



UNIVERSITAS INDONESIA

**ANALISIS SENTIMEN CUITAN MASYARAKAT TERHADAP
PENYERANGAN SIBER KEPADA PEMERINTAH OLEH PERETAS BJORKA**

LAPORAN PENELITIAN

BRYAN TJANDRA

2106706634

JOHANNES SETIAWAN

2106750345

MUHAMAD HILMAN AL AYUBI

2106706653

YUDI PUTRA SABRI

2106706123

**FAKULTAS ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER**

DEPOK

NOVEMBER 2022

ABSTRAK

Pada zaman modern yang penuh akan perkembangan teknologi, semua masyarakat memberikan semua informasi yang mereka punya melalui media sosial. Media sosial Twitter merupakan salah satu platform media sosial yang banyak digunakan oleh masyarakat untuk memberikan tanggapan terhadap suatu fenomena. Fenomena yang dimaksud adalah serangan siber yang dilakukan oleh seseorang bernama Bjorka kepada pemerintah Indonesia. Berdasarkan konteks yang ada, cuitan dari pengguna Twitter dapat dikelompokkan berdasarkan sentimen positif, netral ataupun negatif untuk Pemerintah dan Bjorka. Selain itu, *hate speech* dari suatu cuitan juga dapat diekstrak untuk dilihat apakah cuitan tersebut mengandung kata-kata yang agresif atau kebencian. Berdasarkan hal tersebut, tim peneliti ingin melakukan analisis sentimen dan *hate speech* cuitan masyarakat Indonesia secara objektif untuk melihat tanggapan masyarakat terhadap bjorka. Sebelum melakukan analisis, data diekstrak melalui platform Twitter API. Data yang telah dilakukan *scraping* akan diseleksi dan dilakukan proses *labelling* secara *semi-supervised*. Setelah itu, data akan diproses melalui serangkaian proses, seperti normalisasi, penghilangan *stopwords*, dan tokenisasi. Data yang telah diproses akan dilakukan *training* dengan beberapa model *baseline* seperti BERTweet dan RoBERTa. BERTweet dilatih supaya dapat melakukan *hate speech*, sedangkan RoBERTa dilatih supaya dapat melakukan analisis sentimen. Model *baseline* diseleksi dan dievaluasi, lalu dilakukan *fine-tuning* terhadap model. Setelah itu, peningkatan model dilakukan dengan menggunakan *post evaluation* yang berupa *pseudo-labelling* dan *post processing*. Hasil dari proses tersebut dibandingkan dengan hasil yang didapat dari hasil *baseline* dan hasil *fine-tuning*. Hasil evaluasi model dengan *fine-tuning* dan *post evaluation* sentimen meningkat sebesar kurang lebih 10% dari model *baseline*, yaitu dengan rata-rata F1 dan akurasi sebesar 80%. Sementara untuk *hate-speech* meningkat sebesar 12% menjadi 55%. Hasil ini menunjukkan bahwa teknologi *Deep Learning* dapat digunakan untuk memprediksi sentimen dan *hate-speech* masyarakat. Dari hasil analisis sentimen, sebagian besar masyarakat memberikan sentimen netral, baik ke Pemerintah maupun Bjorka. Selain sentimen netral, jumlah cuitan bersentimen negatif jauh lebih banyak daripada yang bersentimen positif. Melihat persentase cuitan yang bersifat negatif baik terhadap Bjorka maupun Pemerintah, tim peneliti menyimpulkan bahwa sebagian besar pengguna Twitter yang membuat cuitan dengan topik terkait tidak menyukai kedua belah pihak.

Kata Kunci: Analisis Sentimen, *Hate Speech*, BERT, BERTweet, RoBERTa, Cuitan, Bjorka, Peretasan, Pemerintah, *Post Processing*, *Post Evaluation*, *Fine Tuning*, *Pseudo Labelling*, Siber

DAFTAR ISI

ABSTRAK	2
DAFTAR ISI	3
DAFTAR GAMBAR	5
DAFTAR TABEL	6
BAB 1 PENDAHULUAN	7
1.1. Latar Belakang	7
1.2. Pertanyaan Penelitian	8
1.3. Tujuan Penelitian	8
1.4. Manfaat Penelitian	8
1.4.1 Manfaat Praktis	8
1.4.2 Manfaat Akademis	9
1.5. Ruang Lingkup Penelitian	9
1.6. Sistematika Penulisan	9
BAB 2 TINJAUAN PUSTAKA	10
2.1. Analisis Sentimen	10
2.2. Hate Speech	10
2.3. Neural Network	10
2.4. Pysentimiento	11
2.5. BERT	11
2.6. RoBERTa	13
2.7. BERTweet	14
2.8. Pseudo-labeling	14
2.9. Transfer Learning & Pre-trained Model	15
2.10. Rangkuman Literatur	15
BAB 3 METODOLOGI	17
3.1. Pendekatan Penelitian	17
3.2. Tahapan Penelitian	17
3.3. Dataset	18
3.3.1. Data Collection	18
3.3.2. Analisis Dataset	19
3.3.3 Pembagian Dataset	20
3.4. Desain dan Implementasi	21
3.4.1. Software dan Environment	21
3.4.2. Overview Model	21
3.4.3. Matrik Evaluasi	23

3.4.4. Metode Eksperimen	23
3.4.5. Labeling	24
3.4.6. Data Processing	25
BAB 4 HASIL PENELITIAN DAN PEMBAHASAN	26
4.1. Hasil Pengujian	26
4.1.1. Hasil evaluasi baseline sentimen	26
4.1.2. Hasil evaluasi fine-tuning sentimen	26
4.1.3. Hasil post-evaluation sentimen	28
4.1.4. Hasil evaluasi hate speech	29
4.1.5. Analisis Hasil Prediksi	29
4.2. Pembahasan	30
BAB 5 PENUTUP	33
5.1. Kesimpulan	33
5.2. Saran	33
DAFTAR PUSTAKA	35

DAFTAR GAMBAR

Gambar 2.1 BERT	12
Gambar 2.2 RoBERTa	13
Gambar 2.3 BERTweet	14
Gambar 2.4 <i>Pseudo-labelling</i>	15
Gambar 3.1 Grafik Tren Cuitan Terkait Bjorka	19
Gambar 3.2. Banyak Kemunculan Kata	20
Gambar 3.3. Persentase Cuitan Berkonteks Bjorka	20
Gambar 3.4. Diagram Metode Eksperimen	24
Gambar 4.1. Graf <i>fine-tuning</i> Pemerintah	27
Gambar 4.2. Graf <i>fine-tuning</i> Bjorka	27
Gambar 4.3. Persentase Cuitan <i>Hate Speech</i>	30
Gambar 4.4. Persentase Sentimen dari Cuitan terhadap Pemerintah	31
Gambar 4.5. Persentase Sentimen dari Cuitan terhadap Bjorka	31
Gambar 4.6. Sentimen terhadap Bjorka Berdasarkan Kata	32

DAFTAR TABEL

Tabel 3.1 Contoh Cuitan berdasarkan konteks	18
Tabel 3.2. <i>Package</i> yang Digunakan	21
Tabel 3.3. Arsitektur BERTweet dan RoBERTa	22
Tabel 3.4. Hasil Pemrosesan Teks	25
Tabel 4.1. Hasil Evaluasi <i>Baseline</i>	26
Tabel 4.2. Hasil Evaluasi <i>fine-tuning</i> Pemerintah	27
Tabel 4.3. Hasil Evaluasi <i>fine-tuning</i> Bjorka	28
Tabel 4.4. Contoh Hasil Prediksi Sentimen <i>Hate Speech</i>	28
Tabel 4.5. Performa Evaluasi Prediksi <i>Hate Speech</i>	29
Tabel 4.6. Contoh Hasil Prediksi Sentimen dan <i>Hate Speech</i>	29
Tabel 4.7. Contoh Kesalahan Hasil Prediksi Sentimen dan <i>Hate Speech</i>	30

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Industri digital berkembang dengan sangat cepat dan sudah merambat ke berbagai aktivitas manusia. Perkembangan ini membuat sebagian besar informasi lembaga atau individu disimpan dalam bentuk digital. Hal ini mempengaruhi pentingnya peran keamanan siber dalam melindungi sistem digital dari berbagai ancaman atau akses ilegal. Di Indonesia sendiri, beberapa usaha sudah dilakukan dalam menangani kasus penyerangan siber, seperti pembuatan hukum terkait keamanan siber dan pembentukan Kementerian Pertahanan dan Kepolisian. Namun demikian, pada bulan Juli 2022, Bank Siber dan Sandi Negara mengklaim sebanyak 700 juta penyerangan siber telah dilancarkan di Indonesia (CNN Indonesia, 2022). Bahkan, menurut *National Cyber Security Index*, keamanan siber Indonesia menduduki peringkat 83 dari 160 negara (Databooks, 2022).

