

Klasifikasi Opini Terhadap Pertanian Sawit (*Palm Oil*) Indonesia Menggunakan *Naïve Bayes*

Hafiz Irsyad¹, M Rizky Pribady²

^{1,2}STMIK Global InformatikaMDP; Jl. Rajawali No. 14 Palembang Sumatera Selatan, Telp:
(0711) 376400

e-mail: *¹hafizirsyad@mdp.ac.id, ²rizky@mdp.ac.id

Abstrak

Dalam tiga tahun terakhir ini produksi pertanian sawit dinilai sangat meningkat. Indonesia merupakan negara penyumbang sawit terbanyak. Bersama Malaysia yakni 85-90% dari total keseluruhan hasil sawit dunia. Dengan banyaknya informasi terhadap sawit Indonesia di twitter sehingga dapat dimanfaatkan untuk melihat opini masyarakat tentang sawit Indonesia. Pada penelitian ini berhasil mengumpulkan data tweet dari 28 Agustus 2019 sampai dengan 21 Juni 2018 menghasilkan 1015 tweet. Agar bias melihat tweets tersebut maka dikelompokkan perkategori menjadi positif, negatif dan netral, kemudian tweets tersebut diklasifikasikan dengan menggunakan metode naïve bayes dan menggunakan tools Orange. Sedangkan untuk melakukan crawling data memanfaatkan fasilitas API Twitter. Dari 1015 data tweets tersebut 70% digunakan untuk data training dan 30% untuk data testing. Dalam penerapan kalisisifikasi dengan metode naïve bayes ini menghasilkan akurasi rata-rata akurasi 0.83337 % untuk rata-rata seluruh kategori, untuk precision memperoleh 0.80303% untuk rata-rata seluruh kategori, dan untuk recall menghasilkan 0.90853% untuk seluruh rata-rata kategori. Dengan tingkat akurasi tersebut metode naïve bayes berjalan sesuai dengan harapan.

Kata kunci—Naïve Bayes, Sawit Indonesia, Opini, Klasifikasi

Abstract

Last three years the production of oil palm agriculture has been considered to have increased significantly. Indonesia is the largest contributor to palm oil with Malaysia, which is 85-90% of the total world palm oil yield. With so much information on Indonesian oil palm on Twitter so that it can be used to see public opinion about Indonesian oil palm. In this study managed to collect tweet data from 28 August 2019 to 21 June 2018 resulting in 1015 tweets. In order to see the tweets, the categories are categorized into positive, negative and neutral, then the tweets are classified using the naïve Bayes method and using the Orange tools. Meanwhile, to do data crawling using Twitter API facilities. Of the 1015 data tweets 70% is used for training data and 30% for testing data. In the application of calisification with the naïve bayes method it produces an average accuracy of 0.83337% for the average of all categories, for precision obtains 0.80303% for the average of all categories, and for recall produces 0.90853% for all average categories. With this level of accuracy the Naïve Bayes method works in line with expectations.

Keywords—Naïve Bayes, Indonesian Palm Oil, Opinion, Classification

1. PENDAHULUAN

Indonesia mempunyai geografis yang sangat efektif terhadap sumberdaya alam. Sehingga Indonesia merupakan negara yang kaya akan sumberdaya alamnya. Merajut kepada data Badan Pusat Statistika, Indonesia memiliki tiga sektor yang ikut dalam kemajuan ekonominya berupa sektor pertanian, perdagangan dan sektor industri. Pertanian dapat dibagi dalam empat kategori yaitu, Pertanian Hortikultura, Perkebunan, Kehutanan dan Peternakan [1]. Kepala Sawit adalah salah satu hasil pertanian dari sawit. Minyak kelapa sawit merupakan yang paling banyak digunakan di dunia. Dengan hasil dari minyak sawit ini dapat diproduksi untuk bermacam makanan, kosmetik, hasil kebersihan dan dengan penemuan canggih dapat sebagai sumber biofuel atau lebih dikenal dengan biodiesel. Indonesia merupakan negara yang mendominasi produksi hasil pertanian sawitnya dengan menyumbang 85-90% dari semua hasil produksi minyak sawit dunia bersama dengan Malaysia [2]. Pada tiga tahun terakhir Indonesia telah mengeksport kelapa sawit yaitu, 2018 dengan hasil 41.67 juta ton, 2017 dengan hasil 38.17 juta ton dan 2016 dengan hasil 27.0 juta ton dengan hasil yang diperoleh meningkatkan devisa negara [3].

