units: 150 kernel regularizer: 0.001 rec regularizer: 0.01 filters: 300 kernel size: 5 kernel cnn: 0.001 kernel dense: 0.01	88.78442049026480%

RUNNING KE-	HYPERPARAMETER	AKURASI
7	units: 200 kernel regularizer: 0.001 rec regularizer: 0.01 filters: 250 kernel size: 5 kernel cnn: 0.001 kernel dense: 0.01	87.44123578071590%
8	units: 150 kernel regularizer: 0.001 rec regularizer: 0.001 filters: 250 kernel size: 4 kernel cnn: 0.01 kernel dense: 0.001	88.11282515525810%

RUNNING KE-	HYPERPARAMETER	AKURASI
9	units: 100 kernel regularizer: 0.001 rec regularizer: 0.001 filters: 250 kernel size: 4 kernel cnn: 0.001 kernel dense: 0.001	89.18737173080440%
10	units: 200 kernel regularizer: 0.001 rec regularizer: 0.001 filters: 300 kernel size: 3 kernel cnn: 0.001 kernel dense: 0.01	88.85157704353330%

# HYPERPARAMETER TERPILIH DALAM MODEL ANALISIS SENTIMEN (NILAI AKURASI TERTINGGI)

units: 100

kernel regularizer: 0.001

rec regularizer: 0.001

filters: 250

kernel size: 4

kernel cnn: 0.001

kernel dense: 0.001

Dengan nilai akurasi : 89.18737173080440%

# OPTIMASI PARAMETER MODEL DETEKSI TOPIK

Parameter Tuning pada Model Topik Deteksi

PARAMETER	NILAI
Number of Top Words	10; 15; 20
Number of Topics	5; 10; 15; 20
Number of n Components	3; 5
Value of m (Konstanta Fuzziness)	1.1; 1.3

# OPTIMASI HYPERPARAMETER MODEL DETEKSI TOPIK

RUNNING KE-	HYPERPARAMETER	COHERENCE VALUE
1	Number of Top Words : 20 Number of Topics : 20 Number of n components: 5 Value of m : 1.1	0.2868993450246463
2	Number of Top Words : 20 Number of Topics : 20 Number of n components: 5 Value of m : 1.1	0.2883502281000267
3	Number of Top Words : 20 Number of Topics : 20 Number of n components: 5 Value of m : 1.1	0.2862529789624252

RUNNING KE-	HYPERPARAMETER	COHERENCE VALUE
4	Number of Top Words : 20 Number of Topics : 20 Number of n components: 5 Value of m : 1.3	0.2855347397673513
5	Number of Top Words : 20 Number of Topics : 20 Number of n components: 5 Value of m : 1.1	0.2860158775618793
6	Number of Top Words : 15 Number of Topics : 20 Number of n components: 5 Value of m : 1.3	0.2863916631685362
7	Number of Top Words : 15 Number of Topics : 15 Number of n components: 5 Value of m : 1.3	0.2839849667256982

<b>RUNNING KE-</b>	<b>HYPERPARAMETER</b>	<b>COHERENCE VALUE</b>
8	Number of Top Words : 20 Number of Topics : 20 Number of n components: 5 Value of m : 1.1	0.2885334946804027
9	Number of Top Words : 15 Number of Topics : 20 Number of n components: 5 Value of m : 1.1	0.28368135193318794
10	Number of Top Words : 20 Number of Topics : 20 Number of n components: 5 Value of m : 1.3	0.2857051108736706

