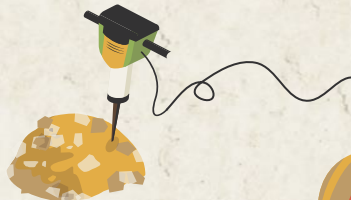
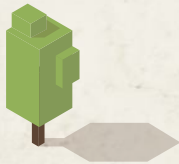


# IMPLEMENTASI PENAMBANGAN DATA UNTUK MENGENALI EMOSI PADA ULASAN PRODUK TOKO ONLINE DENGAN METODE CRISP-DM

Penulis:  
Valentio Stanley Gunadi  
2540096011  
Computer Science (S1)

Dosen Pembimbing:  
Said Achmad, S.Kom., M.Kom.



# Outline



## 01

**Pendahuluan**

## 02

**Tinjauan Pustaka**

## 03

**Metode Penelitian**

## 04

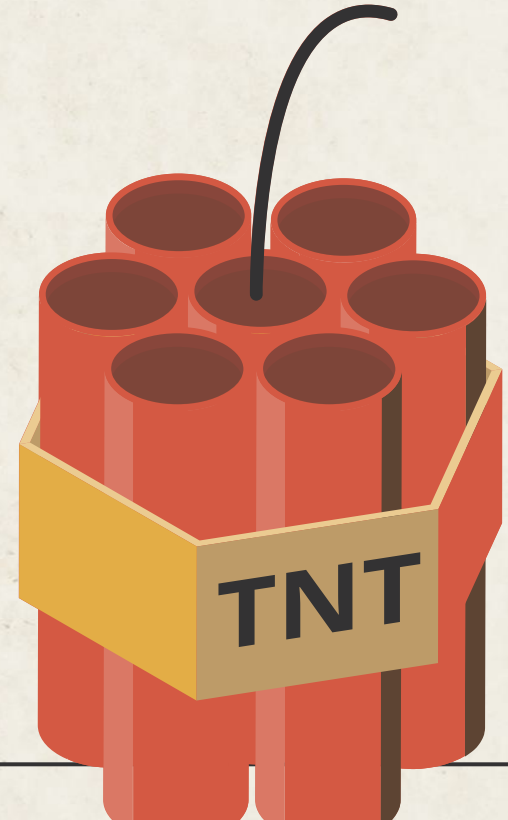
**Kesimpulan**



# 01

# Pendahuluan

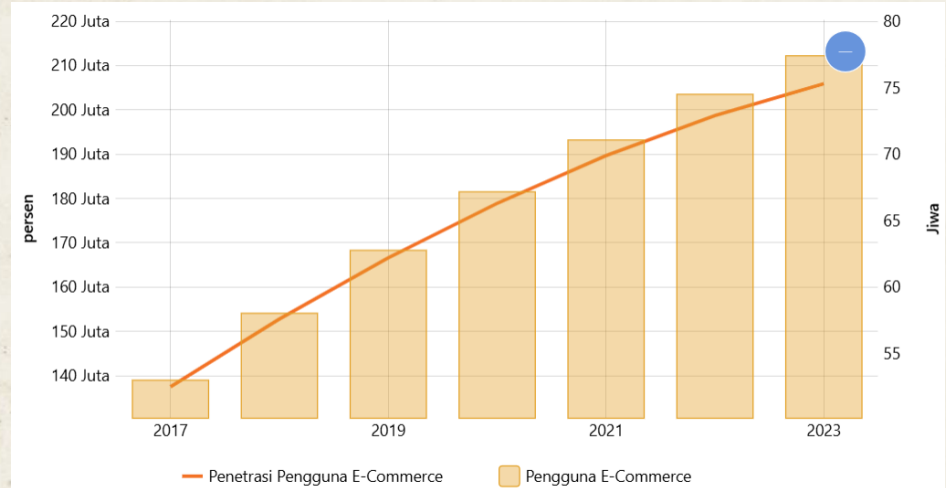
Bab 1



# Latar Belakang

Jumlah pengguna *e-commerce* meningkat secara signifikan pada setiap tahun dan seterusnya, (Jayani, 2019). Beberapa pengguna tidak berbelanja ataupun tidak berjualan melalui online dengan alasan faktor yang mempengaruhi kepercayaan, (Ramadhany et al., 2023).

Beberapa algoritma diterapkan pada penelitian-penelitian untuk menganalisis data ulasan produk dengan *emotion*. Penulis menerapkan algoritma dari penelitian dengan tujuan untuk menemukan model terbaik untuk penelitian ini, maupun memenuhi kebutuhan bisnis penjual dari CRISP-DM.



Sumber:

<https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/be57ed78ccc44e2/tren-pengguna-e-commerce-2017-2023>



# Identifikasi Masalah



## Rumusan Masalah

Bagaimana mengolah dataset ulasan produk toko online dengan menerapkan metode Cross Industry Standard Process for Data Mining (CRISP-DM)?

Bagaimana menganalisis emosi pada dataset ulasan produk toko online dengan menerapkan teknik classification?

## Tujuan Penelitian

Mengolah dataset ulasan produk toko online sehingga menemukan pola baru dengan menerapkan metode Cross Industry Standard Process for Data Mining (CRISP-DM)

Menerapkan teknik classification yang efektif berdasarkan pola klasifikasi yang dihasilkan pada jurnal atau artikel penelitian yang bersangkutan untuk menganalisis emosi pada data ulasan produk toko online.

