

Contents list available at www.jurnal.unimed.ac.id

CESS (Journal of Computing Engineering, System and Science)

journal homepage: <https://jurnal.unimed.ac.id/2012/index.php/cess>



Komparasi Algoritma Machine Learning Dalam Menganalisa Sentimen Opini Publik Terhadap Kenaikan Harga BBM

Comparison of Machine Learning Algorithms in Analyzing Public Opinion Sentiments Against Fuel Price Increases

Hanif Wira Saputra¹, Rahmaddeni^{2*}, Fazri³

^{1,2,3}Teknik Informatika, STMIK Amik Riau

Jl. Purwodadi, Indah No.KM. 10, Sidomulyo Barat, Kecamatan Tampan, Kota Pekanbaru, Provinsi Riau, Indonesia telp. (0761) 589561

email: ¹2110031802063@sar.ac.id, ²rahmaddeni@sar.ac.id, ³2210031802146@sar.ac.id

ABSTRAK

Abstrak *Twitter* merupakan *platform* media sosial yang cukup banyak digunakan masyarakat dunia khususnya masyarakat di Indonesia. *Twitter* menjadi salah satu media sosial yang memberikan suatu informasi, salah satunya yaitu kenaikan harga minyak mentah yang tercatat *105US dollar* per barel. Kenaikan harga BBM memiliki dampak yang buruk terhadap masyarakat sehingga menyebabkan terjadi pro dan kontra. Berdasarkan permasalahan tersebut, penulis bertujuan melakukan perbandingan performa algoritma *artificial neural network* dan *naïve bayes* untuk mengetahui model yang terbaik terhadap analisis sentiment kenaikan BBM. Data yang digunakan berjumlah 1000 dataset dalam bentuk *text document* dengan pelabelan menggunakan *lexicon* dan *split* data 90:10, 80:20, 70:30, dan 60:40 sebagai perbandingan nilai presisi. Penerapan vektorisasi kata memanfaatkan TF-IDF dalam memberikan bobot nilai pada setiap kata. Berdasarkan hasil percobaan yang sudah dilakukan, bahwasanya didapati algoritma terbaik dengan menggunakan *artificial neural network* mampu menghasilkan nilai akurasi sebesar 87% terhadap 1000 data sentimen opini publik kenaikan BBM. Berdasarkan hasil evaluasi, model yang dibangun dapat mengategorikan sentimen opini publik ke dalam sentimen positif, sentimen negatif, dan sentimen netral secara otomatis serta polaritas sentimen publik cenderung positif terhadap isu kenaikan BBM yang terjadi.

Kata Kunci: Analisis Sentimen, Artificial Neural Network, Lexicon, Naïve Bayes, TF-IDF

ABSTRACT

Twitter is a social media platform that is quite widely used by the world community, especially people in Indonesia. Twitter is one of the social media that provides information, one of which

*Penulis Korespondensi:

email: rahmaddeni@sar.ac.id

is the increase in the price of crude oil which was recorded at 105 US dollars per barrel. The increase in fuel prices has a negative impact on society, causing pros and cons. Based on these problems, the authors aim to compare the performance of the artificial neural network and naïve Bayes algorithms to determine the best model for sentiment analysis of fuel price hikes. The data used amounted to 1000 datasets in the form of text documents with labeling using the lexicon and split data 90:10, 80:20, 70:30 and 60:40 as a comparison of precision values. The application of word vectorization utilizes TF-IDF in assigning a weight value to each word. Based on the results of the experiments that have been carried out, it is found that the best algorithm using an artificial neural network is capable of producing an accuracy value of 87% for 1000 data on public opinion sentiment on fuel price hikes. Based on the evaluation results, the model built can categorize public opinion sentiment into positive sentiment, negative sentiment, and neutral sentiment automatically and the polarity of public sentiment tends to be positive towards the issue of the fuel price increase that occurred.

Keywords: *Sentiment Analysis, Artificial Neural Network, Lexicon, Naïve Bayes, TF-IDF*

1. PENDAHULUAN

Pengguna media sosial mengakses akun media sosial mereka masing-masing selama kurang lebih satu jam setiap harinya [1]. Saat ini ada banyak jenis media sosial yang tersebar diseluruh dunia dan dapat digunakan oleh segala khalayak dengan mudah, salah satunya adalah Twitter. *twitter* merupakan *platform* media sosial yang cukup banyak digunakan masyarakat dunia khususnya masyarakat di Indonesia, yaitu sebanyak 56% [2]. Twitter dipergunakan oleh masyarakat untuk membagikan informasi dalam bentuk teks, foto, video dan mendiskusikan hal menarik. Contohnya masalah yang lagi trending topik di *twitter* yaitu Kenaikan harga BBM.