Salah satu kasus penyerangan siber di Indonesia dilakukan oleh seorang peretas yang berinisial Bjorka. Bjorka mengklaim telah membocorkan beberapa data privasi di Indonesia seperti 1.3 miliar data registrasi SIM *card* dan 105 juta data masyarakat terkait KPU (Dewi, 2022). Selain itu, Bjorka juga membocorkan data pemerintah yang disembunyikan ke masyarakat seperti dalang pembunuhan Munir. Hal ini menimbulkan kontroversi bagi masyarakat, baik mendukung atau menolak aksi yang dilakukan oleh Bjorka. Menurut data statistik Twitter, kontroversi ini menimbulkan cuitan tertinggi hingga 1.2 juta cuitan dalam sehari.

Perkembangan teknologi kecerdasan buatan dapat menjadi sebuah peluang dalam memperoleh sentimen dan *hate speech* dari opini cuitan yang dituangkan masyarakat. Analisis sentimen terhadap cuitan dapat diekstrak lebih lanjut untuk memperoleh konteks cuitan. Wawasan yang diperoleh dari analisis ini diharapkan dapat menciptakan peluang bagi pemerintah supaya lebih memperhatikan isu keamanan data dan pengambilan keputusan selanjutnya.

1.2. Pertanyaan Penelitian

Penelitian ini dilakukan karena timbulnya beberapa pertanyaan saat fenomena ini terjadi, diantaranya:

1. Bagaimana sentimen dan *hate speech* yang timbul pada masyarakat Indonesia terhadap Bjorka dan pemerintah melalui cuitan yang diunggah?
2. Apa *keywords* yang sering muncul di masyarakat tentang Bjorka berdasarkan sentimen positif dan negatif?

1.3. Tujuan Penelitian

Penelitian ini bertujuan untuk :

1. Memberikan gambaran secara luas tentang sentimen respons masyarakat Indonesia terhadap konteks yang berkaitan dengan kasus penyerangan siber oleh peretas Bjorka.
2. Mengetahui sentimen dan *hate speech* yang timbul pada masyarakat Indonesia terhadap Bjorka dan pemerintah.
3. Membuka jalan bagi penelitian-penelitian selanjutnya yang terkait dengan deteksi sentimen terhadap kasus keamanan data.

1.4. Manfaat Penelitian

1.4.1 Manfaat Praktis

Manfaat penelitian ini secara praktis adalah sebagai berikut.

1. Pemerintah tidak abai dengan isu-isu yang berpotensi menimbulkan kegaduhan.
2. Membuat pemerintah dan masyarakat lebih sadar akan pentingnya keamanan data.
3. Memberikan timbal balik berupa pengetahuan kepada pemerintah untuk pengambilan keputusan atau kebijakan.

1.4.2 Manfaat Akademis

Manfaat penelitian ini secara akademis adalah sebagai berikut.

1. Model ini dapat digunakan kembali untuk menyelesaikan permasalahan sentimen analisis kedepannya yang serupa dengan kasus Bjorka.
2. Model dapat digunakan untuk menentukan sentimen dan *hate speech* suatu cuitan.

1.5. Ruang Lingkup Penelitian

Metode yang dipaparkan oleh tim peneliti dan data yang digunakan memiliki beberapa batasan sebagai berikut:

1. Cuitan yang digunakan berasal dari negara Indonesia
2. Cuitan berasal dari opini masyarakat yang berbahasa Indonesia
3. Rentang waktu cuitan yang dikumpulkan dengan metode *data scraping* adalah 14 hari, mulai dari tanggal 8 sampai dengan 21 September 2022

1.6. Sistematika Penulisan

Laporan penelitian ini terdiri dari lima bab. Bab I membahas pendahuluan yang mencakup latar belakang dibuatnya laporan penelitian, pertanyaan penelitian, tujuan penelitian, manfaat penelitian secara praktis dan akademis, ruang lingkup penelitian, serta sistematika penulisan laporan penelitian. Bab II membahas tinjauan pustaka yang meliputi rangkuman teori dan hasil penelitian terdahulu yang digunakan sebagai dasar dari penelitian ini. Bab III membahas metodologi penelitian yang mencakup pendekatan penelitian yang digunakan, tahapan penelitian yang dijalankan, *datasets* yang digunakan, serta desain dan implementasi model yang digunakan. Bab IV membahas tentang hasil dari penelitian yang dilakukan beserta pembahasannya. Bab V membahas kesimpulan dari penelitian yang dilakukan, saran untuk penelitian selanjutnya, serta hal yang menghambat penelitian yang dilakukan oleh tim peneliti.

BAB 2

TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen atau *opinion mining* adalah sebuah proses untuk mengidentifikasi atau mengkategorikan pendapat masyarakat terkait suatu topik, layanan, atau produk. Kategori dari suatu pendapat dapat dikategorikan positif, negatif, atau netral. Metode ini biasanya menerima masukan berupa sebuah teks opini atau ulasan dan memberikan hasil berupa prediksi sentimen. Untuk memprediksi sebuah sentimen, pendekatan yang saat ini efektif digunakan berupa pendekatan metode *deep learning* berbasis *neural network* (Mehta, Pooja et al., 2020).

2.2. Hate Speech

Hate speech recognition merupakan proses untuk menentukan ujaran kasar atau kebencian yang terkandung dalam suatu respons masyarakat. *Hate speech* secara umum adalah mengacu kepada wacana ofensif seseorang atau kelompok tergantung pada seperangkat atribut, seperti etnis asal, agama, jenis kelamin, ras, asal kebangsaan, orientasi seksual, identitas gender, dan disabilitas. *Hate speech* dibuat secara sengaja maupun tidak sengaja. *Hate speech* dapat dikategorikan ke dalam beberapa konteks, yaitu konteks kata yang mengandung kebencian, kemarahan, atau yang ditargetkan pada seseorang. Rekognisi *hate speech* dapat dilakukan dengan metode *Natural Language Processing* yang dibantu dengan model *Deep Learning* (Ahlam Alrehili, 2019).

2.3. Neural Network

Neural Network merupakan sebuah model *deep learning* yang terinspirasi dari cara kerja otak manusia. Sebuah *neural network* memiliki tiga *layer* dalam modelnya, yaitu *input layer*, *hidden layer*, dan *output layer*. *Input layer* merupakan titik pertama yang merupakan tempat kita memasukkan data kita untuk diproses pada beberapa *layer* selanjutnya. *Output layer* merupakan tempat dari hasil yang diinginkan setelah data melalui beberapa *layer*. Seluruh *layer* kecuali *input layer* dan *output layer* disebut sebagai *hidden layer* (Wang, 2003).

Salah satu contoh dari *neural networks* yang digunakan dalam analisis sentimen adalah *deep neural networks* atau biasa disingkat menjadi DNN. DNN adalah *artificial neural networks* yang memiliki beberapa *hidden layer* dan berada di banyak arsitektur yang bergantung dengan topologi dari neuron dan koneksinya (Pota et al., 2021). Salah satu contoh dari DNN adalah *recurrent neural networks* atau biasa disingkat menjadi RNN (Ashima dan Vishwakarma, 2020).

Model RNN adalah model yang digunakan dalam berbagai aplikasi untuk memodelkan data yang sambung-menyambung satu sama lain. Model RNN merepresentasikan suatu data yang berupa kalimat menjadi struktur data *tree* dengan representasi *word vector* berada pada *leaf* dari *tree* tersebut. Representasi *parent* dari semua *leaf* tersebut dibuat secara rekursif dari bawah ke atas. Proses tersebut menggabungkan token-token menjadi representasi dari frasa yang akhirnya menjadi kalimat lengkap. Representasi frasa yang dihasilkan akan digunakan untuk klasifikasi sentimen (Ashima dan Vishwakarma, 2020).

2.4. Pysentimiento

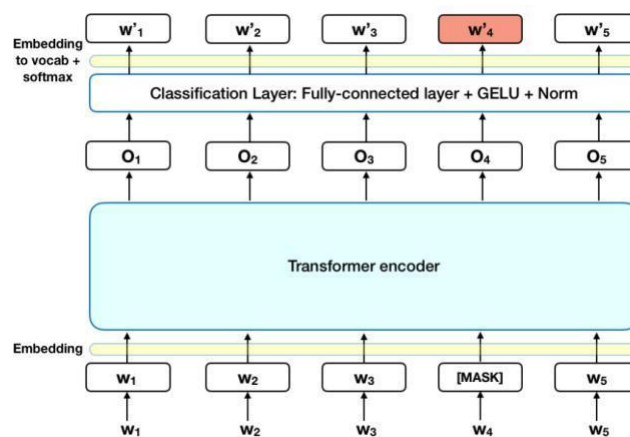
Pysentimiento merupakan sebuah *multilingual toolkit* berbasis bahasa Python yang menyediakan solusi-solusi penyelesaian masalah analisis sentimen dan *Social Natural Language Processing*. *Toolkit* ini dapat digunakan dalam hal pemrosesan teks dan analisis sentimen menggunakan pendekatan *Neural Network*. Pemrosesan teks ini mendukung ekstensi dari cuitan sehingga dapat memproses *tag*, *hashtag*, dan emoji dari cuitan. Model analisis sentimen yang digunakan merupakan *pretrained state-of-the-art* model berbasis BERT, seperti RoBERTa dan BERTweet.

