

ANALIS SENTIMEN MASYARAKAT DI TWITTER TERHADAP PEGAWAI PAJAK DENGAN METODE MULTINOMIAL NAÏVE BAYES

Wildan Syahidillah^{1*}, Grace Gata²

^{1,2}Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta Selatan, Indonesia

Email: ^{1*}wildansyahidillah15@gmail.com, ²grace.gata@budiluhur.ac.id
(* : corresponding author)

Abstrak-Pajak merupakan iuran wajib masyarakat kepada negara yang dapat dikenakan tanpa pelanggaran secara langsung, dan apabila ada masyarakat yang tidak membayar pajak, maka akan dikenakan sanksi oleh negara. Penyalahgunaan kekuasaan dalam pemerintahan pajak terkait dengan praktik-praktik yang melanggar etika atau hukum yang dilakukan oleh pegawai pajak atau pejabat pajak. Hal ini dapat mencakup penyelewengan dalam penerimaan atau pengeluaran pajak, manipulasi data atau informasi keuangan, atau tindakan korupsi. Penyalahgunaan kekuasaan di bidang pajak juga bisa memiliki dampak yang merugikan pada masyarakat dan perekonomian, seperti kehilangan pendapatan negara, ketidakadilan dalam sistem pajak, dan ketidakpercayaan masyarakat terhadap pemerintah dan lembaga pajak. Dalam penelitian ini bertujuan untuk menganalisis opini publik atau masyarakat Indonesia terhadap pegawai pajak. Metode analisis pada penelitian ini menggunakan *Multinomial Naïve Bayes*. Klasifikasi sentimen dibagi menjadi dua yaitu positif dan negatif. Dalam penelitian ini dilatarbelakangi atas penyalahgunaan kekuasaan oleh pegawai pajak dan hilangnya rasa kepercayaan masyarakat terhadap pegawai pajak. Peneliti juga menggunakan CRISP-DM sebagai metode pengolahan data serta menggunakan aplikasi *rapidminer* untuk mendapatkan, mengolah dan menghasilkan klasifikasi positif dan negatif. Selain *rapidminer* peneliti juga menggunakan aplikasi *microsoft excel* dalam membantu menentukan klasifikasi positif dan negatif. Pada penelitian ini mengambil 1.078 data *tweet* yang membahas pegawai pajak pada media sosial *twitter* dengan menggunakan *keywords* “RafaelPajak”, “PejabatPajak”, “PajakIndonesia”, “RafaelAlun”, dan “GratifikasiPajak”. Setelah data diolah maka dihasilkan 363 data *tweet* dari 1.078 data *tweet* yang siap digunakan pada penelitian ini. Sehingga hasil dari penelitian ini menghasilkan nilai *sentiment* positif sebanyak 118 data dan *sentiment* negatif sebanyak 245 data. Sedangkan untuk nilai *accuracy* yaitu 68.52%, *precision* 51.52%, dan *recall* 48.57% dengan pemodelan *naïve bayes* yang menggunakan perbandingan 70:30 untuk data *training* dan data *testing*. Sehingga dapat dihasilkan pegawai pajak masih dinilai negatif oleh masyarakat Indonesia.

Kata Kunci: *Mulinomial Naïve Bayes*, *RapidMiner*, Pajak, CRISP-DM.

