

Analisis Sentimen *Twitter* Pada Cryptocurrency Menggunakan *Bidirectional Encoder Representations From Transformers* Dan *Region-Based Convolutional Neural Network*

Achmad Wafa Prasetyo¹, Ricky Eka Putra²

^{1,2} Jurusan Teknik Informatika Fakultas Teknik Universitas Negeri Surabaya

¹achmad.17051204069@mhs.unesa.ac.id

²rickyeka@unesa.ac.id

Abstrak—Saat ini penyebaran berita dengan mudah dapat di akses kapanpun & dimanapun berada dengan memanfaatkan media sosial, salah satunya twitter. Salah satu topik yang seringkali menjadi *trending* di media sosial terutama twitter yaitu cryptocurrency. Tingginya minat investasi pada generasi milenial menyebabkan cryptocurrency menjadi topik yang hangat dibicarakan di berbagai media sosial. Pengaruh konten pada media sosial terhadap masyarakat memiliki signifikansi dalam kehidupan bersosialisasi antar individu atau kelompok, namun masyarakat dianggap belum mampu memilah informasi secara baik dan benar.

Oleh karena itu diperlukan sebuah sistem untuk mengetahui kebenaran sebuah informasi yang terdapat di media sosial. Analisa terhadap berbagai macam pendapat seseorang dalam memberikan komentarnya merupakan pengertian dari analisis sentimen. Untuk mengklasifikasikan bentuk sentimen yang diberikan pada media sosial khususnya twitter dapat menggunakan text mining, yaitu proses mendapatkan informasi dalam sekumpulan dokumen besar secara otomatis. Tujuannya untuk mengetahui seberapa positif, netral & negatif terhadap hastag tersebut.

Hasil yang diperoleh setelah dilakukan analisis sentimen, review pengguna mayoritas menyatakan netral dengan persentase, sentimen positif 23,9% ; sentimen netral 40,3% dan sentimen negatif 35,8%. Dengan ini dapat ditarik kesimpulan bahwa metode BERT dan RCNN kurang efektif untuk implementasi analisis sentiment.

Kata Kunci - *Twitter, Analisis Sentimen Cryptocurrency, Bidirectional Encoder Representations from Transformer (BERT), Region-based Convolution Neural Network (R-CNN).*

I. PENDAHULUAN

Pesatnya perkembangan teknologi saat ini sangat berpengaruh besar bagi masyarakat. Dengan berkembangnya teknologi dapat memudahkan masyarakat dalam melakukan pekerjaan dan mendapat informasi. Penyebaran berita saat ini telah mengalami banyak kemajuan. Berita dengan mudah dapat diakses kapanpun dan dimanapun berada dengan memanfaatkan mediasosial, salah satunya twitter. Twitter merupakan salah satu mediasosial yang masuk dalam kategori paling sering dipakai oleh masyarakat Indonesia berdasarkan datareportal.com pada artikel yang ditulis oleh . Media sosial tersebut adalah Youtube, Whatsapp, Facebook, Instagram & Twitter. Masing-masing mediasosial tersebut memiliki ciri tersendiri sehingga bisa menarik perhatian penggunanya. You

tube dapat diakses bagi mereka yang lebih menyukai tampilan audio visual, Whatsapp cenderung kearah sebagai media dalam berkomunikasi secara interaktif, Facebook memiliki segmen pasar bagi mereka yang ingin bertransaksi jual beli pada komunitas-komunitas, Instagram menampilkan sisi visualitasnya, dan Twitter menampilkan penggunaanya dengan informasi trending lokal maupun internasional secara tepat guna. Pada twitter juga terdapat fitur Top Trending yang mana merupakan fitur untuk memudahkan pengguna melihat tweet apa yang paling populer saat ini. Selain itu, twitter juga memiliki konten *verified account* yang merupakan lencana akun yang telah terverifikasi oleh twitter dan tidak sembarang orang dapat mempunyai lencana ini. Lencana ini ditandai dengan icon checklist biru di sebelah nama pengguna Twitter [1].

Salah satu topik yang seringkali menjadi *trending* di media sosial terutama twitter yaitu cryptocurrency. Dampak sebagian dari perkembangan teknologi informasi saat ini yaitu berkembangnya jenis instrumen keuangan baru, yaitu *cryptocurrency* yang mana merupakan matauang virtual. Matauang ini dapat dijadikan sebagai alat transaksi elektronik serta pemilik mata uang ini dapat menggunakannya sebagai investasi maupun trading. *Cryptocurrency* dapat digunakan juga dalam transaksi bisnis secara daring tanpa melibatkan pihak ketiga, yaitu bank karena mata uang digital *cryptocurrency* yang digunakan adalah sama dan tanpa adanya perbedaan kurs sehingga transaksi dapat dilakukan seketika dengan cepat meskipun berada dalam lintas negara serta terjamin kerahasiannya [2]. Tingginya minat investasi pada generasi milenial menyebabkan cryptocurrency menjadi topik yang hangat dibicarakan di berbagai media sosial. Pengaruh konten pada media sosial terhadap masyarakat memiliki signifikansi dalam kehidupan bersosialisasi antar individu/kelompok, namun masyarakat dianggap belum bisa memilah informasi dengan baik&benar [3].

Secara umum, penambangan bitcoin adalah proses otomatis untuk menemukan nilai hash tertentu yang memecahkan satu blok data transaksi dan menambahkannya ke rantai blok yang berkembang yang dikenal sebagai blockchain. Penambangan adalah proses yang mengamankan buku besar transaksi yang terdistribusi ini, namun ini dicapai dengan biaya tinggi. Penambang mata uang kripto yang paling sukses menjalankan gudang yang penuh dengan mesin khusus yang memakan daya besar secara terus-menerus. Imbalan untuk memecahkan

masalah matematika yang rumit ini adalah pelepasan beberapa bitcoin baru ke penambang, yang mana menjadikannya usaha yang berpotensi menguntungkan. Penambang Bitcoin memperoleh penghasilan antara US\$7-\$9 untuk memproses setiap transaksi bitcoin [4].

Oleh karena itu dibutuhkan sebuah sistem untuk mengetahui kebenaran sebuah informasi yang terdapat di media sosial. Analisa terhadap berbagai macam pendapat seseorang dalam memberikan komentarnya merupakan pengertian dari analisis sentimen. Untuk mengklasifikasikan bentuk sentimen yang diberikan pada media sosial khususnya twitter dapat menggunakan text mining, yaitu proses mendapatkan informasi dalam sekumpulan dokumen besar secara otomatis. Scrapping data atau dikenal dengan *web scrapping* merupakan proses pengambilan kalimat/bahkan dokumen semiterstruktur dari internet&umumnya berupa halaman web dalam bahasa XHTML/HTML. Scrapping data memungkinkan untuk menganalisis dokumen lalu dipakai lagi. Ini termasuk teknik dalam memperoleh informasi dari web site secara otomatis tanpa harus menyalin secara manual [5].