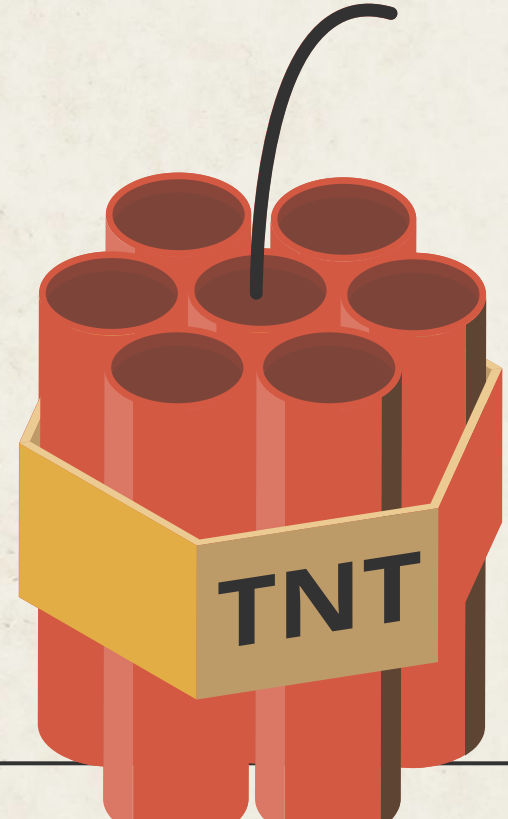




# 02

## Tinjauan Pustaka

Bab 2



# Related Works

Menerapkan Teknik Classification yang efektif berdasarkan pola yang dihasilkan pada jurnal atau artikel penelitian yang bersangkutan untuk *Data Mining*

## Paper 1

**JURNAL RESTI**  
(Rekayasa Sistem dan Teknologi Informasi)  
Vol. 6 No. 4 (2022) 648 - 655 ISSN Media Electronic: 2580-0760

Implementation of BERT, IndoBERT, and CNN-LSTM in Classifying Public Opinion about COVID-19 Vaccine in Indonesia


Siti Saadah<sup>1</sup>, Kaenova Mahendra Auditama<sup>2</sup>, Ananda Affan Fattahila<sup>3</sup>, Fendi Irfan Amorokhman<sup>4</sup>, Annisa Aditsania<sup>5</sup>, Aniq Atiqi Rohmawati<sup>6</sup>

<sup>1,6</sup>Informatics Faculty, Telkom University  
<sup>1</sup>sitisaadah@telkomuniversity.ac.id, <sup>2</sup>kaenova@student.telkomuniversity.ac.id,  
<sup>3</sup>affanfattahila@student.telkomuniversity.ac.id, <sup>4</sup>fendirfan@student.telkomuniversity.ac.id,  
<sup>5</sup>aditsania@telkomuniversity.ac.id, <sup>6</sup>aniqatiq@telkomuniversity.ac.id




IndoBERTweet - Akurasi: 0.73

## Paper 3

BiLSTM - Akurasi: 70.83

ORIGINAL ARTICLE 

Emotion classification of Indonesian Tweets using Bidirectional LSTM

Aaron Glenn<sup>1</sup>  · Phillip LaCasse<sup>1</sup>  · Bruce Cox<sup>1</sup> 

Received: 24 May 2022 / Accepted: 22 November 2022 / Published online: 6 February 2023  
This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2023

## Paper 2

Engineering, Technology & Applied Science Research Vol. 14, No. 5, 2024, 16471-16477 16471

Enhancing Emotion Detection in Textual Data: A Comparative Analysis of Machine Learning Models and Feature Extraction Techniques

Wedad Q. A. Saif  
Faculty of Engineering and Information Technology, Taiz University, Taiz, Yemen  
wedadalshameri49@gmail.com

Majid Khalaf Alshammari  
School of Educational Studies, Universiti Sains Malaysia, 11800 Penang, Malaysia  
m.alnehait@hotmail.com

Badia Abdulcareem Mohammed  
College of Computer Science and Engineering, University of Ha'il, Ha'il 81481, Saudi Arabia  
b.alshaibani@uoh.edu.sa

