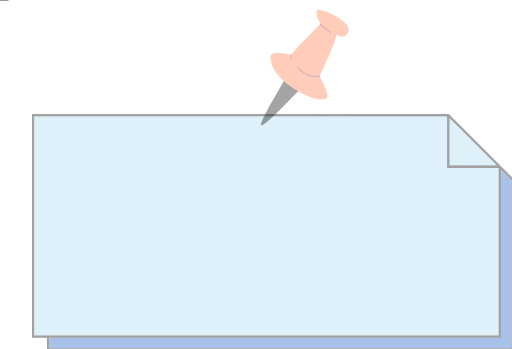
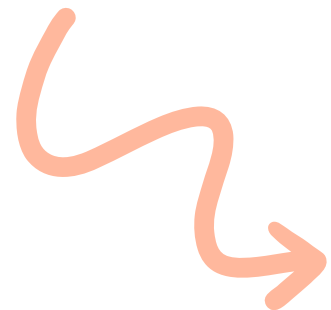
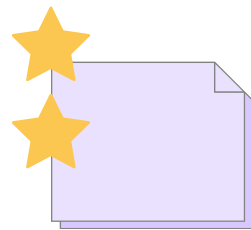


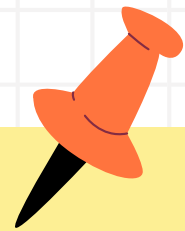
Analisis Sentimen Masyarakat Terhadap Wisatawan Asing di Provinsi Bali Berdasarkan Data Twitter



Ahmad Sirrullah
Agil Akbar Fahrezi
MazkaBuanaHidayat



Agenda



Pendahuluan



Metode



Hasil



Analisis

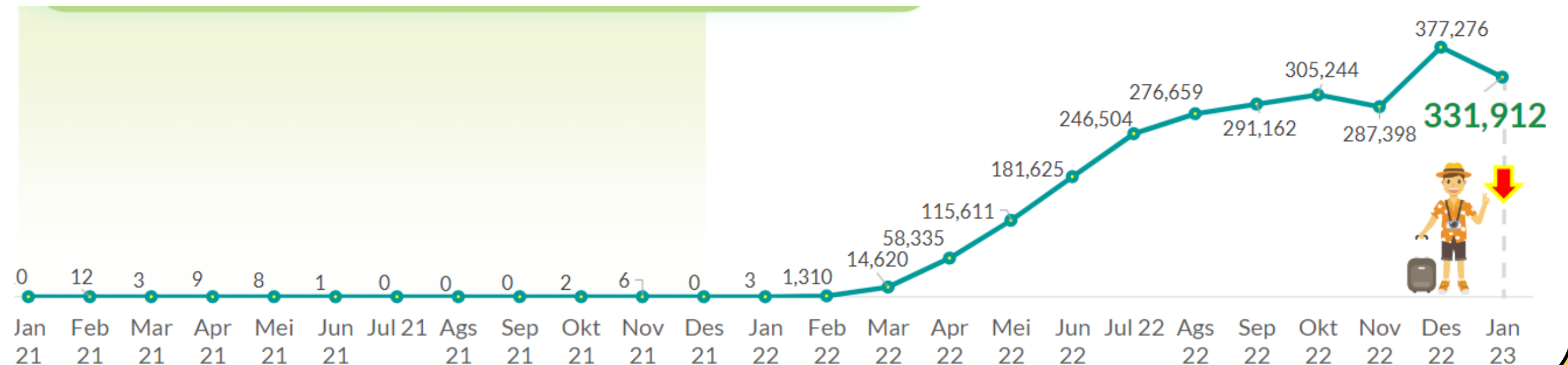
Pendahuluan: Pariwisata Bali

Bali merupakan destinasi yang populer di kalangan wisatawan mancanegara

Pandemi Covid-19 cukup memberikan dampak yang signifikan terhadap pariwisata Bali

Namun dengan kondisi yang semakin membaik, pariwisata Bali berangsur pulih (BPS,2023)