Analisis sentimen dipakai untuk mengidentifikasi bagaimana suatu sentimen diekspresikan memakai teks serta bagaimana dapat dikategorikan sebagai suatu respon positif/negatif. Selain itu, analisis sentimen menjelaskan bahwa merupakan sebuah alat bantu mencari opini dan termasuk pemrosesan sekumpulan hasil pencarian item yang akan menghasilkan kualitas, fitur, dan sebagainya serta juga dapat mengintegrasikan opini misalnya rendah, sedang, atau tinggi. Sentimen sendiri menurut Kamus Besar Bahasa Indonesia yakni: 1) pendapat yang didasarkan pada perasaan berlebihan terhadap sesuatu, misalnya keputusan yang disertai rasa pribadi akan menghasilkan keputusan yang tidak adil; 2) emosi yang berlebihan; 3) iri hati, dendam, tidak senang; 4) reaksi yang merugikan [6]. Model analisis sentimen baru menggunakan BERT & deep CNN dapat dikembangkan oleh para peneliti. Berdasarkan kamus sentimen yang dimasukkan ke dalam vektor kata beberapa model, jaringan saraf berulang dapat mengekstraksi informasi kontekstual maju dan mundur. Oleh karena itu, LSTM mungkin dapat kurang lebih menekankan kata-kata yang berbeda dalam sebuah teks dengan menambahkan mekanisme perhatian pada keluarannya. Ada juga kemungkinan menggunakan model BERT berkualitas tinggi dan ringan untuk melakukan tugas seperti ini [7].

Algoritma BERT (*Bidirectional Encoder Representations from Transformer*) pertama kali diciptakan oleh Jacob Devlin pada tahun 2018 dan mulai diperkenalkan oleh Google pada tahun 2019 yang mana algoritma ini bertugas terhadap hal yang berkaitan dengan pencarian sesuai konteks. Pada mulanya Google membuat sebuah sistem yang berfungsi dalam perangkungan yang mana sistem tersebut membantu pengguna dalam memperoleh informasi yang paling relevan sesuai kata kunci yang pengguna inputkan. Kemudian dalam perkembangannya, Google terus melakukan pembaruan dan perbaikan untuk meningkatkan kinerja dan kualitas yang dimiliki algoritma tersebut yang mana mereka melakukan perubahan hingga sebanyak 3234 kali dalam kurun waktu satu tahun [8]. Merepresentasikan kata sebagai vektor numerik dengan mengandalkan isinya telah menjadi salah satu metode efektif untuk menganalisis teks dalam pembelajaran mesin, sehingga setiap kata adalah diwakili oleh vektor untuk menentukan maknanya atau untuk mengetahui seberapa dekat

atau jauh kata ini dari kata lainnya. BERT (BiDirectional Encoder Representation Transformer) adalah salah satu metode penyematan. Ini dirancang untuk melatih bentuk kiri dan kanan di semua pelatihan mendalam lapisan. Ini adalah model bahasa mendalam yang digunakan untuk berbagai tugas dalam pemrosesan bahasa alami [9].

R-CNN (*Region-based Convolution Neural Network*) merupakan metode dalam ranah *computer vision* yang berfungsi untuk pendeteksian objek yang berbasis CNN (*Convolution Neural Network*). R-CNN pada mulanya diperkenalkan pada tahun 2015 dengan fungsi sebagai metode pendeteksian objek yang berkombinasi dengan algoritma RPN (*Region Proposal Network*) dan CNN. Kemudian seiring dengan berkembangnya teknologi & berjalannya waktu, metode R-CNN ini terus dikembangkan dan diperbaiki untuk meningkatkan performanya dalam tingkat akurasi. Akan tetapi, algoritma R-CNN memiliki kekurangan yang mana salah satunya ialah adanya bottleneck pada RPN, yaitu ketidakmampuan untuk menyamai kecepatan komputasi pada CNN [10].

Preprocessing bertujuan untuk mempelajari data, mendongkrak kualitas data, dan penambahan data fungsional hasil. Tahapan ini memiliki porsi pekerjaan yang komprehensif & menghabiskan hampir 70% data mining proses. Beberapa tugas preprocessing yaitu pembersihan data, integrasi data & normalisasi data [11].

Berdasarkan penelitian yang dilakukan Joloudari J.H., DKK pada tahun 2023 dengan judul “*BERT-Deep CNN: State-of-the-Art for Sentiment Analysis of COVID-19 Tweets*”. Model analisis sentimen baru menggunakan BERT dan deep CNN dapat dikembangkan oleh para peneliti. Berdasarkan kamus sentimen yang dimasukkan ke dalam vektor kata beberapa model, jaringan saraf berulang dapat mengekstraksi informasi kontekstual maju dan mundur. Oleh karena itu, LSTM mungkin dapat kurang lebih menekankan kata-kata yang berbeda dalam sebuah teks dengan menambahkan mekanisme perhatian pada keluarannya. Ada juga kemungkinan menggunakan model BERT berkualitas tinggi dan ringan untuk melakukan tugas seperti ini [11]. Selain itu pada penelitian yang berjudul “*Sentiment analysis of Indonesian reviews using fine-tuning Indo-BERT and R-CNN*”, bertujuan untuk mengotomatisasi analisis sentimen. Hasil mengujian menunjukkan bahwa IndoBERT-RCNN memiliki hasil lebih baik daripada basis IndoBERT. IndoBERT dengan RCNN diperoleh akurasi 95,16%, presisi 94,05%, recall 92,74% dan skor f1 93,27% [12].

Berdasarkan analisa dan penjelasan di atas, penulis ingin melakukan penelitian dengan judul “*Analisis Sentimen Twitter Menggunakan Bidirectional Encoder Representations From Transformers (BERT) dan Region-Based Convolutional Neural Network (R-CNN)*”. Penelitian ini akan melakukan klasifikasi analisis sentimen terhadap opini pengguna twitter dengan menggunakan dua algoritma, yaitu Bidirectional Encoder Representations from Transformers (BERT) dan Region-Based Convolutional Neural Network (R-CNN).

II. METODOLOGI PENELITIAN

A. Studi Literatur

Metode penelitian merujuk pada serangkaian langkah yang dilakukan oleh seorang peneliti guna mengumpulkan informasi & data, serta mengolah data yang telah diperoleh. Dalam pe

nelitian ini, teknik yang dipakai termasuk studiliteratur & ob servasi. Studiliteratur dilakukan guna mencari informasi terka it masalah penelitian, metode yang telah digunakan dalam penelitian sebelumnya, dan penelitian terkait lainnya.

B. Analisis Dan Gambaran Umum Sistem

Pengembangan sistem yang dikembangkan oleh peneliti merupakan implementasi algoritma BERT dan R-CNN pada analisis sentimen terhadap media sosial Twitter. Sistem yang akan dibuat adalah menggunakan Python.