Amer A. Sallam

Dataset ISEAR LinearSVC - Akurasi: 88.63

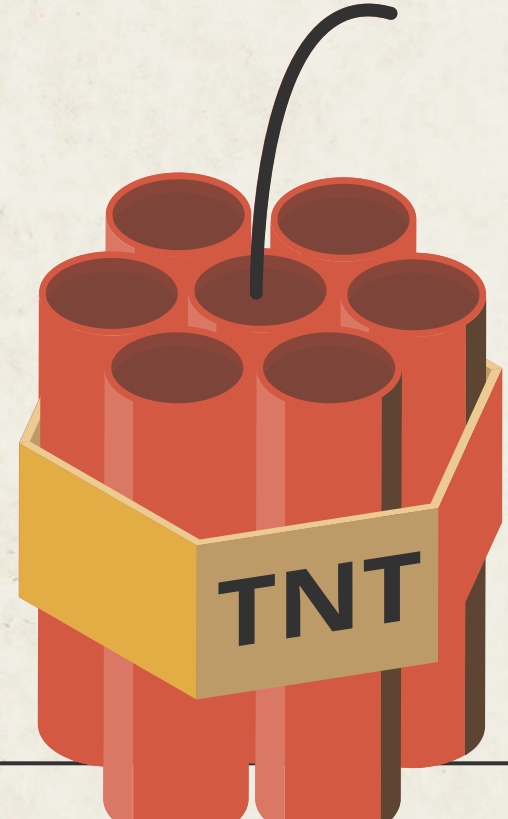
Dataset AIT ERT - Akurasi: 89.14

**More Paper until 5....**

# 03

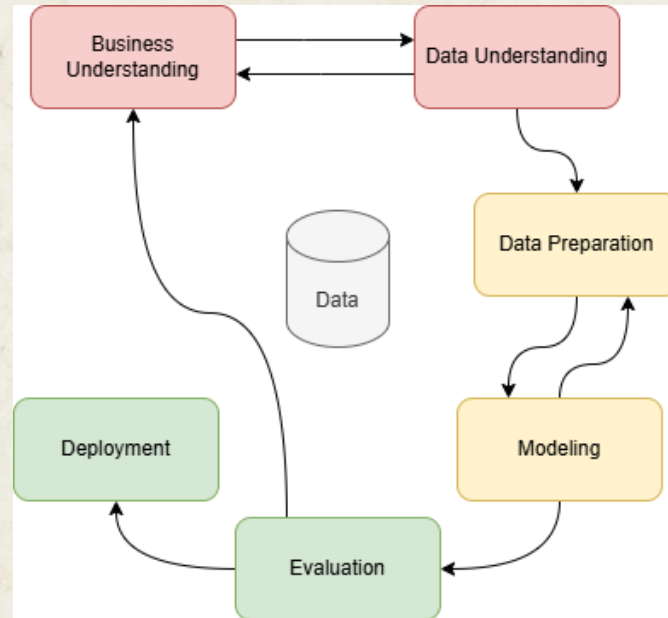
## Metode Penelitian

Bab 3



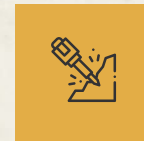


# Metode Penelitian: Cross Industry Standard Process for Data Mining



Tahap-Tahap yang dilaksanakan dari Metode CRISP-DM

**Business  
Understanding**



Start from Business  
Understanding

# Business Understanding: Ulasan Produk

Bagaimana kualitas produk ini secara keseluruhan?

★★★★★ Sangat Baik

Berikan ulasan untuk produk ini

barangnya kerennn saya sukaaa!!

Bagikan foto-foto dari produk yang Anda terima



Pelanggan memasukkan teks ke dalam ulasan produk dengan meluapkan emosi mereka sendiri.

Proses  
Tampil  
Ulasan  
Produk

★★★★★ Lebih dari 1 tahun lalu



Henry

Varian: 4-susun 60cm

Barang sesuai deskripsi. Kiriman super cepat. Packing rapi dan aman (pake kardus dan bubble). Reseller fast respon dan ramah. Produk gampang dirakit, ringan dan setelah dicoba berfungsi baik. Kualitas produk cukup bagus. Semoga awet. Rekomendasi. 👍



Tutup Ulasan



👍 3 orang terbantu

Lihat Balasan ▾

## Hasil Ulasan Produk Pelanggan Puas

★☆☆☆☆ 6 bulan lalu



Taeyoung

Varian: 4-susun putih

Barang nya beda dengan deskripsi. Jelek sekali. Pasang juga susah. Sudah komplain ke orang yg toko ini, tapi tidak peduli. Mereka seperti penipu. Mungkin dari barang Tiongkok, produk Tiongkok yg bagus lumayan banyak pun, mereka hanya dijual barang jelek. Kelihatan hanya cari untung Seperti oranh tamak. Tidak rekomendasi Toko ini



Tutup Ulasan



Kendala: Kualitas barang bermasalah

👍 2 orang terbantu

Lihat Balasan ▾

## Hasil Ulasan Produk Pelanggan Tidak Puas

# Data Understanding



## Product Reviews Dataset for Emotions Classification Tasks - Indonesian (PRDECT-ID) Dataset

Published: 20 May 2022 | Version 1 | DOI: 10.17632/574v66hf2v.1

Contributors: Rhio Sutoyo, Andry Chowanda, Said Achmad, Esther Widhi Andangsari, Sani M. Isa, Jason Pratama Soetandar, Jomari Rasheed Andres, Richard Richard, Michael Andrew Hadiwijaya, Darryl Andrews, Felix Putra Pirdaus

### Description

PRDECT-ID Dataset is a collection of Indonesian product review data annotated with emotion and sentiment labels. The data were collected from one of the giant e-commerce in Indonesia named Tokopedia. The dataset contains product reviews from 29 product categories on Tokopedia that use the Indonesian language. Each product review is annotated with a single emotion, i.e., love, happiness, anger, fear, or sadness. The group of annotators does the annotation process to provide emotion labels by following the emotions annotation criteria created by an expert in clinical psychology. Other attributes related to the product review are also extracted, such as Location, Price, Overall Rating, Number Sold, Total Review, and Customer Rating, to support further research.

Download All 303 KB

### Files

PRDECT-ID Dataset.csv

1.2 MB

Sumber: <https://data.mendeley.com/datasets/574v66hf2v/1>

## Definition

PRDECT-ID merupakan sebuah kumpulan data ulasan produk dengan label emosi / *emotion*, yang dikumpulkan berasal dari Tokopedia.

## Content

Kumpulan data ini mengandung 5400 data barisan dan 11 kolom dengan anotasi emosi dasar, yaitu *Anger*, *Fear*, *Happy*, *Love* & *Sadness*.

## Column

*Category*, *Product Name*, *Location*, *Price*, *Overall Rating*, *Number Sold*, *Total Reviews*, *Customer Rating*, *Customer Review*, *Sentiment*, *Emotion*

## Verify Data Quality

Mencermatkan konten pada *dataset* sehingga menemukan *field Customer Review* terdapat masalah data yang tidak valid, yaitu *noise*, *emoji* dan kata-kata kurang informatif.