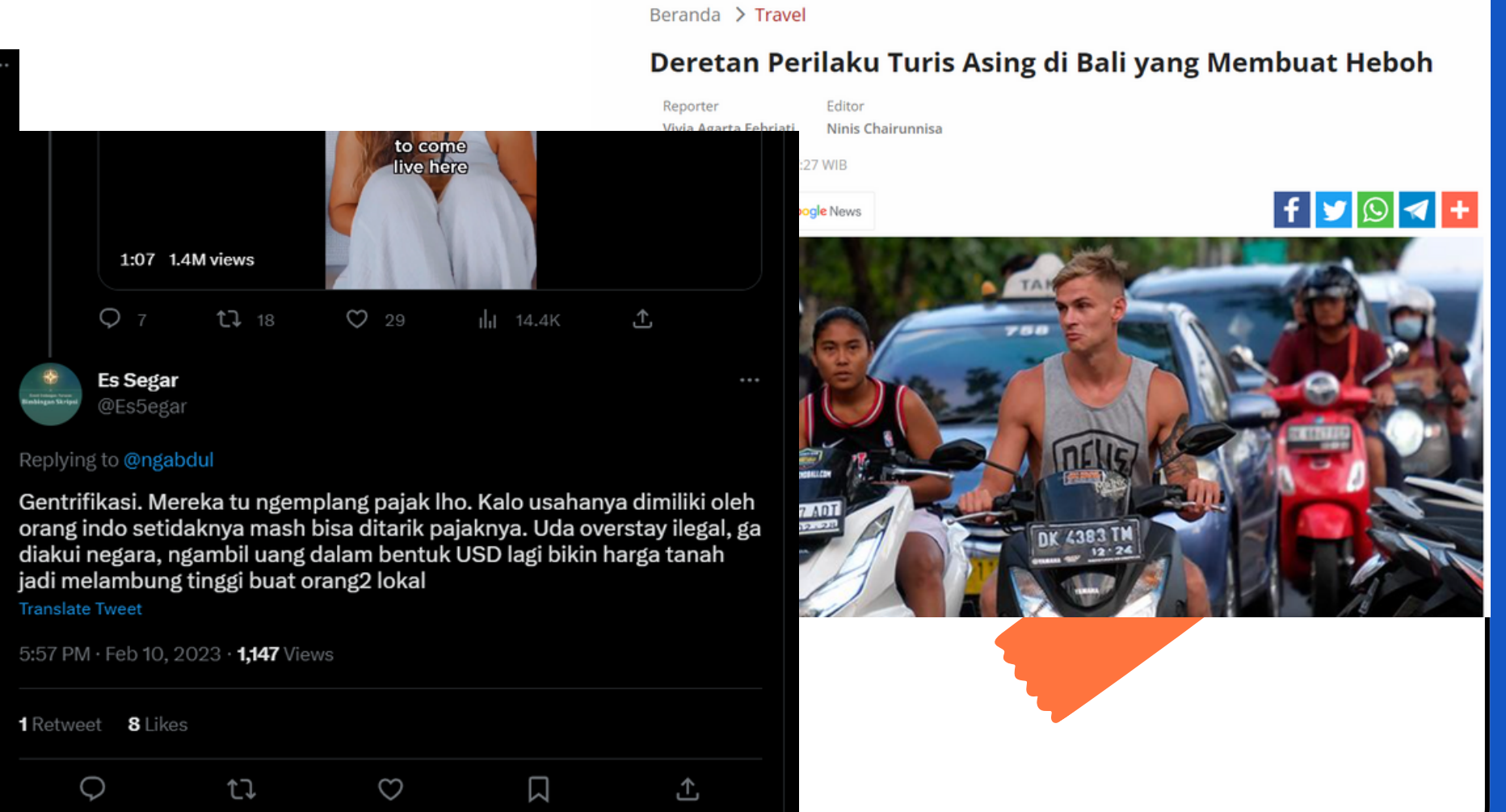
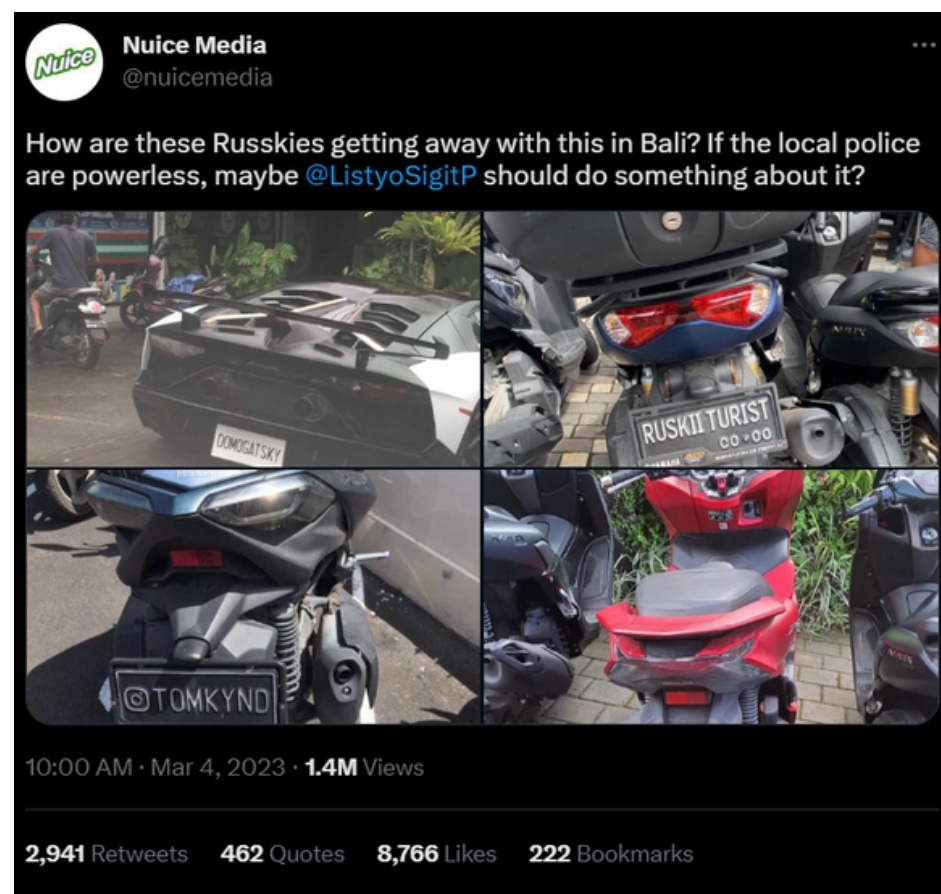
Meskipun pariwisata asing memberikan kontribusi ekonomi yang besar, terdapat suatu permasalahan mengenai sifat dan ulah dari beberapa wisatawan mancanegara



Pendahuluan: Apa yang terjadi di dunia maya?