ANALYSIS OF PUBLIC SENTIMENT ON TWITTER TOWARDS TAX EMPLOYEES USING THE MULTINOMIAL NAÏVE BAYES METHOD

Abstract-Taxes are mandatory contributions from the community to the state that can be imposed without getting direct contravention, and if there are people who do not pay them, they will be subject to sanctions by the state. The abuse of power in the tax administration is related to unethical or legal practices by tax officials or tax officials. This could include misappropriation of tax receipts or disbursements, manipulation of financial data or information, or acts of corruption. Abuse of power in the tax sector can also have detrimental effects on society and the economy, such as loss of state revenues, injustice in the tax system, and public distrust of the government and tax agencies. In this study aims to analyze public opinion or Indonesian society towards tax officials. The analytical method in this study uses *Multinomial Naïve Bayes*. *Sentiment classification* is divided into two, namely positive and negative. In this research, the background is the abuse of power by tax officials and the loss of public trust in tax officials. Researchers also use CRISP-DM as a data processing method and use the *rapidminer* application to obtain, process and produce positive and negative classifications. Apart from *rapidminer*, researchers also used the *Microsoft Excel* application to help determine positive and negative classifications. In this study, 1,078 *tweet* data were taken that discussed tax officials on *Twitter* social media using the *keywords* “RafaelPajak”, “PejuangPajak”, “PajakIndonesia”, “RafaelAlun”, and “Pajak Gratification”. After the data is processed, 363 data *tweets* are generated from 1,078 data *tweets* which are ready to be used in this study. So the results of this study resulted in a positive *sentiment* value of 118 data and a negative *sentiment* of 245 data. As for *accuracy* values, namely 68.52%, *precision* 51.52%, and *recall* 48.57% with *naïve Bayes* modeling using a ratio of 70:30 for *training* data and data *testing*. So that it can be produced tax officials are still considered negative by the people of Indonesia.

Keywords: *Multinomial Naïve Bayes*, *RapidMiner*, Tax, CRISP-DM.

1. PENDAHULUAN

Pajak merupakan iuran masyarakat yang bersifat wajib kepada negara yang dapat dikenakan tanpa pelanggaran secara langsung, dan apabila ada masyarakat yang tidak membayar pajak maka akan dikenakan sanksi oleh negara. Oleh karena itu, pengaturan pajak atas dasar asas dalil dan merata, khususnya pajak langsung sebagai sumber penerimaan APBN merupakan dukungan yang sangat tepat untuk mengatasi masalah pembiayaan negara[1].

Penyalahgunaan kekuasaan dalam pemerintahan pajak terkait dengan praktik-praktik yang melanggar etika atau hukum yang dilakukan oleh pegawai pajak atau pejabat pajak. Hal ini dapat mencakup penyelewengan dalam penerimaan atau pengeluaran pajak, manipulasi data atau informasi keuangan, atau tindakan korupsi. Penyalahgunaan kekuasaan di bidang pajak juga bisa memiliki dampak yang merugikan pada masyarakat dan perekonomian, seperti kehilangan pendapatan negara, ketidakadilan dalam sistem pajak, dan ketidakpercayaan masyarakat terhadap pemerintah dan lembaga pajak. Dalam penelitian ini dilatarbelakangi atas penyalahgunaan kekuasaan oleh pegawai pajak dan hilangnya rasa kepercayaan masyarakat terhadap pegawai pajak

Penelitian ini dibatasi mengambil data dari *twitter* dengan *keywords* tertentu dan hanya menggunakan dua label yaitu label positif dan label negatif. Di penelitian ini *tweet* yang membahas terkait pegawai pajak akan dikumpulkan dan diolah apakah *tweet* tersebut berlabel positif atau negatif. Tujuannya untuk menganalisis opini publik atau masyarakat Indonesia terhadap pegawai pajak setelah banyaknya kasus penerimaan gratifikasi, menemukan klasifikasi atau label positif dan negatif terhadap pegawai pajak di media social *twitter* dan untuk menerapkan algoritma *naïve bayes* dalam melakukan analisis sentimen.

2. METODE PENELITIAN

2.1 *Crawling*

Dalam penelitian ini, pengumpulan data dilakukan dari salah satu platform media sosial *twitter*. Proses *crawling* merupakan sebuah proses dalam mengumpulkan data dan mengindeks data dari suatu laman *website* agar data dapat di analisis lebih lanjut. Proses *crawling* diolah menggunakan *RapidMiner* bertujuan untuk dapat memperoleh data *tweet* pada *twitter*. Data yang diambil adalah *tweet* yang terdapat pada media sosial *twitter* dengan menggunakan *twitter API*[2]. Proses *crawling* ini dilakukan untuk mengumpulkan *text* atau *tweet* dari *twitter* terkait pegawai pajak.

2.2 *Text Preprocessing*

Text Preprocessing berfungsi untuk mengubah data mentah menjadi data yang terstruktur dan dapat dimengerti. Data mentah seringkali tidak lengkap, tidak konsisten, dan mungkin mengandung kesalahan[3].

