



Klasifikasi Opini Masyarakat Terhadap Jasa Ekspedisi JNE dengan Naïve Bayes

Fithri Selva Jumeilah^{*)}

Program Studi Sitem Informasi, STMIK Global Informatika MDP

Naskah Diterima : 1 Maret 2018; Diterima Publikasi : 17 April 2018

DOI : 10.21456/vol8iss1pp92-98

Abstract

The large number of online sales transactions has increased the number of service users. One of the companies engaged in the delivery service in Indonesia is Tiki Nugraha Ekakurir or more known JNE. Currently, JNE service users reach 14.000.000 per month. JNE has used many media communications with its customers one of them with Twitter. The number of followers of JNECare is 108,000 and the number of tweets is 375,000. The number of comments for people who can be used to see what they think of JNE is an inseparable comment is a negative, positive or neutral category. To simplify the grouping of comments, the data will be classified using the Naïve Bayes method present in Rstudio. The amount data used on the internet is 1725 tweets. The data will be divided into allegations of 70% data training as much as 1208 data and 30% data testing or as many as 517 data. Before the data is classified the previous data must go through the process of preprocessing that is changing all the letters into lowercase and other letters other than letters and spaces (case folding), tokenizing words, and the removal of the word common (stopword remove). After the data is cleared the data will be labeled manually one by one and new data can be used for the training process to get the probability model for each category. Probailitas obtained by using Naïve bayes algorithm. Models obtained from the training will be used using data testing. The categories obtained from the test will be used to process the data used by using the confusion matrix and will calculate the accuracy, precision and recall. From the results of the classification of JNE comments obtained that Naïve Bayes was able to classify the data well. This is evidenced by the average percentage accuracy of 85%, 78% precision and 67% recall.

Keywords: Text Mining; Classification; Naïve bayes; Rstudio; Crawling.

Abstrak

Banyaknya jumlah transaksi penjualan online membuat bertambahnya pengguna jasa pengiriman. Salah satu perusahaan yang bergerak di bidang jasa pengiriman di Indonesia adalah Tiki Jalur Nugraha Ekakurir atau lebih dikenal JNE. Saat ini, pengguna jasa JNE mencapai 14.000.000 perbulan. JNE telah memanfaatkan banyak media komunikasi dengan pelanggannya salah satunya dengan Twitter. Jumlah *followers* dari JNECare adalah 108.000 dan jumlah *tweets* sebanyak 375.000. Banyaknya komentar pelanggan dapat dimanfaatkan untuk melihat gambaran opini mereka tentang JNE apakah komentar dominan adalah kategori negatif, positif atau netral. Untuk mempermudah pengelompokan komentar, data akan diklasifikasi menggunakan metode Naïve Bayes yang ada pada Rstudio. Jumlah data yang digunakan dalam penelitian berjumlah 1725 *tweets*. Data akan dibagi dua 70% *data training* sebanyak 1208 data dan 30% *data testing* atau sebanyak 517 data. Sebelum data diklasifikasi data sebelumnya harus melalui proses *preprocessing* yaitu mengubah semua karakter huruf menjadi *lowercase* dan membuang karakter selain huruf dan spasi (*case folding*), pemotongan kata (*tokenizing*), dan penghapusan kata umum (*stopword removing*). Setelah data dibersihkan data akan dilabeli secara manual satu persatu dan baru data dapat digunakan untuk proses *training* untuk mendapatkan model probabilitas untuk setiap kategori. Probailitas diperoleh dengan menggunakan algoritma Naïve bayes. Model yang diperoleh dari tahap *training* akan diuji dengan menggunakan *data testing*. Kategori yang diperoleh dari tahap *testing* akan dibandingkan dengan kategori data yang sebenarnya dengan menggunakan dan disajikan ke dalam *confusion matrix* dan akan dihitung akurasi, *precision* dan *recall*. Dari hasil klasifikasi komentar pelanggan JNE diperoleh bahwa Naïve bayes mampu mengklasifikasikan data dengan baik. Hal ini dibuktikan dari nilai rata-rata presentase akurasi 85%, *precision* 78% dan *recall* 67%.

Keywords: Penambangan Teks; Klasifikasi; Naïve bayes; Rstudio; Penarikan Data.

^{*)} Fithri Selva: fithri.selva@mdp.ac.id

1. Pendahuluan

Tiki Jalur Nugraha Ekakurir atau sering dikenal JNE merupakan perusahaan yang bergerak dibidang jasa pengiriman dan logistik. Semakin majunya e-commerce di Indonesia menyebabkan meledaknya jumlah pengguna jasa pengiriman. Pengguna JNE setiap bulannya dapat mencapai 14 juta pengiriman. Tidak heran jika saat ini banyak sekali bermunculan jasa pengiriman baru di Indonesia. Untuk menjaga pelanggan JNE menyediakan media komunikasi dengan pelanggan salah satunya dengan media sosial Twitter.