Banyaknya perilaku turis asing yang 'nakal' meresahkan banyak pihak, khususnya masyarakat Bali

Mulai dari kelakuan nakal, digital nomad, penyalahgunaan visa, isu gentrifikasi, sampai kampung bule





Pendahuluan: Apa yang bisa dilakukan?

Twitter merupakan salah satu media sosial terpopuler di Indonesia.

Twitter merupakan semacam 'taman' digital dimana masyarakat dari semua lapisan berinteraksi

Data twitter yang sangat besar membuka potensi untuk melakukan berbagai analisis untuk melihat sentimen yang berkembang di masyarakat

Analisis sentimen merupakan analisis yang dilakukan untuk mengetahui opini publik. Melalui metode data mining, analisis sentimen publik terhadap perilaku wisatawan mancanegara dilakukan lewat project ini

Metode



B

Data Preprocessing and Processing (noise removal, normalisasi, transform case, filter stopwords, filter tokenize, and labelling)

C

Visualization (wordcloud, most frequent words)

A

Web Scrapping Data Twitter (Twint library)

D

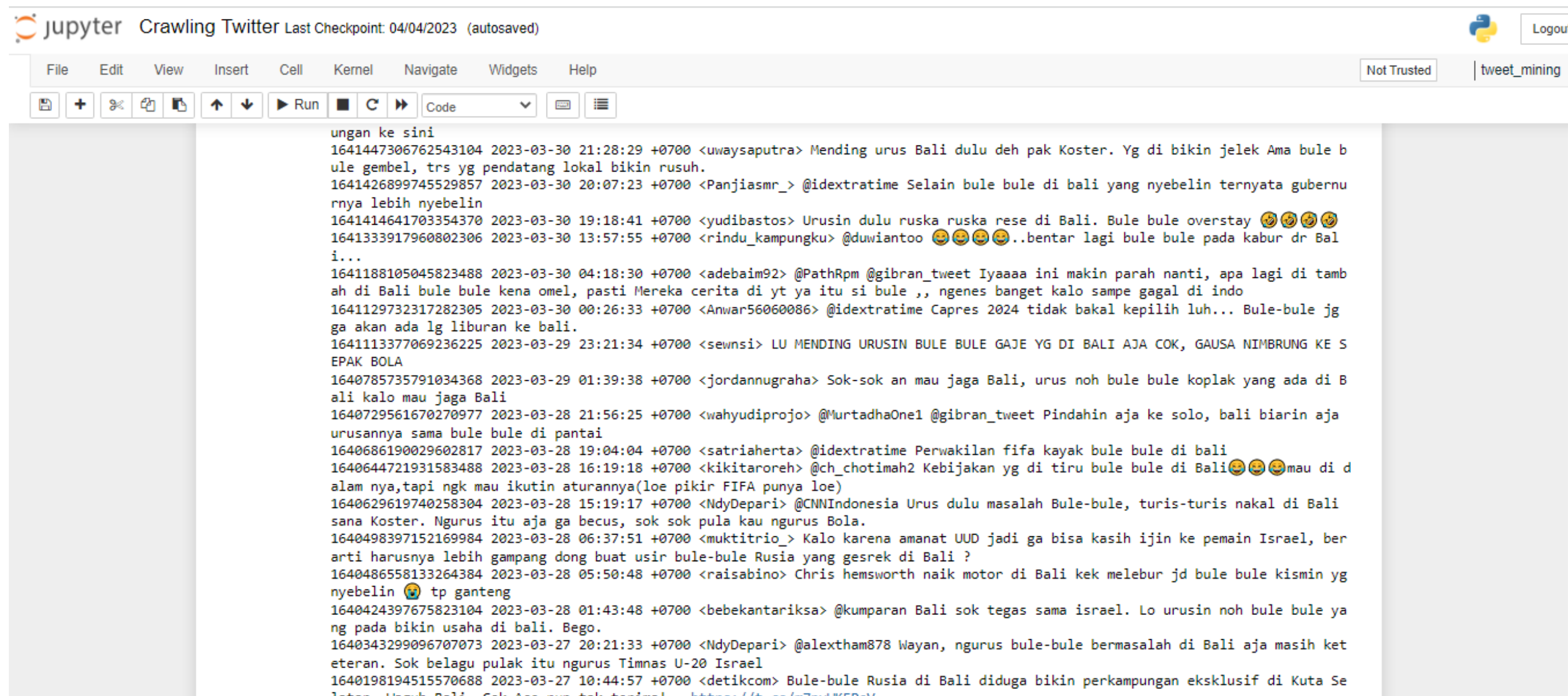
Analysis (Neural Network IndoBERTweet)

Hasil: Web Scrapping

Dalam melakukan analisis sentimen, langkah pertama adalah text mining untuk menggali opini mengenai pendapat masyarakat terhadap turis asing di Pulau Bali. Untuk Text mining dilakukan dengan library Twint. Keyword yang digunakan adalah sebagai berikut