2.3 Pelabelan Data

Pelabelan data adalah proses mengklasifikasikan kalimat atau teks. Dua kategori dinilai dalam proses ini, yaitu sentimen positif dan negatif. Hasil pemerinkatan memberikan data sentimen positif dan negatif.

2.4 Penentuan Data Latih dan Data Uji

Pada tahap ini telah dilakukan identifikasi data *training* dan *testing*, dimana data akan dibagi menjadi dua tahap yaitu split data dan *cross-validation*. Data dibagi dengan rasio data yang diperlukan untuk menguji dan melatih dengan perbandingan yang diperlukan untuk data *testing* dan *training* adalah 60:40, 70:30, dan 80:20 sedangkan untuk *cross-validation* dengan pembagian acak menjadi 10 bagian (*number of folds* = 10).

2.5 Pembobotan TF-IDF

Bobot kata bertujuan untuk memberikan nilai bobot pada setiap kata. Perhitungan bobot ini membutuhkan dua hal, yaitu *term frequency* (TF) dan *inverse document frequency* (IDF). *Term frequency* adalah jumlah kata atau istilah tertentu dalam suatu dokumen. Sedangkan *document inverse frequency* adalah frekuensi kemunculan kata atau istilah dalam semua dokumen. Nilai IDF berbanding terbalik dengan jumlah dokumen yang mengandung *term* tertentu. Istilah yang jarang muncul di semua dokumen memiliki nilai IDF lebih tinggi dari nilai IDF istilah yang sering muncul. Jika setiap dokumen berisi *term* tertentu, nilai *term* IDF adalah 0. Hal ini menunjukkan bahwa istilah yang muncul di semua dokumen bukanlah istilah yang berguna untuk membedakan dokumen dengan topik tertentu. Pada penelitian ini penulis menggunakan kelas *Tfidf Vectorizer* (), pada perhitungan IDF ditambahkan 1 pembilang dan penyebut untuk mencegah pembagian dengan 0. Perhitungan rumus TF-IDF adalah sebagai berikut:

$$W_{t,d} = tf_{t,d} \times idf_t = tf_{t,d} \times LN \frac{(N + 1)}{(df_t + 1)} + 1 \quad (1)$$

Keterangan:

- $W_{t,d}$: Bobot TF-IDF
- $tf_{t,d}$: Jumlah frekuensi kata
- idf_t : Jumlah *inverse* frekuensi dokumen tiap kata
- df_t : Jumlah frekuensi dokumen tiap kata
- N : Jumlah total dokumen

2.6 Multinomial Naïve Bayes

Multinomial Naïve bayes adalah salah satu algoritma klasifikasi yang populer dalam metode pembelajaran mesin (*machine learning*) dengan mempelajari pola-pola dalam data pelatihan menggunakan perhitungan statistik dan perhitungan probabilitas berdasarkan pengalaman *Thomas Bayes*. Rumus untuk mencari peluang bersama menggunakan *Multinomial Naïve Bayes*. Dalam proses perhitungan klasifikasi, probabilitas kemunculan kata w_i dapat dihilangkan karena tidak mempengaruhi perbandingan hasil klasifikasi setiap kelas[4]. Rumus tersebut dapat diamati dengan persamaan sebagai berikut:

$$P(c_j|w_i) = \frac{P(c_j)P(w_i|c_j)}{P(w_i)} \quad (2)$$

Keterangan:

- $P(c_j|w_i)$: Peluang kategori c_j jika terdapat kemunculan kata w_i
- $P(c_j)$: Probabilitas kemunculan kelas prior
- $P(w_i|c_j)$: Peluang sebuah kata w_i yang ada pada kelas c_j
- $P(w_i)$: Peluang kemunculan kata w_i

Pada persamaan diatas, akan terjadi masalah ketika data berisi fitur dengan sampel yang tidak diketahui, sehingga proses *training* akan menyebabkan probabilitas yang diketahui adalah 0 dan kelas tidak dapat diprediksi. Oleh karena itu, metode parameter *laplace smoothing* digunakan untuk menghindari probabilitas nol dengan menambahkan nilai 1 pada setiap perhitungan. Oleh karena itu, persamaan dari *laplace smoothing* adalah sebagai berikut [4].