2.5. BERT

BERT (*Bidirectional Encoder Representations for Transformers*) merupakan *framework deep learning* yang menggunakan *Transformers*. *Transformers* ini digunakan untuk mempelajari hubungan antara dua kata pada suatu tulisan. Mekanisme pada *Transformers* dibagi menjadi dua, yaitu *encoder* untuk membaca tulisan input dan *decoder* sebagai hasil dari prediksi dari tugas yang telah dilakukan (Batra et al., 2021). *Encoder transformers* membaca seluruh urutan kata sekaligus, sehingga model BERT

dianggap sebagai model dua arah. Karakteristik ini yang membuat model BERT mempelajari konteks sebuah kata berdasarkan semua lingkungannya.

Model ini melalui dua tahapan, yaitu *pre-training* dan *fine-tuning*. Pada tahap pertama, model BERT dilatih dengan beragam tugas dengan data yang belum diberikan label. Pada tahap kedua, model BERT yang diinisialisasi dengan parameter dari tahap pertama akan dilakukan proses *fine-tuned* dengan data yang sudah diberi label. (Devlin et al., 2019)



Gambar 2.1. BERT (Sumber:<https://medium.com/@ranihorev>)

Model BERT sudah digunakan di beberapa penelitian sebelumnya. Penelitian yang dilakukan oleh Chiorrini, Mircoli, Diamantini, dan Potena (2021) menguji nilai akurasi, presisi, dan *recall* dari model BERT. Mereka mengevaluasi performa model BERT dalam menganalisis sentimen *datasets* dari cuitan Twitter. Penelitian mereka menghasilkan skor 92% untuk akurasi dari model BERT dalam menganalisis sentimen cuitan dari Twitter. Nilai akurasi yang tinggi membuat mereka menyimpulkan bahwa model BERT memiliki akurasi yang tinggi untuk analisis sentimen. Akurasi yang tinggi membuat model BERT memberikan kontribusi signifikan dalam pengklasifikasian teks yang baik. Hal yang sama juga disampaikan oleh Marco Pota, Mirko Ventura, Rosario Catelli, dan Massimo Esposito (2021) yang mengatakan bahwa model BERT mendapatkan hasil yang memuaskan dalam pengklasifikasian teks.

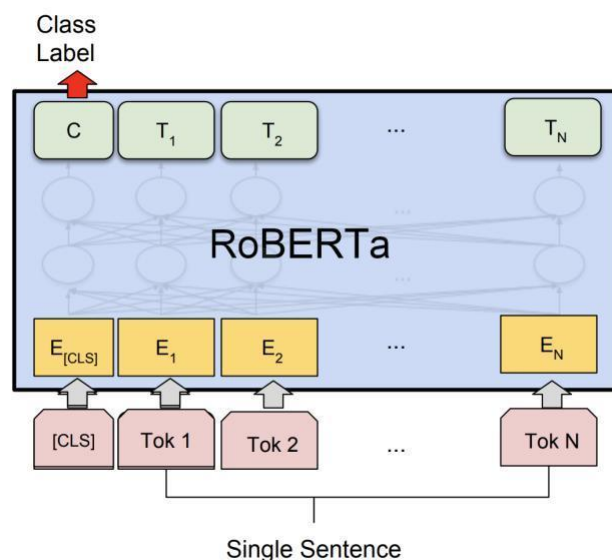
Dalam penelitian lain yang dilakukan oleh Kumar T. S. (2021) dan kawan-kawannya, model BERT mendapatkan skor akurasi dan F1 yang tertinggi jika dibandingkan dengan beberapa model yang lain. Kumar T. S. dan kawan-kawannya

(2021) menguji model BERT dan beberapa model lain seperti *Support Vector Machine* (SVM), *Stochastic Gradient Descent* (SGD), *Logistic Regression*, *Decision-Tree*, *Random Forest Method*, dan *XGBoost* dalam melakukan analisis sentimen terhadap data cuitan Twitter yang didapat dari platform Kaggle.

Performa yang diuji dan dibandingkan dalam penelitian Kumar T. S. dan kawan-kawannya (2021) adalah skor akurasi, presisi, *recall*, dan F1. Hasil akhir dari penelitian tersebut menyatakan model BERT mendapatkan skor tertinggi pada skor akurasi, presisi, *recall*, dan F1 dibanding model lainnya. Berdasarkan hasil ini, Kumar T. S. dan kawan-kawannya (2021) menyimpulkan bahwa model BERT adalah model yang paling cocok untuk analisis sentimen. Kesimpulan tersebut sejalan dengan hasil dari penelitian-penelitian lainnya. Oleh sebab itu, peneliti memilih untuk menggunakan model BERT dalam melakukan sentimen analisis.

2.6. RoBERTa

RoBERTa merupakan salah satu *framework deep learning* yang menggunakan BERT sebagai dasar dari modelnya. RoBERTa memiliki optimisasi pada analisis data yang berukuran besar dengan mengubah β_2 menjadi 0.98. *Training* data yang dilakukan berupa kalimat yang memiliki ukuran maksimal sesuai dengan ketentuan. Optimisasi lain yang dilakukan adalah data yang dilakukan *training* diduplikasi sepuluh kali dengan sepuluh cara analisis yang berbeda untuk menghindari pengelompokan sebuah kata yang sama.



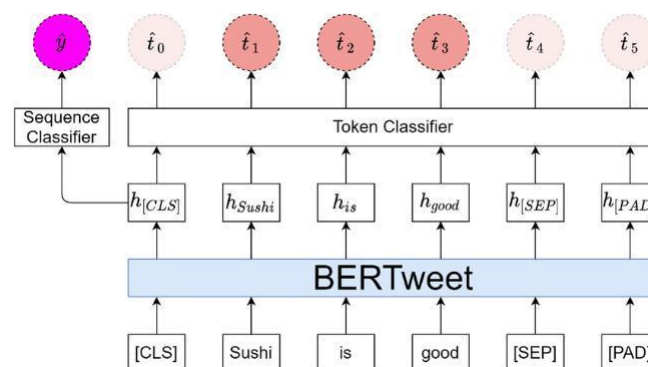
Gambar 2.2 RoBERTa (Sumber:

<https://paperswithcode.com/model/RoBERTa-large-sst>)

Beberapa contoh dari model RoBERTa adalah akahana, ayameRushia, dan w11wo. Model ayameRushia memiliki akurasi sebesar 93,5%. Model w11wo memiliki akurasi sebesar 94,2%. Model akahana memiliki akurasi sebesar 91,6%. Ketiga model tersebut memiliki performa yang tinggi sehingga ketiga model tersebut digunakan dalam penelitian ini untuk mengevaluasi data yang didapat dari metode *data scraping*.

2.7. BERTweet

BERTweet merupakan salah satu model dalam analisis sentimen yang juga merupakan model pertama yang dikembangkan untuk mengolah cuitan berbahasa Inggris. Tim peneliti menggunakan BERTweet karena BERTweet menggunakan arsitektur berbasis BERT yang dikembangkan dengan mengoptimisasi model RoBERTa. Optimisasi dilakukan dengan membatasi panjang cuitan sepanjang 128 kata, menggabungkan implementasi model Adam, dan meningkatkan spesifikasi dari mesin yang digunakan.



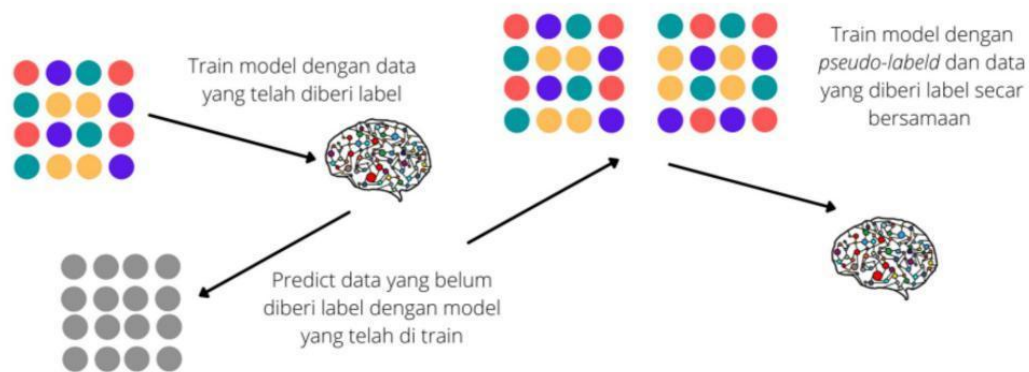
Gambar 2.3. BERTweet (Sumber:

https://www.researchgate.net/figure/Diagram-of-the-dual-task-BERTweet-model_fig4_355918048)

2.8. Pseudo-labeling

Pseudo-labeling merupakan salah satu pendekatan yang dapat dipilih dalam *semi-supervised learning*. Teknik *pseudo-labeling* menggunakan model yang dibangun

dari proses klasifikasi menggunakan data yang sudah ditandai untuk melakukan prediksi terhadap *unlabeled* data. Hasil prediksi dari data yang belum ditandai ini yang dimaksud sebagai *pseudo-label* data dengan label hasil prediksi dianggap sebagai label sebenarnya dari data. Kedua data tersebut kemudian dimanfaatkan untuk membuat model *classifier* baru. (Qiu et al., 2019)



Gambar 2.4. *Pseudo-labelling* (Sumber:

<https://datawhatnow.com/pseudo-labeling-semi-supervised-learning>)

2.9. Transfer Learning & Pre-trained Model

Transfer Learning (TL), merupakan salah satu metode yang umum digunakan pada penelitian berbasis *modelling*. Metode tersebut menyimpan pengetahuan yang dipunyai dan mengaplikasikannya ke masalah lainnya yang memiliki konteks serupa. *Pre-trained* model adalah model yang dibuat oleh orang lain yang dapat digunakan untuk memecahkan sebuah masalah yang memiliki kasus yang memerlukan pendekatan yang serupa. Dibandingkan membuat model dari awal, *pre-trained model* dapat digunakan sebagai langkah awal untuk menganalisis kasus secara efisien. BERT merupakan salah satu *pre-trained model* yang dipakai tim peneliti dengan *Transfer Learning* dalam makalah kali ini.

