

ANALISIS SENTIMEN DENGAN PENDEKATAN ENSEMBLE LEARNING DAN WORD EMBEDDING PADA TWITTER

Erna Daniati ¹⁾, Hastari Utama ²⁾

¹⁾ Sistem Informasi Universitas Nusantara PGRI Kediri

²⁾ Teknik Informatika Universitas AMIKOM Yogyakarta

email : ernadaniati@unpkediri.ac.id¹⁾, utama@amikom.ac.id²⁾

Abstraksi

Twitter merupakan media sosial yang sering digunakan dimana mencapai 284 juta pengguna aktif dan setiap harinya lebih dari 500 juta tweet per hari. Hari ini menjadikan peluang emas bagi perusahaan dan individu untuk meningkatkan relasi sosial, politik, dan ekonomi yang kuat demi meningkatkan reputasi. Pada kumpulan tweet tersebut memuat opini yang berasal dari berbagai pengguna. Hal ini sangat potensial bagi perusahaan untuk menggali dan mendapatkan informasi mengenai jenis tweet ini. Selanjutnya, ekstraksi data yang disusun dalam bentuk Bag Of Word ini terdapat kekurangan. Fitur yang dihasilkan cukup banyak sehingga berpengaruh dalam waktu proses untuk pelatihan data. Metode Word2Vec memiliki keunggulan dalam menangkap hubungan sintaksis dan semantik antar kata. Pada penelitian ini berusaha untuk meningkatkan akurasi yang dicapai dengan penggunaan word embedding sebagai representasi teks dan ensemble learning dari pengklasifikasi yang digunakan. Hasil penelitian ini mampu menunjukkan tingkat akurasi yang lebih tinggi dalam penggunaan algoritma Adaboost dan Word2Vec.

Kata Kunci :

Ensemble Learning, Sentiment Analysis, Word Embedding, Twitter.

Abstract

Twitter is a social media that is often used where it reaches 284 million active users and every day more than 500 million tweets per day. The day presents a golden opportunity for companies and individuals to cultivate strong social, political and economic relations to enhance reputations. The collection of tweets contains opinions from various users. This is very potential for companies to explore and get information about this type of tweet. Furthermore, the extraction of data arranged in the form of a Bag of Word has drawbacks. The resulting features are quite a lot so that they affect the processing time for data training. The Word2Vec method has the advantage of capturing syntactic and semantic relationships between words. This study seeks to improve the accuracy achieved by using word embedding as a text representation and ensemble learning from the classifiers used. The results of this study were able to show a higher level of accuracy in using the Adaboost and Word2Vec algorithms.

Keywords :

Ensemble Learning, Sentiment Analysis, Word Embedding, Twitter

Pendahuluan

Twitter merupakan media sosial yang sering digunakan dimana mencapai 284 juta pengguna aktif dan setiap harinya lebih dari 500 juta tweet per hari. Hari ini menjadikan peluang emas bagi perusahaan dan individu untuk meningkatkan relasi sosial, politik, dan ekonomi yang kuat demi meningkatkan reputasi. Data tweet yang ada merupakan bahan dasar untuk dijadikan analisis dalam mengetahui tren bisnis kedepannya [1].

Pada kumpulan tweet tersebut memuat opini yang berasal dari berbagai pengguna. Opini dapat dikelompokkan dalam beberapa kategori yaitu opini positif, negatif, dan netral. Hal ini sangat potensial bagi organisasi atau individu untuk menggali dan mendapatkan informasi mengenai jenis tweet ini sehingga produk yang dihasilkan atau anggapan tentang organisasi dapat disingkap dengan baik melalui analisis ini. Klasifikasi jenis opini ini disebut dengan analisis sentimen. Hal ini perlu

dibangun perangkat lunak atau mesin analisis sentimen untuk agar dapat mengklasifikasi setiap sentimen yang ditemukan dalam berbagai tweet [2].

Analisis sentiment ini juga mengalami peningkatan baik ditujukan oleh peneliti, perusahaan dan pemerintah [3]. Pada kegiatan ini terdiri dari 2 pendekatan utama yaitu pendekatan leksikon dan machine learning. Pendekatan berbasis leksikon merupakan pendekatan yang pertama kali digunakan dalam analisis sentimen. Selanjutnya disusul dengan pendekatan machine learning dimana sekarang telah banyak yang menggunakan.