$$P(w_i|c_j) = \frac{\text{count}(w_i|c_j) + 1}{(\sum \text{count}(w|c_j) + |V|)} \quad (3)$$

Keterangan:

- $P(w_i|c_j)$: Likelihood atau peluang kata ke- i yang muncul pada kelas- j
- $\text{count}(w_i|c_j)$: Jumlah kemunculan kata w pada kategori c_j
- $\sum \text{count}(w|c_j)$: Peluang sebuah kata w_i yang ada pada kelas c_j
- $|V|$: Jumlah kata unik yang ada pada seluruh kategori

2.7 Confusion Matrix

Confusion matrix adalah alat evaluasi yang digunakan dalam pengolahan kata dan pemodelan statistik untuk mengevaluasi kinerja model klasifikasi atau sistem pengenalan pola. *Confusion matrix* adalah tabel yang mengklasifikasikan jumlah data uji yang benar dan jumlah data uji salah [5]. *Confusion matrix* terdiri dari *recall*, *precision*, dan *accuracy*. Berikut tabel dari *confusion matrix* [6].

Tabel 1. *Confusion Matrix*

	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	TP (<i>True Positive</i>)	FN (<i>False Negative</i>)
<i>Negative</i>	FP (<i>False Positive</i>)	TN (<i>True Negative</i>)

Keterangan:

- True Positive (TP)* : Jumlah data positif yang diklasifikasikan dengan benar oleh sistem.
True Negative (TN) : Jumlah data negatif yang diklasifikasikan dengan benar oleh sistem.
False Positive (FP) : Jumlah data positif yang salah diklasifikasikan oleh sistem.
False Negative (FN) : Jumlah data negatif yang salah diklasifikasikan oleh sistem.

Pada penelitian ini, matriks konfusi digunakan untuk mengevaluasi pengujian klasifikasi dan akurasi dari aplikasi yang dirancang untuk menemukan nilai akurasi, presisi, dan perolehan kembali dengan melakukan persamaan berikut:

- a. Akurasi adalah jumlah prediksi yang diklasifikasikan dengan benar dari seluruh data, baik positif maupun negatif. Nilai ini ditentukan dari persamaan di bawah ini:

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

- b. Presisi adalah jumlah prediksi yang diklasifikasikan dengan benar dari data agregat, baik dari segi presisi positif maupun negatif. Nilai presisi dapat ditentukan menurut persamaan di bawah ini:

$$\text{Presisi} = \frac{TP}{TP + FP} \quad (5)$$

- c. *Recall* adalah prediksi yang diklasifikasikan dengan benar dari data global atau data aktual, baik recall positif maupun *recall* negatif. Pemulihan dapat ditentukan menurut persamaan berikut:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

2.8 Cross Validation

Cross-validation adalah salah satu jenis pengujian *cross-validation* yang berfungsi untuk mengevaluasi kinerja proses suatu algoritma dengan membagi sampel data acak dan kelompok data. Tujuan validasi silang adalah untuk menguji seberapa baik performa model yang dilatih dapat digeneralisasikan terhadap data yang sebelumnya tidak terlihat. Langkah validasi silang:

- Dengan membagi data, data awalnya dibagi menjadi k himpunan bagian yang lebih kecil, biasanya dengan ukuran yang sama. Misalnya, validasi silang *k-fold* membagi data menjadi k himpunan bagian yang berbeda.
- Pelatihan dan pengujian, model dilatih menggunakan subset k-1 sebagai data pelatihan dan subset yang tersisa digunakan untuk menguji data saat menguji kinerja model.
- Evaluasi kinerja dan kinerja model diukur dengan parameter evaluasi yang sesuai seperti akurasi, presisi, dan pemulihan. Metrik ini dianalisis untuk setiap iterasi atau lipatan, lalu dirata-ratakan untuk memberikan gambaran tentang kinerja model secara keseluruhan.
- Validasi silang yang lebih tinggi, jika perlu, langkah validasi silang dapat diulangi dengan skema validasi silang yang lebih tinggi.