C. Tempat Dan Waktu Penelitian

1. Tempat Penelitian

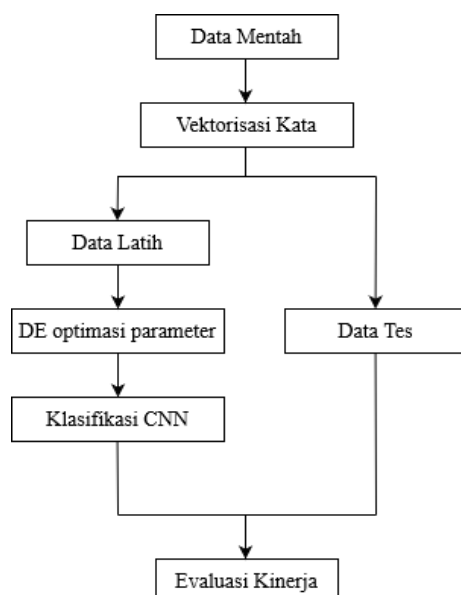
Tempat penelitian merupakan tempat dimana peneliti memperoleh suatu informasi mengenai data-data yang di perlukan saat menjalankan penelitian. Dalam hal ini, pene litian dilaksanakan di berbagai tempat secara fleksibel seperti di kediaman peneliti, *wifi corner*, *café* & kampus UNESA tepatnya di gedung jurusan Teknik Informatika.

2. Waktu Pelaksanaan Penelitian

Waktu pelaksanaan penelitian merupakan rentang wak tu yang dilaksanakan dari dimulainya penelitian hingga berakhirnya penelitian. Meliputi persiapan awal menen tukan topik penelitian, pengajuan judul penelitian, menyu sun proposal penelitian, seminar proposal, revisi proposal, pelaksanaan penelitian, sampai pada pelaksanaan akhir penelitian. Dalam hal ini, waktu pelaksanaan pada peneli tian ini dilakukan selama tahun 2023 hingga tahun 2024.

D. Flowchart

Pada bagian ini akan dijelaskan secara rinci model klasifikasi teks berdasarkan CNN yang digunakan dalam penelitian ini. Bagian alur model pada gambar 3 termasuk tiga modul: teks representasi, pelatihan klasifikasi dan evaluasi kinerja.



Gbr 1. Alur Model Pengembangan Sistem.

Gbr 1. yaitu alur model dari pengembangan sistem yang dikembangkan oleh peneliti. Tahapan pertama adalah raw text data atau teks data mentah yaitu data yang sudah dikumpulkan

sesuai dengan topik tentang cryptocurrency. Tahapan kedua adalah Word vectorization, suatu tahapan metode embedding word yang berfungsi mempresentasikan kata menjadi sebuah vektor. Dari hasil word vectorization nantinya akan mengha silkan data yang akan diproses pada training set dan test set. Pada tahapan training set, data akan diproses sistem memakai algoritma klasifikasi dihasilkan classification rules/aturan kla sifikasi yang disebut dengan classifier (pengklasifikasi). Selan jutnya DE parameter optimization, Differential Evolution adalah algoritma pencarian langsung paralel yang diusulkan oleh Price pada tahun 1996, yang memiliki kemampuan optimasi global yang kuat. Ini mencakup empat langkah: inisialisasi populasi, mutasi, crossover, dan seleksi. Setelah dilakukan parameter, tahapan selanjutnya adalah memasukkan data pada model BERT dan R-CNN yang sudah disusun. Lalu akan dilakukan serangkaian tes dan selanjutnya tahapan terakhir adalah evaluasi performa sistem yang sudah dibuat.

E. Parameter Optimization

Struktur jaringan syaraf tiruan dan parameter pelatihnnya memiliki pengaruh yang besar terhadap kinerja jaringan. Parameternya meliputi ukuran karnel konvolusi, jumlah neuron di lapisan yang terhubung penuh, kecepatan pembelajaran, dll. Maka dari itu, pada penelitian ini rancangan optimasi parameter menggunakan Differential Evolution (DE) untuk mengoptimalkan parameter jaringan syaraf konvolusi.

F. Dataset

Dataset pada penelitian ini diambil dari Twitter, dengan topik tentag cryptocurrency (Bitcoin&Ethereum). Jumlah data set yang dipakai ada 7078 data, dari data set tersebut nantinya akan dibagi menjadi datatraining, datatesting & datavalidation.

G. Presentase Pembagian Data

Dari dataset yang sudah dikumpulkan akan dibagi menjadi 3 yaitu: datatraining, datatesting & datavalidation adalah menggunakan persentase perbandingan 80%, 10%, dan 10%.

H. Rancangan Dan Teknik Pengujian

Pada penelitian ini set data dibagi menjadi datatraining, da ta testing & datavalidation. Lalu akan dilakukan evaluasi kiner ja model berdasarkan matriks kesalahan guna menentukan ke akuratan model. Untuk itu peneliti memakai k-fold cross vali dation sebagai teknik pengujian, dimana teknik ini berfungsi untuk menilai bagaimana sebuah classifier bekerja. Selain itu proses ini berfungsi guna menaikkan tingkat akurasi dari hasil klasifikasi. K-fold yaitu kumpulan data yang diberikan lalu dibagi menjadi sejumlah K-fold dimana setiap fold dipakai sebagai set pengujian beberapa titik. Misalkan validasi silang 5-fold ($K=5$). Kumpulan data dipecah menjadi 5 fold. Pada iterasi pertama, fold pertama dipakai untuk menguji model & sisanya dipakai untuk melatih model. Pada iterasi kedua, fold kedua dipakai untuk set pengujian dan sisanya sebagai set pelatihan. Proses ini akan diulang secara terus menerus setiap fold dari 5 fold yang telah dipakai.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Proses pengumpulan data respon&opini warga Indonesia terhadap cryptocurrency dilakukan dengan memanfaatkan web

Pengumpulan data menggunakan bantuan Twitter scraping tool “twint” yang memungkinkan untuk mengambil komentar dari Twitter.

Gbr 2. Pengumpulan Data Pada Twitter Dengan Twint

```
tweets_data = pd.read_csv('tweet_data.csv')
tweets = tweets_data[['id', 'username', 'created_at', 'tweet', 'replies_count', 'retweets_count', 'likes_count']]
tweets
```

	id	username	created_at	tweet	replies_count	retweets_count	likes_count
0	155735769042132225	tousnesol	2022-08-10 13:26:22 UTC	\$DOGE # 2H!!! Volume candle bull masih kencang...	1	0	3
1	15573575322162945153	kriptokusu	2022-08-10 13:25:42 UTC	BTC takip Bitcoin / TetherUS BINANCE BTCUSDOT ...	0	0	0
2	15573574940783001157	mishack366669993	2022-08-10 13:25:35 UTC	Har ghar tiranga #FlagRan #HarGharTiranga #B...	0	0	0
3	1557356846638732576	trademasteralp	2022-08-10 13:20:25 UTC	#DAX30 vs #NASDAQ 📈 1 lotra 5.793 USD 103...	0	0	0
4	155735769042132225	tousnesol	2022-08-10 13:26:22 UTC	\$DOGE # 2H!!! Volume candle bull masih kencang...	1	0	4
--	--	--	--	--	--	--	--
1006	155725725840379200	cryptosroom	2022-08-10 06:47:10 UTC	CryptoSrooms x AwaikiSaga 🔥🔴🔴🔴🔴🔴 + Proyek ...	20	24	23
1007	1557257217583978680	bukutug	2022-08-10 06:47:07 UTC	@crypto_inex @Tamadogecoin #Tamadoge Tembus L...	1	0	0
1008	155725655690020354	sultan_arsa	2022-08-10 06:46:05 UTC	@crypto_inex @Tamadogecoin #Tamadoge Tembus L...	1	0	0
1009	1557256570551291904	xsaah	2022-08-10 06:44:38 UTC	@inKansha Alhamdulillah baikk!! kemana sajaaaa	1	0	0
1010	1557256286814652417	i.am_bloks	2022-08-10 06:43:28 UTC	@Crypto_Dive Ohayo gozaimasu	1	0	0

