

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini :

- | | |
|--------------------|---|
| 1. Nama | <u>Dr. Liza Wikarsa, BCS., M.Comp</u> |
| 2. Kewarganegaraan | <u>Indonesia</u> |
| 3. Alamat | <u>Jl. Walanda Maramis Lingk IV, RW. 005 Pinaesaan
Wenang</u> |
| 1. Nama | <u>Thomas Ch. Suwanto, S.Kom., M.Mm</u> |
| 2. Kewarganegaraan | <u>Indonesia</u> |
| 3. Alamat | <u>Bahu, Malalayang Lingk VII RW. 007</u> |
| 1. Nama | <u>Frendy Indriawan Pandiono, S.T</u> |
| 2. Kewarganegaraan | <u>Indonesia</u> |
| 3. Alamat | <u>Jl. Mangga Raya No.17 Paniki Dua</u> |

Dengan ini menyatakan bahwa :

1. Karya Cipta yang saya mohonkan :

Berupa : Program Komputer
Berjudul : Aplikasi Analisis Sentimen Pengaruh Teknologi
Blockchain terhadap Netizen Twitter dengan Algoritma
Lexicon-Based Approach

- Tidak meniru dan tidak sama secara esensial dengan Karya Cipta milik pihak lain atau objek kekayaan intelektual lainnya sebagaimana dimaksud dalam Pasal 68 ayat (2);
- Bukan merupakan Ekspresi Budaya Tradisional sebagaimana dimaksud dalam Pasal 38;
- Bukan merupakan Ciptaan yang tidak diketahui penciptanya sebagaimana dimaksud dalam Pasal 39;
- Bukan merupakan hasil karya yang tidak dilindungi Hak Cipta sebagaimana dimaksud dalam Pasal 41 dan 42;
- Bukan merupakan Ciptaan seni lukis yang berupa logo atau tanda pembeda yang digunakan sebagai merek dalam perdagangan barang/jasa atau digunakan sebagai lambang organisasi, badan usaha, atau badan hukum sebagaimana dimaksud dalam Pasal 65 dan;
- Bukan merupakan Ciptaan yang melanggar norma agama, norma susila, ketertiban umum, pertahanan dan keamanan negara atau melanggar peraturan perundang-undangan sebagaimana dimaksud dalam Pasal 74 ayat (1) huruf d Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta.

2. Sebagai pemohon mempunyai kewajiban untuk menyimpan asli contoh ciptaan yang dimohonkan dan harus memberikan apabila dibutuhkan untuk kepentingan penyelesaian sengketa perdata maupun pidana sesuai dengan ketentuan perundang-undangan.
3. Karya Cipta yang saya mohonkan pada Angka 1 tersebut di atas tidak pernah dan tidak sedang dalam sengketa pidana dan/atau perdata di Pengadilan.
4. Dalam hal ketentuan sebagaimana dimaksud dalam Angka 1 dan Angka 3 tersebut di atas saya / kami langgar, maka saya / kami bersedia secara sukarela bahwa :
 - a. Permohonan karya cipta yang saya ajukan dianggap ditarik kembali;
Karya Cipta yang telah terdaftar dalam Daftar Umum Ciptaan Direktorat Hak Cipta, Direktorat Jenderal Hak Kekayaan Intelektual, Kementerian Hukum Dan Hak Asasi Manusia R.I dihapuskan sesuai dengan ketentuan perundang-undangan yang berlaku.
 - b. Dalam hal kepemilikan Hak Cipta yang dimohonkan secara elektronik sedang dalam berperkara dan/atau sedang dalam gugatan di Pengadilan maka status kepemilikan surat pencatatan elektronik tersebut ditangguhkan menunggu putusan Pengadilan yang berkekuatan hukum tetap.

Demikian Surat pernyataan ini saya / kami buat dengan sebenarnya dan untuk dipergunakan sebagaimana mestinya.

Manado, 26 Mei 2023

Yang menyatakan,



(Dr. Liza Wikarsa, BCS., M.Comp)

Yang menyatakan,



(Thomas Ch. Suwanto, S.Kom., M.Mm)

Yang menyatakan,



(Frendy Indriawan Pandiono, S.T)

Aplikasi Analisis Sentimen Pengaruh Teknologi *Blockchain* terhadap *Netizen Twitter* dengan *Algoritma Lexicon-Based Approach*

Dibuat Oleh :

Dr. Liza Wikarsa, B.C.S., M.Comp

Thomas Christian Suwanto, S.Kom., M.Mm

Frendy Indriawan Pandiono, S.T

Web#3

ANALISIS SENTIMEN PENGARUH TEKNOLOGI BLOCKCHAIN TERHADAP NETIZEN TWITTER MENGGUNAKAN ALGORITMA LEXICON-BASED APPROACH

Dashboard

Nov. 24, 2022, 4:59 p.m.

Fetch New Data:

API Key

API Secret Key

Access Token

Access Secret Token

Keywords

Enter your keywords

Topic:

Ecosystem and Metaverse

Number of Tweets:

Enter tweets quantity

Process

Existing Data:

Total Tweets in Existing Data: 1460

Ecosystem: 480

Metaverse: 1000

Show 10 entries

Search:

No.	Tweet ID	Tweet	Topic
-----	----------	-------	-------

MANADO

2023

ABSTRAK

Reaksi dan Respon, serta ungkapan perasaan orang dalam sosial media menjadi tren dan standar untuk penentuan Kebutuhan dan pengambilan keputusan. Informasi tersebut dapat digunakan sebagai dasar untuk membentuk opini tentang kebutuhan dan tren masa depan. Analisis sentimen dapat digunakan untuk menganalisis pendapat yang terkandung dalam teks digital dan menghasilkan wawasan objektif yang berguna dalam membangun produk atau layanan yang lebih baik. Analisis sentimen juga membantu dalam mengidentifikasi tren saat ini.