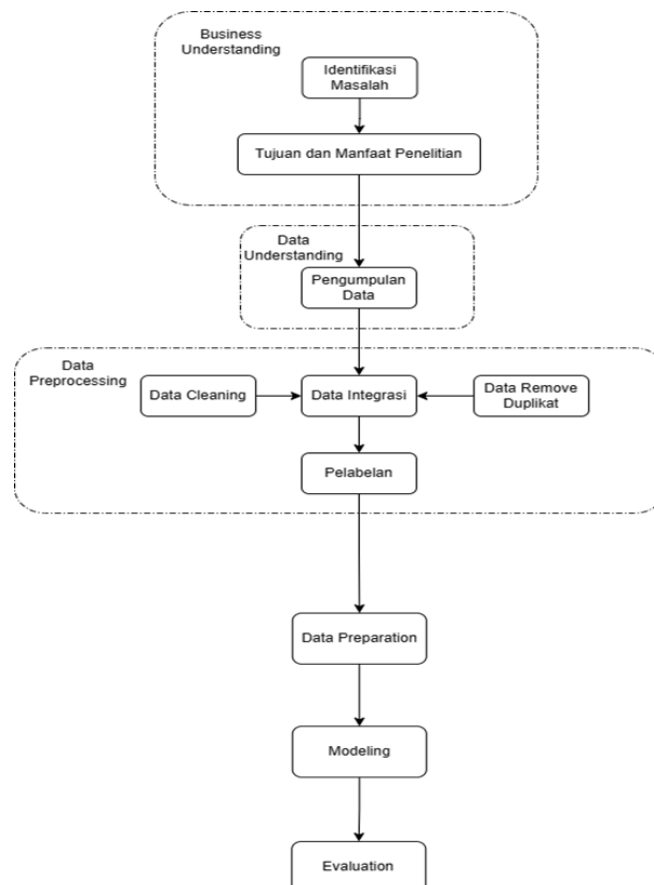
2.9 Cross Industry Standard Process Model for Data Mining (CRISP-DM)

Pendekatan umum yang digunakan dalam penambangan data dikembangkan pada tahun 1996 oleh analis dari beberapa industri seperti Daimler Chrysler, SPSS dan NCR. CRISP-DM adalah metode yang dapat diterapkan pada strategi pemecahan masalah umum serta metode referensi untuk data mining. Salah satu keuntungan menggunakan metode ini adalah menjelaskan langkah-langkah paling umum dalam proses penambangan data. CRISP-DM terdiri dari enam bagian yang saling terkait:

- a. *Business Understanding*

Dari permasalahan yang terjadi tentang pegawai pajak, dalam penelitian ini mencoba untuk memahami permasalahan tersebut dengan mempelajari artikel yang berkaitan dengan pajak, pegawai pajak, dan berita

- dari internet tentang pegawai pajak. Selanjutnya mencari trending topic di media sosial *twitter* yang berkaitan dengan pajak dan pegawai pajak.
- b. *Data Understanding*
Untuk dapat mengetahui pandangan atau opini publik dan masyarakat tentang pegawai pajak dengan pendekatan analisis sentimen diperlukan data respon publik dan masyarakat dari media sosial *twitter* melalui proses *crawling* dengan menggunakan beberapa *keywords* seperti kasus pajak, #RafaelAlun, #PejabatPajak, Indonesia, dan Jakarta.
 - c. *Data Preparation*
Pada tahap ini akan dilakukan preprocessing data. Data yang sudah didapat akan dilakukan pembersihan seperti penghapusan data duplikat, *case folding*, *tokenizing*, dan menghapus *stopword*.
 - d. *Modeling*
Kemudian dilakukan pemodelan terhadap data yang telah melalui tahap preprocessing menggunakan metode *split data*, *multinomial naïve bayes* dan *cross validation*.
 - e. *Evaluation*
Selanjutnya melakukan evaluasi terhadap data yang telah melalui tahap modeling menggunakan metode *confusion matrix*.
 - f. *Deployment*
Pada tahap penyebaran ini hasil penelitian akan ditemukan sesuai tujuan dari penelitian.