Analisis Sentimen yang dilakukan akan melalui tahap ekstraksi data. Pada tahap ini, kumpulan teks berupa tweet akan dilakukan proses menghasilkan fitur. Fitur ini merupakan kolom yang berisi nilai penentu. Fitur yang dihasilkan ini dibentuk dalam bentuk Bag Of Word [4]. Namun, bentuk ini tidak memperhatikan urutan kata. Selain itu, fitur yang dihasilkan cukup banyak sehingga berpengaruh

dalam waktu proses untuk pelatihan data. Oleh karena itu bentuk Bag Of word perlu diperbaiki. Fitur yang dihasilkan dalam jumlah banyak ini dapat diatasi dengan pendekatan Word Embedding. Word Embedding merupakan salah satu pendekatan berbasis deep learning untuk menyusun representasi vektor kata dan dokumen [4]. Metode ini mendapatkan perhatian khusus dalam klasifikasi teks dan analisis sentimen karena kemampuannya dalam menangkap hubungan sintaksis dan semantik antar kata. Selain itu terdapat metode yang sering dipakai dalam pendekatan Word Embedding ini yaitu Word2Vec dan GloVe.

Peningkatan akurasi ini juga dapat dicapai dengan metode ensemble learning [5]. Metode ini berusaha untuk menggabungkan prediksi beberapa model yang dipelajari dari data yang sama. Misalnya, sering kali mengambil data pelatihan dan memperoleh beberapa data pelatihan yang berbeda darinya, serta mempelajari model dari masing-masing, dan menggabungkannya untuk menghasilkan kumpulan model yang dipelajari. Teknik memang melakukan hal yang bisa sangat ampuh. Misalnya, mungkin untuk mengubah skema pembelajaran yang relatif lemah menjadi skema yang sangat kuat.

Tinjauan Pustaka

Analisis sentimen adalah bidang penelitian yang berkembang pesat dalam pemrosesan bahasa alami (NLP) dan klasifikasi teks. Teknik ini telah menjadi bagian penting dari berbagai aplikasi termasuk politik, bisnis, periklanan dan pemasaran [4]. Ada berbagai teknik untuk analisis sentimen, tetapi baru-baru ini metode Word Embeddings telah banyak digunakan dalam tugas klasifikasi sentimen. Word2Vec dan GloVe saat ini merupakan salah satu metode penyematan kata yang paling akurat dan dapat digunakan yang dapat mengubah kata menjadi vektor yang bermakna. Namun, metode ini mengabaikan informasi sentimen teks dan membutuhkan kumpulan teks yang besar untuk pelatihan dan menghasilkan vektor yang tepat. Akibatnya, karena ukuran beberapa korpora yang kecil, peneliti sering kali harus menggunakan Word Embedding terlatih yang dilatihkan pada korpora teks besar lainnya seperti Google News dengan sekitar 100 miliar kata. Akurasi yang meningkat dari Word Embedding yang telah dilatih sebelumnya berdampak besar pada penelitian analisis sentimen. Dalam makalah ini, penelitian ini mengusulkan metode baru, Improved Word Vectors (IWV), yang meningkatkan akurasi Word Embedding yang telah dilatih sebelumnya dalam analisis sentimen. Metode kami didasarkan pada teknik penandaan Part-of-Speech (POS), pendekatan berbasis leksikon, algoritme posisi kata, dan metode Word2Vec / GloVe. Kami menguji keakuratan metode kami melalui berbagai model pembelajaran mendalam dan kumpulan data sentimen benchmark. Hasil percobaan kami menunjukkan bahwa Improved

Word Vectors (IWV) sangat efektif untuk analisis sentimen.

Penambahan sentimen dari sumber seperti Twitter yang berisi teks informal diperlukan karena terdapat informasi yang menonjol dan sejumlah besar data untuk dianalisis, dipahami, dan dicoba. Telah banyak penelitian di bidang ini untuk mendapatkan informasi semantik dari domain ini dan untuk membuat prediksi yang lebih baik dalam hal klasifikasi Sentimen [6]. Penelitian ini menyajikan pendekatan baru yang menyediakan model ensemble untuk Klasifikasi yang menggunakan SVM sebagai pembelajar dasar dan Adaboost sebagai algoritma Ensemble Boosting. Penelitian ini menampilkan skor Precision, Recall, dan F1 dengan membandingkannya dengan algoritme SVM baseline.