# HYPERPARAMETER TERPILIH DALAM MODEL DETEKSI TOPIK (NILAI COHERENCE TERTINGGI)

Number of Top Words : 20

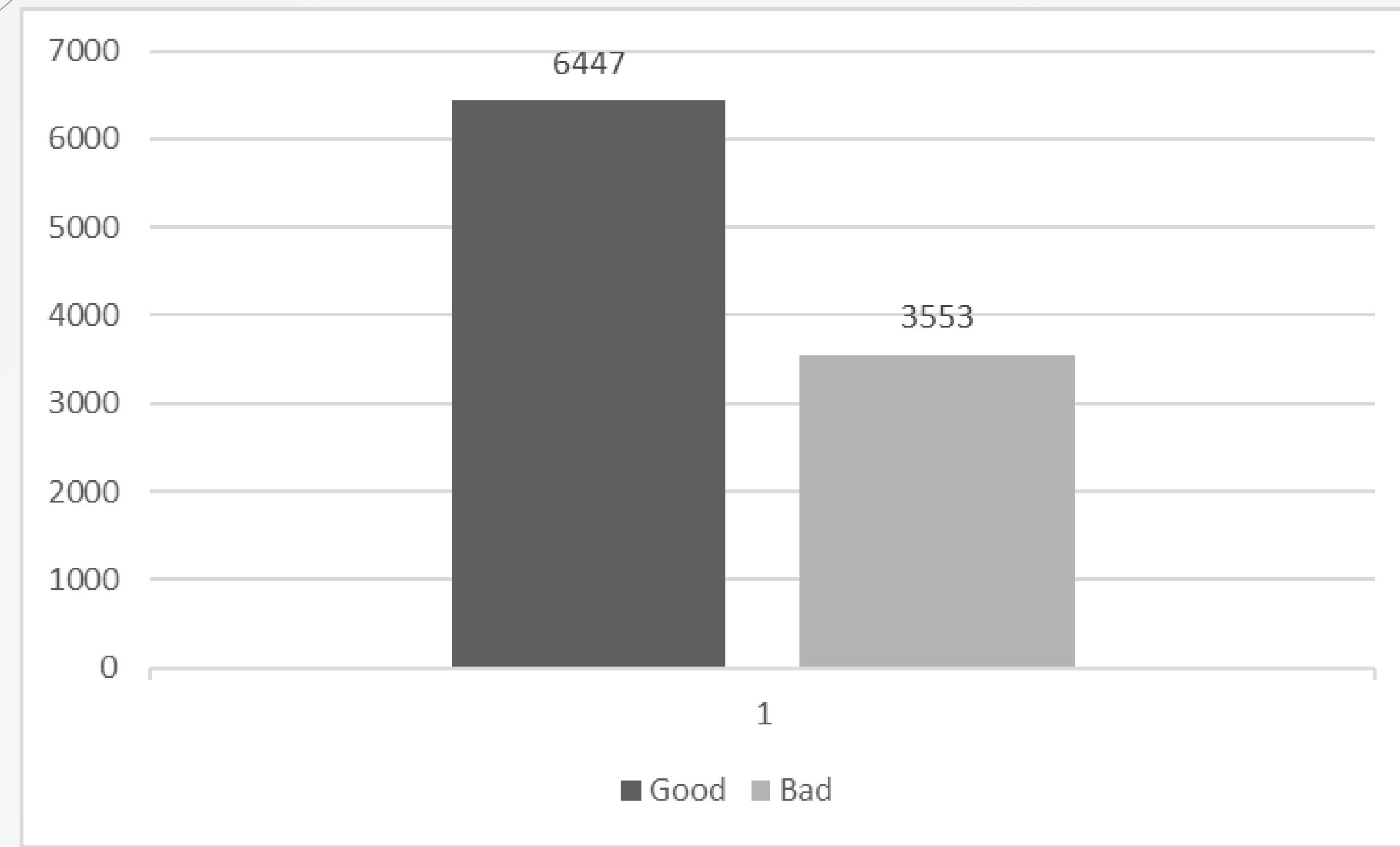
Number of Topics : 20

Number of n components: 5

Value of m : 1.1

Dengan nilai coherence : 0.2885334946804027

# VISUALISASI HASIL PREDIKSI SENTIMEN DATASET PREDIKSI MENGGUNAKAN MODEL ROBERTA+GRU-CNN



# ANALISIS HASIL

Didapat 10 *top words* sebagai representasi topik oleh RoBERTa dari data prediksi sebagai berikut