Gbr 3. Proses Baca Data Hasil Scrapping Twitter

Pada gbr 3. diatas yaitu proses baca data hasil scrapping twitter. Hasil data yang terkumpul 7.078 data, lalu data set tweet yang masih mentah akan disimpan dalam bentuk CSV (*Comma Separated Values*) nantinya akan diproses lebih lanjut. Dari data yang diperoleh kemudian akan dibagi menjadi 2 data yaitu 3034 data sebagai data testing dan 4044 data sebagai data training.

B. Pre-Processing

Setelah proses pengumpulan data & data disimpan dalam bentuk CSV maka langkah selanjutnya ialah preprocessing data, dikarenakan data masih data mentah dan belum terstruktur. Tujuan dari preprocessing data ialah untuk menghi langkan dan mengatasi noise data supaya hasil perhitungan lebih optimal. Preprocessing dilakukan pada data training melalui beberapa tahapan secara berurutan. Adapun tahapan dalam preprocessing data yakni: cleaning data, casefolding, tokenizing, filtering, stemming. Berikut ini adalah dataset

	id	username	created_at	tweet	replies_count	retweets_count	likes_count
0	1787026945415440747	Devil	2024-05-05 07:46:25+00:00	@prabhas_varma 2021 kine apesah di 50k profil...	0	0	0
1	1787012182822388548	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:22+00:00	Mengenal nilai abad emas, Schiff menegaskan b...	1	0	0
2	17870121930081583941	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:21+00:00	'Bitcoin bukan lagi emas digital, melainkan ga...	1	0	0
3	1787012187788813328	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:21+00:00	Kemudian, para pendukung Bitcoin mencoba untu...	1	0	0
4	1787012184574489801	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:20+00:00	Ditup dan https://t.co/8KatyQDQ7T Peter Sch...	1	0	0
--	--	--	--	--	--	--	--
163	1787012178482081213	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:18+00:00	Debat Epik Soal Bitcoin Vs Emas, Mana yang Leb...	1	0	0
164	1787011955650299377	Mr. MIM 🍁	2024-05-05 06:50:05+00:00	#Bitcoin/ USD, 1M 📈 https://t.co/kcnv68M45c	0	0	0
165	1787009993545494705	CryptoBintang.com 🍀	2024-05-05 06:42:31+00:00	@kmswed_Bang, coba lalf disini bang @Covelon...	0	0	0
166	1787008621183603935	Let's Go Up Bitcoin 🍀	2024-05-05 06:38:22+00:00	@adnanradisni Waduh, pengguna mangabat nyimak	0	0	0
167	1787008932855274012	Ada Apa Dengan Bitcoin	2024-05-05 06:30:42+00:00	@mglangianur Kan ga semua soprenreus bikin	0	0	0

Gbr 4. Dataset Sebelum Pre-processing

Pada gbr 4. diatas merupakan dataset sebelum pre-processing.

1. Cleaning

Pada tahap ini semua karakter didalam tweet yang bukan alfabet dihapus sehingga bisa mengurangi karakter yang tidak dikehendaki & tidak memiliki arti dalam analisis sentimen. Contoh karakternya seperti angka, emoji, link, hastag dan mentions (sebutan). Selain itu juga menghapus semua tanda baca, mengganti enter menjadi spasi dan menghapus spasi di depan dan dibelakang teks agar mengembalikan teks utuh untuk analisa sentiment.

```
def cleaningText(text):
    text = re.sub(r'@[A-Za-z0-9]+', '', text) # remove mentions
    text = re.sub(r'#[A-Za-z0-9]+', '', text) # remove hashtag
    text = re.sub(r'RT[\s]', '', text) # remove RT
    text = re.sub(r'http\S+', '', text) # remove link
    text = re.sub(r'[0-9]+', '', text) # remove numbers

    text = text.replace('\n', ' ') # replace new line into space
    text = text.translate(str.maketrans('', '', string.punctuation)) # remove all punctuations
    text = text.strip(' ') # remove characters space from both left and right text
    return text
```

Gbr 5. Cleaning Text

Pada gbr 5. diatas merupakan dataset sebelum dilakukan cleaning text.

	id	username	created_at	tweet	replies_count	retweets_count	likes_count	text_clean
0	178701265045415440747	Devil	2024-05-05 07:49:25+00:00	@prabhas_vamaa 2021 lone apesaa oh 50k profit tho bayata padde entha sep bitcoin midaa run avudi bokke andulo	0	0	0	vamaa lone apesaa oh k profit tho bayata padde entha sep bitcoin midaa run avudi bokke andulo
1	1787012192823288949	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:22+00:00	Mengenai nilai abadi emas, Schiff menegaskan bahwa emas mempertahankan sifat intrinsiknya seiring berjalannya waktu, berfungsi sebagai penyimpan nilai sejati. #viniblockvestglobal #bitcoin #news #crypto	1	0	0	Mengenai nilai abadi emas Schiff menegaskan bahwa emas mempertahankan sifat intrinsiknya seiring berjalannya waktu berfungsi sebagai penyimpan nilai sejati
2	1787012190081593941	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:21+00:00	"Bitcoin bukan lagi emas digital, melainkan gambar hamburger adalah makanan digital," kata Schiff. #vinDia menekankan kegunaan nyata emas dalam industri seperti perhiasan dan elektronik, membandingkannya dengan Bitcoin, yang menurutnya kurang memiliki kegunaan dan kegunaan praktis	1	0	0	Bitcoin bukan lagi emas digital melainkan gambar hamburger adalah makanan digital," kata Schiff. #vinDia menekankan kegunaan nyata emas dalam industri seperti perhiasan dan elektronik, membandingkannya dengan Bitcoin, yang menurutnya kurang memiliki kegunaan dan kegunaan praktis
3	17870121673938813328	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:21+00:00	Kemudian, para pendukung Bitcoin mencoba untuk memposiskannya kembali aset tersebut sebagai emas versi digital. Namun, dalam pandangannya, nilai intrinsik emas masih jauh dari nilai intrinsik emas yang dilihat dari sifat fisiknya. #viniblockvestglobal #bitcoin #news #crypto	1	0	0	Kemudian para pendukung Bitcoin mencoba untuk memposiskannya kembali aset tersebut sebagai emas versi digital. Namun dalam pandangannya nilai intrinsik emas masih jauh dari nilai intrinsik emas yang dilihat dari sifat fisiknya
4	1787012184574498901	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:20+00:00	Dikutip dari https://t.co/BXeqQOQPTI, Peter Schiff, seorang kritikus Bitcoin membuka perdebatan dengan menyatakan bahwa Bitcoin awalnya dibuat sebagai mata uang digital. Namun rencana tersebut gagal	1	0	0	Dikutip dari Schiff seorang kritikus Bitcoin membuka perdebatan dengan menyatakan bahwa Bitcoin awalnya dibuat sebagai mata uang digital. Namun rencana tersebut gagal

Gbr 6. Dataset Setelah Proses Cleaning

Pada gbr 6. diatas ini yaitu dataset yang sudah dilakukan proses cleaning.