Pemerintah menetapkan kebijakan kenaikan harga Bahan Bakar Minyak (BBM) di Indonesia. Kebijakan tersebut merupakan imbas dari peperangan yang terjadi di Eropa antara Rusia dan Ukraina yang menyebabkan terjadinya lonjakan harga minyak dunia [3]. Harga minyak dunia mencapai *105 US dollar* per barel dengan *kurs* Rp 14.700 per *dollar* AS. Akibatnya, terjadi kegoncangan ekonomi secara signifikan terhadap harga minyak bumi, tidak terkecuali di Indonesia. Embargo terhadap perdagangan minyak Rusia memicu *supply shock*, sehingga menimbulkan konsekuensi kenaikan harga di pasar global [4][5]. Meningkatnya suku bunga dan likuiditas tentu saja berdampak pada negara berkembang termasuk Indonesia sehingga menyebabkan lonjakan anggaran subsidi yang telah ada.

Selama beberapa hari sejak ditetapkannya kenaikan harga BBM, *twitter* dipenuhi Pro dan kontra di masyarakat. ini menjadi menarik untuk diteliti dengan mengetahui tanggapan masyarakat. Berbagai metode yang dapat digunakan untuk menganalisis tanggapan masyarakat tersebut khususnya di *twitter*. Banyak metode analisis yang dapat digunakan untuk menganalisis tanggapan masyarakat salah satunya analisis sentimen. Menurut [6], analisis sentimen adalah suatu cara untuk menilai opini tertulis atau lisan untuk menentukan suatu opini bersifat positif, negatif atau netral. Metode analisis sentimen mencari polaritas dari data khususnya melalui internet dengan melakukan analisis[7]. Analisis sentimen dilakukan untuk mengidentifikasi opini atau tren opini tentang masalah atau topik seseorang, apakah mereka memiliki opini atau opini negatif atau positif. Seperti yang telah dilakukan banyak peneliti di berbagai penjuru dunia [8][9][10][11].

Penelitian yang dilakukan oleh kurniasih dan suseno (2022), bertujuan untuk melakukan analisis terhadap sentimen masyarakat Indonesia pada kebijakan pemerintah terkait

pemberian BSU pada kenaikan harga BBM. Hasil analisis menunjukkan bahwa pada kata kunci BBM sentimen positif yang didapatkan yaitu 28,2% dan pada sentimen negatif yaitu 71,8%. Pada kata kunci BSU, sentimen positif sebesar 65,2% dan sentimen negatif 34,8%. Pada tingkat akurasi dengan metode ini mendapatkan hasil 82,64% dan presisi sebesar 92,89% [5].

Penelitian mengenai analisis sentimen masyarakat terhadap kenaikan harga BBM pernah dilakukan oleh Mujahidin, dkk [4]. Penelitian ini dilakukan dengan melihat komentar yang dihasilkan platform *youtube* dengan menggunakan metode *Gaussian Naïve Bayes*. Hasil penelitian tersebut menunjukkan nilai akurasi sebesar 74%. pada dataset tanpa filtering *stopword* dan model *fast textsize* 100. Berdasarkan hasil evaluasi, sistem yang dibangun dapat mengklasifikasikan sentimen atau opini publik ke dalam sentimen positif dan sentiment negatif secara otomatis.

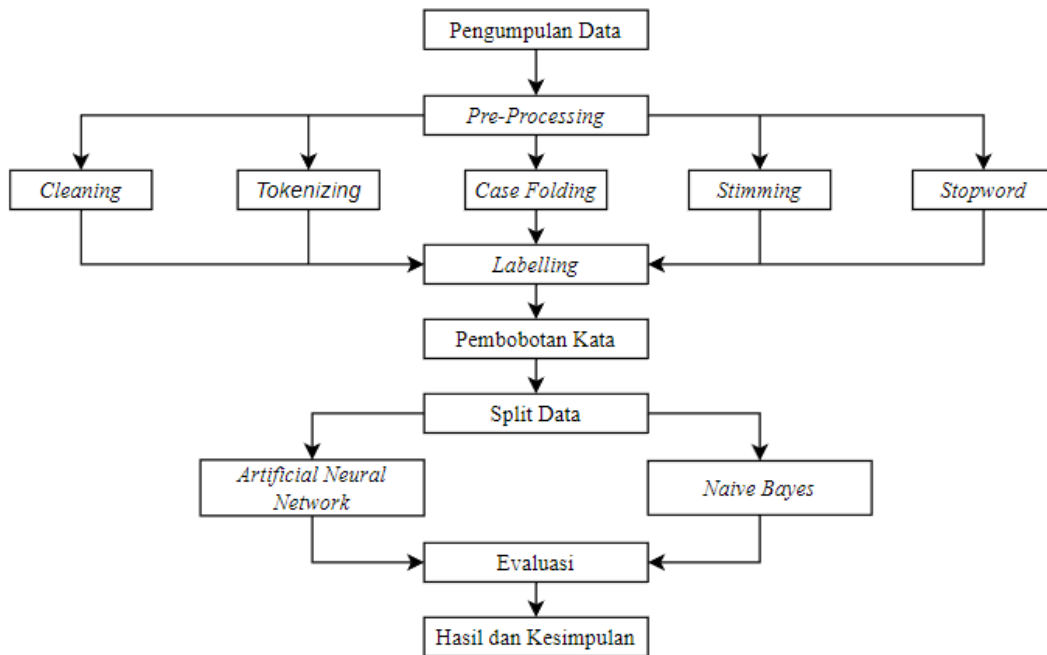
Sementara itu, dari segi dampak ekonomi, penelitian mengenai dampak kenaikan BBM telah diteliti oleh [3]. Berdasarkan hasil penelitian yang telah dilakukan, diperoleh kesimpulan bahwa kenaikan BBM yang terjadi di Indonesia akan berdampak pada inflasi sehingga akan berpengaruh terhadap fundamental makro ekonomi Indonesia.