NOMOR	TOP WORDS	JUMLAH
Topic 18	<code>['chat', 'using', 'use', 'gpt', 'ask', 'gpt4', 'make', 'question', 'anyone', 'new', 'future', 'get', 'answer', 'asked', 'need', 'prompt', 'google', 'wrote', 'via', 'coinex']</code>	682
Topic 1	<code>['chat', 'like', 'asked', 'question', 'one', 'people', 'could', 'student', 'even', 'human', 'think', 'come', 'tool', 'answer', 'amp', 'know', 'say', 'would', 'make', 'way']</code>	674
Topic 8	<code>['gpt', 'gpt4', 'airdrop', 'new', 'nft', '500', 'time', 'powerful', 'chatbot', 'current', 'microsoft', 'artificialintelligence', 'powered', 'tool', 'rsi', 'create', 'use', 'chat', 'free', 'google']</code>	628
Topic 4	<code>['tool', 'like', 'model', 'chat', 'use', 'language', 'way', 'using', 'text', 'gpt', 'human', 'make', 'data', 'new', 'question', 'answer', 'code', 'learning', 'write', 'need']</code>	614

NOMOR	TOP WORDS	JUMLAH
Topic 15	['amp', 'like', 'people', 'think', 'human', 'chat', 'use', 've', 'using', 'would', 'make', 'thing', 'even', 'could', 'answer', 'question', 'need', 're', 'know', 'ca']	567
Topic 5	['artificialintelligence', 'technology', 'microsoft', 'via', 'machinelearning', 'artificial', 'future', 'chatbot', 'tech', 'crypto', 'business', 'intelligence', 'make', 'chat', 'google', 'learn', 'new', 'use', 'gpt4', 'ml']	552
Topic 10	['artificialintelligence', 'stockstobuy', 'yolo', 'daytrading', 'redditarmy', 'bottomfishing', 'tothemoon', 'stockstowatch', 'fomo', 'news', 'investment', 'stock', 'investing', 'amp', 'tech', 'trading', 'option', 'iot', 'look', 'metaverse']	537
Topic 2	['language', 'model', 'gpt', 'new', 'using', 'google', 'help', 'get', 'tool', 'chat', 'use', 'like', 'microsoft', 've', 'code', 'prompt', 'gpt4', 'api', 'time', 'make']	512
Topic 11	['chat', 'google', 'let', 've', 'll', 'know', 'write', 'asked', 'use', 'via', 'good', 'like', 'first', 'prompt', 'bing', 'gpt', 're', 'day', 'one', 'valentine']	506

NOMOR	TOP WORDS	JUMLAH
Topic 3	['06', 'sunrise', 'sunset', 'chat', 'google', 'pm', 'microsoft', '2023', 'via', 'hourlyupdate', 'km', 'tempupdate', 'temp', 'weatherupdate', 'humidity', 'chatbot', 'wind', 'midjourney', 'status', 'speed']	501
Topic 13	['technology', 'new', 'microsoft', 'content', 'amp', 'read', 'artificialintelligence', 'tech', 'google', 'free', 'chatbot', 'ai', 'business', 'like', 'blog', 'gpt', 'learn', 'bing', 'using', 'create']	490
Topic 19	['machinelearning', 'tech', 'artificialintelligence', 'nft', 'python', 'google', 'web3', 'bitcoin', 'marketing', 'chatbot', 'chatgpt3', 'airdrop', 'cybersecurity', 'fintech', 'microsoft', '100daysofcode', 'crypto', 'iot', 'technology', 'bigdata']	486
Topic 6	['chat', 'think', 'asked', 'write', 'question', 'answer', 'know', 'would', 'like', 'use', 'got', 'need', 'thought', 'could', 'good', 'ask', 'human', 'already', 'one', 'people']	467
Topic 16	['google', 'gpt4', 'amp', 'new', 'technology', 'microsoft', 'artificialintelligence', 'bing', 'bard', 'chatbot', 'let', 'tech', 'nft', 'world', 'gpt', 'future', 'search', 'one', 'chat', 'week']	439

NOMOR	TOP WORDS	JUMLAH
Topic 14	['chat', 'write', 'asked', 'know', 'like', 'gt', 'really', 'good', 'te', 'im', 'think', 'thanks', 'ask', 'amo', 'thing', 'love', 'could', 'wonder', 'job', 'prompt']	428
Topic 9	['chat', 'ca', 'like', 'write', 've', 'asked', 'know', 'll', 'think', 'one', 'people', 'better', 'answer', 'would', 'question', 'say', 're', 'get', 'much', 'could']	409
Topic 20	['amp', 'like', 'human', 'one', 'could', 'chat', 'use', 'people', 'many', 'code', 'tool', 'student', 'model', 'also', 'language', 'data', 'potential', 'think', 'good', 'would']	408
Topic 17	['chat', 'gt', 'use', 'ask', 'love', 'mem', 'try', 'thought', 'save', 'friend', 'best', 'thread', 'amo', 'written', 'amazing', 'via', 'time', 'well', 'thank', 'el']	406
Topic 12	['chat', 'gpt', 'using', 'use', 'code', 'write', 'api', 'google', 'check', 'help', 'new', 'writing', 'make', 'via', 'model', 'tool', 'answer', 'result', 'ask', 'get']	380