Bule langgar aturan, Bule nakal, Bule ganggu, Bule bangsat, Bule goblok, Bule kere, Bule ngemis, Gentrifikasi bali, Bule bule bali

ulah bule bali, kampung bule, bali polah bule bali, bule bali berulah, bule bali helm, usaha bule bali



```
jupyter Crawling Twitter Last Checkpoint: 04/04/2023 (autosaved)
File Edit View Insert Cell Kernel Navigate Widgets Help
Not Trusted | tweet_mining
Code
ungan ke sini
1641447306762543104 2023-03-30 21:28:29 +0700 <uwaysaputra> Mending urus Bali dulu deh pak Koster. Yg di bikin jelek Ama bule b
ule gembel, trs yg pendatang lokal bikin rusuh.
1641426899745529857 2023-03-30 20:07:23 +0700 <Panjiasmr_> @idextratime Selain bule bule di bali yang nyebelin ternyata gubernu
rnya lebih nyebelin
1641414641703354370 2023-03-30 19:18:41 +0700 <yudibastos> Urusin dulu ruska ruska rese di Bali. Bule bule overstay 🤔🤔🤔🤔
1641333917960802306 2023-03-30 13:57:55 +0700 <rindu_kampungku> @duwiantoo 🤔🤔🤔🤔..bentar lagi bule bule pada kabur dr Bal
i...
1641188105045823488 2023-03-30 04:18:30 +0700 <adebaim92> @PathRpm @gibran_tweet Iyaaaa ini makin parah nanti, apa lagi di tamb
ah di Bali bule bule kena omel, pasti Mereka cerita di yt ya itu si bule ,, ngenes banget kalo sampe gagal di indo
1641129732317282305 2023-03-30 00:26:33 +0700 <Anwar56060086> @idextratime Capres 2024 tidak bakal kepilih luh... Bule-bule jg
ga akan ada lg liburan ke bali.
1641113377069236225 2023-03-29 23:21:34 +0700 <sewnsi> LU MENDING URUSIN BULE BULE GAJE YG DI BALI AJA COK, GAUSA NIMBRUNG KE S
EPAK BOLA
1640785735791034368 2023-03-29 01:39:38 +0700 <jordannugraha> Sok-sok an mau jaga Bali, urus noh bule bule koplak yang ada di B
ali kalo mau jaga Bali
1640729561670270977 2023-03-28 21:56:25 +0700 <wahyudiprojo> @MurtadhaOne1 @gibran_tweet Pindahin aja ke solo, bali biarin aja
urusannya sama bule bule di pantai
1640686190029602817 2023-03-28 19:04:04 +0700 <satriaherta> @idextratime Perwakilan fifa kayak bule bule di bali
1640644721931583488 2023-03-28 16:19:18 +0700 <kikitaroreh> @ch_chotimah2 Kebijakan yg di tiru bule bule di Bali 🤔🤔🤔 mau di d
alam nya,tapi ngk mau ikutin aturannya(loes pikir FIFA punya loe)
1640629619740258304 2023-03-28 15:19:17 +0700 <NdyDepari> @CNNIndonesia Urus dulu masalah Bule-bule, turis-turis nakal di Bali
sana Koster. Ngurus itu aja ga becus, sok sok pula kau ngurus Bola.
1640498397152169984 2023-03-28 06:37:51 +0700 <muktitrio_> Kalo karena amanat UUD jadi ga bisa kasih ijin ke pemain Israel, ber
arti harusnya lebih gampang dong buat usir bule-bule Rusia yang gesrek di Bali ?
1640486558133264384 2023-03-28 05:50:48 +0700 <raisabino> Chris hemsworth naik motor di Bali kek melebur jd bule bule kismin yg
nyebelin 🤔 tp ganteng
1640424397675823104 2023-03-28 01:43:48 +0700 <bebekantariksa> @kumparan Bali sok tegas sama israel. Lo urusin noh bule bule ya
ng pada bikin usaha di bali. Bego.
1640343299096707073 2023-03-27 20:21:33 +0700 <NdyDepari> @alextham878 Wayan, ngurus bule-bule bermasalah di Bali aja masih ket
eteran. Sok belagu pulak itu ngurus Timnas U-20 Israel
1640198194515570688 2023-03-27 10:44:57 +0700 <detikcom> Bule-bule Rusia di Bali diduga bikin perkampungan eksklusif di Kuta Se
latan. Waguh Bali Cok Ace nun tak terima! https://t.co/m7nvWKEPeV
```

490 tweet

01-01-2022

s.d.

03-04-2023

Source Code:

[https://drive.google.co](https://drive.google.com/file/d/18YWdlIJE27PLAPIZodv3CDQYWW_DnIP7/view?usp=sharing)
[m/file/d/18YWdlIJE27](https://drive.google.com/file/d/18YWdlIJE27PLAPIZodv3CDQYWW_DnIP7/view?usp=sharing)
[PLAPIZodv3CDQYWW](https://drive.google.com/file/d/18YWdlIJE27PLAPIZodv3CDQYWW_DnIP7/view?usp=sharing)
[_DnIP7/view?](https://drive.google.com/file/d/18YWdlIJE27PLAPIZodv3CDQYWW_DnIP7/view?usp=sharing)
[usp=sharing](https://drive.google.com/file/d/18YWdlIJE27PLAPIZodv3CDQYWW_DnIP7/view?usp=sharing)