2.10. Rangkuman Literatur

Analisis sentimen adalah suatu proses untuk mengidentifikasi sentimen masyarakat. Analisis sentimen menggunakan beberapa model *deep learning* berupa

BERTweet dan RoBERTa. Tim peneliti menggunakan BERTweet sebagai model untuk mengenali cuitan yang mengandung *hate speech*, sedangkan RoBERTa digunakan untuk mengekstrak sentimen dari data tersebut. BERTweet dan RoBERTa yang digunakan merupakan *pre-trained model* yang diekstrak dengan *transfer learning*.

Pre-trained model yang digunakan masih membutuhkan proses *fine-tuning* lagi dengan data yang dimiliki oleh tim peneliti. *Deep learning* dilakukan dengan menggunakan model berbasis *Neural Network* dan *toolkit Pysentimiento*. Setelah itu, model akan dilakukan *pseudo-labeling* untuk memperoleh peningkatan performa model. Dengan *pseudo-labeling*, model akan mendapatkan pembelajaran baru yang menyesuaikan data yang telah ditambahkan. Dengan demikian, hasil analisis dari model tersebut dapat menjadi lebih akurat terhadap data yang dikumpulkan.

BAB 3

METODOLOGI

3.1. Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif. Penelitian secara kuantitatif digunakan untuk menghitung serta membandingkan hasil dari skor akurasi dan *weighted F1* yang diperoleh dari data yang sudah dianalisis. Selain itu, pendekatan kuantitatif juga digunakan untuk mengetahui persentase seberapa besar dari akumulasi cuitan yang sudah didapat tersebut mendukung Bjorka atau pemerintah.

3.2. Tahapan Penelitian

Tahapan penelitian mencakup langkah-langkah pelaksanaan dari awal sampai akhir, adapun langkahnya sebagai berikut:

1. Penentuan Topik Analisis Sentimen dan *Hate Speech*

Topik analisis sentimen dan *hate speech* diangkat dari fenomena yang sedang menjadi buah bibir pada sebuah komunitas atau masyarakat. Fenomena yang dijadikan topik juga merupakan fenomena yang kontroversial. Fenomena kontroversial sering kali menimbulkan respons yang tidak dapat dilihat secara jelas, baik yang mendukung, menolak ataupun netral. Dengan begitu, tim peneliti dapat menganalisis sentimen dan *hate speech* dari respons tersebut.

2. Studi Pendahuluan

Pada tahap ini, peneliti mengkaji studi literatur yang berkaitan dengan analisis sentimen dan *hate speech*. Hal ini dilakukan untuk mengkaji dan mengetahui secara teoritis metode-metode yang dipakai pada penelitian-penelitian sebelumnya. Setelah itu, peneliti melakukan perbandingan metodologi yang dipakai oleh penelitian-penelitian sebelumnya. Perbandingan dilakukan untuk mengetahui metodologi yang sesuai dengan penelitian yang akan dilakukan.

3. Perumusan Masalah

Pada tahap selanjutnya akan dilakukan perumusan masalah. Perumusan masalah dilakukan agar alur penelitian menjadi terarah dan tidak keluar dari tujuan utama. Hal ini bertujuan untuk menghasilkan hasil yang sesuai dengan ekspektasi tujuan penelitian.

4. Pengumpulan Data

Pengumpulan data dilakukan pada salah satu media sosial Twitter. Twitter dipilih berdasarkan beberapa alasan yang sudah dijelaskan sebelumnya. Pengumpulan data dilakukan dengan metode *data scraping* melalui Twitter API.

5. Pemodelan data

Data yang sudah terkumpul akan diolah dengan menggunakan model pembelajaran mesin BERTweet dan RoBERTa dengan metode yang dijelaskan pada bagian 3.4.4. Data yang sudah diolah akan dilihat berdasarkan akurasi dan bobot skor F1-nya.

3.3. Dataset

3.3.1. Data Collection

Untuk pengumpulan data, tim peneliti melakukan pengumpulan data dengan metode *data scraping* melalui platform Twitter dengan kriteria rentang waktu 14 hari, mulai dari tanggal 8 sampai 21 September 2022. *Scraping* yang dilakukan memiliki rata-rata cuitan sebanyak 9 ribu cuitan per harinya dan mengandung kata kunci bjorka. Hasil proses *data scraping* menunjukkan perolehan sebanyak 131 ribu cuitan dari kata kunci tersebut pada rentang tanggal yang sudah ditentukan sebelumnya. Berikut beberapa contoh cuitan yang memiliki bobot #negatif, #positif, ataupun #netral

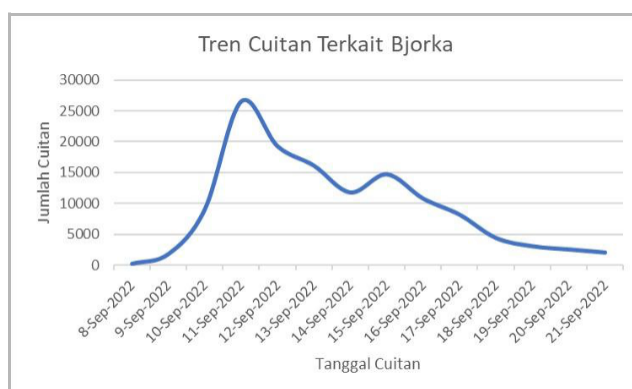
Tabel 3.1. Contoh Cuitan berdasarkan konteks

Cuitan	Konteks
“...Heker bjorka itu ternyata orang cirebon? Kok kayak masi bocah sih...”	Bjorka
“.. dia dianggap pahlawan hanya karena bobol data-data milik pemerintah.”	Pemerintah

“... Kalo sistem yang di retas, sudah bisa jadi amburadul semua.. Bjorka bagiku ada sisi White Hacker, Black Hacker ... “	Bjorka
“...bjorka susah2 buat ngeretas data penduduk,eh pemerintah kita segampang itu nyebarin data nya...”	Pemerintah

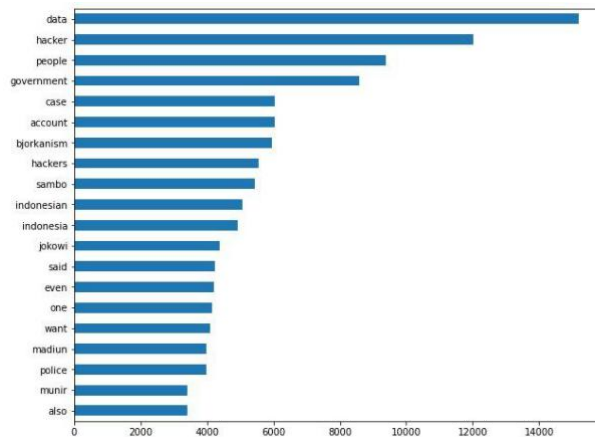
3.3.2. Analisis Dataset

Dataset yang sudah terkumpul akan diorganisir berdasarkan kapan waktu cuitan tersebut diunggah. Cuitan tersebut kemudian dikelompokkan berdasarkan hari. Pengelompokan tersebut akan menghasilkan sebuah grafik yang menunjukkan jumlah cuitan yang diunggah per hari. Grafik jumlah cuitan yang diunggah per hari menunjukkan bahwa terjadi dua puncak jumlah cuitan yang terkait dengan Bjorka, yaitu pada tanggal 11 September 2022 dan tanggal 15 September 2022.