NOMOR	TOP WORDS	JUMLAH
Topic 7	<p>[chat', 'asked', 'write', 'artificialintelligence', 'midjourney', 'like', 'know', 'via', 'think', 'say', 'could', 'aiart', 'use', 'make', 'google', 'come', 'would', 'tweet', 'bard', 'talk']</p>	307

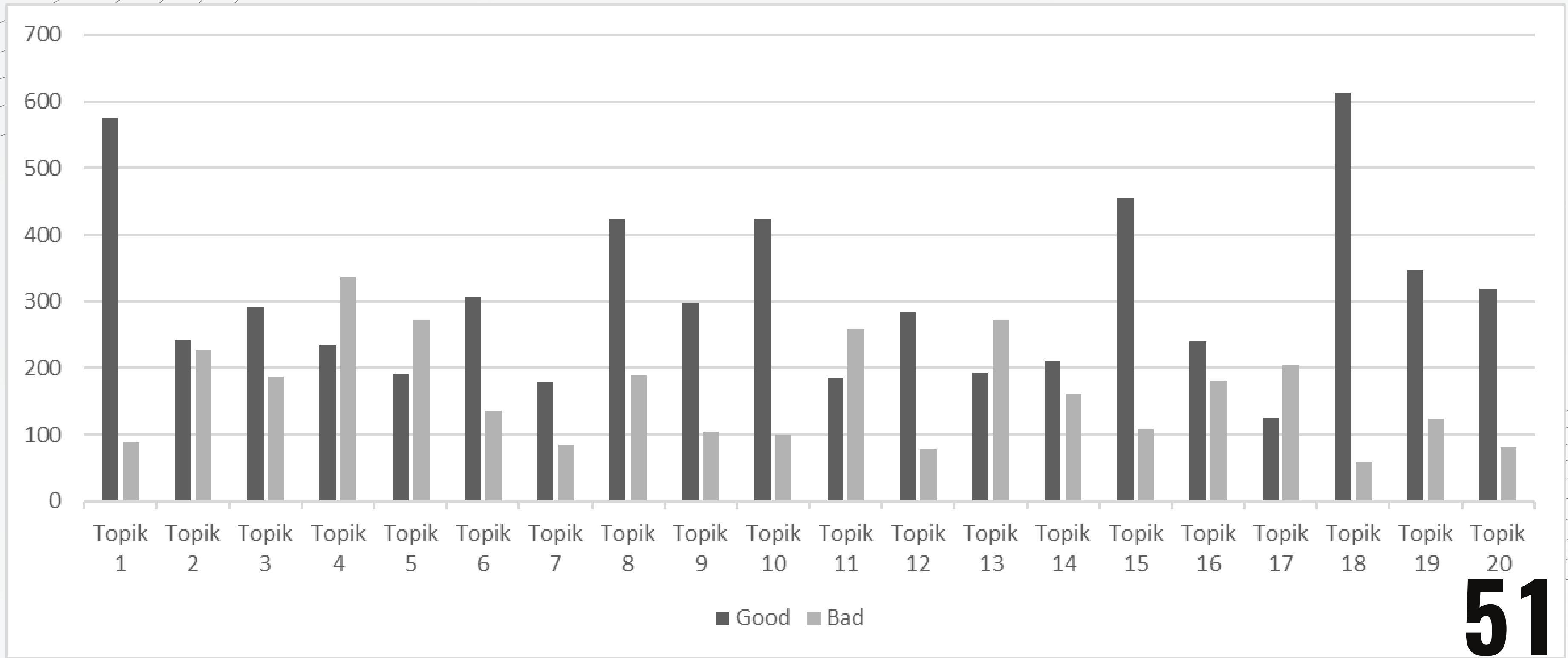
# INTERPRETASI TOPIK

No Topik	Interpretasi
Topik 18	Anyone can ask ChatGPT anything about the future of cryptocurrency
Topik 1	ChatGPT as a question-answering tool for student
Topik 8	People excitement about ChatGPT chatbot that uses artificial intelligence
Topik 4	ChatGPT uses human-like language when answering people's question
Topik 15	People curiosiy about how ChatGPT works
Topik 5	ChatGPT uses machine learning and artifial intelligence to generates answers in various domains, such as tech, crypto, business, etc