Hasil: Data Pre Processing

Preprocessing data disini dilakukan untuk menghilangkan noise, meningkatkan akurasi, memperjelas fitur, dan mengubah data untuk visualisasi dan analisis



| # | tweet # | EYD # | remove_stopword # | Stemming # | Token # |
|---|---|---|---|---|--|
| 0 | pangerannya kya bule yg suka berulah di bali | pangerannya kayak bule yang suka berulah di bali | pangerannya bule suka berulah bali | pangeran bule suka ulah bali | [pangeran, bule, suka, ulah, bali] |
| 1 | 380h m ago marak bule rusia ukraina berulah d... | 380h sama ago marak bule rusia ukraina berula... | 380h ago marak bule rusia ukraina berulah bali... | 380h ago marak bule rusia ukraina ulah bali ko... | [380h, ago, marak, bule, rusia, ukraina, ulah,...] |
| 2 | menkumham sebut bule di bali naik motor tanpa ... | menteri hukum dan hak asasi manusia sebut bule... | menteri hukum hak asasi manusia bule bali moto... | menteri hukum hak asasi manusia bule bali moto... | [menteri, hukum, hak, asasi, manusia, bule, ba...] |
| 3 | video chris hemsworth pakai helm di bali jadi ... | video chris hemsworth pakai helm di bali jadi ... | video chris hemsworth pakai helm bali viral wa... | video chris hemsworth pakai helm bali viral wa... | [video, chris, hemsworth, pakai, helm, bali, v...] |
| 4 | heran bule di bali pakai helm trus viral kenap... | heran bule di bali pakai helm terus viral kena... | heran bule bali pakai helm viral mendasar vira... | heran bule bali pakai helm viral dasar viral p... | [heran, bule, bali, pakai, helm, viral, dasar,...] |
| 5 | naik motor di bali helm bule ini kacanya hampi... | naik motor di bali helm bule ini kacanya hampi... | motor bali helm bule kacanya lepas random | motor bali helm bule kaca lepas random | [motor, bali, helm, bule, kaca, lepas, random] |
| 6 | mas keren mau disiplin pake helm di bali ngga ... | mas keren mau disiplin pakai helm di bali tida... | mas keren disiplin pakai helm bali bule bali k... | mas keren disiplin pakai helm bali bule bali b... | [mas, keren, disiplin, pakai, helm, bali, bule...] |
| 7 | ini juga bule di bali naek motor pake helm | ini juga bule di bali naik motor pakai helm | bule bali motor pakai helm | bule bali motor pakai helm | [bule, bali, motor, pakai, helm] |
| 8 | askrl thor ke bali aja pake helm emang bule y... | askrl thor ke bali saja pakai helm memang bul... | askrl thor bali pakai helm bule berduit taat a... | askrl thor bali pakai helm bule duit taat atur... | [askrl, thor, bali, pakai, helm, bule, duit, t...] |
| 9 | rame grombolan oknum bule pekok dri negri anta... | ramai gerombolan oknum bule bodoh dari negeri ... | ramai gerombolan oknum bule bodoh negeri antah... | ramai gerombol oknum bule bodoh negeri antah a... | [ramai, gerombol, oknum, bule, bodoh, negeri, ...] |

- Noise removal (url, @, emoticon, tanda baca, tagar, angka, strip berlebih)
- Transform case (menjadikan huruf kecil)
- Normalisasi (mengubah kata slang menjadi baku (EYD))
- Stopword filtering (klasifikasi kata - kata yang tidak berhubungan)
- Stemming (merubah ke kata dasar)
- Token filtering (menghilangkan tanda baca, dan membagi kalimat)

[Source Code](#)

Hasil: Labelling

Untuk lebih mengetahui sentimen yang muncul dari hasil crawling, dilakukan labeling untuk tweet positif dan negatif. Tweet positif diberi nilai 1 dan tweet negatif diberi nilai 0

Labelling dilakukan menggunakan TextBlob dan Vader. Namun, hasilnya sangat tidak akurat sehingga dilakukan secara manual.

Data positif: 61 tweet

Data negatif: 429 tweet

| | Stemming | label |
|----|---|-------|
| 0 | pangeran bule suka ulah bali | 0 |
| 1 | 380h ago marak bule rusia ukraina ulah bali ko... | 0 |
| 2 | menteri hukum hak asasi manusia bule bali moto... | 0 |
| 3 | video chris hemsworth pakai helm bali viral wa... | 0 |
| 4 | heran bule bali pakai helm viral dasar viral p... | 0 |
| 5 | motor bali helm bule kaca lepas random | 1 |
| 6 | mas keren disiplin pakai helm bali bule bali b... | 0 |
| 7 | bule bali motor pakai helm | 1 |
| 8 | askri thor bali pakai helm bule duit taat atur... | 1 |
| 9 | ramai gerombol oknum bule bodoh negeri antah a... | 0 |
| 10 | mantap gubernur bali the best tampan berani bu... | 1 |
| 11 | artis hollywood kelas chris hemsworth harga at... | 1 |
| 12 | bule bali motor pakai helm | 1 |
| 13 | bali langsung lihat bule motor pakai baju helm... | 0 |
| 14 | thor pakai helm bali contoh bule bule motor pa... | 1 |
| 15 | heboh lihat bule pakai helm jakarta lokal paka... | 0 |
| 16 | pandang turis asing kendara sepeda motor makan... | 0 |
| 17 | model bule bali motor minimal pakai helm | 1 |
| 18 | ya bali polisi jalan atur lintas bule gapakai ... | 0 |
| 19 | tarik perhati mingil lihat video libur chris h... | 1 |