Gambar 1. Diagram Alir Tahapan Penelitian

3. HASIL DAN PEMBAHASAN

3.1 *Crawling* Data

Proses *crawling* data dari sebuah *website* agar dapat dianalisis lebih lanjut. Proses *crawling* diolah menggunakan *RapidMiner* sehingga data *tweet* dapat diperoleh dari *twitter*. Data yang diambil adalah *tweet* yang terdapat pada media sosial *twitter* dengan menggunakan *twitter* API [2]. Data yang digunakan 1.078 data dengan

kata kunci yaitu #RafaelPajak, #PejabatPajak #PajakIndonesia #RafaelAlun #GratifikasiPajak dengan tools *RapidMiner*.

3.2 Tahap Preprocessing

- a. **Case Folding** mengubah huruf besar menjadi huruf kecil (*lowercase*) [7]. Proses ini menggunakan bantuan tools *RapidMiner* dan operator *transform case*.

Tabel 1. Tahap *Case Folding*

Sebelum	Sesudah
Yg sabar yah jadi rakyat bayar pajak terus buat ngidupin kehidupan HEDON para PEJABAT	yg sabar yah jadi rakyat bayar pajak terus buat ngidupin kehidupan hedon para pejabat
Bayar pajak tepat waktu untuk kemajuan Indonesia	bayar pajak tepat waktu untuk kemajuan indonesia

- b. **Tokenize** menurut [6] proses yang dilakukan untuk memotong atau mebagi kalimat menjadi bagian-bagian atau kata-kata. Hasil dari proses ini disebut *tokenize*. Teks yang masih berupa kalimat akan dipisahkan kata per kata sehingga setiap kata dapat diperiksa.

Tabel 2. Tahap *Tokenize*

Sebelum	Sesudah
yg sabar yah jadi rakyat bayar pajak terus buat ngidupin kehidupan hedon para pejabat	“yg”, “sabar”, “yah”, “jadi”, “rakyat”, “bayar”, “pajak”, “terus”, “buat”, “ngidupin”, “kehidupan”, “hedon”, “para”, “pejabat”

- c. **Filter Token** pada tahap ini hasil *token* akan disaring berdasarkan panjang karakter untuk menyaring kata-kata yang tidak perlu atau tidak berarti. Proses pemilihan *token* sesuai dengan yang kita butuhkan, pada penelitian ini ukuran *token* dibatasi minimal 4 karakter (huruf) dan batas maksimal yang digunakan 25 karakter (huruf) [8].

Tabel 3. Tahap *Filter Token*

Sebelum	Sesudah
“yg”, “sabar”, “yah”, “jadi”, “rakyat”, “bayar”, “pajak”, “terus”, “buat”, “ngidupin”, “kehidupan”, “hedon”, “para”, “pejabat”	“sabar”, “jadi”, “rakyat”, “bayar”, “pajak”, “terus”, “buat”, “ngidupin”, “kehidupan”, “hedon”, “para”, “pejabat”

- d. **Stopwords** menurut [9] stopwords adalah proses menghilangkan kata-kata umum yang tidak memiliki arti. Pada tahap ini akan dilakukan *filtering* untuk menghilangkan kata-kata yang mungkin dihilangkan atau tidak memiliki arti.

Tabel 4. Tahap *Stopwords*

Sebelum	Sesudah
“sabar”, “jadi”, “rakyat”, “bayar”, “pajak”, “terus”, “buat”, “ngidupin”, “kehidupan”, “hedon”, “para”, “pejabat”	“sabar”, “rakyat”, “bayar”, “pajak”, “ngidupin”, “kehidupan”, “hedon”, “pejabat”

- e. **Stemming** merupakan proses untuk mengubah kata-kata berimbuhan menjadi kata dasar [10]. Tahapan mengubah kata menjadi kata dasar menurut kaidah bahasa Indonesia yang benar [6].