# Dataset PRDECT-ID



Category	Product Name	Location	Price	Overall Rating	Number Sold	Total Review	Customer Rating	Customer Review	Sentiment	Emotion
Computers and Laptops	Wireless Keyboard i8 Mini TouchPad Mouse 2.4G Handheld PC Android TV	Jakarta Utara	53500	49	5449	2369	5	Alhamdulillah berfungsi dengan baik. Packaging aman. Respon cepat dan ramah. Seller dan kurir an	Positive	Happy
Computers and Laptops	PAKET LISENSI WINDOWS 10 PRO DAN OFFICE 2019 ORIGINAL + BONUS	Kota Tangerang Selatan	72000	49	2359	1044	5	barang bagus dan respon cepat, harga bersaing dengan yg lain.	Positive	Happy
Computers and Laptops	SSD Midesforce 128 Gb - Tanpa Caddy	Jakarta Barat	213000	5	12300	3573	5	barang bagus, berfungsi dengan baik, seller ramah, pengiriman cepat	Positive	Happy
Computers and Laptops	ADAPTOR CHARGER MONITOR LCD LED TV LG merek LG 19V ORIGINAL	Jakarta Timur	55000	47	2030	672	5	bagus sesuai harapan penjual ya juga ramah, trimakasih pelapak ??	Positive	Happy
Computers and Laptops	ADAPTOR CHARGER MONITOR LCD LED TV LG merek LG 19V ORIGINAL	Jakarta Timur	55000	47	2030	672	5	Barang Bagus, pengemasan Aman, dapat Berfungsi dengan Baik	Positive	Happy
Computers and Laptops	ADAPTOR CHARGER MONITOR LCD LED TV LG merek LG 19V ORIGINAL	Jakarta Timur	55000	47	2030	672	5	barang bagus, seller ramah..	Positive	Happy
Computers and Laptops	Kepala Colokan Listrik Charger Macbook untuk colokan Indonesia	Jakarta Selatan	85000	49	1339	718	5	mantap paten joss	Positive	Happy
Computers and Laptops	AFL Bidirectional HDMI Switcher 1-In 2-Out & 2-In 1-Out- AFU	Jakarta Barat	136500	5	1201	632	5	Works fine. Respon seller cepat, barang berfungsi dengan baik. Recommended.	Positive	Happy
Computers and Laptops	AFL Bidirectional HDMI Switcher 1-In 2-Out & 2-In 1-Out- AFU	Jakarta Barat	136500	5	1201	632	5	barang bagus.. sejel.. utuh.. original.. berfungsi dengan benar.. seller respon nya ajib.. cepat amir...	Positive	Happy
Computers and Laptops	Adaptor Charger Original Laptop Asus X441 X441U X441UV X441UA X441S	Jakarta Barat	113777	49	1630	777	5	Barang datang dan berfungsi dengan baik dan terlihat bagus kali ada lampu birunya.	Positive	Happy
Computers and Laptops	Adaptor Charger Original Laptop Asus X441 X441U X441UV X441UA X441S	Jakarta Barat	113777	49	1630	777	5	Pengiriman super cepat, barang bagus, recommended seller, thanks	Positive	Happy
Computers and Laptops	Adaptor Charger Laptop Asus A455LA A455LD A455LB A455 A455L Original	Jakarta Utara	66000	48	2702	882	5	alhamdulillah barangnya udah nyampe...d coba bagus... semoga awet... terima kasih seller kurir dar	Positive	Happy
Computers and Laptops	Kabel Power Adaptor Charger Laptop 3 Lubang	Jakarta Pusat	25000	49	2122	828	5	Penjual amarah, responsif. Kabel berfungsi baik. Pengiriman cepat. Terima kasih.	Positive	Happy
Computers and Laptops	Charger Adaptor Laptop Acer Aspire E14 E5-421 E5-431 E5-471 E5-4750	Jakarta Pusat	55000	49	2010	970	5	Alhamdulillah barang sudah sampai, bisa dipakai semoga berfungsi dengan baik dan awet, kurirnya	Positive	Happy
Computers and Laptops	Cooling Pad Laptop Gaming 6 Fan Kipas Pendingin Laptop Kipas Cooler Q3	Kota Bandung	127880	49	1257	700	5	bagus, berjalan baik dan cukup silent. kokoh dan kuat untuk laptop 16 inch	Positive	Happy
Computers and Laptops	Cooling Pad Laptop Gaming 6 Fan Kipas Pendingin Laptop Kipas Cooler Q3	Kota Bandung	127880	49	1257	700	5	Packing aman, pengiriman cepat dan barang berfungsi dengan baik	Positive	Happy
Computers and Laptops	kipas laptop cooling pad Ace cooling ergostand 15.6inch 1 fan	Jakarta Utara	80000	48	3982	1890	5	Produknya baik dan kualitasnya good banget, produk sampai ke konsumen dalam kondisi baik, sesu	Positive	Happy
Computers and Laptops	kipas laptop cooling pad Ace cooling ergostand 15.6inch 1 fan	Jakarta Utara	80000	48	3982	1890	5	barang udah sampai ya dan berfungsi. terima kasih	Positive	Happy
Computers and Laptops	CoolCold Laptop Vacuum Vacuum Cooler Cooling Pad Hisap Panas Notebook	Kota Depok	55000	48	13900	6671	5	barang mantap sesuai deskripsi laptop jadi adem n sejuk	Positive	Happy
Computers and Laptops	CoolCold Laptop Vacuum Vacuum Cooler Cooling Pad Hisap Panas Notebook	Kota Depok	55000	48	13900	6671	5	produk saya terima dengan kondisi baik, terima kasih ^^	Positive	Happy
Computers and Laptops	CoolCold Laptop Vacuum Vacuum Cooler Cooling Pad Hisap Panas Notebook	Kota Depok	55000	48	13900	6671	5	belanja kedua kalinya, selalu puas, pengiriman lumayan cepat, overall bagus banget	Positive	Happy