Aplikasi ini bertujuan untuk menganalisis sentimen yang berkaitan dengan pengaruh teknologi blockchain terhadap netizen Twitter dengan menggunakan pendekatan berbasis leksikon. Aplikasi ini akan memberikan informasi dan kebutuhan bagi para pengembang teknologi blockchain yang akan membangun teknologi baru di Indonesia. Aplikasi ini berhasil dibangun dengan baik dengan menggunakan pendekatan berbasis leksikon dan menghasilkan kesimpulan yang sesuai dengan hasil pengujian yang dilakukan dengan akurasi 72,7%. Hasil analisis menunjukkan bahwa emosi senang adalah sentimen yang paling dominan. Aplikasi ini dapat dijadikan referensi dan tren positif bagi para pengembang.

DAFTAR ISI

DAFTAR ISI.....	ii
DAFTAR TABEL	iii
DAFTAR GAMBAR	iv
BAB I_PENDAHULUAN.....	1
BAB II_METODE PENELITIAN.....	2
BAB III_HASIL DAN PEMBAHASAN	5
BAB IV_KESIMPULAN	11
DAFTAR PUSTAKA_.....	11

DAFTAR TABEL

Tabel 2.1 Pengambilan <i>Tweet</i>	3
Tabel 2.2 <i>Candidate Word</i>	3
Tabel 3.1 Analisis Hasil Pengujian	9

DAFTAR GAMBAR

Gambar 2.1 Alur Penerapan Algoritma <i>Lexicon-based</i> [5].....	2
Gambar 2.2 <i>Pairing Words</i>	4
Gambar 3.1 <i>Use Case Diagram</i> Aplikasi Analisis Sentimen.....	6
Gambar 3.2 <i>Class Diagram</i> Aplikasi Analisis Sentimen	7
Gambar 3.3 Pengaturan Penarikan Data <i>Tweet</i>	8
Gambar 3.4 Proses <i>Pre-processing</i>	8
Gambar 3.5 Validasi Hasil di <i>Topic Level</i>	9

BAB I

PENDAHULUAN

Teknologi yang disebut sebagai blockchain merupakan suatu teknologi pencatatan dan penyimpanan data transaksi yang memberikan akses jaminan keamanan berbasis kode unik yang saling berhubungan dan bersifat terdistribusi antar blok transaksi. Data disimpan dalam blok yang terkoneksi satu sama lainnya dan terenkripsi dengan aman[1]. Blockchain digunakan dalam berbagai bidang, termasuk bidang kesehatan di mana teknologi ini digunakan untuk memastikan keamanan data dan konsistensi riwayat medis dan pemberian obat-obatan[2]. Selain itu, teknologi blockchain juga dapat digunakan dalam bidang keuangan[3].

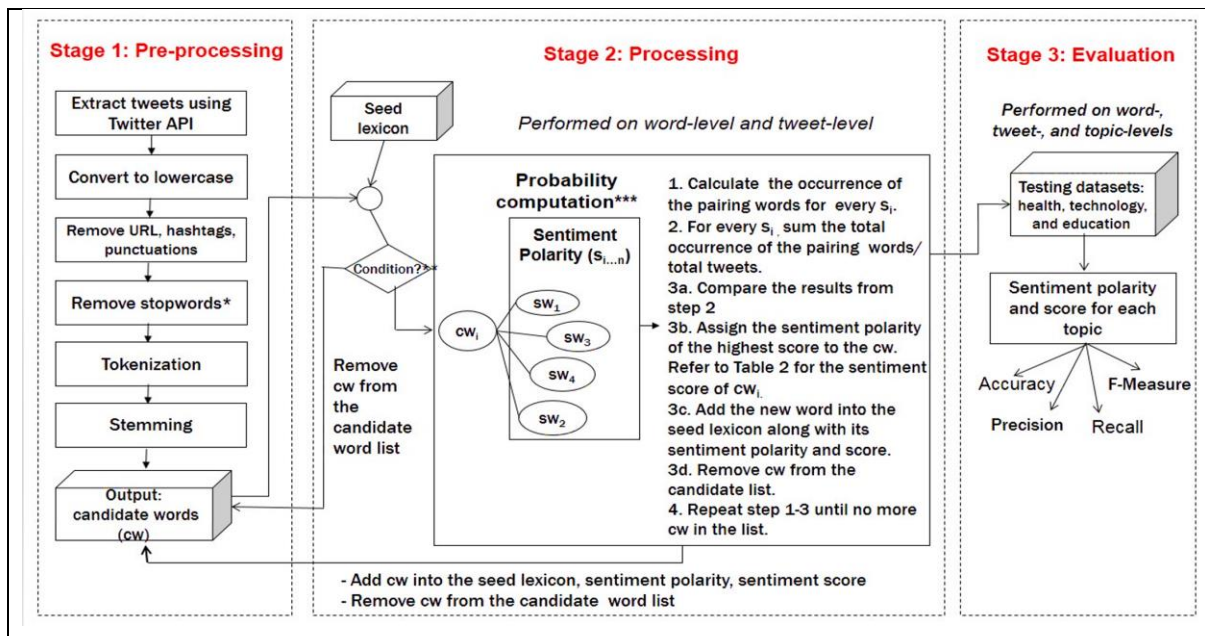
Penggunaan teknologi *blockchain* di Indonesia menarik perhatian dan dapat diekspresikan melalui sentimen mereka terhadap penerapan dan penggunaan teknologi *blockchain*, sebagai pendapat terkait respon atau tanggapan publik pada Twitter. Twitter merupakan media sosial yang paling sering digunakan dalam melihat tanggapan dan komentar karena pengembang teknologi *blockchain* lebih banyak berinteraksi pada Twitter [4]. Analisis sentimen saat ini banyak digunakan oleh publik dan sektor komersial sebagai bahan dasar untuk mengetahui pendapat atau opini terkait produk dan layanan, perancangan strategi politik, peramalan keuangan, dan lainnya [5].