No Topik	Interpretasi
Topik 10	Artificial intelligence is used in trading to look for option stocks to buy
Topik 2	A new tool to make language model is using ChatGPT for code prompt
Topik 11	ChatGPT asked to write prompts for Valentine day
Topik 3	Hourly updates on the weather, humidity, wind speed, and temp can be tracked by ChatGPT via Google
Topik 13	ChatGPT is a new Artificial Intelligence technology and a free-to-use chatbot
Topik 19	Machine learning and Artificial Intelligence is used in ChatGPT3 bringing the topic of cybersecurity, bitcoin, marketing, crypto, and IoT
Topik 6	Human could use ChatGPT to ask any question and the it would know the answer

No Topik	Interpretasi
Topik 16	ChatGPT as the future of chatbots and artificial intelligence in a tech-driven world.
Topik 14	ChatGPT and its capability to make well-written responses in chats, understand what people like, and provide good answers that leave a positive impression.
Topik 9	ChatGPT improved chat experiences with better responses and understanding of user needs.
Topik 20	ChatGPT and its potential to enhance human-like chat interactions and the code tools it offers.
Topik 17	Public excitement about the positive impact of ChatGPT.
Topik 12	ChatGPT is useful as an API for writing, checking, and generating content.
Topik 7	Artificial intelligence through chat and discussing its potential applications.

# VISUALISASI KOLABORATIF HASIL PREDIKSI SENTIMEN MENGGUNAKAN MODEL ROBERTA+GRU-CNN DENGAN HASIL DETEKSI TOPIK MENGGUNAKAN ROBERTA+EFCM DARI DATASET PREDIKSI



# **ANALISIS VISUALISASI HASIL KOLABORATIF**

**Topik 1: ChatGPT as question-answering tool for student**  
didominasi dengan sentimen positif, dengan jumlah 575 data positif

**Topik 2: A new tool to make language model is using ChatGPT for code prompt**  
didominasi dengan sentimen positif, dengan jumlah 243 data positif

**Topik 3: Hourly update on the weather, humidity, wind speed, and temperature can be tracked by ChatGPT via Google**  
didominasi dengan sentimen positif, dengan jumlah 292 data positif

**Topik 4: ChatGPT uses human-like language when answering people's question**  
didominasi dengan sentimen negatif, dengan jumlah 337 data negatif

# **ANALISIS VISUALISASI HASIL KOLABORATIF**

**Topik 5: ChatGPT uses machine learning and artificial intelligence to generates answers in various domains, such as tech, crypto, business, etc**

didominasi dengan sentimen negatif, dengan jumlah 272 data negatif

**Topik 6: Human could use ChatGPT to ask any question and the it would know the answer**

didominasi dengan sentimen positif, dengan jumlah 308 data positif

**Topik 7: Artificial intelligence through chat and discussing its potential applications**

didominasi dengan sentimen positif, dengan jumlah 180 data positif

**Topik 8: People excitement about ChatGPT's chatbot that uses artificial intelligence**

didominasi dengan sentimen positif, dengan jumlah 424 data positif

# **ANALISIS VISUALISASI HASIL KOLABORATIF**

**Topik 9: ChatGPT improved chat experiences with better responses and understanding of user needs**

didominasi dengan sentimen positif, dengan jumlah 298 data positif

**Topik 10: Artificial intelligence is use in trading to look for option stocks to buy**

didominasi dengan sentimen positif, dengan jumlah 423 data positif

**Topik 11: ChatGPT asked to write prompt for valentine day**

didominasi dengan sentimen negatif, dengan jumlah 258 data negatif

**Topik 12: ChatGPT is useful as an API for writing, checking, and generating content**

didominasi dengan sentimen positif, dengan jumlah 283 data positif

# **ANALISIS VISUALISASI HASIL KOLABORATIF**

**Topik 13: ChatGPT is a new artificial intelligence technology free-to-use chatbot**

didominasi dengan sentimen negatif, dengan jumlah 272 data negatif

**Topik 14: ChatGPT and its capability to make well-written responses in chats, understand what people like, and provide good answers that leave a positive impression**