Computers and Laptops	CoolCold Laptop Vacuum Vacuum Cooler Cooling Pad Hisap Panas Notebook	Kota Depok	55000	48	13900	6671	5	Kualitas barang bagus, harga relatif murah, kiriman super cepat, response penjual baik, recommen	Positive	Happy
Computers and Laptops	Cooling Pad Laptop Gaming Ergostand Kipas Pendingin Laptop Big Fan Cooler Pad - KIPAS ERGOSTAND	Jakarta Barat	74900	47	3713	2153	5	mantab barang nya oke punya	Positive	Happy
Computers and Laptops	Cooler Master NotePal CMC3 Silent Fan Laptop Cooling Pad	Jakarta Barat	112000	49	5988	2969	5	pengiriman cepat, packing rapi, kualitas mantap	Positive	Happy
Computers and Laptops	Cooler Master NotePal CMC3 Silent Fan Laptop Cooling Pad	Jakarta Barat	112000	49	5988	2969	5	silent dan bagus untuk pekerjaan ringan	Positive	Happy
Computers and Laptops	Cooler Master NotePal CMC3 Silent Fan Laptop Cooling Pad	Jakarta Barat	112000	49	5988	2969	5	Barang cepat sampai. Kualitas bagus dan pengiriman aman tidak ada yang pecah. Manstapu ?????	Positive	Happy
Computers and Laptops	Cooler Master NotePal CMC3 Silent Fan Laptop Cooling Pad	Jakarta Barat	112000	49	5988	2969	5	Bagus, berkualitas, sesuai gbr. Makasih Seller, Tokped dan Kurir	Positive	Happy
Computers and Laptops	cooling pad laptop tahan 24 jam 6.000 Rpm angin TOPAN - Gambar no 1. Bening	Kota Surabaya	385000	5	2716	1843	5	mantap kualitasnya kenceng, barangnya berkualitas sesuai sama harga	Positive	Happy
Computers and Laptops	cooling pad laptop tahan 24 jam 6.000 Rpm angin TOPAN - Gambar no 1. Bening	Kota Surabaya	385000	5	2716	1843	5	mantap proses pengerjaan dan pengiriman juga barang sesuai dengan deskripsi	Positive	Happy
Computers and Laptops	cooling pad laptop tahan 24 jam 6.000 Rpm angin TOPAN - Gambar no 1. Bening	Kota Surabaya	385000	5	2716	1843	5	mantap kipasnya kenceng, barangnya berkualitas sesuai sama harga	Positive	Happy
Computers and Laptops	cooling pad laptop tahan 24 jam 6.000 Rpm angin TOPAN - Gambar no 1. Bening	Kota Surabaya	385000	5	2716	1843	5	Penjualnya Ramah Melayani Pembeli Dengan Sabar dan Memberikan Berbagai Saran Yang Pembeli T	Positive	Happy
Computers and Laptops	cooling pad laptop tahan 24 jam 6.000 Rpm angin TOPAN - Gambar no 1. Bening	Kota Surabaya	385000	5	2716	1843	5	barang sampai dengan cepat, packingnya rapi dan amat sekali, tidak ada bagian acrylic yang pecah. I	Positive	Happy
Computers and Laptops	MacBook Skin Keyboard Protector TRANSPARANT BENING - A2141 A2289	Jakarta Utara	15000	49	690	338	5	Sangat presisi dan pas	Positive	Happy
Computers and Laptops	MacBook Skin Keyboard Protector TRANSPARANT BENING - A2141 A2289	Jakarta Utara	15000	49	690	338	5	Puas banget, seller helpfull, sangat responsif dan cepet banget sampai nya, recommended...	Positive	Happy
Computers and Laptops	MacBook Skin Keyboard Protector TRANSPARANT BENING - A2141 A2289	Jakarta Utara	15000	49	690	338	5	Mantappp, responnya juga memuaskan, sudah dicoba dan dipasang presisi dan nyaman. CAs juga	Positive	Happy
Computers and Laptops	Logitech Bluetooth Keyboard K380, K 380, K-380 - Hitam	Jakarta Barat	365000	5	2554	1396	5	Mantap barangnya cepet sampai dan enak digunainnya!	Positive	Happy
Computers and Laptops	Logitech Bluetooth Keyboard K380, K 380, K-380 - Hitam	Jakarta Barat	365000	5	2554	1396	5	Barang bagus, oke, cepet pengirimannya juga	Positive	Happy
Computers and Laptops	Logitech Bluetooth Keyboard K380, K 380, K-380 - Hitam	Jakarta Barat	365000	5	2554	1396	5	Pengiriman cepat, seller responsif ??	Positive	Happy
Computers and Laptops	Wireless Keyboard i8 Mini TouchPad Mouse 2.4G Handheld PC Android TV	Jakarta Utara	53500	49	5449	2369	5	Barang berfungsi baik. Pemilik toko sangat responsif dan membantu saat ada sedikit kendala. Suku	Positive	Happy
Computers and Laptops	Wireless Keyboard i8 Mini TouchPad Mouse 2.4G Handheld PC Android TV	Jakarta Utara	53500	49	5449	2369	5	seller luar biasa.. packing baik, pengiriman cepat, penting nya barang sesuai deskripsi dan berfungsi	Positive	Happy
Computers and Laptops	Logitech M191 Mouse Wireless Full Size - Mid Grey	Jakarta Timur	119000	49	9777	5546	1	Sangat kecewa. Baru 4 bulan scroll sudah rusak.	Negative	Sadness
Computers and Laptops	Adaptor Charger Original Laptop Asus X441 X441U X441UV X441UA X441S	Jakarta Barat	113777	49	1630	777	1	jelek, yg pasti jauh dari ori. mousepad jd ga bs dipakai kl di charger, jd getar parah. barang ga sesu	Negative	Sadness
Computers and Laptops	Kabel Power Adaptor Charger Laptop 3 Lubang	Jakarta Pusat	25000	49	2122	828	1	gak bisa di pakai??	Negative	Sadness