Penelitian tentang sentimen analisis terhadap *blockchain* sudah ada sebelumnya, namun perbedaan dengan aplikasi yang akan dibangun terletak pada fokus pada konsep penggunaan teknologi yang ditinjau dan pemakaian algoritmanya [6]. Terdapat juga penelitian yang menerapkan algoritma *Lexicon-based*, yang terbagi atas *Dictionary-based approach* dan *Corpus-based approach*, untuk mengenali sentimen berdasarkan hubungan timbal balik satu sama lainnya [7]. Pembahasan di atas memberikan kesempatan untuk membangun aplikasi analisis sentimen terhadap pengaruh teknologi *blockchain* terhadap *netizen* Twitter yang dapat menerapkan algoritma *Lexicon-based approach*, penerapan tersebut akan memberikan hasil untuk riset penelitian dalam mengimplementasikan teknologi tersebut pada masa yang akan mendatang, terlebih digunakan oleh pengembang yang ada di Indonesia, serta melihat tanggapan orang Indonesia akan masuknya teknologi yang terbaru.

BAB II

METODE PENELITIAN

Algoritma *lexicon-based* memiliki 2 bagian yaitu *Dictionary-based approach* dan *Corpus-based approach*. *Dictionary-based approach* pada dasarnya memiliki inputan kata yang akan dibandingkan pada dasar sinonim dan antonim kata yang memiliki nilai masing-masing, sentimen yang mengandung kata baru dimasukan menjadi kandidat, kata akan ditambahkan pada daftar kata[5].



Gambar 2.1 Alur Penerapan Algoritma *Lexicon-based* [5]

Cara kerja *lexicon based* terdapat beberapa langkah [5]:

1. Fase awal dilakukan pengambilan data dan proses pembersihan data untuk memudahkan dalam pengolahan data pada fase ke-2. Selanjutnya Mengambil *tweet* yang sudah bersih dan sudah dibagi menjadi kata, kemudian memasukan pada *candidate word*. Jika kata pada *candidate word* ada pada kamus (*seed lexicon*) maka kata-kata tersebut akan dihapus dalam *candidate word*, sebaliknya jika tidak ada maka akan tetap pada *candidate word*.
2. Fase kedua yaitu *processing* atau bagian penerapan algoritma, melakukan proses *pairing* atau pasang kata pada *candidate word* dan *seed lexicon* untuk mencari kata-kata yang ada pada *training tweet*. Menghitung pasangan kata yang memiliki kata pada *tweet* untuk dihitung total kemunculannya. Memasukkan data *candidate word* pada *seed lexicon* yang diambil nilai kemunculan tertinggi untuk dimasukan nilai emosi.
3. Fase ketiga yaitu melakukan pengujian pada data yang uji yang meliputi *tweet level* dan *topik level*, hasil dari pengujian memberikan data keseluruhan emosi pada *tweet* yang sudah dilabel dan

memberikan kesimpulan pada topik yang diambil, juga dapat diolah mencari akurasi, *precision*, *recall* dan lainnya.

Berikut merupakan alur algoritma yang meliputi 3 fase yang akan digunakan:

1. Fase awal adalah Pengumpulan data *tweet* yang diambil pada *Stream API (Application Programming Interface)* Twitter berjumlah 1460 data. Kemudian, *tweet* tersebut akan dibersihkan pada fase *preprocessing* berikutnya.

Tabel 2.1 Pengambilan Tweet

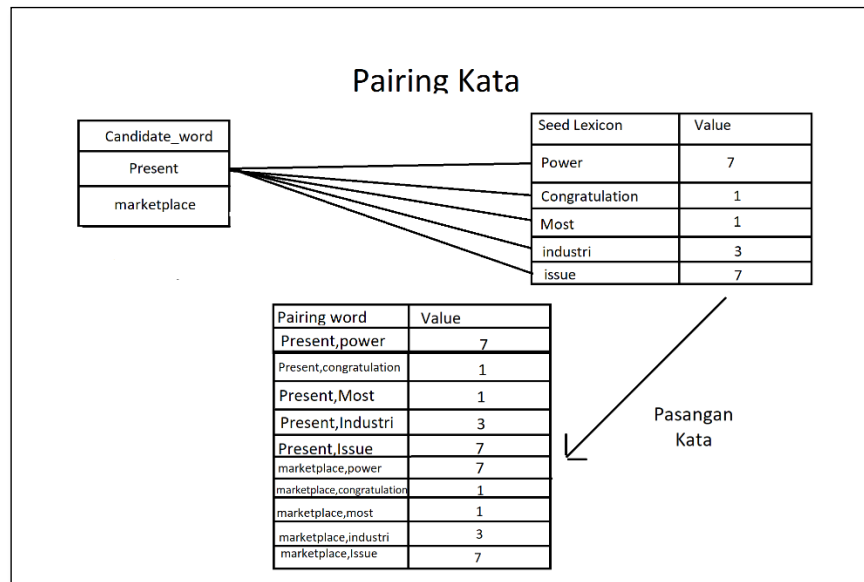
Data Tweet	Clean Tweet	Topic
@cookieslovedone @cehlos #Cookies is NOT just a meme coin! ???? @cookieslovedone Staking \$COOKIES will earn you \$USDT rewards. It will also solve a variety of #blockchain key tech and #ecosystem problems. please make Trending & Hashtag #COOKIES ???????? in #Twitter #Cookies #iwonwithcookies	cookies meme coin staking cookies earn usdt rewards also solve variety blockchain key tech ecosystem problems please make trending hashtag cookies twitter cookies iwonwithcookies.	Ecosystem
#GTC 952% PROFITS CONGRATULATIONS #BTC #ETH #TRX #ELON #Metaverse #BinanceTTC #BINANCE #Bybit https://t.co/zzQBBet2Bp .	gtc profits congratulations btc eth trx elon metaverse binancettc binance bybit.	Metaverse