Berbagi informasi melalui jejaring sosial online telah menjadi mekanisme utama orang berbagi informasi dengan teman-temannya melalui perilaku retweet, yang dapat mengakibatkan berbagai aliran penyebaran informasi di media sosial seperti Facebook, Twitter, dan Weibo [7]. Memprediksi perilaku retweet pengguna di jejaring sosial tersebut sangat menantang. Untuk menyelesaikan tugas prediksi, diperlukan identifikasi faktor-faktor yang mempengaruhi perilaku retweet dan pembuatan model yang efisien. Dalam tulisan ini, kami mempelajari jaringan hubungan heterogen dengan mempertimbangkan berbagai interaksi sosial, yang mencerminkan bagaimana tindakan retweet tertentu dipengaruhi oleh perilaku sosial yang dilakukan oleh pengirim dan penerima retweet. Kami kemudian menghasilkan berbagai fitur dari faktor yang kami identifikasi milik tiga dimensi - semantik konten, perilaku difusi pengguna, dan struktur jaringan. Selain itu, kami memasukkan masalah prediksi kami sebagai masalah pembelajaran ensemble dan mengusulkan pendekatan pembelajaran ensemble baru untuk memecahkan masalah. Menyisir fitur yang dihasilkan dan pendekatan pembelajaran ensemble novel. Eksperimen pada kumpulan data nyata yang diekstrak dari situs jejaring sosial Weibo menunjukkan keefektifan model yang kami usulkan, yang menunjukkan bahwa fitur yang kami hasilkan dan pendekatan yang kami usulkan dapat secara signifikan meningkatkan kinerja dalam memprediksi perilaku retweet pengguna yang terjadi dalam proses penyebaran informasi di jejaring sosial.

Automatic Text Summarization System (ATS) memungkinkan pengguna mendapatkan inti informasi dan pengetahuan dalam waktu singkat untuk membuat keputusan penting dengan cepat [8]. Jaringan neural dalam telah membuktikan kemampuannya untuk mencapai kinerja yang sangat baik di banyak aplikasi Pemrosesan Bahasa Alami dan visi komputer di dunia nyata. Namun, ATS masih kurang diperhatikan. Masalah utama dari aplikasi tradisional adalah aplikasi tersebut melibatkan data berdimensi tinggi dan jarang, yang membuatnya sulit untuk menangkap informasi yang

relevan. Salah satu teknik untuk mengatasi masalah ini adalah mempelajari fitur melalui reduksi dimensionalitas. Di sisi lain, penyematan kata adalah teknik jaringan saraf lain yang menghasilkan representasi kata yang jauh lebih ringkas daripada pendekatan Bag-of-Words (BOW) tradisional. Dalam penelitian ini, berusaha untuk meningkatkan kualitas ATS dengan mengintegrasikan teknik jaringan saraf dalam tanpa pengawasan dengan pendekatan embedding kata. selain itu, juga mengusulkan tiga teknik ansambel. Ansambel pertama menggabungkan BOW dan word2vec menggunakan teknik pemungutan suara mayoritas. Ansambel kedua mengumpulkan informasi yang disediakan oleh pendekatan BOW dan jaringan saraf tanpa pengawasan. Ansambel ketiga mengumpulkan informasi yang disediakan oleh Word2Vec dan jaringan saraf tanpa pengawasan. Hal ini menunjukkan bahwa metode ensemble meningkatkan kualitas ATS, khususnya ensemble berdasarkan pendekatan word2vec memberikan hasil yang lebih baik. Akhirnya, kami melakukan eksperimen yang berbeda untuk mengevaluasi kinerja model yang diselidiki. Hal ini menggunakan dua jenis kumpulan data yang tersedia untuk umum untuk mengevaluasi tugas ATS. Hasil studi statistik menegaskan bahwa model berbasis embedding kata mengungguli tugas peringkasan dibandingkan dengan model berbasis BOW. Secara khusus, teknik pembelajaran ansambel dengan representasi Word2Vec melampaui semua model yang diselidiki. Setiap hari, analisis kuantitatif berusaha keras untuk mendapatkan akurasi yang lebih baik dari model pembelajaran mesin mereka untuk memperkirakan pengembalian dari saham. Support Vector Machine (SVM) dan model regresi berbasis Random Forest dikenal karena keefektifannya dalam memprediksi harga penutupan secara akurat [9]. Dalam pekerjaan ini, kami mengusulkan teknik untuk menganalisis dan memprediksi harga saham perusahaan menggunakan algoritma yang disebutkan di atas sebagai ansambel. Kumpulan data dari Bursa Efek Nasional India yang berisi informasi harga pasar dasar diproses sebelumnya untuk menyertakan indikator teknis terkemuka yang terkenal sebagai fitur. Pemilihan fitur, yang memberi peringkat fitur berdasarkan tingkat pengaruhnya pada harga penutupan akhir telah dimasukkan untuk mengurangi ukuran set data pelatihan. Selain itu, kami mengevaluasi efektivitas mempertimbangkan opini publik suatu perusahaan dengan menggunakan analisis sentimen. Dengan menggunakan model Word2Vec yang terlatih, kiriman dengan tanda pagar khusus perusahaan dari Twitter diklasifikasikan sebagai positif atau negatif. Model ansambel yang kami usulkan kemudian dilatih pada kumpulan data baru yang menggabungkan data indikator teknis bersama dengan jumlah agregat tweet positif / negatif perusahaan dari waktu ke waktu. Eksperimen kami menunjukkan bahwa dalam beberapa skenario, model ansambel berkinerja lebih