Twitter merupakan salah satu dari sekian banyak media egati yang difungsikan sebagai pemasaran produk, dimana twitter sebagai media egati yang sangat populer setelah Facebook, Instagram dan lain-lain. Pada tahun 2016 Indonesia memiliki pengguna aktif sebanyak 24.340 juta orang. Twitter dapat menampung sebanyak 280 karakter, kemudian diolah menjadi sebuah *statement* [4]. Dewasa ini media twitter dimanfaatkan sebagai perluasan bisnis, dimana pengguna bisa menemukan pelaku bisnis lain sehingga bisa menjadi teman atau pengikut (*followers*) dan tentu bisa saling berinteraksi.

Twitter menyediakan fasilitas-fasilitas agar bias dimanfaatkan oleh para pengguna akun twitter, salah satunya adalah fasilitas *retweet*. Banyaknya *retweet* atau komentar masyarakat tentang perkembangan maupun penyebaran informasi sektor pertanian di twitter, *retweet* tersebut dapat diolah kemudian disajikan dalam bentuk informasi sehingga dapat mengetahui pendapat masyarakat tentang perkembangan sektor pertanian. *Retweet* dari *followers* akan dapat mempengaruhi pola egati dari *followers* sehingga bias membuat *retweet* bisa memungkinkan terjadinya *overload retweet*. Dengan demikian *retweet* dapat di klasifikasikan kalimat yang berhubungan dengan *retweet* perkembangan pada sektor pertanian, sehingga dapat menemukan pola dalam pengelompokan kalimat yang positif, negatif dan netral. Pengumpulan data *tweet* dari twitter dapat dilakukan dengan menghubungkan Twitter API dan Orange Tools [5].

Berdasarkan penelitian yang telah dituliskan diatas, penelitian ini akan melakukan sebuah dengan klasifikasi opini data pada twitter. Data yang telah diperoleh dari twitter akan diproses dengan *textmining* untuk menghindari data yang tidak diinginkan, setelah itu dikelompokan data *tweet* atau komentar berdasarkan klasifikasi positif, negatif dan netral. Pengelompokan ini menggunakan perbandingan metode *Naïve Bayes*.

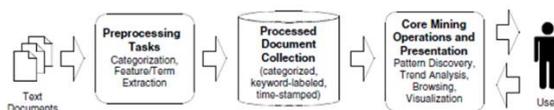
2. LANDASAN TEORI

2.1 Text Mining

Text mining adalah proses mengekstrak informasi yang berguna dari sekumpulan dokumen dari waktu ke waktu melalui identifikasi pola yang menarik [4]. *Text mining* adalah proses penemuan informasi baru dengan mengekstrak pola secara otomatis dari berbagai sumber teks [6].

Text mining mempunyai beberapa tingkatan fungsional, *text mining* mengikuti model yang telah disediakan oleh beberapa aplikasi data mining sehingga dapat terbagi menjadi empat

tingkatan fungsional utama, yakni *preprocessing task*, *processed document collection*, dan *core mining operations* [4]. Tingkatan arsitektur dari *text mining* dapat dilihat pada gambar. 1.



Gambar 1. Arsitektur *Text Mining* [4].