Tabel 5. Tahap *Stemming*

Sebelum	Sesudah
“sabar”, “rakyat”, “bayar”, “pajak”, “ngidupin”, “kehidupan”, “hedon”, “pejabat”	“sabar”, “rakyat”, “bayar”, “pajak”, “hidup”, “hidup”, “kaya”, “pejabat”

3.3 Pelabelan Data

Pelabelan adalah langkah yang dilakukan hanya setelah satu atau lebih data *clean text* dalam basis data (*database*) yang dihasilkan dari tahapan *preprocessing*. Pada tahap *labelling* dari 717 data *text* menghasilkan

sebanyak 118 data merupakan sentimen positif, 245 data dengan sentimen negatif, dan 354 data dengan sentimen netral. Hasil data *labelling* yang akan digunakan hanya *text* dengan sentimen positif dan negatif, setelah dilakukan *sort & filter* pada *tools Microsoft Excel* data berkurang menjadi 363 dari 717 data.

Tabel 6. Pelabelan Data

<i>Text</i>	Label
Kantor Wilayah Direktorat Jenderal Pajak Jakarta Timur menolak Gratifikasi dalam bentuk apapun	Positif
pajak untuk Ngasih Makan Para Pejabat Dan Jadi Ladang Korupsi Pejabat Di Indonesia pajak hiburan itu besar sekali 35%	Negatif
	Netral

3.4 Penentuan Data Latih dan Data Uji

Selama fase identifikasi data, *tweet* yang diberi tag akan dibagi menjadi data uji dan data latih. Pada penelitian ini pembagian data dilakukan dengan cara membagi dataset menjadi 70% data latih dan 30% data uji.

Tabel 7. Penentuan Data Latih dan Data Uji

Jenis Data	Jumlah
Data Latih	254
Data Uji	109
Data Set	363

3.5 Pengujian dan Nilai Akurasi

Selanjutnya dilakukan tahap pengujian dataset dengan menggunakan *confusion matrix* yang diperoleh dari tahapan *split* data. Pada gambar 2 dibawah ini merupakan hasil dari perhitungan dari *tools RapidMiner* dengan perbandingan 70:30.

accuracy: 68.52%

	true Negatif	true Positif	class precision
pred. Negatif	57	18	76.00%
pred. Positif	16	17	51.52%
class recall	78.08%	48.57%	

Gambar 2. Hasil Perhitungan dengan Perbandingan 70:30

Pada Tabel 8 dibawah ini merupakan hasil dari perhitungan terbaik menggunakan algoritma *Naïve Bayes* dengan *split* data perbandingan 70:30 dari *tools RapidMiner*. Dimana memiliki data *training* sebanyak 254 data, yang didalamnya terdapat 83 data sentimen positif dan 172 data sentimen negatif. Kemudian pada data *testing* sebanyak 109 data, yang terdapat 35 data sentimen positif dan 71 data sentimen negatif.

Tabel 7. Confusion Matrix Naïve Bayes dengan Split Data 70:30

	<i>True Negative</i>	<i>True Positive</i>	<i>Class Precision</i>
<i>Pred. Negative</i>	57	18	76.00%
<i>Pred. Positive</i>	16	17	51.52%
<i>Class Recall</i>	78.08%	48.57%	

Perhitungan dari Tabel 3 adalah sebagai berikut:

- $Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{17+57}{108} = \frac{74}{108} = 0.6852 * 100\% = 68.52\%$
- $Precision = \frac{TP}{TP+FP} = \frac{17}{17+16} = \frac{17}{33} = 0.5152 * 100\% = 51.52\%$
- $Recall = \frac{TP}{TP+FN} = \frac{17}{17+18} = \frac{17}{35} = 0.4857 * 100\% = 48.57\%$

Berdasarkan pada perhitungan gambar 3 dengan menggunakan algoritma *naïve bayes*, maka dapat disimpulkan bahwa *precision* dan *recall* menunjukkan tingkat ketepatan antara informasi yang diminta dengan jawaban yang diberikan oleh sistem yaitu negatif. Sedangkan untuk *accuracy* menunjukkan nilai keakuratan yang sama dalam memprediksi suatu sentimen dengan sebuah algoritma. Dan setelah dilakukan pencarian terhadap ketiga nilai *confusion matrix* tersebut dengan menggunakan *tools RapidMiner*, maka hasil yang diperoleh adalah nilai *accuracy* 68.52%, *precision* 51.52%, dan *recall* 48.57%. Dapat diartikan data dan metode yang digunakan mampu memberikan hasil yang akurat.