2. Pada fase kedua data yang sudah dibersihkan kemudian akan dipisahkan menjadi data *training* dan *testing* dengan perbandingan 60:40.
3. Data tersebut kemudian ditokenisasi menjadi kata untuk dimasukkan pada daftar *candidate word* seperti di Tabel 2.2.

Tabel 2.2 Candidate Word

Candidate Word
<i>increased</i>
<i>destruction</i>
<i>present</i>
<i>surrounding</i>
<i>marketplace</i>

4. kata tersebut dimasukkan emosi manual dan diproses untuk dibandingkan pada kamus *lexicon* pada proses *pairing* kata.



Gambar 2.2 Pairing Words

5. Kata tersebut yang memiliki nilai kemunculan terbanyak pada pairing word akan dimasukan pada kamus sebagai kata baru.
6. Fase ketiga pengujian pada data yang uji yang meliputi *tweet level* dan *topik level*, hasil dari pengujian memberikan data keseluruhan emosi pada *tweet* yang sudah dilabel dan memberikan kesimpulan pada topik yang diambil, juga dapat diolah mencari akurasi, *precision*, *recall* dan lainnya.

BAB III

HASIL DAN PEMBAHASAN

1. Tempat Pelaksanaan Kegiatan : Universitas Katolik De La Salle Manado
2. Waktu Pelaksanaan Kegiatan : Februari 2022 Sampai Agustus 2022
3. Hasil Pelaksanaan Kegiatan : Hak cipta untuk Aplikasi Analisa Sentimen

A. Aturan Penggunaan Sistem

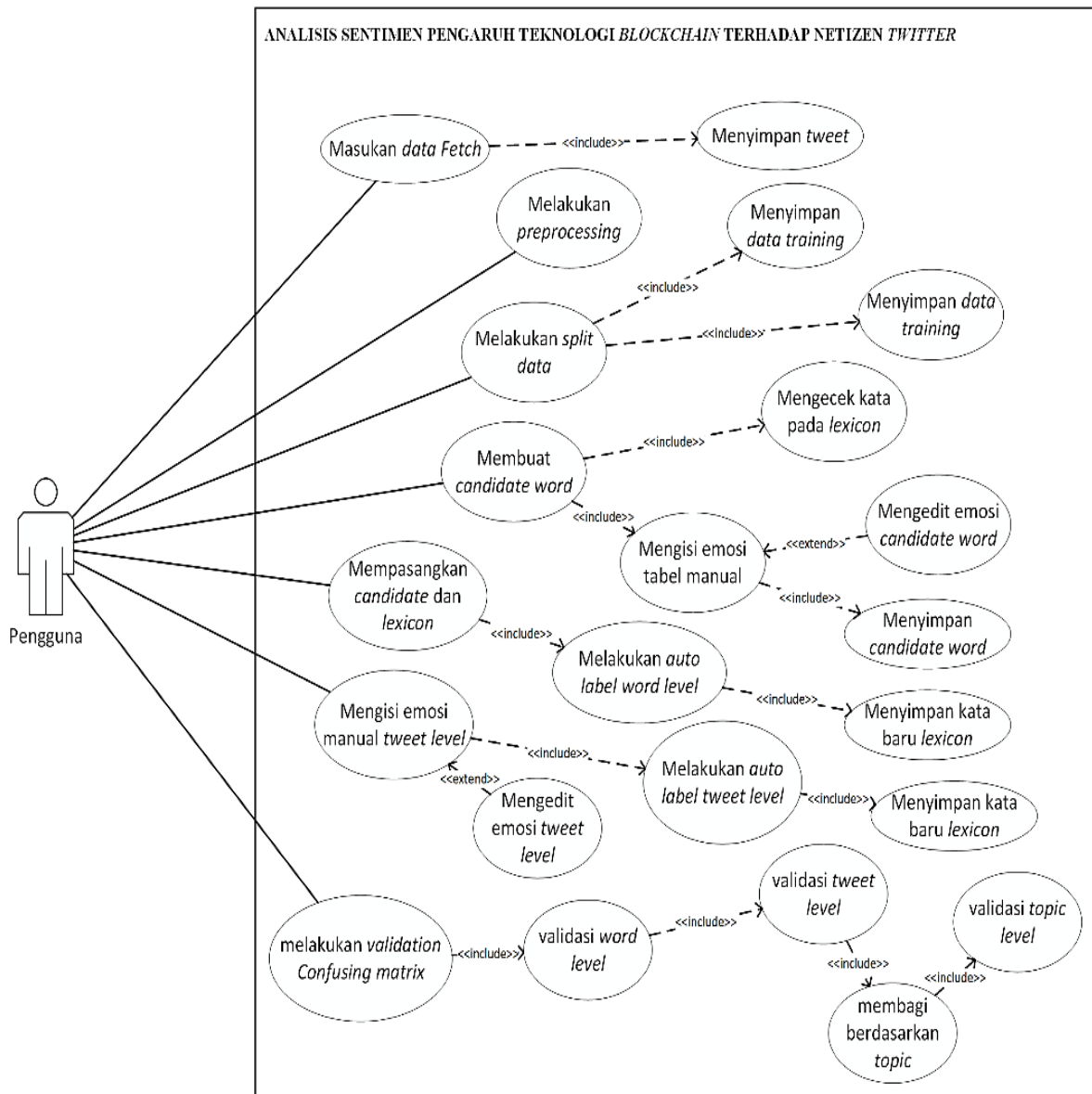
Berikut merupakan aturan yang akan diterapkan pada aplikasi yang akan dibangun, yaitu:

1. Total data yang diambil adalah 1460 data *tweets*, pembagian data latih dan uji adalah 60:40.
2. Pada aplikasi topik teknologi *blockchain* web3, terdapat 2 parameter yang digunakan, yaitu ekosistem dan *metaverse*.
3. Terdapat 6 emosi yang diambil sebagai nilai sentimen emosi, frekuensi tingkatan emosi dari 1-6, dimulai pada emosi sedih (1), takut (2), netral (3), marah (4), terkejut (5), dan senang (6).
4. Proses memasang *candidate word* dan kata dalam kamus *seed lexicon* dilakukan untuk mengecek kemungkinan persamaan kata dari keduanya. Jika *candidate word* ini sudah ada dalam kamus *seed lexicon*, maka *candidate word* tersebut akan dihapus secara otomatis oleh sistem. Akan tetapi, jika tidak ditemukan, maka *candidate word* ini akan diberikan label emosi yang mempunyai nilai kemunculan tertinggi (paling dominan) dan kemudian ditambahkan ke dalam kamus *seed lexicon*. Lalu, *candidate word* ini akan dihapus dari daftar *candidate word*.
5. Apabila emosi yang dimasukkan secara manual pada tabel *seed lexicon* memiliki nilai yang lebih besar, maka emosi yang diambil adalah berasal dari manual.
6. Untuk proses *training* dan *testing* pada fase *tweet level*, pengguna diharuskan untuk mengisi label emosi untuk tiap *tweet* secara manual terlebih dahulu. Sesudahnya, sistem akan mengaplikasikan algoritma *Lexicon-based* pada data tersebut sehingga setiap *tweet* diberikan label emosi yang diisi secara otomatis oleh *classifier*.
7. Hasil data yang sudah diproses akan tetap dibagi menjadi dua topik untuk hasil akhir penelitian.
8. Untuk *testing*, sistem akan menggunakan data latih yang sudah disiapkan di mana pengguna juga harus mengisi label emosi secara manual. Kemudian, *classifier* memberikan label emosi secara otomatis dan membandingkan hasil pelabelan emosi manual dan otomatis untuk menghitung tingkat akurasi, presisi, dan *recall*.