2.2 Preprocessing

Text preprocessing merupakan awal mula dari *text mining*. Tujuan dari *text preprocessing* adalah membersihkan data dari *noise* sehingga data menjadi lebih kecil dan lebih terstruktur. Berikut ini adalah tahapan proses dari *text preprocessing* [4].

a. Tokenizing

Tokenizing merupakan proses pemecahan teks menjadi kata tunggal dan menghapus tanda baca serta angka, sesuai dengan kamus yang telah ditentukan.

b. Stopword Removing

Stopword removing merupakan proses menghilangkan kata tidak penting dalam teks. Hal ini dilakukan untuk memperbesar akurasi dari pembobotan. Dalam penelitian ini *Stopword removing* digunakan untuk menghilangkan kata-kata seperti: dan, atau, mungkin, ini, itu dan sebagainya merupakan kata yang dapat dihilangkan.

c. Stemming

Stemming merupakan perubahan berbagai kata imbuhan menjadi sebuah kata dasar. Dalam penelitian ini kata imbuhan yang dihilangkan adalah teks yang berbahasa Inggris, dikarenakan teks bahasa Inggris memiliki bentuk lampau.

2.3 Naïve Bayes

Naïve Bayes adalah salah satu algoritma klasifikasi yang ditemukan oleh Tomas Bayes. Klasifikasi adalah proses untuk memperoleh aturan atau model yang dapat mengklasifikasikan data baru yang belum pernah dipelajari dengan mempelajari sekumpulan data yang lama. *Naïve Bayes* adalah algoritma klasifikasi probabilistik yang sederhana berdasarkan Teorema Bayes [7]. Prinsip umumnya adalah mengasumsikan bahwa nilai suatu atribut tidak bergantung dan mempengaruhi atribut yang lainnya. Model *Naïve bayes* memungkinkan setiap atribut memiliki kontribusi yang sama terhadap keputusan akhir dan komputasinya lebih efisien bila dibandingkan dengan algoritma pengklasifikasi teks lainnya. Model yang diperoleh dari proses *training* berisikan kumpulan konstanta untuk setiap data *train*. Model tersebut akan digunakan pada data *testing* untuk melihat seberapa akurat model yang telah diperoleh [8]. *Naïve bayes* menjumlahkan frekuensi dan kombinasi nilai dari dataset untuk menghitung probabilitas. Probabilitas dihitung dengan menggunakan persamaan 1 [7].

$$P(H|X)^{\wedge} = (P(X|H) \cdot P(H)) / (P(X)) \quad (1)$$

Pada persamaan diatas adalah:

X = Data dengan *class* yang belum diketahui

H = Hipotesis data merupakan suatu *class* spesifik

$P(H|X)^{\wedge}$ = Probabilitas hipotesis berdasar kondisi

P(H) = Probabilitas hipotesis

P(X|H) = Probabilitas berdasarkan kondisi pada hipotesis

P(X) = Probabilitas H

2.4 Sentiment Analysis

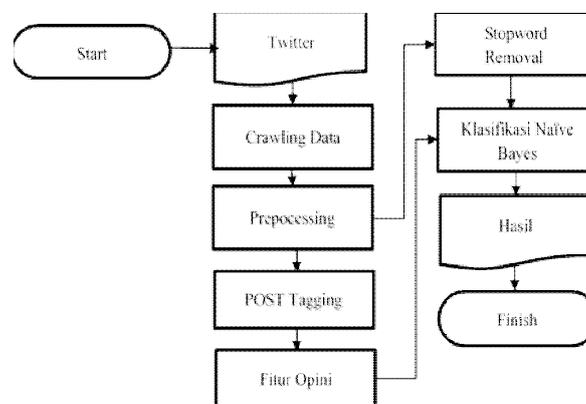
Analisis sentimen, juga disebut penambangan opini, adalah bidang studi itu menganalisis pendapat, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya. Itu mewakili besar ruang masalah. Ada juga banyak nama dan tugas yang sedikit berbeda, misalnya, analisis sentimen, penambangan opini, ekstraksi pendapat, sentimen penambangan, analisis subjektivitas, analisis pengaruh, analisis emosi, peninjauan penambangan, dan lain-lainya. Namun, sekarang semuanya berada di bawah payung analisis sentimen atau penambangan opini. Sementara di industri, analisis sentimen jangka lebih umum digunakan, tetapi di akademisi baik analisis sentimen dan penambangan opini sering digunakan.[9]