Hasil: Visualisasi Wordcloud



Wordcloud positif



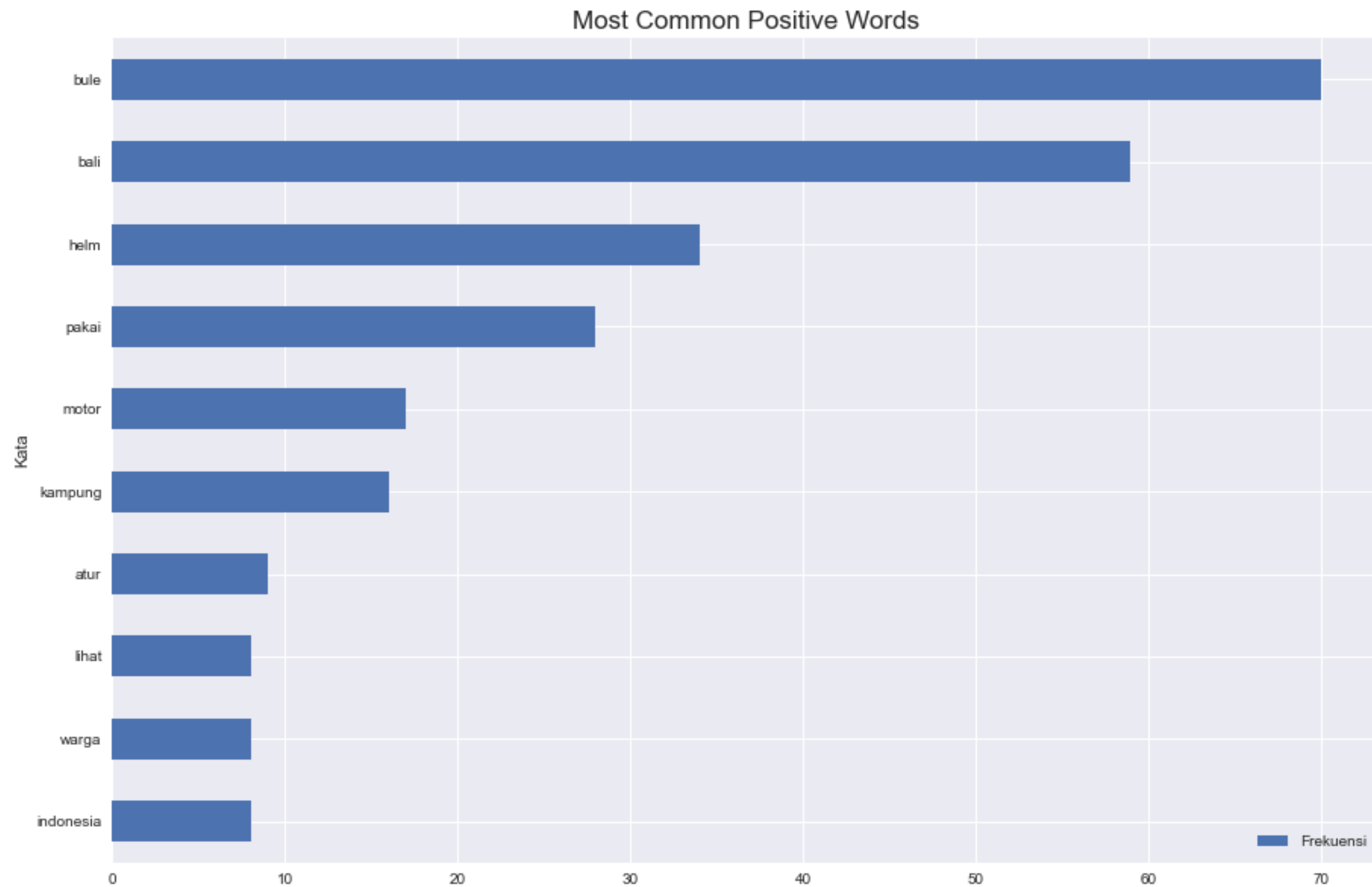
Wordcloud Negatif

Hasil dari labelling kemudian divisualisasikan sebagai berikut
https://colab.research.google.com/drive/1rB52oD8B3K5hAul98s7_WP
[W4mdkqF7jX?usp=sharing](https://colab.research.google.com/drive/1rB52oD8B3K5hAul98s7_WP)