baik daripada model konstituen dan sangat bergantung pada sifat dan ukuran data pelatihan. Namun, menggabungkan data indikator teknis dengan jumlah tweet positif / negatif agregat memiliki efek yang dapat diabaikan pada kinerja model ansambel.

Analisis sentimen Twitter memberi organisasi pemantauan waktu nyata atas perasaan publik terhadap produk dan acara tertentu yang terkait dengannya. Sebagian besar penelitian yang ada difokuskan pada ekstraksi fitur sentimen melalui analisis fitur leksikal dan sintaksis yang diekspresikan secara eksplisit melalui kata-kata, emotikon, tanda seru, dan lainnya [1]. Pengklasifikasi pembelajaran mesin tunggal biasanya digunakan oleh pendekatan ini untuk klasifikasi sentimen tweet. Pada penelitian ini memperkenalkan model fitur semantik yang memanfaatkan statistik kejadian bersama dan hubungan semantik kontekstual laten antar kata dalam tweet. Fitur semantik ini digabungkan dengan fitur skor polaritas sebelumnya dan fitur n-gram sebagai kumpulan fitur sentimen tweet. Kumpulan fitur digabungkan ke dalam pengklasifikasi ensemble yang dibentuk oleh Support Vector Machines, Regresi Logistik, dan Random Forest untuk melatih dan memprediksi label klasifikasi sentimen. Lima set data Twitter digunakan dalam evaluasi kami dan performa model dibandingkan dengan model kata n-gram sebagai baseline. Hasil eksperimen menunjukkan bahwa metode kami memiliki kinerja yang unggul dalam hal akurasi klasifikasi sentimen dan pembelajaran ensemble dapat digunakan untuk meningkatkan akurasi klasifikasi.

Dengan meningkatnya layanan mikroblog dalam beberapa tahun terakhir, klasifikasi sentimen mikroblog telah banyak dipelajari dan diterapkan di banyak bidang seperti pemantauan opini publik, evaluasi komoditas, dan prakiraan pasar. Metode ensemble telah banyak digunakan dalam konstruksi fitur dan tahap klasifikasi sentimen mikroblog karena kinerjanya yang sangat baik [3]. Dalam konstruksi fitur, sebagian besar peneliti menggunakan penggabungan fitur atau metode ansambel untuk menggabungkan fitur yang berbeda, sedangkan penggabungan dua metode diabaikan. Dalam klasifikasi, sebagian besar metode klasifikasi ensemble menggabungkan pengklasifikasi berdasarkan voting mayoritas atau rata-rata tertimbang, dan metode tersebut tidak sepenuhnya mempertimbangkan perbedaan informasi yang terdapat dalam pengklasifikasi. Dalam makalah ini, metode pembelajaran ensemble multi-view baru diusulkan untuk menggabungkan informasi yang terkandung dalam fitur yang berbeda untuk klasifikasi sentimen mikroblog yang lebih baik. Metode ini terdiri dari dua tahap: tahap fusi lokal dan tahap fusi global. Dalam tahap fusi lokal, fitur mentah dan fitur penggabungan digunakan untuk membuat pengklasifikasi dasar, dan pengklasifikasi

dasar ini digabungkan menjadi lima grup pengklasifikasi untuk mengidentifikasi sentimen mikroblog di antara semua informasi fitur mentah. Dalam tahap fusi global, grup pengklasifikasi ini dengan tampilan global diintegrasikan lebih lanjut untuk membuat prediksi yang lebih akurat dan komprehensif. Dua set data tolok ukur mikroblog publik yang disediakan oleh Sina Weibo digunakan dalam eksperimen, dan hasil eksperimen menunjukkan bahwa metode kami mengungguli metode perbandingan lainnya dalam mengidentifikasi polaritas pos mikroblog.