Penelitian lainnya dilakukan oleh Fungsiawan (2022) yang menyoroti kenaikan tarif ojek online yang merupakan imbas dari kenaikan BBM [12]. Dalam penelitian tersebut dinyatakan bahwa kenaikan tarif ojek online akan mendorong masyarakat untuk pindah ke moda transportasi lain bahkan kendaraan pribadi. Jika hal tersebut terus berlanjut maka dapat menambah tingkat kemacetan dan kerugian ekonomi di Indonesia.

Berdasarkan penelitian yang sudah dilakukan para peneliti sebelumnya, Peneliti tertarik untuk melakukan penelitian mengenai analisis sentimen opini publik di *twitter* terhadap kenaikan harga BBM dengan membandingkan performa model algoritma *artificial neural network dan naïve bayes*. Penelitian ini bertujuan untuk mengetahui polaritas opini pengguna *twitter* di indonesia apakah masyarakat lebih beropini positif, negatif atau netral terhadap kebijakan pemerintah tentang kenaikan harga BBM. Penelitian ini menganalisis opini publik dengan *tweet* yang di posting pada media sosial *twitter*. Data yang merupakan *tweet* opini publik tersebut dikelompokkan dengan label positif, negatif, maupun netral menggunakan metode *artificial neural network dan naïve bayes*.

2. METODE PENELITIAN

Dalam penelitian ini dimulai dengan beberapa Langkah, mengumpulkan data yang diperlukan untuk proses analisis. Mendistribusikan data latih dan data uji menggunakan rasio 90:10, 80:20, 70:30, dan 60:40. Lalu merancang model algoritma *artificial neural network dan naïve bayes*. Selanjutnya melatih dan menganalisa performa algoritma *artificial neural network dan naïve bayes* yang dalam penelitian ini. Adapun proses tahapan penelitian ini dapat ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Adapun dataset dalam Studi ini yaitu kumpulan data *tweet* kenaikan harga BBM yang diperoleh dari media sosial twitter sebanyak 1000 data. Teknik pengumpulan data dilakukan menggunakan Teknik *crawling* data *twitter* dengan menggunakan *python* dengan rentang waktu 10 desember 2022 hingga 27 desember 2022.

2.2. Pre-Processing Data

Jika ada *Pre-Processing* merupakan Langkah awal sebelum melakukan pengujian model algoritma, dataset yang digunakan akan diolah dengan membuang data yang tidak sempurna yang dapat menyebabkan gangguan pada data. *Pre-processing* bertujuan agar data siap untuk diuji pada model yang digunakan dengan menghilangkan elemen yang tidak berarti [13]. Pada penelitian ini *pre-processing* yang digunakan adalah sebagai berikut:

1. *Cleaning* data tahap menghilangkan karakter dan tanda baca yang tidak perlu pada teks. Pembersihan berfungsi untuk mengurangi noise pada dataset[14]
2. Tokenisasi adalah proses memotong atau memecah kalimat menjadi beberapa kata[14]
3. *Case Folding* adalah proses untuk mengubah semua teks dokumen menjadi huruf kecil[15]
4. *Stemming* adalah tahapan untuk membuat akhiran kata menjadi kata-kata esensial sesuai dengan aturan yang benar[14]
5. *Stopword* adalah kata umum yang biasanya muncul dalam jumlah besar dan dianggap tidak berarti[16]

2.3. Labelling

Labeling merupakan pemberian label berupa bobot pada setiap kata yang ada didalam data. pelabelan menggunakan *lexicon* untuk menentukan polaritas agar mendapatkan nilai yang pasti. Tujuan pelabelan menggunakan *lexicon* agar prosesnya menjadi cepat dan dilakukan secara otomatis [17].