2.5 Twitter Rest API

Twitter adalah jejaring sosial populer di mana pengguna dapat berbagi pesan singkat seperti SMS (*Short Message Service*) yang disebut *tweet*. Pengguna berbagi pemikiran, tautan, dan gambar di twitter, jurnalis mengomentari secara langsung, perusahaan mempromosikan produk, dan terlibat dengan pelanggan. Berbagai cara dalam menggunakan twitter, dengan 500 jutaan *tweet* perhari, ada banyak data untuk dianalisis dan dimainkan. Twitter Rest APIs menyediakan akses terprogram untuk membaca dan menulis data twitter. Buat *tweet* baru, baca profil penulis dan data pengikut, dan banyak lagi. REST API mengidentifikasi aplikasi Twitter dan pengguna yang menggunakan OAuth; tanggapan tersedia di JSON [10].

3. METODE PENELITIAN

Gambaran umum dalam proses penelitian dapat dilihat pada gambar 2 sebagai berikut:



Gambar 2. Gambaran Umum Penelitian

Skenario dari penelitian ini adalah sebagai berikut:

- Tools* Orange akan menerima inputan berupa dataset *twitter*. Data tersebut diambil berdasarkan inputan *tweet* dari akun “@infosawit”.
- Melakukan *preprocessing* terhadap dataset agar dapat memenuhi klasifikasi dan mempermudah dalam memproses data.
- POST Tagging* berfungsi sebagai pencarian 233 egati opini. *POST Tagging* berfungsi juga sebagai ekstraksi fitur opini. *POST Tagging* dilakukan secara manual dalam menentukan fitur opini masuk kategori positif, 233 egati atau netral.
- Setelah fitur opini didapatkan, maka akan dilakukan langkah selanjutnya dengan menggunakan *stop word removal*. *Stop word removal* sendiri berfungsi untuk menghapus kata-kata yang sering muncul akan tetapi tidak memiliki makna dalam penelitian ini.

- e) Proses selanjutnya adalah melakukan pengklasifikasian negatif data dengan menggunakan *Naïve Bayes*.

Dataset yang digunakan dalam penelitian ini merupakan opini-opini masyarakat tentang sawit yang ada di Indonesia pada saat sekarang. Algoritma yang digunakan untuk klasifikasi adalah *Naïve Bayes* dengan menggunakan *tools* Orange [5].

a. *Crawling Data Tweets*

Dataset yang digunakan dalam penelitian ini adalah data dari berbagai akun yang selalu mentweetkan pertanian dan sawit. Dataset yang diambil menggunakan *Application Programming Interface (API) Twitter* menggunakan *Tools Orange*.

Keyword yang digunakan adalah pertanian, penarikan data mulai dari tanggal 28 Agustus 2019 sampai dengan 21 Juni 2018. Dari hasil *crawling* data diperoleh *tweet* sebanyak 1015 *tweet*. Adapun hasil dari perolehan *crawling* data adalah:

Tabel 1. Hasil *Crawling* dari *Twitter*

| No | Komentar Opini |
|----|--|
| 1. | RT @infosawit: Petani Sawit Minta Jokowi Serius Tingkatkan Kualitas SDM Petani #beritasawit #sawit #palmoil #infosawit #Smallholders @InfoSawit |
| 2. | Harga TBS Sawit Jambi Periode 23 – 29 Agustus 2019 Naik Rp 39,2/Kg #Beritasawit #sawit #palmoil #infosawit #HargaTBS #HargaSawit #CPOPrice https://t.co/dpNvkLA8tR |
| 3. | Perkuat Pasar Sawit Domestik, Hilangkan Ketergantungan Pasar Ekspor #beritasawit #sawit #palmoil #infosawit #PasarCPO https://t.co/EKHkmRP9t2 |
| 4. | Presiden Awasi Langsung Mandatori Biodiesel Sawit (B30) di Tahun 2020 #Beritasawit #sawit #palmoil #infosawit #aprobi #biodiesel @jokowi https://t.co/16drVJK190 |
| 5. | Harga TBS Sawit Kaltim Periode Agustus 2019 Turun Rp 24,01/Kg #beritasawit #Sawit #palmoil #infosawit #hargatbs #hargasawit https://t.co/mRvZlqbPAd |
| 6. | Harga TBS Sawit Jambi Periode 12-18 Juli 2019 Turun Rp 21,31/Kg #beritasawit #sawit #palmoil #infosawit #hargaTBS #HargaSawit @PerekonomianRI @KeuanganRI https://t.co/HieqAzbgPn |
| 7. | Harga TBS Sawit Riau Periode 3-9 Juli 2019 Turun Rp. 44,64/Kg #beritasawit #Sawit #palmoil #infosawit #hargatbs #hargasawit https://t.co/FijqgQZWKO |
| 8. | RSPO Resmi Luncurkan Pedoman “Upah Hidup Layak” untuk Para Pekerja Sawit #beritasawit #sawit #palmoil #infosawit #rspo #sustainable #sustainability https://t.co/lpmPAGHeIL |