2. Case Folding

Tahapan yang bertujuan untuk mengubah semua karakter atau bentuk huruf di teks menjadi huruf kecil.

```
def casefoldingText(text): # Converting all the characters in a text into lower case
    text = text.lower()
    return text
```

Gbr 7. Mengubah Huruf Menjadi Huruf Kecil

Pada gbr 7. diatas merupakan gambar source code untuk melakukan perubahan huruf menjadi huruf kecil semua.

	id	username	created_at	tweet	replies_count	retweets_count	likes_count	text_clean
0	178701265045415440747	Devil	2024-05-05 07:49:25+00:00	@prabhas_vamaa 2021 lone apesaa oh 50k profit tho bayata padde entha sep bitcoin midaa run avudi bokke andulo	0	0	0	vamaa lone apesaa oh k profit tho bayata padde entha sep bitcoin midaa run avudi bokke andulo
1	1787012192823288949	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:22+00:00	Mengenai nilai abadi emas, Schiff menegaskan bahwa emas mempertahankan sifat intrinsiknya seiring berjalannya waktu, berfungsi sebagai penyimpan nilai sejati. #viniblockvestglobal #bitcoin #news #crypto	1	0	0	mengenai nilai abadi emas schiff menegaskan bahwa emas mempertahankan sifat intrinsiknya seiring berjalannya waktu berfungsi sebagai penyimpan nilai sejati
2	1787012190081593941	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:21+00:00	"Bitcoin bukan lagi emas digital, melainkan gambar hamburger adalah makanan digital," kata Schiff. #vinDia menekankan kegunaan nyata emas dalam industri seperti perhiasan dan elektronik, membandingkannya dengan Bitcoin, yang menurutnya kurang memiliki kegunaan dan kegunaan praktis	1	0	0	bitcoin bukan lagi emas digital melainkan gambar hamburger adalah makanan digital" kata schiff dia menekankan kegunaan nyata emas dalam industri seperti perhiasan dan elektronik membandingkannya dengan bitcoin yang menurutnya kurang memiliki kegunaan dan kegunaan praktis
3	17870121673938813328	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:21+00:00	Kemudian, para pendukung Bitcoin mencoba untuk memposiskannya kembali aset tersebut sebagai emas versi digital. Namun, dalam pandangannya, nilai intrinsik emas masih jauh dari nilai intrinsik emas yang dilihat dari sifat fisiknya. #viniblockvestglobal #bitcoin #news #crypto	1	0	0	kemudian para pendukung bitcoin mencoba untuk memposiskannya kembali aset tersebut sebagai emas versi digital namun dalam pandangannya nilai intrinsik emas masih jauh dari nilai intrinsik emas yang dilihat dari sifat fisiknya
4	1787012184574498901	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:20+00:00	Dikutip dari https://t.co/BXeqQOQPTI, Peter Schiff, seorang kritikus Bitcoin membuka perdebatan dengan menyatakan bahwa Bitcoin awalnya dibuat sebagai mata uang digital. Namun rencana tersebut gagal	1	0	0	dikutip dari schiff seorang kritikus bitcoin membuka perdebatan dengan menyatakan bahwa bitcoin awalnya dibuat sebagai mata uang digital namun rencana tersebut gagal

Gbr 8. Dataset Setelah Proses Casefolding

Pada gbr 8. diatas merupakan hasil dataset setelah melakukan proses casefolding.

3. Tokenizing

Ini termasuk sebuah tahap yang mencakup proses se derhana dimana urutan teks dipecah menjadi bagian-bagian yang lebih kecil atau token. Bagian tersebut berupa kata sebelum dianalisa lebih lanjut.

```
def tokenizingText(text): # Tokenizing or splitting a string, text into a list of tokens
    text = word_tokenize(text)
    return text
```

Gbr 9. Tokenizing Text

Pada gbr 9. merupakan gambar source code untuk tokenizing text.

	id	username	created_at	replies_count	retweets_count	likes_count	text_clean	text_preprocessed
0	17870121628045415440747	Dewi	2024-05-05 (07:48:25+00:00)	0	0	0	varmas lone apesaa th k profit tho bayata pada entha sep bitcoin midaa run audu bokke andulo	(varmas, lone, apesaa, th, k, profit, tho, bayata, pada, entha, sep, bitcoin, midaa, run, audu, bokke, andulo)
1	1787012162823288949	#BLOKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 (06:51:22+00:00)	1	0	0	mengena nilai abadi emas schiff menapakai bahwa emas mempertahankan sifat intrinsiknya seiring berjalannya waktu berfungsi sebagai penyimpan nilai sejati	(mengena, nilai, abadi, emas, schiff, menapakai, bahwa, emas, mempertahankan, sifat, intrinsiknya, seiring, berjalannya, waktu, berfungsi, sebagai, penyimpan, nilai, sejati)
2	17870121630081583941	#BLOKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 (06:51:21+00:00)	1	0	0	bitcoin bukan lagi emas digital melainkan gambar hamburger adalah makanan digital kata schiff dia menekankan kegunaan nyata emas dalam industri seperti perhiasan dan elektronik membandingkannya dengan bitcoin yang menurutnya kurang memiliki kegunaan dan kegunaan praktis	(bitcoin, bukan, lagi, emas, digital, melainkan, gambar, hamburger, adalah, makanan, digital, "kata, schiff, dia, menekankan, kegunaan, nyata, emas, dalam, industri, seperti, perhiasan, dan, elektronik, membandingkannya, dengan, bitcoin, yang, menurutnya, kurang, memiliki, kegunaan, dan, kegunaan, praktis)
3	1787012167388813328	#BLOKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 (06:51:21+00:00)	1	0	0	kemudian para pendukung bitcoin mencoba untuk memposiskannya kembali aset tersebut sebagai emas versi digital namun dalam pandangannya nilai intrinsik emas masih jauh dari nilai intrinsik emas yang dilihat dari sifat fisiknya	(kemudian, para, pendukung, bitcoin, mencoba, untuk, memposiskannya, kembali, aset, tersebut, sebagai, emas, versi, digital, namun, dalam, pandangannya, nilai, intrinsik, emas, masih, jauh, dari, nilai, intrinsik, emas, yang, dilihat, dari, sifat, fisiknya)
4	1787012164574489801	#BLOKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 (06:51:20+00:00)	1	0	0	dikup dari schiff seorang kritikus bitcoin membuka perdebatan dengan mengatakan bahwa bitcoin awalnya dibuat sebagai mata uang digital namun rencana tersebut gagal berfungsi dengan baik dan efektif karena sifatnya yang lambat dan mahal	(dikup, dari, schiff, seorang, kritikus, bitcoin, membuka, perdebatan, dengan, mengatakan, bahwa, bitcoin, awalnya, dibuat, sebagai, mata, uang, digital, namun, rencana, tersebut, gagal, berfungsi, dengan, baik, dan, efektif, karena, sifatnya, yang, lambat, dan, mahal)

Gbr 10. Dataset Setelah Proses Tokenizing

Pada gbr 10. diatas merupakan hasil dataset setelah melakukan proses tokenizing text.