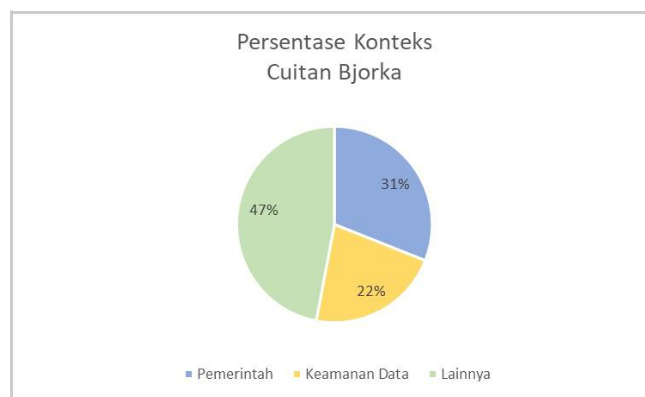
Gambar 3.1. Grafik Tren Cuitan Terkait Bjorka

Dataset juga diorganisir berdasarkan kata. *Dataset* dikelompokkan berdasarkan jumlah kemunculan kata dalam *dataset* tersebut. Kata “Bjorka” akan dikecualikan untuk perhitungan kali ini. Kata “data” menjadi kata yang paling banyak muncul diikuti oleh kata “hacker”, “people”, dan “government”.



Gambar 3.2. Banyak Kemunculan Kata

Selain itu, *dataset* yang ada dapat dikelompokkan berdasarkan konteks tweet tersebut. *Dataset* tersebut dikelompokkan berdasarkan konteks dan dihitung persentase jumlah cuitannya. Berdasarkan grafik tersebut, 47% cuitan membicarakan atau membahas pemerintah, 22% membahas keamanan data, dan sisanya tidak dapat dikelompokkan secara spesifik karena terlalu bervariasi.



Gambar 3.3. Persentase Cuitan Berkonteks Bjorka

3.3.3 Pembagian *Dataset*

Sebelum dilakukan *training* pada data, data dibagi terlebih dahulu berdasarkan konteks cuitan yang mengarah ke pemerintah dan bjorka. Pembagian konteks pemerintah tersebut didasarkan oleh *keywords* yang berhubungan dengan pemerintah, seperti “pemerintah”, “menteri”, “presiden”, dan lainnya. Sementara, untuk konteks bjorka, data dibagi berdasarkan *keywords* “keamanan”, “peretas”, “siber”, dan lainnya.

Data yang telah dibagi ini akan dilakukan analisis sentimen dan *hate speech* untuk melihat komposisi sentimen dan *hate speech* dari masing-masing konteks. Tim peneliti juga melakukan evaluasi pada pembagian konteks yang kurang tepat dengan cara melakukan *filter* pada cuitan yang kurang sesuai.

3.4. Desain dan Implementasi

3.4.1. Software dan Environment

Penambahan dan analisis data yang dilakukan menggunakan bahasa pemrograman Python melalui platform Google Colab dengan GPU Nvidia T4 Tensor Core beserta *package* penambahan data yang tersedia oleh komunitas Python. Berikut adalah *package* yang tim peneliti gunakan.

Tabel 3.2. *Package* yang Digunakan

Transformer	Pandas	Tqdm	Datasets	Scikit-learn	NLTK	GoogleTrans
Pysentimiento	Numpy	Tensor Flow	WordCloud	Seaborn	Regex	TextBlob

3.4.2. Overview Model

Eksperimen yang kami lakukan adalah untuk membuat model klasifikasi sentimen dan *hate speech* berdasarkan konteks dengan menggunakan model *deep learning* berbasis BERT yaitu BERTweet dan RoBERTa. Sebelum dilakukan *training*, dilakukan proses *splitting* data terlebih dahulu menjadi *train set* dan *validation set* dengan perbandingan 80:20. Untuk *learning rate*, tim peneliti menggunakan *learning rate* sebesar $1e-05$ yang dioptimasi dengan Adam W. Kemudian, semua model dilakukan *train* sampai maksimal 20 *epochs* dan diadakan *early stopping callback* sampai *convergence*.

Model BERTweet menggunakan arsitektur BertForMaskedLM, sedangkan model RoBERTa menggunakan arsitektur RoBERTaForSequenceClassification dengan parameter dari masing-masing *requirements* sebagai tabel berikut.

Tabel 3.3. Arsitektur BERTweet dan RoBERTa

Parameter	BERTweet	RoBERTa
_name_or_path	"bert-base-cased"	"pysentimiento/bertweet-hate-speech"
architectures	["BertForMaskedLM"]	["RoBERTaForSequenceClassification"]
attention_probs_dropout_prob	0.1	0.1
classifier_dropout	null	null
gradient_checkpointing	false	false
hidden_act	"gelu"	"gelu"
hidden_dropout_prob	0.1	0.1
hidden_size	768	768
initializer_range	0.02	0.02
intermediate_size	3072	3072
layer_norm_eps	1e-12	1e-05
max_position_embeddings	512	130
model_type	"bert"	"RoBERTa"
num_attention_heads	12	12
pad_token_id	0	1
position_embedding_type	"absolute"	"absolute"
transformers_version	"4.13.0"	"4.13.0"

type_vocab	2	2
use_cache	true	true
vocab_size	28996	64001

3.4.3. Matrik Evaluasi

Evaluasi hasil klasifikasi sentimen dan *hate speech* menggunakan *metric* F1 *Score*. F1 dipilih menjadi matrik evaluasi karena adanya *imbalanced* pada target. Selain F1, tim peneliti juga menggunakan matrik evaluasi akurasi untuk melihat ketepatan model dalam melakukan klasifikasi. Rumus dari akurasi dan F1 yang digunakan adalah

$$Accuracy = \frac{\sum_{i=1}^n \frac{1}{2} (1 + \frac{TP_i + TN_i}{TP_i + FP_i + FN_i + TN_i})}{n}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

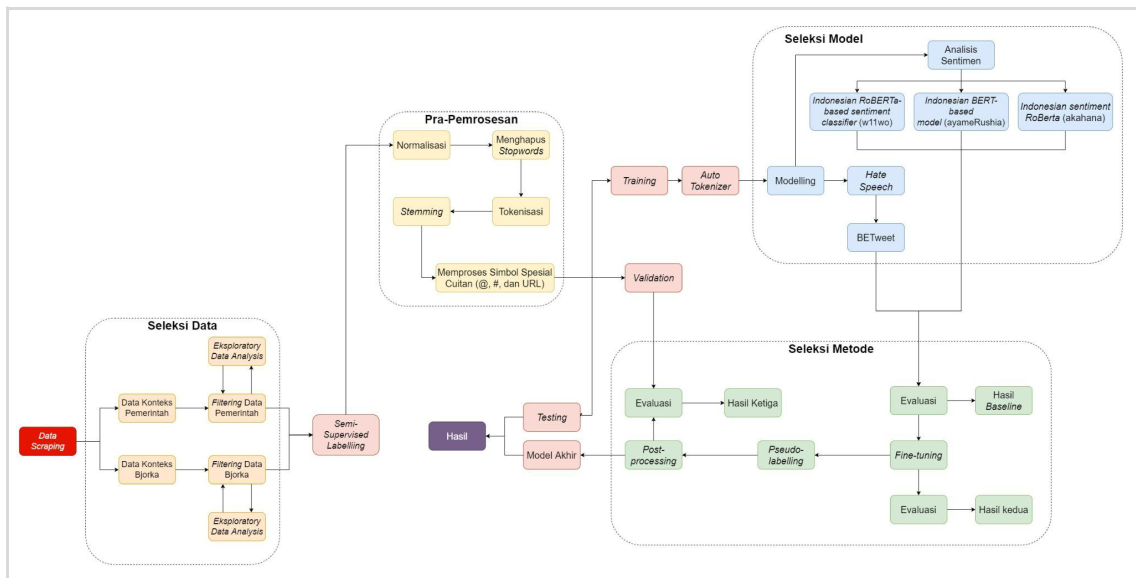
Dengan kedua matriks evaluasi F1 dan akurasi, model dapat mengevaluasi ketepatan prediksi suatu label, sekaligus tidak *bias* terhadap *imbalance* sentimen negatif yang cukup banyak.

3.4.4. Metode Eksperimen

Pada awalnya, data cuitan yang akan digunakan dilakukan *scrap* terlebih dahulu melalui Twitter API. Cuitan yang telah dilakukan *scrap* akan dibagi berdasarkan konteks menjadi data untuk konteks cuitan pemerintah dan konteks cuitan bjorka. Setelah itu, dilakukan *labelling* secara *semi-supervised* dan analisis *dataset* untuk melakukan *filter* terhadap data dengan konteks yang kurang relevan. Data yang sudah dilakukan *filter* dan dibagi akan dikonversi menjadi bahasa Indonesia dan diproses terlebih dahulu dengan melakukan normalisasi dan menghilangkan *stopwords*, *hashtag*, dan lainnya.

Data yang telah diproses akan dilakukan *train* dengan model *baseline* BERTweet dan RoBERTa untuk klasifikasi sentimen dan *hate speech*. Tim peneliti membandingkan beberapa mode *baseline* RoBERTa untuk klasifikasi sentimen yang akan diseleksi berdasarkan evaluasi menggunakan *weighted F1* dan akurasi. Model yang sudah dievaluasi akan di *fine-tuning* sampai proses *training* mencapai *convergence*.