Analisis sentimen juga diterapkan untuk mengidentifikasi komentar negatif dari ojek online. Hal ini diterapkan pada perusahaan gojek [10]. Metode klasifikasi sentimen yang digunakan pada penelitian ini menggunakan SVM dan Decision Tree. Klasifikasi sentimen ini ditujukan pada ulasan yang ada pada Google Playstore. Kelas sentimen yang dijadikan label hanya dua yaitu positif dan negatif. Uji akurasi menunjukkan bahwa penggunaan SVM mencapai 90.20% dan menggunakan Decision Tree sebesar 89.80%. Penggunaan Algoritma SVM juga dapat diterapkan dalam melakukan analisis sentimen pada Twitter. Penelitian ini melakukan analisis sentimen terhadap opini para pengguna Bukalapak [11]. Polaritas sentimen yang digunakan dalam bentuk positif, negatif netral. Proses penentuan multi kelas pada penelitian ini menggunakan one Againts One dan One Againts All dengan representasi dokumen menggunakan TF IDF. Hasil uji akurasi menggunakan K Fold menunjukkan 93% dengan menggunakan metode SVM.

Metode Penelitian

Penelitian ini bersifat eksperimental dimana terdapat parameter yang diujikan. Parameter yang diuji ini merupakan jenis-jenis Ensemble Learning yang dikombinasikan dengan pendekatan Word Embedding. Hasil akurasi tertinggi dan waktu proses pelatihan data paling rendah adalah tujuan utama dalam penelitian ini. Jadi, dalam hasil uji akurasi ini, apabila kombinasi metode tersebut menghasilkan parameter yang ditentukan tersebut maka hal ini menjadi rekomendasi untuk diterapkan.

Model pembelajaran tradisional mengasumsikan bahwa data kelas terdistribusi dengan baik. Namun, dalam banyak domain data dunia nyata, datanya tidak seimbang kelas, di mana kelas minoritas diwakili oleh hanya beberapa tupel. Ini dikenal sebagai masalah ketidakseimbangan kelas [12]. Hal ini merupakan isu utama dalam proses mencari model. Apabila dioptimalkan kinerja pengklasifikasi maka hasilnya masih kurang relevan. Oleh karena itu peran ensemble learning diperlukan untuk mengatasi ketidakseimbangan kelas ini.

Selain itu, akurasi dari pengklasifikasi juga didukung dengan penyediaan data latih atau representasi teks dengan metode word embedding. Metode ini menyusun fitur dari kata-kata pada

corpus berupa matriks. Hal ini merupakan perbaikan dari bentuk sebelumnya yaitu Bag Of Word yang disusun menggunakan pembobotan TF-IDF. Urutan kata dalam word embedding benar-benar diperhatikan sehingga model yang terbentuk memiliki akurasi yang lebih bagus.

Pada penelitian ini memuat beberapa kegiatan-kegiatan. Beberapa kegiatan yang dilalui pada penelitian ini ditunjukkan pada Gambar 1. Kegiatan penelitian ini diawali dengan studi literatur dimana mencari sumber referensi baik dari makalah maupun buku teks yang berkaitan dengan bidang penambangan teks. Langkah selanjutnya adalah pengumpulan data. Data yang digunakan bersifat sekunder. Data ini diperoleh dari situs kaggle dengan alamat:

<https://www.kaggle.com/crowdfLOWER/twitter-airline-sentiment>

Data ini memuat dataset dari twitter yang berkaitan dengan penerbangan yang terjadi bulan Februari 2015 di Amerika Serikat. Kumpulan data ini perlu dianalisis untuk dilakukan pemrosesan teks.



Gambar 1. Alur Penelitian

Kemudian, terdapat kegiatan merancang dan membangun sistem analisis sentimen. Tujuan membangun sistem ini adalah menghasilkan prototipe sistem analisis sentimen dimana merupakan usulan terhadap kontribusi hasil penelitian ini. Prototipe ini perlu diuji dengan menyiapkan test case atau rencana pengujannya. Rencana pengujian ini ditentukan parameter uji akurasi dengan Word Embedding yang dikombinasikan dengan Bagging, Boosting, dan Adaboost.