2.4. Pembobotan Kata

Menghitung bobot setiap kata yang paling sering digunakan dalam pengambilan informasi. Cara ini juga efisien, mudah, dan memiliki hasil yang akurat. Ini metode yang digunakan untuk menghitung *Term Frequency (TF)* dan *Inverse Document Frequency (IDF)* untuk setiap token (kata) di setiap dokumen dalam korpus [18].

2.5. Splitting Data

Data dapat dibagi menjadi dua. Dalam hal ini, data digunakan untuk pengujian data sedangkan sisanya digunakan untuk pelatihan data. Dengan menggunakan pemisahan data, uji coba pemisahan rasio 80:20 akan dilakukan, dan pemisahan rasio dari sisa data pelatihan digunakan sebagai data pengujian. Data yang belum pernah digunakan dalam suatu penelitian tetapi berguna untuk mengevaluasi keberhasilan atau kegagalan suatu penelitian dikenal sebagai pengujian data, Sedangkan pelatihan data adalah data yang digunakan untuk melakukan suatu penelitian [19].

2.6. Artificial Neural Network

Metode ini terdiri dari beberapa komponen yaitu serangkaian *input*, *layer* dan juga *output*. *Input* hanya bisa memproses data numerik, sehingga apabila data terdiri dari grafik, gambar ataupun *signal* maka harus terlebih dahulu diubah ke data numerik. Layer terdiri dari kumpulan *neuron* yang terhubung dan dikelompokkan dalam lapisan. Lapisan tersebut antara lain *input layer*, *hidden layer* dan *output layer*. Layer bisa berupa *single* ataupun *multi-layer*. Sedangkan *output* adalah solusi dari masalah yang berbentuk data numerik [20]. Kemudian *artificial neural network* juga dapat digunakan untuk mendeteksi ketidakstabilan debit yang tidak terkait dengan bentuk model [21]. *Artificial neural network* belajar berdasarkan contoh dan performanya sangat dipengaruhi oleh kualitas dan kuantitas dari contoh. Semakin banyak jumlah contoh maka nilai kesalahannya akan semakin rendah.

2.7. Naïve Bayes

Algoritma klasifikasi yang dikenal sebagai *naïve bayes* adalah salah satu dari algoritma prediksi dan klasifikasi dimana *naïve bayes* memiliki kelebihan yang mencakup efektifitas (menemukan hasil yang valid) dan efisiensi (komputasi input sangat cepat). *Naïve bayes* dilandaskan dengan dugaan simplifikasi bahwa nilai atribut secara tentatif saling bebas jika ada output. Menurut definisi ini, *output* merupakan produk dari probabilitas individu. Salah satu kelebihan metode *naïve bayes* adalah menggunakan sejumlah data pelatihan (*Training Data*) untuk men estimasi parameter yang digunakan prosedur analisis statistik. *Naïve bayes* sering bekerja jauh meningkat umumnya dalam situasi di kehidupan nyata yang berbelit-belit ketimbang yang diharapkan [22]. Teorema bayes dapat diterapkan dengan persamaan dibawah ini.

$$P(H|X) = \frac{(P(X|H) \times P(H))}{(P(X))} \quad (1)$$

2.8. Evaluasi

Jika *Cross validation* atau dapat disebut estimasi rotasi yang merupakan suatu teknik konfirmasi model yang digunakan dalam menentukan dengan cara menerapkan hasil analisis statistik terhadap sejumlah besar data bebas. Teknik ini umumnya diaplikasikan untuk menghasilkan model prediktif dan memverifikasi ketepatan beberapa model saat digunakan. *K-fold cross validation* ialah satu-satunya metode estimasi rotasi yang mengubah data menjadi

k bagian set data yang identik [23]. Persamaan berikut digunakan untuk mengukur kinerja prediksi tiap-tiap model algoritma yang digunakan.

$$accuracy = \frac{\sum \text{klasifikasi benar}}{\sum \text{data uji}} \times 100\% \quad (2)$$

3. HASIL DAN PEMBAHASAN

Pada hasil dan pembahasan penelitian ini analisis sentimen dilakukan beberapa tahap yaitu pengumpulan data, *pre-processing*. Setelah dilakukan *pre-processing* maka akan dilakukan klasifikasi dengan perbandingan menggunakan algoritma *artificial neural network* dan *naïve bayes*. Pada tahap evaluasi menggunakan beberapa perhitungan yang digunakan yaitu *Accuracy* dan *cross validation*. Berikut rincian tahap dalam menganalisis sentimen opini publik terhadap kenaikan BBM:

3.1 Pengumpulan Data

Data diambil dari tweet menggunakan teknik *crawling* data yang berjumlah 1000 *tweet* dengan periode 10 desember 2022 sampai 27 desember 2022 pada media sosial *twitter*. Data tersebut termasuk opini di media massa *twitter* ataupun respons dari publik terhadap kenaikan harga BBM. Data yang didapat yaitu *tweet* yang berisi name merupakan nama akun *twitter* dan content merupakan opini dari akun *twitter*. Berikut pengumpulan data yang di ambil menggunakan teknik *crawling* data ditunjukkan pada Tabel 1.