9. Pengambilan data *Twitter* pada penelitian waktu pengambilan data dimulai pada bulan September tahun 2022.

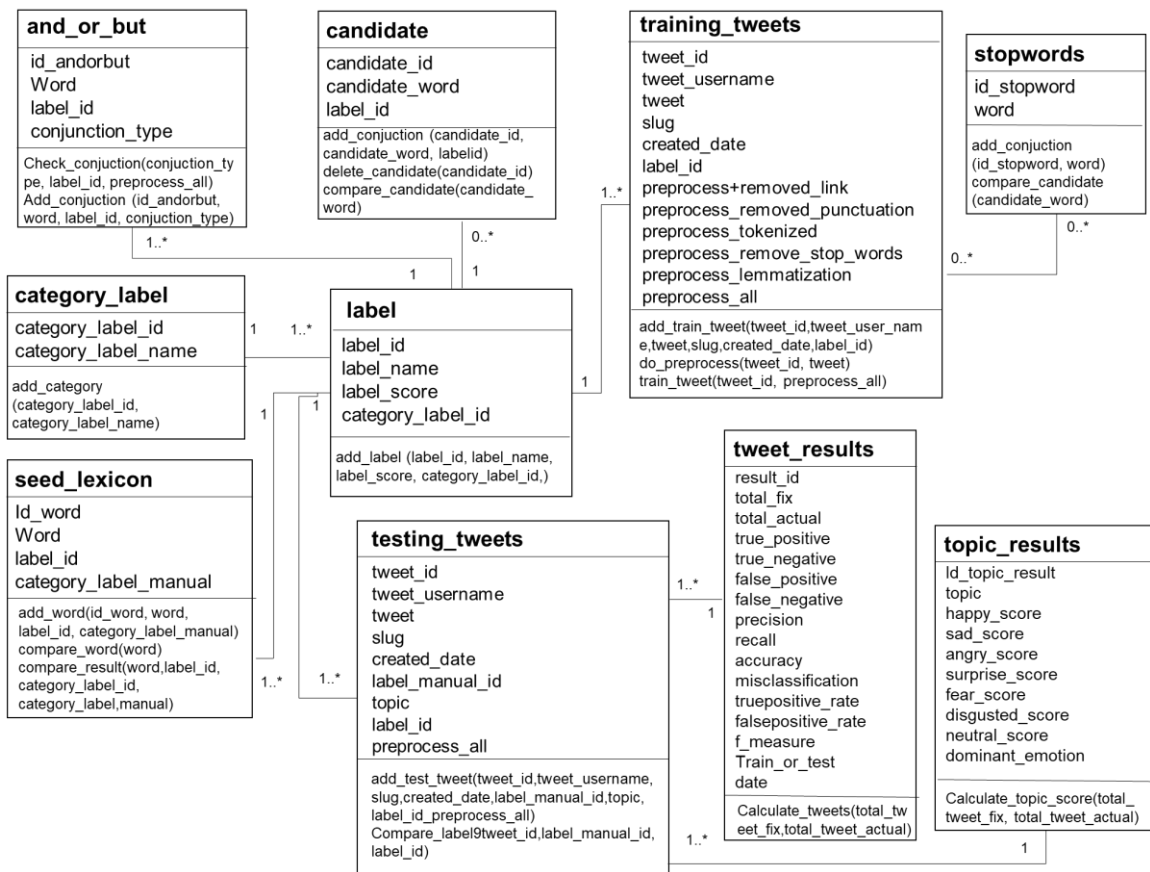
B. Desain

Diagram *use case* di bawah ini menggambarkan hubungan aktor dengan sistem yang akan dibuat, ditunjukkan pada gambar 3.1.



Gambar 3.1 Use Case Diagram Aplikasi Analisis Sentimen

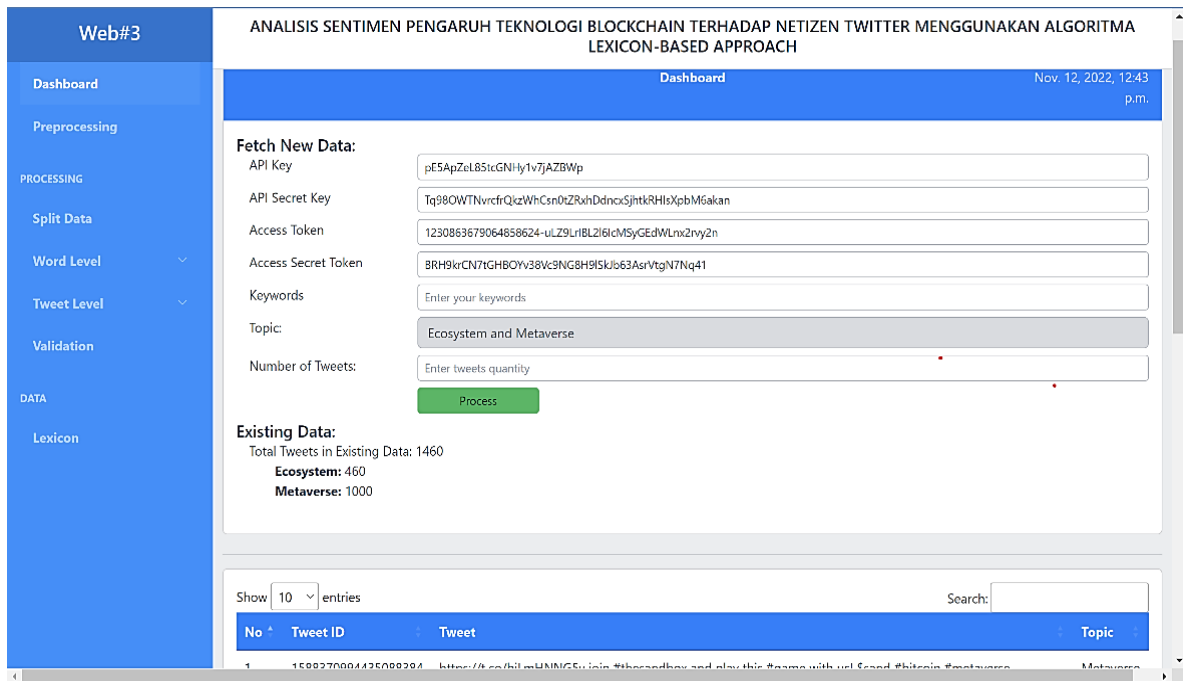
Class diagram pada Gambar 3.2 menggambarkan desain basis data untuk Aplikasi Analisis Sentimen yang akan dibangun.



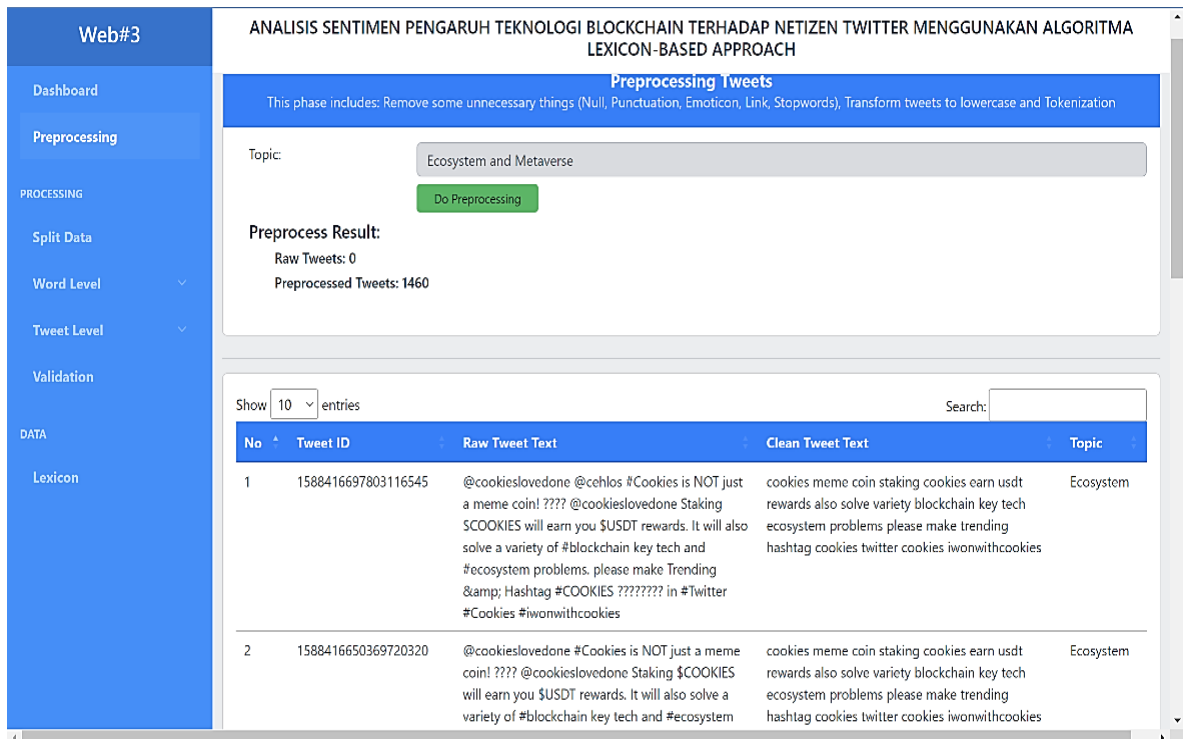
Gambar 3.2 Class Diagram Aplikasi Analisis Sentimen