4. Filtering

Ini termasuk tahapan untuk menghapus kata yang tidak bermemiliki/stopword. Sedangkan stopwords ialah kata-kata yang bukan ciri/unit dari suatu dokumen. Stop word bahasa Indonesia dibuat & disimpan dalam format teks (.txt). Filter stopwords ini digunakan untuk mengurangi jumlah kata yang dianalisa yang nantinya akan mempengaruhi kecepatan dan performa pada NLP.

```
def filteringText(text): # Remove stopwords in a text
    listStopwords = set(stopwords.words('indonesian'))
    filtered = []
    for txt in text:
        if txt not in listStopwords:
            filtered.append(txt)
    text = filtered
    return text
```

Gbr 11. Filtering Stopword

Pada gbr 11. diatas merupakan source code dari filtering stopwords.

	id	username	created_at	replies_count	retweets_count	likes_count	text_clean	text_preprocessed
0	17870121628045415440747	Dewi	2024-05-05 (07:48:25+00:00)	0	0	0	varmas lone apesaa th k profit tho bayata pada entha sep bitcoin midaa run audu bokke andulo	(varmas, lone, apesaa, th, k, profit, tho, bayata, pada, entha, sep, bitcoin, midaa, run, audu, bokke, andulo)
1	1787012162823288949	#BLOKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 (06:51:22+00:00)	1	0	0	mengena nilai abadi emas schiff menegakan bahwa emas mempertahankan sifat intrinsiknya seiring berjalannya waktu berfungsi sebagai penyimpan nilai sejati	(nilai, abadi, emas, schiff, emas, mempertahankan, sifat, intrinsiknya, seiring, berjalannya, waktu, berfungsi, sebagai, penyimpan, nilai, sejati)
2	17870121630081583941	#BLOKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 (06:51:21+00:00)	1	0	0	bitcoin bukan lagi emas digital melainkan gambar hamburger adalah makanan digital kata schiff dia menekankan kegunaan nyata emas dalam industri seperti perhiasan dan elektronik membandingkannya dengan bitcoin yang menurutnya kurang memiliki kegunaan dan kegunaan praktis	(bitcoin, emas, digital, gambar, hamburger, makanan, digital, "kata, schiff, menekankan, kegunaan, nyata, emas, industri, perhiasan, elektronik, membandingkannya, bitcoin, menurutnya, memiliki, kegunaan, kegunaan, praktis)
3	1787012167388813328	#BLOKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 (06:51:21+00:00)	1	0	0	kemudian para pendukung bitcoin mencoba untuk memposiskannya kembali aset tersebut sebagai emas versi digital namun dalam pandangannya nilai intrinsik emas masih jauh dari nilai intrinsik emas yang dilihat dari sifat fisiknya	(pendukung, bitcoin, mencoba, memposiskannya, aset, emas, versi, digital, pandangannya, nilai, intrinsik, emas, nilai, intrinsik, emas, sifat, fisiknya)
4	1787012164574489801	#BLOKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 (06:51:20+00:00)	1	0	0	dikup dari schiff seorang kritikus bitcoin membuka perdebatan dengan mengatakan bahwa bitcoin awalnya dibuat sebagai mata uang digital namun rencana tersebut gagal berfungsi dengan baik dan efektif karena sifatnya yang lambat dan mahal	(dikup, schiff, kritikus, bitcoin, membuka, perdebatan, bitcoin, mata, uang, digital, rencana, gagal, berfungsi, efektif, sifatnya, lambat, mahal)

Gbr 12. Dataset Setelah Proses Filtering Stopwords

Pada gbr 12. diatas merupakan dataset setelah melakukan proses filtering stopwords.

5. Stemming

Ini merupakan tahapan yang bertujuan untuk mereduksi kata atau mengubah kata yang berimbuhan menjadi kata dasar. Fungsi ini digunakan untuk mencocokkan semua variasi kata untuk memunculkan yang paling relevan.

```
def stemmingText(text): # Reducing a word to its word stem that affixes to suffixes and prefixes or to the roots of words
    factory = StemmerFactory()
    stemmer = factory.create_stemmer()
    text = [stemmer.stem(word) for word in text]
    return text
```

Gbr 13. Stemming Text

Pada gambar 13 diatas merupakan source code dari stemming text.

	id	username	created_at	replies_count	retweets_count	likes_count	text_clean	text_preprocessed
0	17870126045415440747	Devil	2024-05-05 07:46:25+00:00	0	0	0	vamaa lona apesaa oh k profit tho bayata padda entha sep bitcoin midaa run avudi bokke andulo	[vamaa, lona, apesaa, oh, k, profit, tho, bayata, padda, entha, sep, bitcoin, midaa, run, avudi, bokke, andulo]
1	1787012192622388849	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:22+00:00	1	0	0	mengenal nilai abadi emas schiff menegaskan bahwa emas mempertahankan nilai intrinsiknya seiring berjalannya waktu berfungsi sebagai penyimpan nilai sejati	[nilai, abadi, emas, schiff, emas, tahan, nilai, intrinsik, ring, jalan, fungsi, simpan, nilai, sejati]
2	1787012190081568841	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:21+00:00	1	0	0	bitcoin bukan lagi emas digital melainkan gambar hamburger adalah makanan digital kata schiff dia menekankan kegunaan nyata emas dalam industri seperti perhiasan dan elektronik membandingkannya dengan bitcoin yang menurutnya kurang memiliki kegunaan dan kegunaan praktis	[bitcoin, emas, digital, gambar, hamburger, makan, digital, schiff, tekan, guna, nyata, emas, industri, hias, elektronik, banding, bitcoin, turut, milk, guna, guna, praktis]
3	1787012167368813328	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:21+00:00	1	0	0	kemudian para pendukung bitcoin mencoba untuk memposiskannya kembali aset tersebut sebagai emas versi digital namun dalam pandangannya nilai intrinsik emas masih jauh dari nilai intrinsik emas yang dilihat dari sifat fisiknya	[dukung, bitcoin, coba, posisi, aset, emas, versi, digital, pandang, nilai, intrinsik, emas, nilai, intrinsik, emas, sifat, fisik]
4	1787012164574488801	#BLOCKCHAIN #INVESTMENT #FESTIVAL	2024-05-05 06:51:20+00:00	1	0	0	dikutip dari schiff seorang kritikus bitcoin membuka perdebatan dengan menyatakan bahwa bitcoin awalnya dibuat sebagai mata uang digital namun rencana tersebut gagal berfungsi dengan baik dan efektif karena sifatnya yang lambat dan mahal	[ikutip, schiff, kritikus, bitcoin, buka, debat, bitcoin, mata, uang, digital, rencana, gagal, fungsi, efektif, lambat, mahal]
		#BLOCKCHAIN	2024-05-05				mengenal apakah bitcoin itu atau emas berfungsi sebagai lindung nilai inflasi yang lebih baik daripada hamparan dolar dan yen	[bitcoin, btc, emas, fungsi, lindung, nilai, inflasi, gabung, neo]

Gbr 14. Dataset setelah proses stemming

Pada gbr 14. diatas merupakan dataset setelah melakukan proses stemming.