Tabel 1. Hasil Pengumpulan Data

| Nama | Tweet |
|---------------|---|
| Mas4ziz | Dengan kenaikan harga BBM yang menyebabkan meroketnya harga kebutuhan sembako di masyarakat |
| inicialA_ | @yusuf_dumdum @habibthink Kenaikan BBM kemarin aja pada jerit semua |
| missdiah2 | @jokowi Kenaikan harga, ppn, bbm, naik nya tidak setipis amplop yang kau beri, Pak Presiden |
| Satsamaptakbr | Polsek Bentian Besar Bagikan Sembako Untuk Warga Terdampak Kenaikan BBM |
| jawapos | Pemkab Dairi Salurkan BLT Sebesar Rp450 Ribu Bagi Warga Terdampak Kenaikan BBM |

3.2 Pre-Processing

Setelah melakukan pengumpulan data, data akan melalui tahap *pre-processing* dengan membuang data yang tidak sempurna sebelum melakukan pengujian model. Terdapat 5 tahapan *pre-processing* yaitu *cleaning data*, *tokenizing*, *case folding*, *stimming*, *stopword*. Berikut hasil *pre-processing* ditunjukkan pada Tabel 2.

Tabel 2. Hasil Pengumpulan Data

| Pre-Processing | Hasil |
|----------------|--|
| Tweet awal | @jokowi Kenaikan harga, ppn, bbm, naik nya tidak setipis amplop yang kau beri, Pak Presiden |
| Cleaning | jokowi Kenaikan harga, ppn, bbm, naik nya tidak setipis amplop yang kau beri, Pak Presiden |
| Tokenizing | 'jokowi' 'kenaikan' 'harga' 'ppn' 'bbm' 'naik' 'nya' 'tidak' 'setipis' 'amplop' 'yang' 'kau' 'beri' 'pak' 'presiden' |
| case folding | jokowi kenaikan harga, ppn, bbm, naik nya tidak setipis amplop yang kau beri, pak presiden |
| Stimming | 'jokowi' 'naik' 'harga' 'ppn' 'bbm' 'setip' 'amplop' 'kau' 'presiden' |

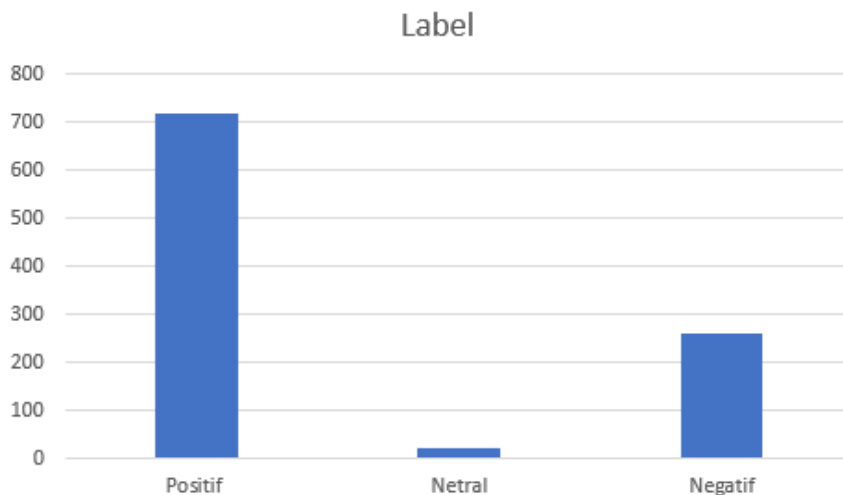
3.3 Labelling

Setelah melakukan *pre-processing* langkah selanjutnya *labelling* data. pelabelan dengan popularitas positif, negatif, dan netral dilakukan menggunakan *lexicon*. Berikut pelabelan data ditunjukkan pada Tabel 3.