sampai 5400 data....



# Data Preparation



## Data Selection

Memilih atribut pada *dataset* berdasarkan ketentuan tujuan penelitian *data mining* dan bisnis penjual adalah *Customer Review* alias *text* dan *Emotion* alias *label*

Hasil

	text	label
1	Alhamdulillah berfungsi dengan baik. Packaging...	Happy
2	barang bagus dan respon cepat, harga bersaing ...	Happy
3	barang bagus, berfungsi dengan baik, seller ramah...	Happy
4	bagus sesuai harapan penjual nya juga ramah. t...	Happy
5	Barang Bagus, pengemasan Aman, dapat Berfungsi...	Happy
...	...	...
5396	Harga bersaing, barang sesuai pesanan. Saya na...	Love
5397	Beli ini krn Anak & Istri mau liburan di Jakar...	Love
5398	pengemasan barang baik, kondisi barang jg utuh...	Happy
5399	Mungil tapi bekerja dng baik. Dan murahh terja...	Happy
5400	Produk sesuai deskripsi, packing aman terlindu...	Love



## Data Cleaning

Memperbaiki kualitas data yang diinstruksikan oleh laporan verifikasi kualitas data untuk mengoptimalkan penerapan *data mining*. Langkah-langkah yang perlu dilewati:

1. *Whitespace*
2. *Punctutation*
3. *Lower Case*
4. *Tokenization*
5. *Normalization*
6. *Stopword*

Hasil

	Customer Review	Emotion	cleaned
1	Alhamdulillah berfungsi dengan baik. Packaging...	Happy	alhamdulillah berfungsi packaging aman respon ...
2	barang bagus dan respon cepat, harga bersaing ...	Happy	barang bagus respon cepat harga bersaing
3	barang bagus, berfungsi dengan baik, seller ramah...	Happy	barang bagus berfungsi seller ramah pengiriman ...
4	bagus sesuai harapan penjual nya juga ramah. t...	Happy	bagus sesuai harapan penjual ramah terima kasi...
5	Barang Bagus, pengemasan Aman, dapat Berfungsi...	Happy	barang bagus pengemasan aman berfungsi
6	barang bagus, seller ramah...	Happy	barang bagus seller ramah
7	mantap paten joss	Happy	mantap paten joss
8	Works fine. Respon seller cepat, barang berfun...	Happy	works fine respon seller cepat barang berfun...
9	barang bagus.. segel.. utuh.. original.. berfu...	Happy	barang bagus segel utuh original berfungsi sel...
10	Barang datang dan berfungsi dengan baik dan te...	Happy	barang berfungsi bagus lampu birunya
11	Pengiriman super cepat, barang bagus, recommen...	Happy	pengiriman super cepat barang bagus recommende...
12	alhamdulillah barangnya udah nyampe... d coba ...	Happy	alhamdulillah barangnya coba bagus semoga awet...
13	Penjual amanah, responsif. Kabel berfungsi bai...	Happy	penjual amanah responsif kabel berfungsi pengi...
14	Alhamdulillah barang sudah sampai, bisa dipaka...	Happy	alhamdulillah barang dipakai semoga berfungsi ...
15	bagus, berjalan baik dan cukup silent. kokoh d...	Happy	bagus berjalan silent kokoh muat laptop inch

# Data Preparation



## Data Construction

Dapat dikenal sebagai *Transformation*, artinya memodifikasi nilai pada atribut yang ada.

### Emotion Attribute

Data Emosi diubahkan menjadi dalam bentuk numerik.

### Customer Review Attribute

Data Ulasan diubahkan menjadi dalam bentuk numerik bertipe vektor yang mengukur kata-kata dengan menggunakan Teknik *Feature Extraction*.

```
Counter({'Happy': 1770, 'Sadness': 1202, 'Fear': 920, 'Love': 809, 'Anger': 699})
Counter({'2': 1770, 4: 1202, 1: 920, 3: 809, 0: 699})
0      2
1      2
2      2
3      2
4      2
..
5395    3
5396    3
5397    2
5398    2
5399    3
Name: emotion, Length: 5400, dtype: int64
```

### Emotion Attribute

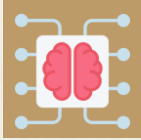
Sebelum mengubah data train	Setelah mengubah data train menggunakan Count Vectorizer
123 terima kasihh bos mantappss semoga lancarr ber...	x-train
3959 salah kirim judul bolak kembalikan salah kirim...	(0, 5500) 1
2785 gambar	(0, 2478) 1
446 kecewa berat salahnya dimana packing kemasan p...	(0, 727) 2
4781 recommended komunikasi lancar instruksi game g...	(0, 3223) 1
...	(0, 4993) 1
4280 barangnya bagus sesuai harapan terima kasih	(0, 2944) 1
2304 order ketiga gerak cepat barang oke pelayanan oke	(0, 562) 1
5295 recommended pelayanan memuaskan respon pengirim...	(0, 4722) 1
4595 barang bagus mengecewakan pesan rok kirim kema...	(0, 5157) 1
2337 berguna bermanfaat lancar swiss	(1, 4993) 1
Name: cleaned, Length: 4320, dtype: object	(1, 4808) 2
Sebelum mengubah data test	(1, 2736) 2
4262 gonta ganti main sehari beli ratusan ribu mend...	(1, 2376) 1
2537 produk bagus pas sesuai harga	(1, 706) 1
2272 kulit cerah cepat meresap cocok kulit kombinas...	(1, 2600) 1
2789 barang sesuai foto leggingnya sobek	(1, 2545) 1
4232 ukuran sesuai bahannya berkualitas nyaman dipa...	(1, 2557) 1
...	(1, 1062) 1
157 barangnya mantap berkualitas kokoh bergaransi ...	(2, 1846) 1
559 beli menyesal kualitas oke pengirimannya bagus...	(3, 2526) 1
1330 cepatt ahlakkk fast respon barang besttt semog...	(3, 510) 1
4714 brgnya cepat packing maxi terima kasih seller ...	(3, 4809) 1
362 tolong cek bungkus	(3, 1361) 1
Name: cleaned, Length: 1080, dtype: object	(3, 3955) 1
	(3, 2595) 1

### Customer Review Attribute

# Modelling

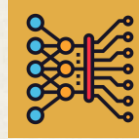


Menerapkan beragam Teknik Klasifikasi pada tahap *data mining* untuk menguji dan meng-*train* model sehingga menemukan sebuah model yang lebih mampu mengenali data ulasan produk.