b. *Preprocessing*

Pada tabel 1 diatas telah didapatkan hasil dari *crawling* data dari *twitter*, maka selanjutnya kita akan ketahap *preprocessing* agar data-data yang telah *crawling* tersebut dapat menjadi terstruktur. Berikut ini adalah tahapan dari proses *preprocessing*:

- *Tokenizing*

Sebagaimana telah dijelaskan di atas, Tokenizing merupakan proses pemecahan teks menjadi kata tunggal dan menghapus tanda baca serta angka, sesuai dengan kamus yang telah ditentukan.

Maka berdasarkan dataset di *twitter*, pada umumnya pemilik akun hanya menggunakan kata-kata baku bahkan menggunakan kata-kata gaul untuk membuat sebuah ciutan atau *tweet*. Tugas pada Tokenizing ini adalah mengubah kata-kata baku tersebut sehingga lebih mudah dimengerti.

Tabel 2. Contoh *Tokenizing*

| Komentar Opini |
|---|
| Urai Masalah Tata Ruang, Diusulkan ATR/BPN Diberikan Kewenangan Lebih #beritasawit #sawit #palmoil #infosawit @Atr_BPN @KemenkoRIhttps://t.co/7EmyufmdMd |
| Menjadi |
| Urai Masalah Tata Ruang Diusulkan ATR atau BPN Diberikan Kewenangan Lebih |

- *Stopword Removing*

Stopword removing merupakan proses menghilangkan kata tidak penting dalam teks. Hal ini dilakukan untuk memperbesar akurasi dari pembobotan. Dalam penelitian ini *Stopword removing* digunakan untuk menghilangkan kata-kata seperti: dan, atau, mungkin, ini, itu dan sebagainya merupakan kata yang dapat dihilangkan.

Tabel 3. Contoh *Stop Removing*

| Komentar Opini |
|---|
| Urai Masalah Tata Ruang Diusulkan ATR atau BPN Diberikan Kewenangan Lebih |
| Menjadi |
| Urai Masalah Tata Ruang usul ATR BPN beri wewenang Lebih |

- *Post Tagging Tweet*

Selanjutnya adalah melakukan pelabelan secara manual dari beberapa proses yang telah diciptakan tadi. Dengan adanya sebuah label maka tweet akan mudah dikelompokkan, adapun label yang digunakan sebanyak tiga label yakni Positif, Negatif dan Netral.

Tabel 4. Contoh *Post Tagging*

| No | Komentar Opini | Tagging |
|----|--|---------|
| 1. | Urai Masalah Tata Ruang usul ATR BPN berikan wewenang Lebih | Netral |
| 2. | Harga TBS Sawit Riau Periode 21 sampai 27 Agustus Naik Rp 47,75/Kg | Positif |

| | | |
|----|---|---------|
| 3. | BPK Temukan Sejumlah Pelanggaran Izin Sawit | Negatif |
| 4. | Sawit Watch: Pemerintah Mesti Tegas Atasi Karhutla | Positif |
| 5. | 2023 RSPO Targetkan Produksi Minyak Sawit Berkelanjutan 23 Juta Ton | Positif |

Setelah *preprocessing* dilaksanakan berdasarkan *point-point* diatas tidak seluruh data yang akan digunakan akan tetapi ada beberapa data digunakan dan ada data yang di simpan atau tidak digunakan. Adapun data yang tidak digunakan seperti *tweet* yang mengandung unsure *emoticon*, *url*, *retweet*, gambar dan *video* [5]. Hasil dari *crawling* data terdapat data sebanyak 1015 *tweet*. Untuk pengklasifikan data kelas positif adalah 520 data, kelas negatif 218 data dan kelas netral sebanyak 276 data.