C. Metode Analisis Data

Tahapan ini tujuannya adalah untuk melakukan pengklasifikasian tweets/analisis sentimen, pada penelitian ini, analisis data menggunakan metode lexicon-based / berbasis kamus positif dan negative.

Dimana polaritas positif merupakan bobot dari kalimat yang diperoleh melalui penjumlahan n skor polaritas kataa opini positif & polaritas negatif merupakan bobot dari kalimat yang diperoleh melalui penjumlahan n skor polaritas kata opini negatif. Polaritas netral adalah kondisi diluar polaritas positif dan negative.

```
# Loads lexicon positive and negative data
lexicon_positive = dict()
import csv
with open('lexicon_positive.csv', 'r') as csvfile:
    reader = csv.reader(csvfile, delimiter=',')
    for row in reader:
        lexicon_positive[row[0]] = int(row[1])

lexicon_negative = dict()
import csv
with open('lexicon_negative.csv', 'r') as csvfile:
    reader = csv.reader(csvfile, delimiter=',')
    for row in reader:
        lexicon_negative[row[0]] = int(row[1])

# Function to determine sentiment polarity of tweets
def sentiment_analysis_lexicon_indonesia(text):
    #for word in text:
    score = 0
    for word in text:
        if (word in lexicon_positive):
            score = score + lexicon_positive[word]
    for word in text:
        if (word in lexicon_negative):
            score = score + lexicon_negative[word]
    polarity=''
    if (score > 0):
        polarity = 'positive'
    elif (score < 0):
        polarity = 'negative'
    else:
        polarity = 'neutral'
    return score, polarity
```

Gbr 15. Klasifikasi Polaritas Kata Dan Score

Pada gbr 15. merupakan source code dari klasifikasi polaritas kata dan score.

D. Proses Visualisasi Data

Proses visualisasi data memakai wordcloud, di proses ini mengklasifikasikan data positif& negatif. Berikut ini adalah script untuk proses wordcloud.

```
list_words=''
for tweet in tweets['text_preprocessed']:
    for word in tweet:
        list_words += ' '+word

wordcloud = WordCloud(width = 600, height = 400,
background_color = 'black', min_font_size =
10).generate(list_words) fig, ax = plt.subplots(figsize =
(8, 6))

ax.set_title('Word Cloud of Tweets Data', fontsize = 18)
ax.grid(False)
ax.imshow((wordcloud))
fig.tight_layout(pad=0)
ax.axis('off')
plt.show()
```

Dari script diatas, kemudian akan dilanjutkan ke tahap wordcloud atau mencari kata yang positif dan negative.

1. Wordcloud Positif

Untuk tahap wordcloud positif menggunakan script berikut.

```

sentiment_words =
tweets['text_preprocessed'].apply(words_with_sentimen
t)
sentiment_words = list(zip(*sentiment_words))
positive_words = sentiment_words[0]
negative_words = sentiment_words[1]
neutral_words = sentiment_words[2]

fig, ax = plt.subplots(1, 2, figsize = (12, 10))
list_words_postive=''
for row_word in positive_words:
    for word in row_word:
        list_words_postive += '+'+(word)
wordcloud_positive = WordCloud (width = 800, height
= 600, background_color = 'black', colormap =
'Greens', min_font_size =
10).generate(list_words_postive)

ax[0].set_title('Word Cloud of Positive Words on Tweets
Data \n (based on Indonesia Sentiment Lexicon)',
fontsize = 14)
ax[0].grid(False)
ax[0].imshow((wordcloud_positive))
fig.tight_layout(pad=0)
ax[0].axis('off')

```

Dari script diatas merupakan wordcloud positif. Hasil dari wordcloud positif dapat dilihat pada gbr 16.



Gbr 16. Hasil Wordcloud Positif

Gbr 16. merupakan wordcloud positif dari dataset. Pada wordcloud diatas terdapat beberapa kata menonjol seperti “pas”, “harga” “naik”, “jual”, “untung”, “nilai”, “moga”, “ketemu”. Kalimat tersebut menggambarkan respon positif dari pengguna twitter dengan topik cryptocurrency, bitcoin, dan ethereum.

2. Wordcloud Negatif

Setelah mencari wordcloud positif selanjutnya wordcloud negatif. Berikut script untuk melakukan -wordcloud negatif.

```
list_words_negative=""

for row_word in negative_words:
    for word in row_word:
        list_words_negative += '(word)

wordcloud_negative = WordCloud (width = 800, height =
600, background_color = 'black', colormap
'Reds' ,min_font_size = 10).generate(list_words_negative)

ax[1].set_title('Word Cloud of Negative Words on Tweets
Data \n (based on Indonesia Sentiment Lexicon)', fontsize
= 14) ax[1].grid(False)

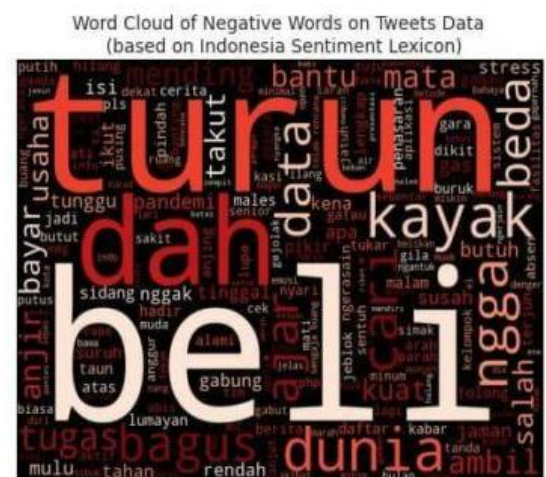
ax[1].imshow((wordcloud_negative))

fig.tight_layout (pad=0)

ax[1].axis('off')

plt.show()
```

Dari script diatas merupakan wordcloud negatif. Hasil dari wordcloud negatif dapat dilihat pada gambar 4.17.



Gbr 17. Hasil Wordcloud Negatif

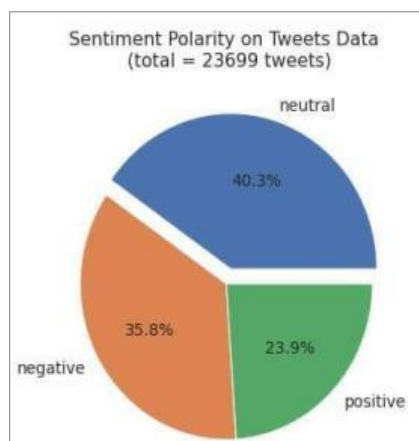
Gbr 17. merupakan wordcloud negatif dari dataset. Pada wordcloud diatas terdapat beberapa kata menonjol yang menggambarkan respon negatif dari pengguna twitter dengan topik cryptocurrency, bitcoin, dan ethereum. Contohnya seperti “turun” , “dah” , “beli” , “dunia” , “bayar” , “beda” , “takut” , “kayak” , “salah.