didominasi dengan sentimen positif, dengan jumlah 210 data positif

**Topik 15: People curious about how ChatGPT works**

didominasi dengan sentimen positif, dengan jumlah 456 data positif

**Topik 16: ChatGPT as the future of chatbots and artificial intelligence in a tech-driven world**

didominasi dengan sentimen positif, dengan jumlah 240 data positif

# **ANALISIS VISUALISASI HASIL KOLABORATIF**

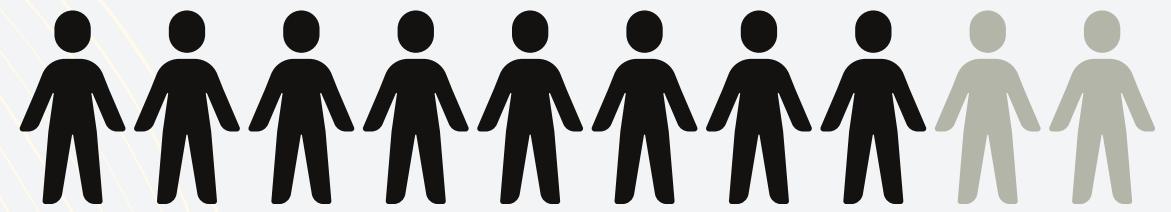
**Topik 17: Public excitement about the positive impact of ChatGPT**  
didominasi dengan sentimen negatif, dengan jumlah 204 data negatif

**Topik 18: Anyone can ask ChatGPT anything about the future of cryptocurrency**  
didominasi dengan sentimen positif, dengan jumlah 613 data positif

**Topik 19: Machine learning and artificial intelligence is used in ChatGPT bringing the topic of cybersecurity, bitcoin, marketing, crypto, and IoT**  
didominasi dengan sentimen positif, dengan jumlah 347 data positif

**Topik 20: ChatGPT and its potential to enhance human-like chat interactions and the code tools it offers**  
didominasi dengan sentimen positif, dengan jumlah 319 data positif

# 4. KESIMPULAN



# KESIMPULAN

- Tweet pada dataset mengenai ChatGPT yang kami gunakan memiliki 3 topik dengan jumlah data terbanyak, yaitu:
  1. **Topik 18:** Anyone can ask ChatGPT anything about the future of cryptocurrency
  2. **Topik 1:** ChatGPT as a question-answering tool for student
  3. **Topik 8:** People excitement about ChatGPT chatbot that uses artificial intelligence
- Sebanyak **15 topik** didominasi oleh **sentimen positif** dengan persentase terbesar terdapat pada Topik 18, yaitu "Anyone can ask ChatGPT anything about the future of cryptocurrency"
- Sebanyak **5 topik** didominasi oleh **sentimen negatif** dengan persentase terbesar terdapat pada Topik 4, yaitu "ChatGPT uses human-like language when answering people's question"

# DAFTAR PUSTAKA

Mubin uL Haque, isuru dharmadasa, zarrin t. sworna, roshan n. rajapakse, hussain ahmad. (2022). Exploring Sentiments of ChatGPT Early Adopters using Twitter Data.

Kian Long Tan, Chin Poo Lee, dan Kian Ming Lim. (2023). RoBERTa-GRU: A Hybrid Deep Learning Model for Enhanced Sentiment Analysis.

Syamsuriani. (2022). Analisis Kinerja BERT sebagai Metode Representasi Teks pada Model Hybrid Deep Learning untuk Analisis Sentimen Berbahasa Indonesia.

Nicholas Ramos Ricardo. (2022). Analisis Performa EFCM dengan BERT sebagai Representasi Teks pada Pendekripsi Topik.

Laure thompson, david mimno. (2020). Topic Modeling with Contextualized Word Representation Clusters



**TERIMA  
KASIH**