Hasil pengujian tersebut menandai berakhirnya penelitian ini. Namun, perlu adanya evaluasi hasil penelitian supaya ada masukan terhadap perbaikan penelitian ini sehingga penyelenggaraan berikutnya akan lebih baik. Akhirnya, luaran utama dari penelitian adalah publikasi karya ilmiah.

Pada penelitian ini terdapat usulan arsitektur dari sistem analisis sentimen yang ditunjukkan pada Gambar 2. Usulan ini diawali dengan pengambilan dataset di situs kaggle. Kemudian dilanjutkan dengan pemrosesan teks yang terdiri dari tokenizing, stop word removal, stemming dan labeling. Hal ini dilanjutkan dengan ekstraksi fitur untuk membentuk

vektor kata dengan Word Embedding. Vektor kata yang terbentuk tersebut dijadikan data latih dan dilakukan pelatihan data dengan ensemble Learning yang dilakukan dengan metode bagging, boosting, dan adaboost. Hasil penggunaan ketiga metode tersebut dibandingkan dan dicari yang memiliki akurasi tertinggi dan waktu pelatihan data yang singkat.

Penelitian ini diharapkan dapat meningkatkan akurasi dari klasifikasi teks pada analisis sentimen Twitter. Kontribusi penggunaan ensemble learning diharapkan mampu mempertahankan penggunaan pengklasifikasi yang ada dengan mendayagunakan kombinasi model yang disediakan dari hasil proses metode ini. Selain itu, penyusunan representasi teks untuk data latih menggunakan word embedding juga diharapkan meminimalkan waktu proses pelatihan data dan meningkatkan akurasinya. Hal ini merupakan kontribusi utama dalam bidang analisis sentimen mendatang.

Hasil dan Pembahasan

a) Pengumpulan Data

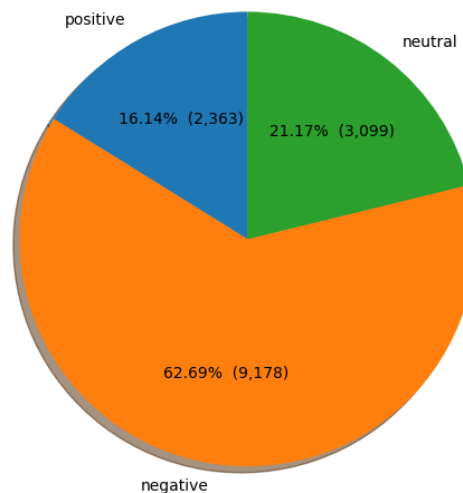
Pada penelitian ini menggunakan kumpulan data atau disebut dengan dataset. Dataset yang digunakan mengenai data tweet penerbangan di Amerika Serikat. Dataset ini dibentuk pada tahun 2015 kemudian diunggah pada situs kaggle.com dengan detail alamat URL sebagai berikut:

<https://www.kaggle.com/datasets/crowdflower/twiter-airline-sentiment>

Dataset ini memiliki kolom atau atribut sebanyak 15 dan memiliki baris data sebanyak 14.640. Dari 15 atribut tersebut yang digunakan dalam penelitian ini hanya atribut `airline_sentiment` dan `text`. `airline_sentiment` berisi sentiment dari tweet maskapai penerbangan yang memuat data positif, negatif, dan netral. Selanjutnya, atribut `text` berisi tweet dari setiap maskapai penerbangan. Jadi, data tweet tersebut telah dilabeli saat dataset didapatkan. Kedua atribut tersebut akan diekstrak menjadi beberapa fitur menggunakan metode `word2vec`. Deskripsi perbandingan sentimen dari masing-masing tweet ditunjukkan pada Gambar 3.

b) Analisis Data

Pada Gambar 3 ditunjukkan jumlah total dari setiap tweet yang memuat sentimen positif, negative, dan netral. Jumlah masing-masing sentimen tidak sama jadi dapat dikatakan bahwa jumlah data sentiment tidak berimbang. Sentimen positif memiliki 2.363 baris data sehingga persentasenya 16,14%. Sentimen negatif memiliki 9.178 baris data sehingga persentasenya 62,69%. Sentimen netral memiliki 3.099 baris data sehingga persentasenya 21,17%. Dengan demikian sentimen negative memiliki jumlah dan persentasi paling besar diantara sentiment lainnya.