Tweet yang telah didapatkan maka akan dibagi menjadi dua data yakni, data training sebanyak 711 data *tweet* dari 70% *crawling* data dan untuk data testing sebanyak 304 data *tweet* dari 30% *crawling* data. Maka dari itu dapat kita lihat pada tabel 5.

Tabel 5. Data Training dan Testing

| Data | Negatif | Netral | Positif |
|----------|---------|--------|---------|
| Training | 200 | 134 | 377 |
| Testing | 87 | 70 | 146 |

Untuk mengukur keakuratan dari hasil porses data training dan data testing maka akan diujikan untuk mengukur dari keakuratan data tersebut. Untuk hasil data *testing* akan menghasilkan *confution matrix*. Hasil dari *confution matrix* dapatlah nilai akurasi, *precision* dan *recall*. Akurasi tersebut dapat dihitung dengan menggunakan persamaan 2.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

Untuk menghitung *precision* menggunakan persamaan 3.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

Untuk menghitung nilai *recall* dapat dilihat pada Persamaan 4.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

Dimana:

TP = True Positif Count.

TN = True Negative Count.

FP = False Positive Count.

FN = False Negative Count.

4. HASIL DAN PEMBAHASAN

Setelah *preprocessing* selesai dilakukan, selanjutnya kita akan mengambil data dari *wordcloud* sehingga dapat diketahui opini yang paling dominan pada *tweet*. Sehingga hasil dari *wordcloud* nantinya dapat digunakan pengkategorian yakni, positif, negatif dan netral. Hasil dari *wordcloud* dapat dilihat pada gambar 3 berikut ini.



Gambar 3. Hasil *Wordcloud*

Dari hasil *wordcloud* tersebut, terdapat 8 kata dominan dari setiap masing-masing opini yang sudah melalui klasifikasi yakni:

Tabel 6. *Wordcloud* yang Dominan

| Weight | Word |
|---------------|----------------------|
| 180 | Harga |
| 177 | Sustainable Palm Oil |
| 159 | Perekonomian RI |
| 152 | Sawit Baik |
| 145 | Harga Sawit |
| 134 | Sustain Ability |
| 127 | Harga TBS |
| 90 | Sawit adalah Kita |

Dari hasil *wordcloud* diatas maka akan dilakukan pengujian terhadap hasil *tweet* tersebut. Hasil pengujian dari *tweet* tersebut terdapat pengkategorian positif 147 data *tweet*, untuk negatif 87 data *tweet* dan untuk netral sebanyak 69 data *tweet*. Hasil pengujian tersebut dapat kita lihat pada tabel 7.

Tabel 7. Pegujian Data *Tweet*

| Kategori | Negatif | Netral | Positif |
|-----------------|----------------|---------------|----------------|
| Negatif | 62 | 1 | 24 |
| Netral | 14 | 22 | 33 |
| Positif | 24 | 31 | 92 |

Berdasarkan tabel 7 diatas, maka akan direalisasikan menjadi *confusion matrix* sehingga dapat digunakan untuk menghitung tingkat akurasi, *precision* dan *recall*, sehingga menghasilkan setiap kategori yang dapat dilihat pada tabel 7.

Tabel 8. Hasil Akurasi, *Precision* dan *Recall*

| Kategori | Akurasi | Precision | Recall |
|----------|---------|-----------|---------|
| Negatif | 0.87129 | 0.64935 | 0.98850 |
| Netral | 0.77723 | 0.80521 | 0.80232 |
| Positif | 0.85149 | 0.95454 | 0.93478 |
| Average | 0.83337 | 0.80303 | 0.90853 |

Dari hasil tabel 8 diatas bahwasan kategori yang memiliki tingkat akurasi yang lebih tinggi adalah *negative* dengan nilai 0.871%, untuk kategori positif 0.851% dan untuk kategori yang rendah adalah netral 0.777%, sehingga mendapatkan nilai rata-rata dari seluruh akurasi adalah 0.833%. selisih antara nilai Negatif dan Positif adalah sebanyak 0.02% dengan hasil selisih yang sedikit dapat dibuktikan terdapat beberapa kesalahan dari tweet, mungkin berupa kata-kata yang sama atau berupa kata-kata singkatan yang sulit *didetect* oleh *tools*, sehingga ada kendala dalam melakukan pengkategorian.