Kemudian, model tersebut akan dilakukan *pseudo-labelling*, yaitu proses prediksi pada *unlabelled data* oleh model dan dilakukan *training* ulang dengan gabungan data tersebut. Selain *pseudo-labelling*, proses *post-processing* juga akan dilakukan pada hasil prediksi yang tidak sesuai dan memiliki *confidence score* rendah. Hasil *training* pada data tersebut dievaluasi untuk dibandingkan dengan hasil *baseline* dan *fine-tuning* sederhana.



Gambar 3.4. Diagram Metode Eksperimen

3.4.5. Labeling

Data yang diperoleh dari Twitter berupa data yang *unlabelled*. Oleh karena itu, diperlukan adanya *labeling* pada data cuitan yang sudah dilakukan *scrap*. *Labeling* dilakukan secara *semi-supervised* melalui *pretrained* BERTweet dan RoBERTa. Setelah label diprediksi, tim peneliti juga melakukan evaluasi untuk memperbaiki label yang masih kurang tepat, sehingga tidak terjadi distraksi pada model. Label yang dievaluasi berjumlah 2000 cuitan yang nantinya akan dilakukan *feed* ke model *training*.

3.4.6. Data Processing

Hasil cuitan yang diperoleh belum sepenuhnya dapat langsung dilakukan *train* ke model dikarenakan masih terdapat data yang belum bersih. Maka dari itu, diperlukan adanya pemrosesan data untuk menghasilkan data yang lebih sederhana. Cuitan yang telah dilakukan *scrap* akan dinormalisasi dengan dibuat *case insensitive* dan emoji serta tanda baca dihilangkan semua dari cuitan. Setelah itu, bahasa campuran dalam cuitan akan dikonversi sepenuhnya menjadi bahasa Indonesia. Lalu, tim peneliti menghilangkan *stopwords* yang ada pada cuitan, misalnya “saya”, “pula”, “sampai”, dan lainnya.

Cuitan yang sudah diproses tersebut akan dilakukan tokenisasi dan *stemming* pada masing-masing kata. Tujuan proses ini adalah untuk membagi teks menjadi token-token dan menghilangkan imbuhan, misalkan “menyukai” menjadi “suka”. Tim peneliti juga memproses cuitan yang mengandung *tag user*, *hashtag*, URL, dan kata yang berulang dengan *framework* pysentimiento. Berikut merupakan beberapa hasil teks yang telah diproses.

Tabel 3.4. Hasil Pemrosesan Teks

Sebelum	Sesudah
Bjorka Ternyata Seorang Hacker Pembobol Data Pemerintah	bjorka <i>hacker</i> bobol data pemerintah
@tvOneNews Tandanya polisi kita tidak mampu melacak Bjorka	tanda polisi tidak mampu lacak bjorka
Tegaskan Tak Ada Data Negara Bocor, Mahfud MD: Bjorka Ngarang Terus Disebar	tegas tidak ada data negara bocor mahfud md bjorka ngarang sebar

BAB 4

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Hasil Pengujian

4.1.1. Hasil evaluasi *baseline* sentimen

Model *baseline* yang akan dibandingkan merupakan model *pretrained* berbasis RoBERTa berbahasa Indonesia, yaitu model milik w11wo, ayameRushia, dan akahana. Berikut hasil pengujian beberapa model *baseline* tersebut pada konteks pemerintah dan konteks Bjorka.

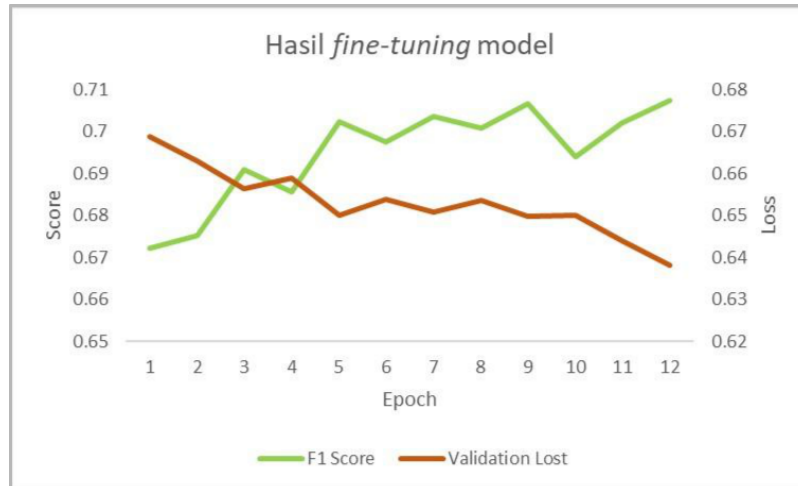
Tabel 4.1. Hasil Evaluasi Baseline

Model	Pemerintah		Bjorka	
	Akurasi	Weighted F1-score	Akurasi	Weighted F1-score
ayameRushia	0.588	0.577	0.560	0.559
w11wo	0.659	0.658	0.670	0.657
akahana	0.588	0.580	0.510	0.535

Berdasarkan hasil evaluasi model *baseline* di atas, model *baseline* yang memperoleh performa paling baik adalah model RoBERTa Indonesia yang dikembangkan oleh w11wo. Model w11wo memperoleh performa rata-rata *weighted F1-Score* dan akurasi berturut-turut sebesar 0.657 dan 0.665. Model ini akan dilakukan *tuning* dan dilakukan *post-processing* lebih lanjut untuk meningkatkan performa model.

4.1.2. Hasil evaluasi *fine-tuning* sentimen

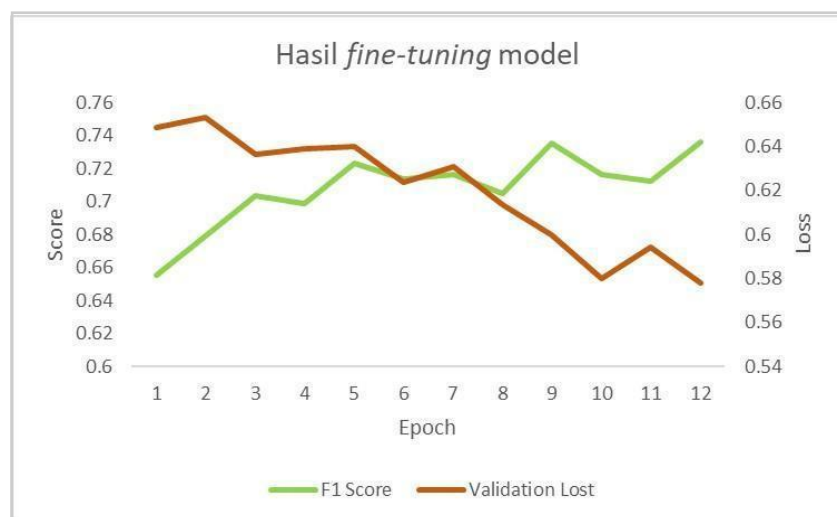
Model *baseline* RoBERTa Indonesia oleh w11wo dilakukan *fine-tuning* sesuai dengan konteks pemerintah dan bjorka. Berikut hasil dari graf *loss* dan skor *fine tuning* dari setiap *epochs* serta evaluasinya untuk cuitan konteks pemerintah dan Bjorka.



Gambar 4.1. Graf *fine-tuning* Pemerintah

Tabel 4.2. Hasil Evaluasi *fine-tuning* Pemerintah

	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>Support</i>
Sentimen Negatif	0.6688	0.7048	0.6863	2707
Sentimen Netral	0.7475	0.7435	0.7455	4078
Sentimen Positif	0.5762	0.4577	0.5102	603
Akurasi	-	-	0.7060	7388
<i>Weighted F1</i>	0.7047	0.7060	0.7046	7388



Gambar 4.2. Graf *fine-tuning* Bjorka

Tabel 4.3. Hasil Evaluasi *fine-tuning* Bjorka

	<i>Precision</i>	<i>Recall</i>	F1 <i>Score</i>	<i>Support</i>
Sentimen Negatif	0.6921	0.7555	0.7341	4112
Sentimen Netral	0.7614	0.7717	0.7633	6953
Sentimen Positif	0.5996	0.5520	0.5752	1005
Akurasi	-	-	0.7432	9515
<i>Weighted F1</i>	0.7442	0.7541	0.7510	9515

Dapat dilihat skor setelah *fine-tuning* meningkat untuk kedua konteks pemerintah dan Bjorka. Akurasi dan F1 *score* untuk konteks pemerintah berturut-turut meningkat dari 0.659 dan 0.658 menjadi 0.7060 dan 0.7046. Sementara itu, akurasi dan F1 *score* untuk konteks Bjorka berturut-turut meningkat dari 0.670 dan 0.657 menjadi 0.7432 dan 0.7510. Selanjutnya, akan dilakukan *post evaluation* berupa *post processing* dan *pseudolabelling*.