Gambar 2. Perbandingan jumlah sentimen

c) Pengujian Hasil

Awal percobaan pada penelitian ini adalah melakukan pelatihan data dimana terjadi ekstraksi tweet menjadi beberapa fitur menggunakan pendekatan *Bag Of Word (BOW)* dengan metode TFIDF. Percobaan ini menghasilkan fitur sejumlah 10.029 atribut. Fitur yang dihasilkan ini berdasarkan pemrosesan teks yang dilakukan pembobotan dengan metode TFIDF. Fitur ini digunakan untuk proses pelatihan data menggunakan beberapa metode klasifikasi seperti Naïve Bayes, SVM, Logistic Regression dan Pohon Keputusan. Hasil dari proses pelatihan data tersebut menghasilkan model dari masing-masing metode klasifikasi. Model ini diuji menggunakan metode *cross validation* yang ditunjukkan pada Gambar 1.

Tabel 1. Evaluasi model hasil pelatihan data dengan BOW-TFIDF

No.	Metode	Akurasi	Kemunculan	Presisi	Skor F1	Waktu Proses (ms)
1	Naive Bayes	0.67978	0.43709	0.74420	0.44161	2.03436
2	Logistic Regression	0.76735	0.64474	0.74250	0.67725	18.75231
3	Decision Tree	0.68661	0.60717	0.61280	0.61065	25.33119
4	SVM Linear	0.76926	0.66308	0.73314	0.68845	174.50046
5	SVM RBF	0.76175	0.62454	0.74676	0.65800	257.52735
6	SVM Polynomial	0.62691	0.33333	0.20897	0.25689	194.33711

Pada tabel 1 ditunjukkan bahwa data latih telah dilakukan pelatihan data dengan pengklasifikasi yang berbeda-beda. Namun, pendekatan ekstraksi data masih menggunakan BOW dengan metode TFIDF. Metode pengklasifikasi menggunakan Naive Bayes, Logistic Regression, Decision Tree, SVM Linear, SVM RBF, dan SVM Polynomial. Hal ini menghasilkan data yang dikelompokkan berdasarkan akurasi, kemunculan (Recall), Presisi, dan Skor F1 yang berbeda-beda. Data tersebut diambil saat melakukan pelatihan data menggunakan pustaka program Scikit Learn yang berbasis Pemrograman

Python. Pustaka program tersebut digunakan untuk melakukan proses penambangan data berbasis machine learning.

Berdasarkan data pada Tabel 1, metode SVM Linear memiliki akurasi tertinggi daripada metode lainnya yaitu sebesar 0,76926 atau dapat dibulatkan dan diubah dalam bentuk prosentasi menjadi 76,92%. Hal ini sebanding dengan nilai kemunculan (recall) dan skor F1 yang juga tertinggi dengan masing-masing nilai 0.66308 dan 0.68845. Selain itu, terdapat metode SVM dengan kernel Polynomial memiliki akurasi dan kemunculan paling rendah yaitu 0,62691 atau 62,69% dengan kemunculan sebesar 0,33333. Selanjutnya, terdapat metode Naive Bayes yang memiliki waktu proses paling cepat yaitu sebesar 2,03436 ms. Selain itu, metode ini memiliki nilai presisi paling tinggi yaitu 0.74420 dengan akurasi sebesar 0.67978 atau 67,98%. Karena memiliki waktu proses paling cepat maka metode Naive Bayes ini digunakan untuk pengklasifikasi dengan pendekatan Word Embedding menggunakan WOrd2Vec dan penggunaan metode Ensemble Learning.

Pada Tabel 2 ditunjukkan data hasil penggunaan metode Bayes dengan pendekatan Word Embedding menggunakan metode WOrd2Vec dan metode Ensemble Learning. Adanya metode Word2Vec ini dapat menurunkan waktu proses dimana pada TF-IDF dihasilkan 2.03436 ms sedangkan penggunaan WOrd2Vec menghasilkan 0.07006 ms. Selanjutnya, peningkatan akurasi terjadi saat penggunaan Naive Bayes dan WOrd2Vec dilakukan metode Adaboost. Hal ini menghasilkan akurasi sebesar 0.69201 atau 69,2% sehingga akurasi meningkat sebesar 1,22%.