Tabel 3. Pelabelan Data

| Pre-Processing | Popularitas Skor | Popularitas |
|---|------------------|-------------|
| ['naik','harga','sebab','butuh','simpul','dapat'] | 8 | Positif |
| ['naik','harga','tidak','tipis','amplop','beri'] | 4 | Positif |
| ['naik','harga','pengaruh','paksa','minum'] | 1 | Positif |
| ['naik','berapa','syarat','dampak','telah','usaha'] | -2 | Negatif |
| ['dampak','bantuan'] | -9 | Negatif |

Popularitas positif dan negatif ditentukan dari total popularitas skor yang digabungkan, jika ≥ 0 disebut label positif, jika < 0 maka disebut label negatif dan jika $= 0$ maka disebut label netral. Hasil labelling data ditunjukkan pada Gambar 2.



Gambar 2. Hasil Pelabelan Data

3.4 Pembobotan Kata

Setelah melakukan pelabelan data, pada tahap ini dilakukan pembobotan kata dimana merubah semua data menjadi bentuk angka dan memberikan nilai bobot disetiap kata. Berikut pembobotan kata ditunjukkan pada Gambar 3.

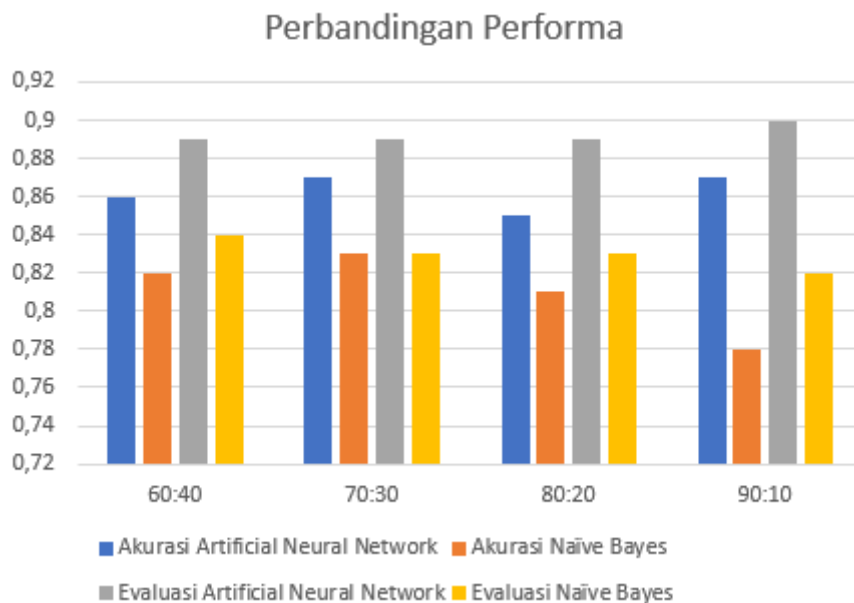
1000 Number of reviews has 2575 words

| | |
|-------------|----------------------|
| (0, 128) | 0.4191770121731385 |
| (0, 905) | 0.4191770121731385 |
| (0, 2189) | 0.40704474338166813 |
| (0, 1424) | 0.22758579168239157 |
| (0, 2134) | 0.30220343467642086 |
| (0, 456) | 0.317628267249666 |
| (0, 2021) | 0.3879014435390901 |
| (0, 2098) | 0.20769898411353008 |
| (0, 298) | 0.06768306069878494 |
| (0, 827) | 0.19014204011067315 |
| (0, 1533) | 0.06741310799907076 |
| (1, 1887) | 0.34708750197377614 |
| (1, 1083) | 0.446730540958422 |
| (1, 170) | 0.446730540958422 |
| (1, 2164) | 0.446730540958422 |
| (1, 1878) | 0.42162750637597446 |
| (1, 1013) | 0.2955206813054529 |
| (1, 298) | 0.06297217934580498 |
| (1, 827) | 0.0884538799472577 |
| (1, 1533) | 0.06272101591368806 |
| (2, 308) | 0.526543841608869 |
| (2, 1487) | 0.5000225385195014 |
| (2, 1699) | 0.5168306949202388 |
| (2, 1777) | 0.3802421242333374 |
| (2, 298) | 0.09376873274565006 |
| : | : |
| (998, 1533) | 0.060828683573912766 |
| (999, 1133) | 0.3096608267735618 |
| (999, 16) | 0.23548156631671202 |
| (999, 2199) | 0.23051387284634503 |
| (999, 700) | 0.18063041487559273 |

Gambar 3. Pembobotan Kata

3.5 Hasil Perbandingan

Langkah terakhir melakukan implementasi data dengan membandingkan performa model algoritma *artificial neural network* dan *naïve bayes* yang diuji berdasarkan akurasi dan hasil evaluasi. Berikut perbandingan performa terhadap empat rasio pemisahan data yaitu 60:40, 70:30, 80:20 dan 90:10 ditunjukkan pada Gambar 4.