4.1.3. Hasil *post-evaluation* sentimen

Setelah dilakukan *post-processing* dan *pseudo-labelling*, performa akurasi dan F1 model meningkat untuk konteks pemerintah dari 0.7060 dan 0.7046 menjadi 0.7727 dan 0.7780. Sementara itu, performa akurasi dan F1 model meningkat untuk konteks bjorka meningkat dari 0.7432 dan 0.7510 menjadi 0.8182 dan 0.8379. Berikut merupakan ringkasan dari evaluasi performa model *baseline*, *fine-tuning*, dan *post-evaluation*.

Tabel 4.4. Performa Evaluasi Prediksi Sentimen

Metode	Pemerintah		Bjorka	
	Akurasi	Weighted F1-score	Akurasi	Weighted F1-score
Model <i>baseline</i>	0.659	0.658	0.670	0.657

Model <i>fine-tuning</i>	0.7060	0.7046	0.7432	0.7510
Model hasil <i>post evaluation</i>	0.7727	0.7780	0.8182	0.8379

4.1.4. Hasil evaluasi *hate speech*

Model yang digunakan untuk *hate speech* menggunakan model BERTweet untuk *baseline*. Setelah itu, dilakukan *fine-tuning* pada model dan *post evaluation* dengan teknik yang sama seperti prediksi sentimen, seperti *pseudo-labelling* dan *post-processing*. Hasil evaluasi F1 dan akurasi ditunjukkan pada tabel berikut

Tabel 4.5. Performa Evaluasi Prediksi *Hate Speech*

Metode	Akurasi	<i>Weighted F1-score</i>
Model <i>baseline</i>	0.4483	0.4328
Model <i>fine-tuning</i>	0.4805	0.4842
Model hasil <i>post evaluation</i>	0.5624	0.5576

4.1.5. Analisis Hasil Prediksi

Tabel 4.6. Contoh Hasil Prediksi Sentimen dan *Hate Speech*

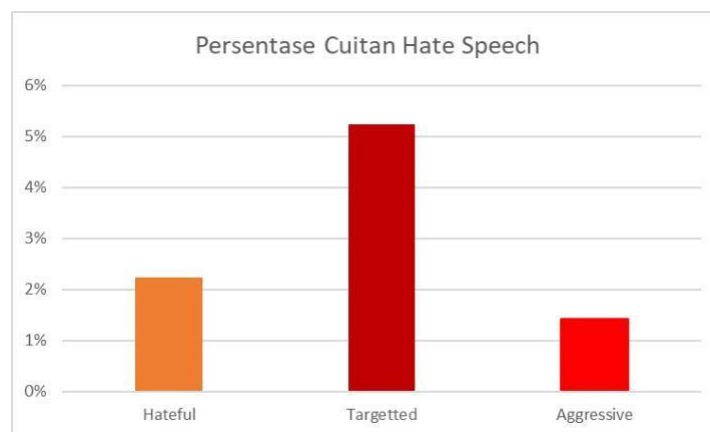
Cuitan	Sentimen	Konteks
“...Heker bjorka itu ternyata orang cirebon? Kok kayak masi bocah sih...”	Negatif	Bjorka
“.. dia dianggap pahlawan hanya karena bobol data-data milik pemerintah.”	Positif	Pemerintah
“... Kalo sistem yang di retas, sudah bisa jadi amburadul semua.. Bjorka bagiku ada sisi White Hacker, Black Hacker ... “	Netral	Bjorka
“...bjorka susah2 buat ngeretas data penduduk,eh pemerintah kita segampang itu nyebarin data nya...”	Negatif	Pemerintah

Tabel 4.7. Contoh Kesalahan Hasil Prediksi Sentimen dan *Hate Speech*

Cuitan	Prediksi Sentimen	Sentimen Asli	Prediksi <i>Hate Speech</i>	<i>Hate Speech</i> Asli	Konteks
“Bjorka emang paling demen indonesia ini bikin isu2 baru. Xixixi”	Positif	Negatif	<i>Hateful</i>	<i>None</i>	Bjorka
“... bengi paling bjorka nyoto ng samping polines.”	Negatif	Netral	<i>Hateful</i>	<i>None</i>	Bjorka
“kaget remaja ini di tuduh sebagai Bjorka “	Negatif	Netral	<i>Hateful</i>	<i>None</i>	Bjorka
“...Ayo, dimulai dengan mengganti orang orang tidak kompeten di @kemmkominfo...”	Positif	Negatif	<i>Hateful</i>	<i>Hateful dan Targetted</i>	Pemerintah

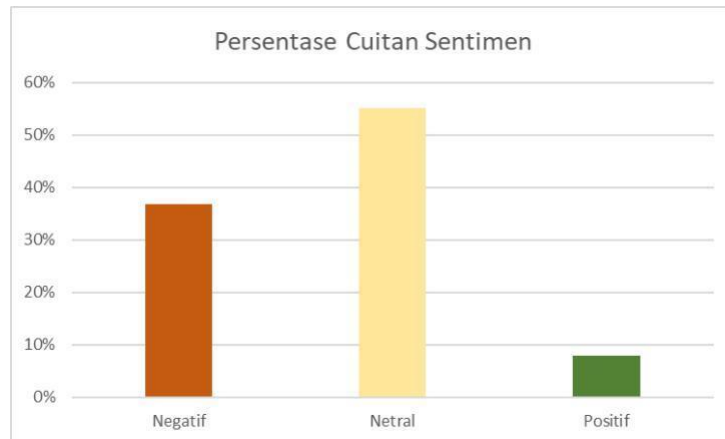
4.2. Pembahasan

Hasil dari analisis sentimen dan *hate speech* menunjukkan data yang tim peneliti peroleh, cuitan yang memiliki konteks *hateful* baik ke pemerintah maupun ke Bjorka memiliki persentase sebesar 2,1%, dengan konteks *targetted* mencapai 5%, dan dengan konteks *aggressive* sebesar 1,4%.

**Gambar 4.3.** Persentase Cuitan *Hate Speech*

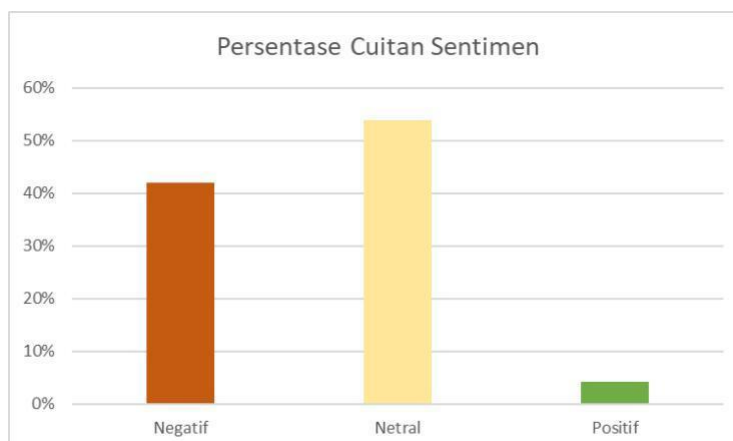
Berdasarkan hasil data dari grafik di bawah, sebanyak lebih dari 50% cuitan masyarakat yang menuju ke pemerintah bersifat netral. Sebanyak lebih dari 40% cuitan

masyarakat memberikan sentimen negatif ke pemerintah. Hanya kurang dari 10% cuitan masyarakat yang memberikan sentimen positif ke pemerintah. Dari hasil tersebut, dapat dilihat bahwa cuitan masyarakat lebih condong untuk memberikan sentimen yang netral dan diikuti dengan sentimen negatif. Hanya sedikit cuitan masyarakat yang memberikan sentimen positif dan mendukung pemerintah.



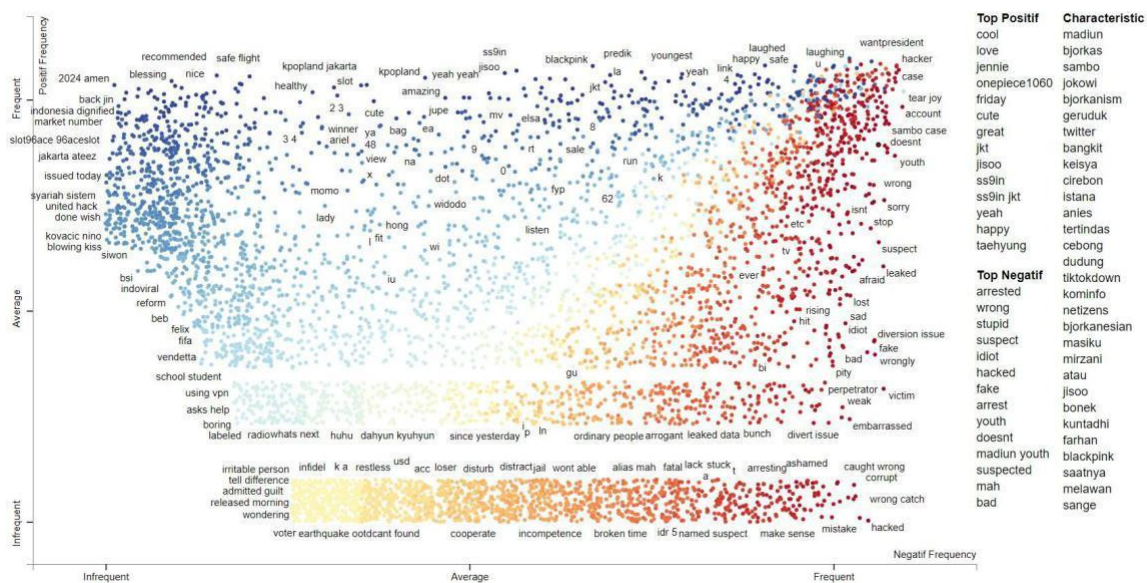
Gambar 4.4. Persentase Sentimen dari Cuitan terhadap Pemerintah

Persentase hasil sentimen terhadap Bjorka juga tidak memiliki perbedaan jumlah yang signifikan dari sentimen terhadap pemerintahan. Sentimen netral masih mendominasi dengan jumlah sebanyak 53%. Diikuti oleh sentimen negatif sebanyak 42% dan sentimen positif sebanyak 5%. Jika dibandingkan dengan persentase sentimen negatif terhadap pemerintah, persentase sentimen negatif terhadap Bjorka ternyata lebih banyak. Dengan demikian, lebih banyak masyarakat yang tidak menyukai Bjorka dibandingkan dengan pemerintah.