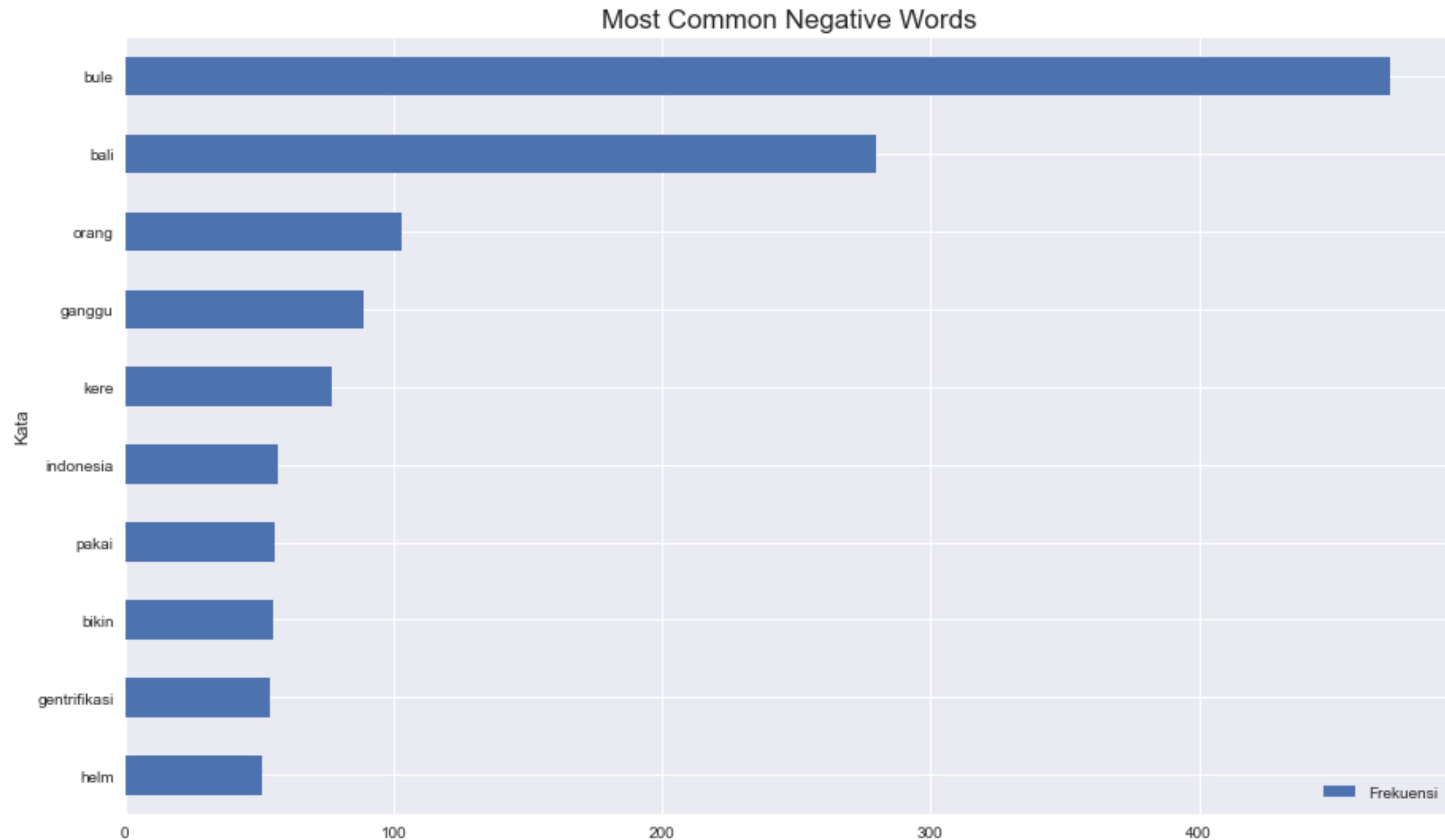
Hasil: Most Frequent Word

Analisis dilakukan dengan melihat kata - kata terbanyak yang muncul. Grafik kemudian digunakan sebagai bentuk dari visualisasi dari kata - kata tersebut. Terdapat 2 grafik yang digunakan yaitu kata positif dan negatif



Hasil: Most Frequent Word

Analisis dilakukan dengan melihat kata - kata terbanyak yang muncul. Grafik kemudian digunakan sebagai bentuk dari visualisasi dari kata - kata tersebut. Terdapat 2 grafik yang digunakan yaitu kata positif dan negatif



Hasil: Train Test Split Data

Selanjutnya, menggunakan library sklearn dengan fungsi train test split dengan proporsi data train dan test sebesar 80% dan 20%. Data terbagi sebanyak

```
train, test = train_test_split(df, test_size=0.2)
```

```
train_label = np.array(train['label'].to_list())  
train_text = np.array(train['stemming'].to_list())
```

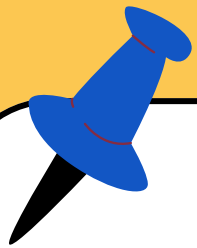
```
test_label = np.array(test['label'].to_list())  
test_text = np.array(test['stemming'].to_list())
```

Hasil: Auto Tokenizing

Setelah preprocessing data, pada tahapan ini data diberi token oleh tokenizer BERT Indonesia (AutoTokenizer) yang membagi teks menjadi token. Bentuk token atau kata ini bertujuan untuk memudahkan dan mencocokkan arsitektur data sebelum dilakukan pemodelan IndoBERTweet. Terdapat penambahan token [CLS] dan [SEP] dengan tujuan untuk memisahkan kalimat sehingga masing-masing token. Setelah itu, token dikonversi menjadi token menggunakan indeks yang telah ditentukan sebelumnya dari kosakata tokenizer BERT. Selain itu, dilakukan embeddings juga diberi attention mask yang menentukan seberapa besar perhatian yang harus diberikan sistem ke masing-masing kata dengan paddim max length sebanyak 256 dengan truncation true.

```
[11] from transformers import AutoTokenizer  
  
tokenizer = AutoTokenizer.from_pretrained("indolem/indobertweet-base-uncased", use_fast=True)  
  
[12] tokenizer
```