4. KESIMPULAN

Berdasarkan hasil analisis yang didapat setelah melakukan penelitian, maka dapat diambil kesimpulan untuk memprediksi klasifikasi opini masyarakat terhadap *tweet* pegawai pajak yaitu untuk mendapatkan sentimen pada dataset yaitu dilakukan proses *labelling* dengan 717 data oleh pakar dalam bidang bahasa Indonesia, sehingga didapatkan sebanyak 118 data sentimen positif, 245 data sentimen negatif, dan 354 data sentimen netral. Penerapan algoritma *naïve bayes* terbukti dapat mengklasifikasikan sentimen masyarakat terhadap pegawai pajak dengan nilai *accuracy* 68.52%. Penerapan algoritma *naïve bayes* terbukti dapat menentukan tingkat perbandingan sentimen positif dan negatif terhadap pegawai pajak dengan hasil 118 data sentimen positif, dan 245 data sentimen negatif. Sehingga secara umum masyarakat Indonesia memiliki respon negatif terhadap pegawai pajak. Langkah utama dalam penelitian ini meliputi: *crawling*, *preprocessing*, pelabelan, pembagian data, dan klasifikasi *naïve bayes*. Langkah *preprocessing* yang baik akan menjadi faktor penentu dalam membentuk hasil yang optimal untuk langkah selanjutnya. Dapat disimpulkan nilai sentimen pegawai pajak di masyarakat yaitu negatif.

DAFTAR PUSTAKA

- [1] S. ANGGRIANI, *SISTEM ADMINISTRASI PERHITUNGAN PAJAK PENGHASILAN PASAL 21 ATAS PEGAWAI TETAP PADA PT. PERTAMINA RETAIL Unit SPBU Coco 11201107 MEDAN*. 2022.
- [2] P. Arsi and R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.0813944.
- [3] L. A. Andika, P. A. N. Azizah, and R. Respatiwulan, "Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier," *Indones. J. Appl. Stat.*, vol. 2, no. 1, p. 34, 2019, doi: 10.13057/ijas.v2i1.29998.
- [4] A. A. Nabhan, B. Rahayudi, and D. E. Ratnawati, "Klasifikasi Tweets Masyarakat yang Membicarakan Layanan GoFood dan GoRide pada GoJek Dimedia Sosial Twitter Selama Masa Kenormalan Baru (New Normal) dengan Metode Naïve Bayes," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 7, pp. 3018–3025, 2021.
- [5] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021, [Online]. Available: <http://ejournal.tunasbangsa.ac.id/index.php/jsakti/article/view/369>
- [6] W. Yulita, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier," *J. Data Min. dan Sist. Inf.*, vol. 2, no. 2, p. 1, 2021, doi: 10.33365/jdmsi.v2i2.1344.
- [7] H. Rhomadhona and J. Permedi, "Klasifikasi Berita Kriminal Menggunakan Naïve Bayes Classifier (NBC) dengan Pengujian K-Fold Cross Validation," *J. Sains dan Inform.*, vol. 5, no. 2, pp. 108–117, 2019, doi: 10.34128/jsi.v5i2.177.
- [8] A. Hendra and F. Fitriyani, "Analisis Sentimen Review Halodoc Menggunakan Naïve Bayes Classifier," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 6, no. 2, pp. 78–89, 2021, doi: 10.14421/jiska.2021.6.2.78-89.
- [9] M. Christianto, J. Andjarwirawan, and A. Tjondrowiguno, "Aplikasi analisa sentimen pada komentar berbahasa Indonesia dalam objek video di website YouTube menggunakan metode Naïve Bayes classifier," *J. Infra*, vol. 8.1, pp. 255–259, 2020.
- [10] A. Guterres, Gunawan, and J. Santoso, "Stemming Bahasa Tetun Menggunakan Pendekatan Rule Based," *Teknika*, vol. 8, no. 2, pp. 142–147, 2019, doi: 10.34148/teknika.v8i2.224.