Gambar 4. Perbandingan performa

Setelah dilakukan pengujian, hasil performa membandingkan model algoritma *artificial neural network* dan *naïve bayes* dirangkum dalam bentuk tabel perbandingan. Berikut hasil perbandingan model algoritma *artificial neural network* dan *naïve bayes* ditunjukkan pada Tabel 4.

Tabel 4. Hasil performa algoritma

| Split data | Akurasi | | Evaluasi | |
|------------|----------------------------------|--------------------|----------------------------------|--------------------|
| | <i>Artificial neural network</i> | <i>Naïve bayes</i> | <i>Artificial neural network</i> | <i>Naïve bayes</i> |
| 60:40 | 86% | 82% | 89% | 84% |
| 70:30 | 87% | 83% | 89% | 83% |
| 80:20 | 85% | 81% | 89% | 83% |
| 90:10 | 87% | 78% | 90% | 82% |

Dari grafik batang dan tabel hasil perbandingan di atas dapat dibuktikan pada penelitian ini bahwa algoritma *artificial neural network* memiliki nilai presisi akurasi yang lebih besar dari pada algoritma *naïve bayes* dengan selisih 4-9% pada tiap split data serta nilai evaluasi dengan selisih 5-8% pada split data yang sama. Hal ini menunjukkan jika algoritma *artificial neural network* unggul dalam menganalisis dan mengklasifikasi sentimen opini publik terhadap kenaikan harga BBM. Adapun polaritas sentimen publik mengarah ke arah positif dengan jumlah 718 sentimen, negatif dengan jumlah 259 sentimen dan netral dengan jumlah 23 sentimen.

4. KESIMPULAN

Terbukti perbandingan pada penelitian ini bahwa model algoritma *artificial neural network* memiliki tingkat performa yang cukup tinggi dibandingkan algoritma *naïve bayes* pada split data 90:10 yaitu akurasi 87% dan cross validasi 90% dan memiliki selisih 4-5% untuk akurasi selisih 5-8% untuk evaluasi pada setiap split data. Penelitian ini mendapat reaksi terhadap kebijakan kenaikan harga BBM pada tweet di twitter dengan hasil sentimen positif berjumlah 718 sentimen, negatif berjumlah 259 sentimen dan netral berjumlah 23 sentimen. Dapat disimpulkan dengan hasil sentimen positif lebih banyak dari negatif dan netral menunjukkan masyarakat indonesia belum merasakan keresahan-keresahan terhadap kenaikan harga BBM yang ditetapkan pemerintah. Tidak menutup kemungkinan jika kenaikan harga BBM lebih lama lagi hal ini dapat berubah. Dengan hasil sentimen opini publik ini, dapat menjadi saran bagi pemerintah sebelum menetapkan kebijakan menaikkan harga BBM di Indonesia. Penelitian selanjutnya diharapkan dapat menambahkan model algoritma klasifikasi lainnya sehingga mengetahui model algoritma mana yang memiliki performa terbaik. Lalu sumber data yang digunakan tidak hanya media sosial *twitter* saja media sosial lainnya seperti *Youtube*, *Facebook*, *Instagram* dan sumber data lain sehingga data lebih bervariasi dan analisis sentimen opini publik terhadap kenaikan BBM lebih akurat.