C. Tampilan Aplikasi

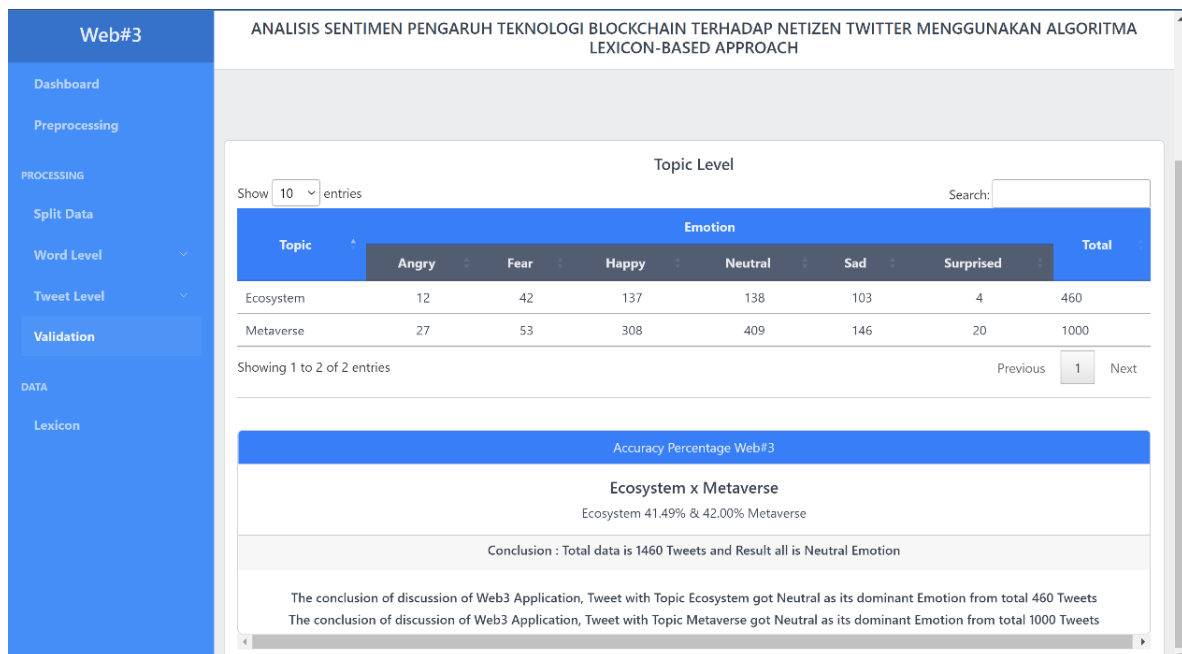
Berikut ini adalah tampilan-tampilan hasil implementasi antarmuka. Gambar 3.3 memperlihatkan halaman pengaturan penarikan data *tweet* menggunakan *stream API* Twitter. Gambar 3.4 menunjukkan halaman *pre-processing* dimana tempat dilaksanakannya pembersihan pada *tweet* yang diambil. Proses pembersihan, meliputi *lowercase*, tokenisasi, pembersihan data kosong, penghapusan simbol dan gambar, penghapusan *stopwords*, dan *stemming*. Gambar 3.5 menampilkan hasil dari permintaan kata kunci yang kita masukkan di awal dalam bentuk kesimpulan emosi mana yang paling dominan terhadap suatu kata kunci di level topik.



Gambar 3.3 Pengaturan Penarikan Data Tweet



Gambar 3.4 Proses Pre-processing



Gambar 3.5 Validasi Hasil di Topic Level

Hasil pengujian dari aplikasi yang dibuat bisa dilihat pada tabel 3.1 yang berisi rincian hasil analisis sentimen pengaruh teknologi *blockchain* terhadap netizen *Twitter* yang memiliki hasil dari 1,460 total keseluruhan data yang diuji.