Gambar 4.5. Persentase Sentimen dari Cuitan terhadap Bjorka

Dengan bantuan *library* ScatterText, *keywords* dari hasil analisis sentimen dapat diekstrak pada kata yang sering muncul dari cuitan-cuitan yang ada terhadap Bjorka, seperti ditampilkan pada grafik di bawah. Dalam grafik tersebut, kata yang memiliki warna semakin merah memaknai kata yang memiliki sentimen semakin negatif. Kata yang memiliki warna semakin biru menandakan bahwa kata tersebut memiliki sentimen yang semakin positif. Dari hasil analisis sentimen tersebut terdapat sebanyak 3.978 kata memiliki sentimen positif dari 41.732 kata. Sebaliknya, 11.859 kata memiliki sentimen negatif dari 146.306 kata. Dari grafik tersebut pula, dapat dilihat peringkat teratas kata-kata sentimen positif atau negatif. Akan tetapi, kata-kata yang merupakan peringkat teratas sentimen positif banyak yang tidak relevan dengan konteks. Berdasarkan hasil analisis sentimen berdasarkan kata dan grafik dibawah terbukti bahwa sentimen negatif terhadap Bjorka lebih banyak daripada sentimen positif.



Gambar 4.6. Sentimen terhadap Bjorka Berdasarkan Kata

BAB 5

PENUTUP

5.1. Kesimpulan

Implementasi *deep learning* dengan memanfaatkan *pretrained model* BERTweet dan RoBERTa dapat menghasilkan analisis sentimen dengan baik. Dari eksperimen yang telah dilakukan, disimpulkan bahwa metode *fine-tuning* serta *post-evaluation* dapat meningkatkan performa hasil prediksi dari analisis sentimen yang dilakukan. Hasil prediksi meningkat sebesar kurang lebih 10%. Hal ini menunjukkan bahwa *fine-tuning* serta *post-evaluation* merupakan metode yang sesuai untuk analisis sentimen dengan konteks yang sedang tim peneliti analisis.

Setelah dilakukan analisis sentimen, sebagian besar cuitan masyarakat memberikan sentimen yang netral, baik ke pemerintah maupun Bjorka. Dan juga sentimen berdasarkan kata juga dihasilkan dari analisis sentimen tersebut. Dapat disimpulkan bahwa persentase cuitan yang bersifat negatif terhadap Bjorka lebih tinggi dibandingkan terhadap pemerintah. Akan tetapi, persentase cuitan yang bersifat negatif baik terhadap Bjorka maupun Pemerintah hampir sama. Tim peneliti menyimpulkan bahwa sebagian besar pengguna Twitter yang membuat cuitan tentang penyerangan siber terhadap pemerintah tidak menyukai kedua belah pihak.

Selain itu, berdasarkan analisis lanjutan yang dilakukan, hasil menunjukkan bahwa terdapat sekitar 9% cuitan yang tergolong *hate speech*. Dari 9% *hate speech* tersebut, komposisi paling besar dengan bobot 5% terdapat pada cuitan yang bersifat ditujukan kepada sebuah pihak. Selanjutnya, komposisi yang bersifat *aggressive* yang memiliki persentase 1.5%, serta *hateful* dengan persentase 2.5%.

5.2. Saran

Analisis sentimen sangat bergantung dengan akurasi model yang digunakan. Oleh karena itu, model yang dipilih oleh peneliti harus sesuai dengan konteks atau tema dari data yang akan dianalisis. Selain itu, karena semakin kompleksnya bahasa Indonesia yang digunakan, seperti bahasa daerah, bahasa gaul, dan bahasa *slang*, belum ada model publik yang dapat menangani keseluruhan kasus tersebut. Di sisi lain, tim peneliti tidak dapat menjalankan program untuk menganalisis data dengan maksimal

karena kekurangan sumber daya *hardware* yang digunakan baik secara lokal maupun *cloud* dalam mengolah data.

DAFTAR PUSTAKA

- Thapa, B. (2022). Sentiment analysis of cybersecurity content on twitter and reddit. Ithaca: Cornell University Library, arXiv.org. Retrieved from <https://www.proquest.com/working-papers/sentiment-analysis-cybersecurity-content-on/docview/2655916995/se-2>
- Samuels, A., & Mcgonical, J. (2020). Sentiment analysis on social media content. Ithaca: Cornell University Library, arXiv.org. Retrieved from <https://www.proquest.com/working-papers/sentiment-analysis-on-social-media-content/docview/2420906348/se-2>
- Sayyida Tabinda Kokab, Sohail Asghar, Shehneela Naz, Transformer-based deep learning models for the sentiment analysis of social media data, Array, Volume 14, 2022, 100157, ISSN 2590-0056, <https://doi.org/10.1016/j.array.2022.100157>.
- Sai Kumar, T. S., Arunaggiri Pandian, K., Thabasum Aara, S., & Nagendra Pandian, K. (2021). A reliable technique for sentiment analysis on tweets via machine learning and Bert. 2021 Asian Conference on Innovation in Technology (ASIANCON). <https://doi.org/10.1109/asiancon51346.2021.9545013>
- Wang, S.-C. (2003). Artificial Neural Network. Dalam S.-C. Wang (Ed.), Interdisciplinary Computing in Java Programming (hlm. 81–100). Springer US. https://doi.org/10.1007/978-1-4615-0377-4_5
- Pota, M., Ventura, M., Catelli, R., & Esposito, M. (2020). An Effective BERT-Based Pipeline for Twitter Sentiment Analysis: A Case Study in Italian. Sensors, 21(1), 133. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/s21010133>
- Ashima, Y., & Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures: A review. The Artificial Intelligence Review, 53(6), 4335-4385. doi:<https://doi.org/10.1007/s10462-019-09794-5>

- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805 [cs]. <http://arxiv.org/abs/1810.04805>
- Batra, H., Narinder, S. P., Sonbhadra, S. K., & Agarwal, S. (2021). BERT-based sentiment analysis: A software engineering perspective. Ithaca: Cornell University Library, arXiv.org. doi:https://doi.org/10.1007/978-3-030-86472-9_72
- Chiorrini, A., Diamantini, C., Mircoli, A., & Potena, D. (2021). Emotion and sentiment analysis of tweets using BERT. EDBT/ICDT Workshops. Retrieved from <https://www.semanticscholar.org/paper/Emotion-and-sentiment-analysis-of-tweets-using-BERT-Chiorrini-Diamantini/30159e346d89c1ce21275e2814e3c3f284991472>
- Qiu, Z., Cho, E., Ma, X., & Campbell, W. (2019). Graph-Based Semi-Supervised Learning for Natural Language Understanding. Proceedings of the Thirteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-13), 151–158. <https://doi.org/10.18653/v1/D19-5318>
- Alrehili, A. (2019). Automatic Hate Speech Detection on Social Media: A Brief Survey. IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA), 2019, pp. 1-6, doi: 10.1109/AICCSA47632.2019.9035228.
- Indonesia, CNN (2022, July). RI Dihantam 700 Juta Serangan Siber di 2022, Modus Pemerasan Dominan. Teknologi. <https://www.cnnindonesia.com/teknologi/20220701164212-192-816150/ri-dihanta-m-700-juta-serangan-siber-di-2022-modus-pemerasan-dominan>

Databooks (2022, July). Keamanan Siber Indonesia Peringkat KE-6 di asia tenggara:
Databoks. Pusat Data Ekonomi dan Bisnis Indonesia.
<https://databoks.katadata.co.id/datapublish/2022/03/07/keamanan-siber-indonesia-peringkat-ke-6-di-asia-tenggara>

Dewi, I. R. (2022, September). Bikin Heboh RI, data Apa Saja Yang dibocorkan hacker Bjorka? CNBC Indonesia.
<https://www.cnbcindonesia.com/tech/20220914095826-37-371939/bikin-heboh-ri-data-apa-saja-yang-dibocorkan-hacker-bjorka>