Banyak sekali pelaku bisnis memanfaatkan media sosial Twitter untuk mengetahui pendapat pelanggan tentang produknya karena Twitter merupakan salah satu media sosial yang populer di Indonesia. Pengguna Twitter dapat berbagi informasi melalui pesan singkat hingga 280 karakter yang sering disebut *tweet*. JNE memiliki 2 akun Twitter yaitu @JNE_ID sebagai akun perusahaan dan akun @JNECare sebagai akun untuk melayani pelanggan. JNECare dibuat pada tahun 2014 yang telah memiliki 375.000 *tweets* dan 108.000 *followers*.

Banyaknya komentar dari pelanggan JNE pada Twitter dapat digunakan untuk mengetahui pendapat masyarakat tentang JNE. Pendapat pelanggan akan menggambarkan pandangan mereka terhadap JNE dan mempengaruhi pandangan orang lain. Namun mudahnya pelanggan memberikan pendapat dengan media Twitter menyebabkan ledakan komentar. Untuk mengolah komentar tersebut, tentunya membutuhkan waktu yang banyak jika dilakukan secara konvensional. Dengan demikian dibutuhkan sebuah sistem yang mampu mengklasifikasi kalimat terkait komentar pelanggan JNE dan menemukan pola untuk kelompok positif, negatif dan netral. Pengelompokan data dilakukan dengan menerapkan teknik *text mining* yaitu klasifikasi.

Penelitian terkait klasifikasi teks sudah banyak dilakukan diantaranya: analisis data twitter “Make-In-India” dengan menggunakan R yang lakukan oleh Rani dan Rani (2016) untuk memvisualisasikan pendapat dari pengguna Twitter, klasifikasi analisis sentimen terhadap tokoh publik dengan menggunakan metode klasifikasi Naïve Bayes dan Support Vector Machine (SVM) (Hidayatullah dan Azhari, 2014), pengklasifikasian penelitian dengan SVM untuk memudahkan pencarian penelitian yang memiliki kategori yang sama (Jumeilah, 2017) dan analisis sentimen masyarakat berdasarkan *emoticon* pada data *tweets* tentang film (Kiruthika *et al.*, 2016).

Naïve Bayes telah berhasil diterapkan di beberapa penelitian salah satunya untuk memprediksi besarnya penggunaan listrik untuk tiap rumah tangga (Saleh, 2015). Dari hasil pengujian yang menggunakan 60 data diperoleh 47 data yang mampu diklasifikasikan dengan benar atau dengan kata lain nilai akurasi 78,333% untuk keakuratan prediksi. Dari hasil

penelitian Ting *et al.* (2011), diperoleh Naïve Bayes mampu mengklasifikasikan data lebih akurat dibanding algoritma lainnya dengan nilai akurasi sebesar 97%. Dengan demikian, penelitian ini akan menggunakan algoritma Naïve bayes untuk mengklasifikasikan komentar pelanggan JNE menjadi tiga kategori yaitu: negatif, positif atau netral.

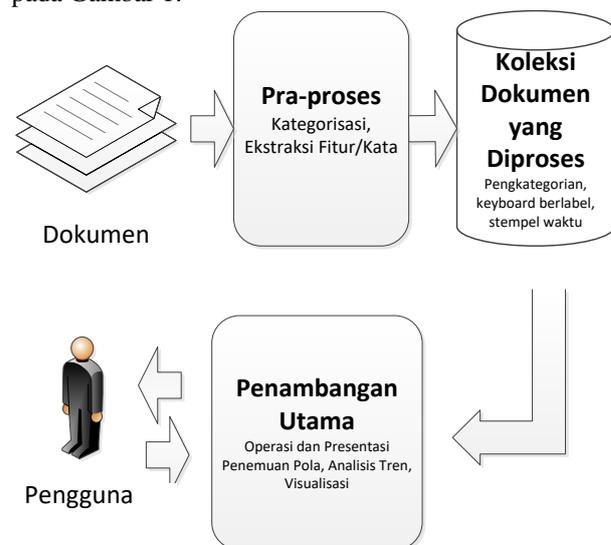
2. Kerangka Teori

2.1. Text Mining

Menurut Feldman dan Sanger (2007), *text mining* adalah proses mengekstrak informasi yang berguna dari sekumpulan dokumen dari waktu ke waktu melalui identifikasi pola yang menarik. Sedangkan menurut Kini *et al.* (2015), *text mining* adalah proses penemuan informasi baru dengan mengekstrak pola secara otomatis dari berbagai sumber teks. *Text mining* sering juga disebut analitik teks. Document sumber data dari *text mining* dapat berasal dari mana saja seperti sosial media, buku, koran, email dan lain-lain.