Tabel 3.1 Analisis Hasil Pengujian

Pengujian	Emosi	Label	Total data
1	Marah	<i>Negative</i>	47
2	Takut	<i>Negative</i>	142
3	Senang	<i>Positive</i>	506
4	Netral	<i>Neutral</i>	271
5	Sedih	<i>Negative</i>	375
6	Terkejut	<i>Positive</i>	24

Berdasarkan tabel diatas dapat dilihat untuk emosi yang mendominasi pada *tweet* topik web3 yang lebih spesifik menjadi *ecosystem* dan *metaverse* adalah 1000 dan 460 yang terdiri atas:

1. Semua fungsi pada website sudah sesuai dengan tujuan pembuatan website, akan tetapi memiliki waktu proses yang cukup lama.
2. Data yang diambil pada *tweet* masih banyak mengandung emosi netral (271). pembahasan yang biasa dimasukan adalah promosi dan keuangan yang spesifik pada produk, akibatnya banyak kata tidak memiliki emosi karena bukan merupakan kata sifat.

3. Klasifikasi yang dilakukan pada penelitian menggunakan memori yang tinggi dalam memproses data yang banyak.
4. Waktu yang lama dibandingkan dengan word level
5. Semakin banyak leksikon maka semakin tinggi tingkat akurasi sentimen
6. Jika sentimen netral dimasukkan ke sentimen positif maka hasil akurasi akan lebih tinggi.
7. Semakin data training dengan menggunakan leksikon banyak maka semakin baik model atau tier yang terbentuk
8. Hasil perbandingan menggunakan *confusion matrix* menghasilkan akurasi 72.7 % pada total keseluruhan data yang dibandingkan.
9. Hasil emosi dengan jumlah paling sedikit yaitu emosi *surprised* sebanyak 24 sentimen.
10. Hasil emosi dengan jumlah terbanyak yaitu emosi *happy* sebanyak 506 sentimen.
11. Jika ada kata dalam daftar *candidate word* yang ternyata sudah terdaftar di *lexicon* maka aplikasi akan menampilkan sentimen dari kata tersebut sesuai ada di *lexicon*. kemudian kata ini akan dihapus dari daftar *candidate word* dan tidak diproses lebih lanjut.
12. Pada fitur pencarian hanya menampilkan pencarian tweet, untuk tabel yang berisi auto label, manual label dan emosi dapat digunakan fitur sort berdasarkan abjad.
13. Berdasarkan hasil pengujian ini maka aplikasi yang dibuat cukup baik karena akurasi yang dibandingkan sudah tinggi, dan memberikan hasil untuk penerapan teknologi web3 di Indonesia sangat terbuka bagi para pengembang. Dampak positif yang dibahas pada *Twitter* sangat berperan penting untuk pengembangan industri tersebut.

BAB IV

KESIMPULAN

Berdasarkan hasil pembahasan serta dilakukannya pengujian pada aplikasi analisis sentimen yang sudah dibangun, diperoleh beberapa kesimpulan sebagai berikut:

1. Berdasarkan hasil yang didapat pada pengujian, hasil tersebut menjadikan emosi senang (34,6%) *sentimen* terbanyak dan yang paling sedikit adalah terkejut (1,6%) *sentimen* pada penelitian. Dengan demikian maka menjawab pertanyaan yang ada pada awal penelitian ini dibuat yaitu, membantu dalam mengetahui pendapat orang terhadap teknologi blockchain yang sedang berkembang ini.
2. Penerapan algoritma *Lexicon-based Approach* berjalan sesuai dengan aturan yang dibuat dan memberikan jawaban yang sesuai.
3. Hasil analisis data sentimen dapat digunakan untuk pertimbangan pengembang untuk mengikuti tren dan perkembangan kedepan dalam mengembangkan aplikasi web3 di Indonesia, melalui data penelitian tersebut dapat dilihat respon reaksi orang terhadap teknologi web3 dapat diterima dengan baik.
4. klasifikasi ini harus menggunakan memori yang tinggi, waktu yang lebih lama pada tweet level dibandingkan dengan *word level*
5. Semakin banyak leksikon pada penyimpanan maka semakin tinggi tingkat akurasi sentimen yang dihasilkan.

DAFTAR PUSTAKA

- [1] I. Bhiantara, "Teknologi Blockchain Cryptocurrency Di Era Revolusi Digital," *Prosiding Seminar Nasional Pendidikan Teknik Informatika*, vol. 9, no. 1, pp. 173-176, 2018.
- [2] E. Setiawan, et al., "Pengembangan Sistem Rekam Medis Rumah Sakit dengan Multi UserRest Serverberbasis Permissioned Blockchain menggunakan Framework Hyperledger," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 4, no. 1, pp. 1-10, 2020.
- [3] E. Harahap, et al., "Pemanfaatan Teknologi Blockchain Pada Platform Crowdfunding," *Technomedia Journal (TMJ)*, vol. 4, no. 2, pp. 199 - 210, 2020.
- [4] A. Prasetya, et al., "Sentiment Analisis Terhadap Cryptocurrency Berdasarkan Comment Dan Reply Pada Platform Twitter," *Journal of Information Systems and Informatics*, vol. 3, no. 2, pp. 268 - 277, 2021.
- [5] L. Wikarsa and M. Kim, "Automatic Generation Of Word-Emotion Lexicon For Multiple Sentiment Polaritieson Social Media Texts," *ICIC Express Letters*, vol. 13, no. 4, pp. 317 - 324, 2019.
- [6] A. Prasetya, et al., "Sentiment Analisis Terhadap Cryptocurrency Berdasarkan Comment Dan Reply Pada Platform Twitter," *Journal of Information Systems and Informatics*, vol. 3, no. 2, pp. 268 - 277, 2021.
- [7] B. Gupta, et al., *Data Mining Approaches for Big Data and Sentiment Analysis in Social*, Hershey: IGI Global, 2021.