5. KESIMPULAN

Hasil opini dari masyarakat terhadap sawit Indonesia (*Palm Oil*) dengan menggunakan dataset pada twitter dengan menerapkan metode *naïve bayes* dapat menghasilkan tingkatan akurasi 0.83337 % untuk rata-rata seluruh kategori, untuk precision memperoleh 0.80303% untuk rata-rata seluruh kategori, dan untuk *recall* menghasilkan 0.90853% untuk seluruh rata-rata kategori. Dengan hasil ini *tools orange* sudah mulai berjalan dengan baik dari pada sebelumnya.

Dari hasil opini yang sering muncul untuk dataset dari tweet adalah masalah harga. *Sustainable Palm Oil*, Perekonomian RI, Sawit Baik, Harga Sawit, *Sustainability*, harga TBS dan Sawit adalah kita.

6. SARAN

Penelitian ini diharapkan dapat dikembangkan dengan menggunakan *tools-tols* yang lain sehingga menghasilkan hasil yang lebih baik lagi.

UCAPAN TERIMA KASIH

Terimakasih kepada TIM LPPM STMIK Global Informatika MDP dan Kemenristek DIKTI atas dukungannya dalam penelitian ini.

DAFTAR PUSTAKA

- [1] B. P. Statistik, "www.bps.go.id," 5 2 2018. [Online]. Available: <https://www.bps.go.id/pressrelease/2018/02/05/1519/ekonomi-indonesia-triwulan-iv-2017--tumbuh-5-19-persen.html>. [Accessed 10 8 2018].
- [2] Indonesia Invesment, "www.indonesia-investment.co"m 29 08 2019 [online]. Available: [https://www.indonesia-investments.com/id/bisnis/komoditas/minyak-sawit/item166?.](https://www.indonesia-investments.com/id/bisnis/komoditas/minyak-sawit/item166?)

- [3] Gatra, "www.gatra.com", 29 8 2019 [Online]. Available: <https://www.gatra.com/detail/news/416887/economy/ekspor-minyak-sawit-indonesia-capai-296-juta-ton-pada-maret-2019>. [Accessed 29 08 2019].
- [4] M. Rani and A. J 2016, "Twitter Data Predicting Geolocation Using Data Mining Techniques," *International Journal of Innovative Research in Computer*, Vol. 4, No. 6, p. 10446.
- [5] Hafiz I, . M. Rizky 2019, "Klasifikasi Opini Masyarakat Terhadap Jasa ISP MYRepublic dengan Naïve Bayes," *Jurnal JNTETI*, Vol. 8, No. 1.
- [6] A. Saleh 2015, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Creative Information Technology Journal* , Vol. 2, No. 3, pp. 207-2017.
- [7] S. Ting, A. H. Tsang and W. H. Ip 2011, "Is Naïve Bayes a Good Classifier for Document Classification?," *International Journal of Software Engineering and its Applications*, Vol. 5, No. 3, p. 398.
- [8] M. S. Kini, Devi, D. PG and N. Chiplunkar 2015, "Text mining Approach to Classify Technical Research Document using Naïve Bayes," *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 4, No. 7, pp. 386-391.
- [9] A. Valdivia, M. V. Luzón and F. Herrera 2017, "Sentiment analysis in tripadvisor," *IEEE Intelligent Systems*, Vol. 4, No. 72-77, p. 32.
- [10] S. Budi 2017, "Text Mining untuk Analisis Sentimen Review Film Menggunakan Algoritma K-Means," *Techno.COM*, Vol. 16, No. 1, pp. 1-8.