Text mining dan *data mining* memiliki perbedaan di sumber datanya. Sumber data dari *texts mining* adalah tekstual yang tidak ada format dan tidak terstruktur. Berbeda dengan *data mining*, dimana sumber datanya memiliki format terstruktur dan tidak hanya teks.

Text mining memiliki beberapa tahapan utama: *preprocessing task*, *processed document collection*, dan *core mining operations* (Feldman dan Sanger, 2007). Adapun arsitektur dari *text mining* dapat dilihat pada Gambar 1.



Gambar 1. Arsitektur *Text Mining* (Feldman dan Sanger, 2007)

Tujuan utama *text mining* sama saja dengan tujuan manusia membaca dokumen yaitu mencari intisari teks. Adapun beberapa teknik dari *text mining* yang dapat digunakan untuk mengekstrak pola meliputi: *classification*, *clustering*, *information extraction*,

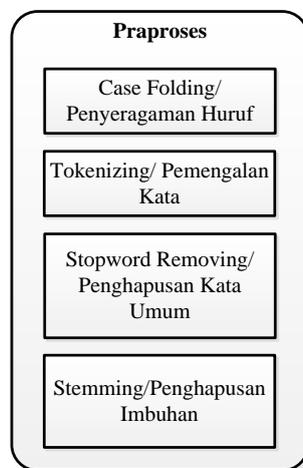
trend analysis, distribution analysis, dan association rules.

2.2. Preprocessing

Data yang digunakan untuk *text mining* memiliki format yang tidak terstruktur sehingga tidak bisa digunakan pada tahap berikutnya. Agar dapat digunakan, dibutuhkan proses *preprocessing*. Tujuan dari *preprocessing* adalah membersihkan data dari noise sehingga data menjadi lebih kecil dan lebih terstruktur (Feldman dan Sanger, 2007).

Dokumen memiliki beraneka ragam karakter mulai dari huruf sampai tanda baca maka dari itu diperlu di bersihkan. Untuk membersihkan data dibutuhkan penyeragaman semua huruf menjadi kecil semua ('a'-'z') atau sebaliknya ('A'-'Z'), penghapusan karakter yang tidak valid dan penghapusan kata yang umum digunakan.

Menurut Feldman dan Sanger (2007), *preprocessing* data tekstual memiliki beberapa tahapan, yaitu: *case folding, tokenizing, stopword removing* dan *stemming* yang dapat dilihat pada Gambar 2.



Gambar 2. Tahapan *Preprocessing*

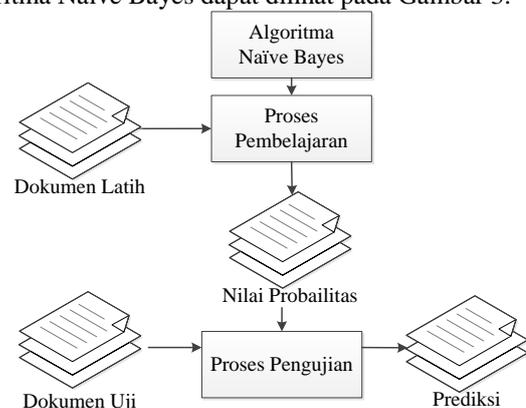
Tahapan pertama dari *preprocessing* adalah *case folding* yaitu menyeragamkan semua karakter huruf menjadi kecil atau kapital semua dan menghilangkan semua karakter yang tidak valid seperti *Uniform Resources Locator (URL)*, tanda baca dan angka akan dihilangkan dan diganti menjadi pemisah antar kata atau spasi. Tahapan kedua adalah pemengalan teks menjadi token-token dengan acuan pemisah setiap token adalah spasi atau dengan kata lain memecah teks menjadi kata. Tahapan pemengalan kata ini sering dikenal dengan *tokenizing*.

Setelah setiap kata dipenggal, maka dapat dilakukan tahapan penghapusan kata yang umum atau yang sering dikenal dengan *stopword removing*. Contoh kata yang termasuk kata umum adalah saya, kamu, adalah, ini, yang, serta, pada, akan, di, dengan, tidak dan lain sebagainya. Tahapan yang terakhir

adalah proses *stemming* yang bertujuan menghapus imbuhan kata baik itu awalan ataupun akhiran, sehingga hanya tersisa *stem* atau akar kata. Pada tahapan *stemming* biasanya digunakan pada data teks Inggris (Indraloka dan Santosa, 2017).