REFERENSI

- [1] A. W. Kurniawan, Z. F. Nurhadi, H. Hendrawan, R. P. Damayanti, and D. Hidayat, "Pengaruh Kalimat 'Twitter Please Do Your Magic' Terhadap Sikap Pengguna Twitter," *Jurnal Digital Media dan Relationship*, vol. 3, no. 1, pp. 8–23, 2021.
- [2] Rinanda T. G., "Literasi Media Di Twitter (Study Deskriptif Remaja Melalui Gerakan 'Twitter Please Do Your Magic')," UNIVERSITAS ISLAM RIAU, Pekanbaru, 2021.
- [3] W. Wardani, S. Suriana, S. U. Arfah, Z. Zulaili, and P. S. Lubis, "Dampak kenaikan Bahan Bakar Minyak (BBM) Terhadap Inflasi dan Implikasinya Terhadap Makroekonomi di Indonesia," *AFoSJ-LAS (All Fields of Science Journal Liaison Academia and Society)*, vol. 2, no. 3, pp. 63–70, 2022.
- [4] S. Mujahidin, B. Prasetio, and M. C. C. Utomo, "Implementasi Analisis Sentimen Masyarakat Mengenai Kenaikan Harga BBM Pada Komentar Youtube Dengan Metode Gaussian naïve bayes," *Voteteknika (Vocational Teknik Elektronika dan Informatika)*, vol. 10, no. 3, pp. 17–24, 2022.
- [5] U. Kurniasih and A. T. Suseno, "Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) pada Kenaikan Harga Bahan Bakar Minyak (BBM)," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 4, pp. 2335–2340, 2022.
- [6] A. Alsaeedi and M. Z. Khan, "A study on sentiment analysis techniques of Twitter data," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, 2019.
- [7] V. K. S. Que, A. Iriani, and H. D. Purnomo, "Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi | Vol*, vol. 9, no. 2, 2020.
- [8] K. Dashtipour, M. Gogate, A. Gelbukh, and A. Hussain, "Extending persian sentiment lexicon with idiomatic expressions for sentiment analysis," *Soc Netw Anal Min*, vol. 12, no. 1, pp. 1–13, 2022.
- [9] N. Sucahyo, L. Nurlaela, and R. R. Waryono, "Analisis Sentimen Masyarakat Jakarta Terhadap Kebijakan Perluasan Dan Perpanjangan Ganjil Genap di Media Sosial Twitter," *Jurnal Teknologi Informatika dan Komputer*, vol. 7, no. 1, pp. 97–111, 2021.
- [10] R. Bensoltane and T. Zaki, "Towards Arabic aspect-based sentiment analysis: a transfer learning-based approach," *Soc Netw Anal Min*, vol. 12, no. 1, pp. 1–16, 2022.
- [11] L. Nemes and A. Kiss, "Social media sentiment analysis based on COVID-19," *Journal of Information and Telecommunication*, vol. 5, no. 1, pp. 1–15, 2021.
- [12] F. Fungsiawan, "Kenaikan Tarif Ojek Online Berdampak Inflasi dan Penurunan Pendapatan Domestik Bruto (PDB): Increase in Online Ojek Tariffs Impact Inflation and Decrease in Gross Domestic Product (GDP)," *CEMERLANG: Jurnal Manajemen dan Ekonomi Bisnis*, vol. 2, no. 3, pp. 268–274, 2022.
- [13] T. N. Fatyanosa and F. A. Bachtiar, "Classification method comparison on Indonesian social media sentiment analysis," in *2017 International Conference on Sustainable Information Engineering and Technology (SIET)*, 2017, pp. 310–315.
- [14] R. Rasenda, H. Lubis, and R. Ridwan, "Implementasi K-NN Dalam Analisa Sentimen Riba Pada Bunga Bank Berdasarkan Data Twitter," *Jurnal Media Informatika Budidarma*, vol. 4, no. 2, pp. 369–376, 2020.
- [15] E. S. Romaito and M. K. Anam, "Rahmaddeni, and AN Ulfah," "Perbandingan Algoritma SVM Dan NBC Dalam Analisa Sentimen Pilkada Pada Twitter," *CSRID J*, vol. 13, no. 3, pp. 169–179, 2021.

- [16] A. N. Ulfah and M. K. Anam, "Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM)," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 7, no. 1, pp. 1–10, 2020.
- [17] R. Mahendrajaya, G. A. Buntoro, and M. B. Setyawan, "Analisis Sentimen Pengguna Gopay Menggunakan Metode Lexicon Based Dan Support Vector Machine," *KOMPUTEK*, vol. 3, no. 2, pp. 52–63, 2019.
- [18] M. K. Anam, M. I. Mahendra, W. Agustin, R. Rahmaddeni, and N. Nurjayadi, "Framework for Analyzing Netizen Opinions on BPJS Using Sentiment Analysis and Social Network Analysis (SNA)," *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 6, no. 1, pp. 11–28, 2022.
- [19] F. Akbar, H. W. Saputra, A. K. Maulaya, M. F. Hidayat, and R. Rahmaddeni, "Implementasi Algoritma Decision Tree C4. 5 dan Support Vector Regression untuk Prediksi Penyakit Stroke: Implementation of Decision Tree Algorithm C4. 5 and Support Vector Regression for Stroke Disease Prediction," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, no. 2, pp. 61–67, 2022.
- [20] S. Rajasekaran and G. A. V. Pai, *Neural networks, fuzzy systems and evolutionary algorithms: Synthesis and applications*. PHI Learning Pvt. Ltd., 2017.
- [21] L. Khanady, "Prediksi Harga Saham Dengan Menggunakan JST (Jaringan Syaraf Tiruan)," *JURNAL ILMIAH INFORMATIKA*, vol. 7, no. 01, pp. 1–4, 2019.
- [22] M. Dennis, R. Rahmaddeni, F. Zoromi, and M. K. Anam, "Penerapan Algoritma Naïve Bayes Untuk Pengelompokan Predikat Peserta Uji Kemahiran Berbahasa Indonesia," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 2, pp. 1183–1190, 2022.
- [23] M. M. Firman Tempola and A. Khairan, "Perbandingan Klasifikasi Antara Knn Dan Naive Bayes Pada Penentuan Status Gunung Berapi Dengan K-Fold Cross Validation," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 5, no. 5, 2018.