Tabel 2. Evaluasi model dengan Naive Bayes, Word Embedding, dan Ensemble Learning

No	Metode	Akurasi	Kemunculan	Presisi	Skor F1	Waktu Proses (ms)
1	Naive Bayes, Word2Vec	0.67760	0.45863	0.62538	0.47277	0.07006
2	Naive Bayes, Word2Vec Adaboost	0.69201	0.50058	0.64876	0.52387	8.2629
3	Naive Bayes, Word2Vec, Bagging	0.67739	0.45784	0.62560	0.47182	0.42504

Kesimpulan dan Saran

Pada penelitian ini dapat disimpulkan bahwa penggunaan metode Naive Bayes memiliki waktu proses paling cepat saat digunakan untuk melakukan pelatihan data dimana ekstraksi data menggunakan TF-IDF. Hal ini dibandingkan dengan metode Logistic Regression, Decision Tree, SVM Linear, SVM RBF, dan SVM Polynomial. Selanjutnya, metode Naive Bayes ini dilakukan Ensemble Learning menggunakan metode Adaboost yang

menghasilkan nilai akurasi sebesar 0.69201 atau 69,2% dimana hal ini telah meningkatkan akurasi sebesar 1,22% dengan waktu proses yang meningkat juga yaitu sebesar 6,23ms. Jadi, penggunaan Word2Vec dapat menurunkan waktu proses pelatihan data sedangkan metode Adaboost dapat meningkatkan akurasi model yang dihasilkan.

Daftar Pustaka

- [1] J. Zhao, "Combing semantic and prior polarity features for boosting twitter sentiment analysis using ensemble learning," *Proc. - 2016 IEEE 1st Int. Conf. Data Sci. Cyberspace, DSC 2016*, pp. 709–714, 2017, doi: 10.1109/DSC.2016.124.
- [2] M. Rathi, A. Malik, D. Varshney, R. Sharma, and S. Mendiratta, "Sentiment Analysis of Tweets Using Various Machine Learning Techniques," *2018 Int. Conf. Adv. Comput. Telecommun. ICACAT 2018*, pp. 2–4, 2018, doi: 10.1109/ICACAT.2018.8933612.
- [3] X. Ye, H. Dai, L. an Dong, and X. Wang, "Multi-view ensemble learning method for microblog sentiment classification," *Expert Syst. Appl.*, vol. 166, p. 113987, 2021, doi: 10.1016/j.eswa.2020.113987.
- [4] S. M. Rezaenia, R. Rahmani, A. Ghodsi, and H. Veisi, "Sentiment analysis based on improved pre-trained word embeddings," *Expert Syst. Appl.*, vol. 117, pp. 139–147, 2019, doi: 10.1016/j.eswa.2018.08.044.
- [5] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining Practical Machine Learning Tools and Techniques Fourth Edition*, 4th ed. Montreal, Canada: Morgan Kaufman, 2011. doi: 10.1016/C2009-0-19715-5.
- [6] C. Dedhia and J. Ramteke, "Ensemble model for Twitter sentiment analysis," *Proc. Int. Conf. Inven. Syst. Control. ICISC 2017*, pp. 1–5, 2017, doi: 10.1109/ICISC.2017.8068711.
- [7] L. Chen and H. Deng, "Predicting User Retweeting Behavior in Social Networks with a Novel Ensemble Learning Approach," *IEEE Access*, vol. 8, pp. 148250–148263, 2020, doi: 10.1109/ACCESS.2020.3015397.
- [8] N. Alami, M. Meknassi, and N. En-nahnahi, "Enhancing unsupervised neural networks based text summarization with word embedding and ensemble learning," *Expert Syst. Appl.*, vol. 123, pp. 195–211, 2019, doi: 10.1016/j.eswa.2019.01.037.
- [9] U. Pasupulety, A. Abdullah Anees, S. Anmol, and B. R. Mohan, "Predicting stock prices using ensemble learning and sentiment analysis," *Proc. - IEEE 2nd Int. Conf. Artif. Intell. Knowl. Eng. AIKE 2019*, pp. 215–222, 2019, doi: 10.1109/AIKE.2019.00045.
- [10] K. A. Rokhman, B. Berlilana, and P. Arsi, "Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online," *J. Inf. Syst. Manag.*, vol. 3, no. 1, pp. 1–7, Jan. 2021, doi: 10.24076/JOISM.2021v3i1.341.
- [11] Z. Alhaq, A. Mustopa, S. Mulyatun, and J. D. Santoso, "Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter," *J. Inf. Syst. Manag.*, vol. 3, no. 2, pp.

- 44–49, Jul. 2021, doi:
10.24076/joism.2021v3i2.558.
- [12] S. Sumathi and S. N. Sivanandam, *Introduction to Data Mining and its Applications*, vol. 29. 2006. doi: 10.1007/978-3-540-34351-6.