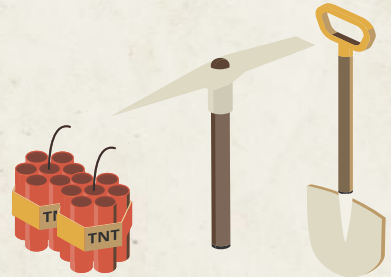
## Machine Learning

1. *Support Vector Machine*
2. *Random Forest*
3. *Logistic Regression*
4. *Decision Tree*
5. *Naïve Bayes*
6. *K-Nearest Neighbor*



## Deep Learning

1. *Bidirectional Encoder Representations from Transformers (BERT)*

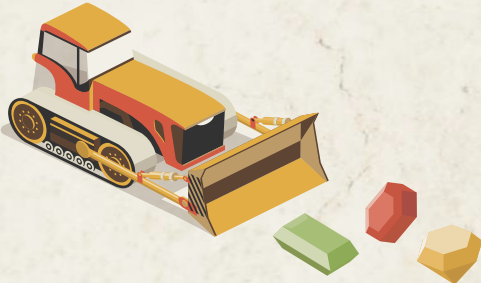


# Evaluation

## Tabel Hasil Komputasi Metrik (Machine Learning)

Model	Vectorizer	Accuracy	Precision	Recall	F1-Score
Random Forest	TF-IDF	58,9815	56,1073	51,7021	52,2719
	Count	58,2407	54,9135	51,2531	51,8526
	Hashing	58,1481	55,9343	50,1692	50,6545
Support Vector Machine	TF-IDF	60,8333	60,5620	53,4797	54,4721
	Count	60,1852	60,2672	52,1755	52,7339
	Hashing	60,6481	61,0026	53,2421	54,1931
Logistic Regression	TF-IDF	<b>61,7593</b>	<b>61,1124</b>	<b>55,3479</b>	<b>56,5082</b>
	Count	60,7407	57,5854	55,3971	56,1124
	Hashing	60,7407	59,7814	54,0473	55,2027
Decision Tree	TF-IDF	54,2593	50,2557	49,6883	49,8006

	Count	54,4444	50,3627	50,0794	50,1590
	Hashing	53,6111	49,4629	48,6353	48,8343
Nearest Neighbors Classifier	TF-IDF	52,7778	55,5003	42,8944	43,3391
	Count	50,0000	44,0218	40,9307	39,4929
	Hashing	51,6667	52,4439	41,8715	41,8961
Multinomial Naïve Bayes	TF-IDF	51,9444	66,4081	39,5784	36,6820
	Count	57,6852	58,2508	48,9655	49,7937
	Hashing				
Complement Naïve Bayes	TF-IDF	57,1296	54,8612	49,3347	50,1443
	Count	58,4259	56,2408	50,7629	51,6577
	Hashing				
Bernoulli Naïve Beys	TF-IDF	50,6481	61,2892	38,4105	35,2244
	Count	50,6481	61,2892	38,4105	35,2244
	Hashing				





# Evaluation

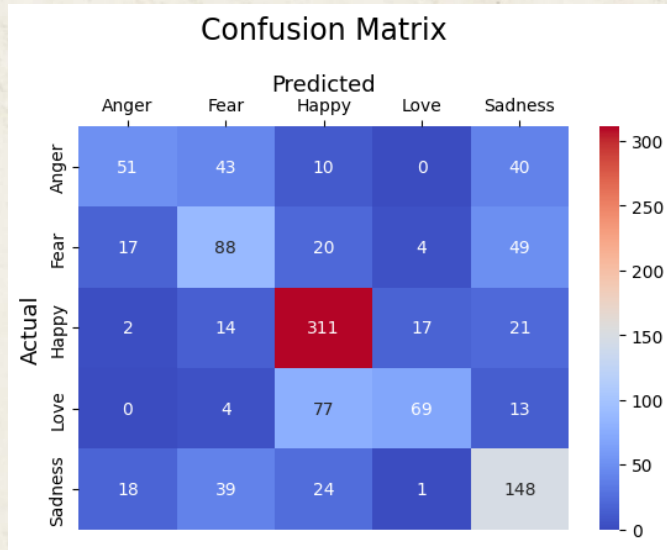
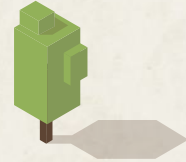
## Tabel Hasil Komputasi Metrik (Deep Learning)

Model	E-pochs	Accuracy	Precision	Recall	F1-Score
indobenchmark/indobert-base-pl	8	61,7593	58,7396	58,8024	58,6251
indolem/indobert-base-uncased	8	60.7407	57.0028	57.2609	57.0539
Wikidepia/indobert-lite-squad	8	60,9259	56,7272	56,6159	56,6340
<b>indolem/indobertweet-base-uncased</b>	<b>8</b>	<b>62,5926</b>	<b>59,7879</b>	<b>59,2317</b>	<b>59,1396</b>

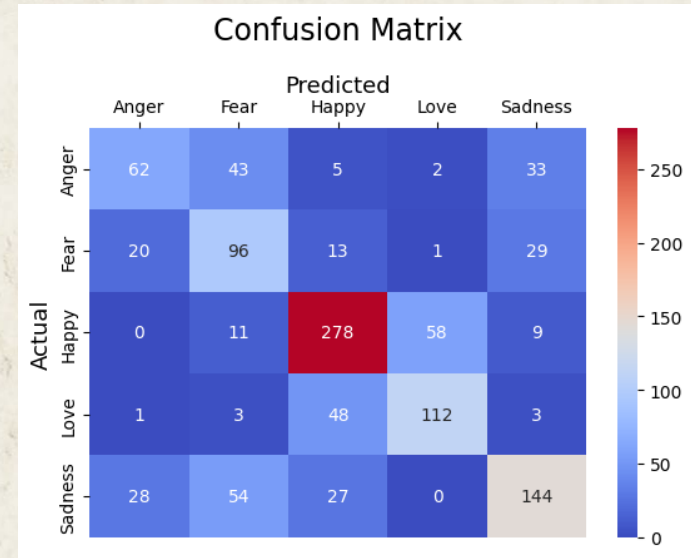


# Evaluation

## Confusion Matrix



Hasil Confusion Matrix Terbaik: Logistic Regression dengan TF-IDF Vectorizer



Hasil Confusion Matrix Terbaik: IndoBERTweet



# Deployment



## Deploy Model Website Application

127.0.0.1:5000/api/result

### Emotion Recognition based on Text of Product Review

Text Review:

bajunya baguss banget!!