3. Wordcloud Neutral

Untuk mendapatkan list wordcloud neutral, dapat mengambil text yang mendapatkan hasil polarity 0. Berikut script untuk mendapatkan wordcloud neutral.

19

netral 40,3%. Berikut adalah diagram hasil sentimen pada penelitian ini.



Gbr 20. Hasil Analisis Sentimen

Berdasarkan diagram gbr 20. diatas dapat dilihat bahwa banyak masyarakat pengguna twitter yang memiliki opini netral terhadap topik cryptocurrency, bitcoin & ethereum.

IV. KESIMPULAN

Berdasarkan hasil pengujian & analisis yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

1. Algoritma BERT mempelajari kategori kata yang termasuk dalam sentimen positif & negatif pada kumpulan data yang divalidasi oleh kamus lexicon bahasa indonesia Pada penelitian ini hasil analisis sentimen terhadap 23.699 tweet dengan metode BERT menghasilkan sentimen negatif 35,8%, sentimen positif 23,9% dan sentimen netral 40,3%.
2. Algoritma RCNN bekerja dengan cara pengguna memasukkan kedalam sistem yang kemudian sistem akan mengekstrak dengan menggunakan fitur CNN dan akan kemudian akan diklasifikasikan sesuai inputan pengguna. Pembentukan model yang digunakan untuk mengolah kata & mengklasifikasikan sentimen mendapatkan skor akurasi rata-rata sebesar 0.5609 atau 56.09.
3. Analisis sentimen bisa dipergunakan untuk mengetahui sentiment warga tentang cryptocurrency khususnya bitcoin, dan ethereum. Tujuannya untuk mengetahui seberapa positif, netral dan negatif terhadap hashtag tersebut. Setelah dilaksanakan analisis sentimen, review pengguna mayoritas menyatakan netral dengan persentase, sentimen positif 23,9% ; sentimen netral 40,3% dan sentimen negatif 35,8%. Dengan ini bisa ditarik kesimpulan bahwa metode BERT dan RCNN kurang efektif untuk implementasi analisis sentiment.

V. SARAN

Berdasarkan hasil analisis & kesimpulan, peneliti dapat memberikan saran untuk peneliti selanjutnya sebagai berikut:

1. Konsisten terhadap data yang digunakan sebagai bahan analisa sentimen.
2. Lebih luas dalam memperoleh data yang lebih bervariasi atau rentang waktu diperpanjang.
3. Menggunakan metode klasifikasi yang lain sehingga dapat diperoleh hasil yang lebih efektif.

UCAPAN TERIMA KASIH

Peneliti senantiasa mengucap rasa syukur yang sangat besar kepada Tuhan YME atas segala berkah, rahmat & juga pertolonganNya, sehingga peneliti mampu menyelesaikan proyek & artikel ilmiah ini dengan baik. Terimakasih penulis ucapkan juga kepada Dosen Pembimbing & Koordinator Program Studi yang telah memberi arahan & masukan yang sangat bermanfaat buat peneliti. Kedua orangtua peneliti yang tak henti-hentinya mendoakan & membantu dalam segala hal. Semua saudara-saudara Jurusan Teknik Informatika Angkatan 2017 terutama kelas TI2017B yang selalu memberi dukungan & semangat hingga peneliti dapat menyelesaikan skripsi ini dengan lancar.

REFERENSI

- [1] Rezeki, S. R. (2020). "Pengguna Sosial Media Twitter dalam Komunikasi Organisasi". *Journal of Islamic and Law Studied*. pp 63-78.
- [2] CNBC Indonesia. (2021, November 6). "China & India Jadi Penentang Keras Bitcoin C, Kenapa?" <https://www.cnbcindonesia.com>
- [3] Ferdiawan, Y. I., Nurjanah, P. A., Krisdyan, E. P., Hidayatullah, A., Sirait, H. J., & Rakhmawati, N. A. (2019). "HOAX Impact to Community Through Social Media Indonesia". *Cakrawala*. Vol. 19, No. 1. pp 121-124.
- [4] Egiyi Modesta A. & Ofoegbu Grace N. (2020). "Cryptocurrency And Climate Change: An Overview". *International Journal of Mechanical Engineering and Technology (IJMET)*. Vol. 11, Issue 3, pp. 15-22.
- [5] A. Yani, Dhita Deviacita, Helen Sasty Pratiwi, and Hafiz Muhandi. (2019). "Implementasi Web Scraping Untuk Pengambilan Data Pada Situs Marketplace." *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*. Vol. 7, No. 4.
- [6] Tuhuteru, H. (2020). "Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berskala Besar Menggunakan Algoritma Support Vector Machine". *Information System Development*. pp 7-13.
- [7] M. Umer, Imran Ashraf, Arif Mehmood, Saru Kumari, Saleem Ullah & Gyu Sang C. (2021). "Sentiment Analysis of Tweets Using a Unified Convolutional Neural Network-Long Short-Term Memory Network Model". *Computational Intelligence*. (2021). 37, 409-434.
- [8] Julian, E. (2021, Februari 23). "Mengenal Algoritma Google BERT & 5 Strategi Utama untuk Mengatasinya". <https://www.garuda.website/blog/algoritma-google-bert/>
- [9] Athar Hussein Mohammed & Ali H. Ali. (2021). "Survey of BERT (Bidirectional Encoder Representation Transformer) types". *Journal of Physics: Conference Series*. doi:10.1088/1742-6596/1963/1/012173
- [10] Rizki, Y. (2021). "Klasifikasi Pola Kain Tenun Melayu Menggunakan Faster R-CNN". *IT Journal Research and Development*. pp 215-225. doi:10.25299/itjrd.2021.
- [11] Mardhiya Hayaty, Siti Muthmainah, Syed Muhammad Ghufuran. (2020), "Random and Synthetic Over-Sampling Approach to Resolve Data Imbalance in Classification". *International Journal Of Artificial Intelligence Research*. Vol 4, No 2, pp. 86-94. DOI: 10.29099/ijair.v4i2.152.
- [12] Joloudari J.H., Hussain, S., Nematollahi, M.A. et al. (2023). "BERT-deep CNN: state of the art for sentiment analysis of COVID-19 tweets". *Soc. Networking Analysis*. Min. Vol. 13, 99. <https://doi.org/10.1007/s13278-023-01102-y>.
- [13] Jayadiani, Herlina, Wilis Kaswidjanti, Agung Tri Utomo, and Shoffan Saifullah. (2022). "Sentiment Analysis of Indonesian Reviews Using Fine-Tuning IndoBERT and R-CNN." *ILKOM*

Jurnal Ilmiah. Vol. 14, No. 3, pp 348–54.
<http://dx.doi.org/10.33096/ilkom.v14i3.1505.348-354>