BertTokenizerFast(name_or_path='indolem/indobertweet-base-uncased', vocab_size=31923,
model_max_length=1000000000000000019884624838656, is_fast=True, padding_side='right', truncation_side='right',
special_tokens={'unk_token': '[UNK]', 'sep_token': '[SEP]', 'pad_token': '[PAD]', 'cls_token': '[CLS]', 'mask_token':
'[MASK]'}, clean_up_tokenization_spaces=True)



Hasil Pemodelan IndoBERTweet dan Evaluasi Error

setelah melakukan *auto tokenizing*, maka dilakukan pemanggilan fungsi *AutoModelForSequenceClassification* dengan jumlah input label sebanyak 2. Pada train preparation digunakan fungsi *TrainingArguments* yang berguna untuk mengatur *tuning* hingga *fine tuning* parameter dengan menggunakan *evaluation strate "steps"*, *save strategy "epoch"*, dan *mengaktifkan fitur drop out*.

Learning Rate : 5×10^{-8}

Epoch : 3

```
accuracy score: 0.8734177215189873
```

```
confusion matrix:
```

```
[[65  1]
 [ 9  4]]
```

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.88 | 0.98 | 0.93 | 66 |
| 1 | 0.80 | 0.31 | 0.44 | 13 |
| accuracy | | | 0.87 | 79 |
| macro avg | 0.84 | 0.65 | 0.69 | 79 |
| weighted avg | 0.87 | 0.87 | 0.85 | 79 |

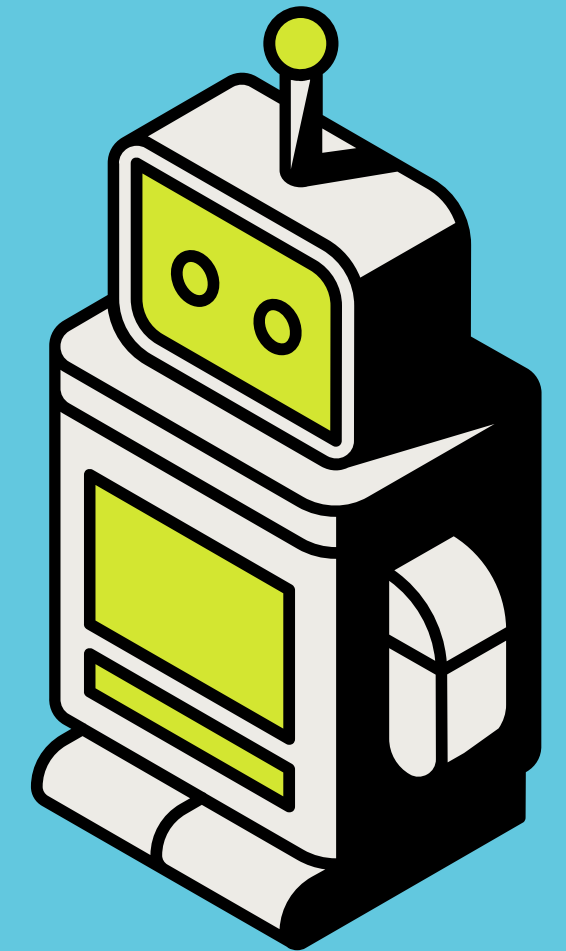
Pemodelan dilakukan dengan menggunakan model IndoBERTweet. Berikut adalah source code yang digunakan

https://colab.research.google.com/drive/100-sRXJ4aIEukCACviDAd_eMkSFIKBfm?usp=sharing

Kesimpulan

- Pada hasil visualisasi WordCloud, baik tweet bernilai positif dan negatif, sentimen masyarakat yang banyak berfokus terkait ketaatan dalam berkendara, tepatnya memakai Helm. Terkait adanya kampung bule di Bali, termasuk dalam sentimen positif. Namun, gentrifikasi bule di Bali masuk kedalam sentimen negatif. Bule bali mendapatkan sentimen negatif karena banyak memberikan gangguan dan berulah. Sebagaimana juga hasil visualisasi most common word. Bahwa untuk kata yg positif menyatakan bahwa bule bali tidak pakai helm adalah meniru warga Indonesia yang tidak taat juga. Sedangkan untuk most common negative word berfokus bahwa kebanyakan bule yang datang adalah bule kere, serta juga berfokus pada permasalahan gentrifikasi dan taat menggunakan helm.
- Klasifikasi menggunakan model Neural Network IndoBERTweet Fine Tuning dengan tuning parameter learning rate = 5×10^{-8} , epochs = 3, dan save steps = 200, menghasilkan accuracy sebesar 87.34%, dengan weighted average recall 87%, precision 87% dan F1-Score 85%. Model klasifikasi sentimen analisis review ini dapat digunakan pemerintah maupun pengelola wisata sebagai acuan dalam mengambil suatu keputusan atau kebijakan dalam meningkatkan kenyamanan tempat wisata di Bali

Thank you.



Social Media Policy: Guidance For Responsible Online Conduct
September 30, 2023