Submit

😊

Emotion: Love

Emosi: Suka

Guide:

1. Just add a review text (Only Indonesian) into textarea
2. Click the submit button, then it will displays emotion.

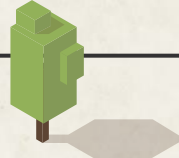
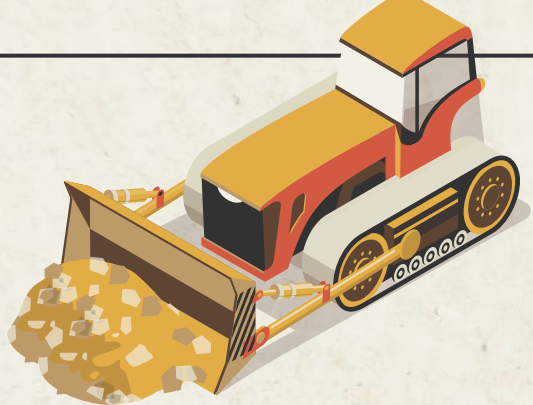
Panduan:

1. Cukup masukkan teks ulasan (Bahasa Indonesia saja) ke textarea
2. Klik tombol submit, lalu muncul hasil emosi.

Created by Valentino Stanley Gunadi  
I developed this application for my thesis, please click here [link](#) to score the functions.  
If you need a hand, please do not hesitate to hit me up!  
Thank you!

Emotion Source

Emotion Recognition Website Application: Recognize Emotions based on the Product Review of Online Shop



# Live Demonstration

[https://youtu.be/FiQw\\_Smqats](https://youtu.be/FiQw_Smqats)

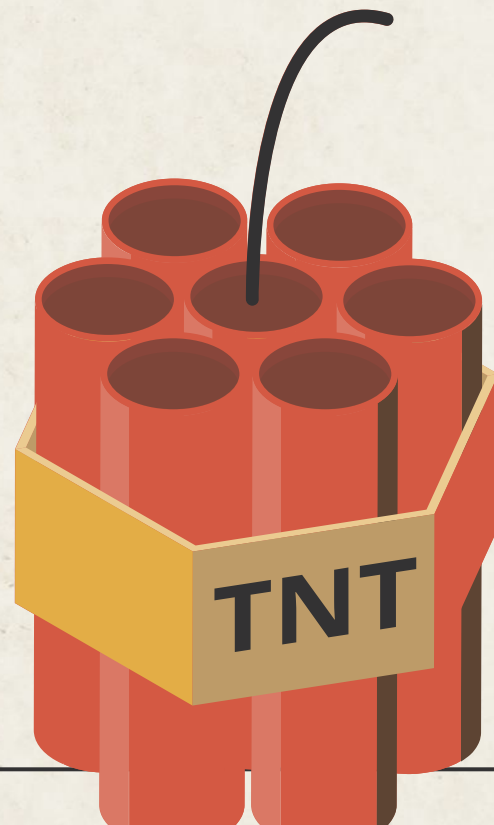




# 04

## Kesimpulan

Bab 5





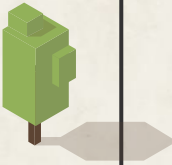
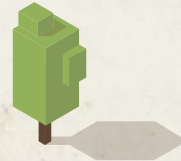
# Simpulan & Saran

## Simpulan:

1. Penelitian ini menerapkan metode *Cross-Industry Standard Process For Data-Mining* (CRISP-DM).
2. Model terbaik adalah *Bidirectional Encoder Representations from Transformers* (BERT) dengan *Fine-Tuning indolem/indobertweet-base-uncased* yang berasal dari Hugging Face, dengan akurasi 62.6%.
3. Beberapa tahapan penting dari CRISP DM adalah bagian *Normalization* lebih penting untuk meningkatkan akurasi sebesar 2% pada tahapan *data cleaning*. Selain ini, membuat laporan mengenai hasil penelitian beserta dengan *deploy website application* pengenalan emosi berbasis ulasan produk pada tahapan *deployment* akan mencapai tujuan bisnis penjual.

## Saran:

1. Mengeksplorasi algoritma klasifikasi lanjutan.
2. Mengumpulkan data yang lebih banyak pada *dataset* agar mengoptimalkan kinerja model saat mengaplikasikan *data mining*, dan juga memastikan *balanced dataset*.
3. Menyempurnakan *User Interface Website* dan mengembangkan *function* pada fitur lanjutan sesuai dengan kebutuhan *user* agar *user* mudah menggunakannya.





# Manfaat



1. Memperoleh pengetahuan dari penelitian ini yang memanfaatkan *dataset* ulasan produk di bidang *data mining*.
2. Memperoleh teknik *classification* yang efektif berdasarkan pembuktian hasil model untuk mengolah data ulasan produk.
3. Menjadikan penelitian ini sebagai tinjauan pustaka untuk memanfaatkan *dataset* ulasan produk di bidang *Data Mining*.
4. Mengakses aplikasi *website* pengenalan emosi berbasis teks ulasan dari ulasan produk untuk mengetahui emosi yang diprediksi oleh model terbaik.





# Thank you!

Author: Valentio Stanley Gunadi