2.3. Naive Bayes

Naive Bayes adalah salah satu algoritma klasifikasi yang ditemukan oleh Thomas Bayes. Klasifikasi adalah proses untuk memperoleh aturan atau model yang dapat mengklasifikasikan data baru yang belum pernah dipelajari dengan mempelajari sekumpulan data yang lama (Han *et al.*, 2012). Naive Bayes adalah algoritma klasifikasi probabilistik yang sederhana berdasarkan Teorema Bayes. Prinsip umumnya adalah mengasumsikan bahwa nilai suatu atribut tidak bergantung dan mempengaruhi atribut yang lainnya. Model Naive Bayes memungkinkan setiap atribut memiliki kontribusi yang sama terhadap keputusan akhir dan komputasinya lebih efisien bila dibandingkan dengan algoritma pengklasifikasi teks lainnya (Ting *et al.*, 2011). 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 (Kini *et al.*, 2015). Skema dari algoritma Naive Bayes dapat dilihat pada Gambar 3.



Gambar 3. Skema Naive Bayes

Naive Bayes mampu memprediksi masa depan dengan mempelajari data pada masa sebelumnya (Ahmed *et al.*, 2014). Naive Bayes adalah metode yang tidak rumit dan banyak digunakan dengan nilai akurasi yang tinggi dengan waktu proses yang cepat (Duan Li-guo *et al.*, 2014). Algoritma Naive Bayes adalah algoritma tercepat dengan model yang sederhana dan *noise* yang sedikit tidak dapat mempengaruhi hasil akhirnya (Kini *et al.*, 2015). Naive Bayes adalah algoritma klasifikasi probabilistik yang sederhana berdasarkan Teorema Bayes. Naive Bayes menjumlahkan frekuensi dan kombinasi nilai dari *data set* untuk menghitung probabilitas (Saleh, 2015). Probabilitas dihitung dengan menggunakan Persamaan 1 (Han *et al.*, 2012).

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

Dimana $P(H|X)$ adalah probabilitas posterior H berdasarkan syarat X yang diperoleh dari probabilitas hipotesis H ($P(H)$) dikali dengan probabilitas X berdasarkan kondisi hipotesis H ($P(X|H)$) dan dibagi dengan probabilitas X ($P(X)$).

2.4. Sentiment Analysis

Sentiment analysis adalah proses penentuan bagaimana perasaan seseorang terhadap topik tertentu. *Sentiment analysis* sering juga dikenal sebagai penambangan opini masyarakat tentang suatu entitas yang diungkapkan dalam teks. *Sentiment analysis* bertujuan untuk mengidentifikasi suasana hati, sikap individu dan pendapat masyarakat tentang suatu entitas. (Pang dan Lee, 2008) Informasi tekstual dapat dibagi menjadi dua jenis : fakta dan opini. Fakta diperoleh dari ekspresi obyektif sedangkan opini adalah ungkapan subjektif yang menggambarkan penilaian seseorang terhadap suatu. *Sentiment analysis* banyak diterapkan pada ulasan pelanggan dan media sosial untuk melihat gambaran pendapat pelanggan. Sentimen analisis adalah salah satu bagian dari klasifikasi teks yang bertujuan menemukan sentimen penulis dalam teks.

3. Metode

Klasifikasi data *tweets* pelanggan JNE dapat diperoleh dengan beberapa tahapan mengikuti tahapan yang ada pada Gambar 1 dan Gambar 3. Dimana tahapannya diawali dari identifikasi masalah, *crawling data tweets*, *preprocessing*, pelabelan data *tweets* secara manual satu persatu, kemudian tahap *training* untuk mendapatkan model probabilitas, proses *testing* model dengan menggunakan data *testing* dan terakhir analisis hasil dengan menghitung nilai akurasi hasil prediksi yang dihasilkan.

3.1. Identifikasi Masalah

Meledaknya jumlah opini pelanggan pada akun Twitter JNECare akan menyulitkan JNE untuk melihat gambaran opini masyarakat tentang JNE apakah negatif, positif atau netral karena opini tersebut belum terklasifikasi. Dibutuhkan sebuah sistem yang mampu mengklasifikasikan opini masyarakat yang dapat digunakan oleh JNE untuk memperbaiki kinerjanya. Salah satu algoritma yang dapat digunakan untuk klasifikasi adalah Naïve bayes dengan menggunakan *tools* Rstudio.

3.2. Crawling Data Tweets

Sumber data yang digunakan dalam penelitian ini adalah data *mention* dari akun JNECare. Pengambilan data menggunakan *Application Programming Interfaces* (API) "twitterR" menggunakan RStudio.

Untuk pengambilan data dari Twitter harus memiliki *customer key*, *customer secret*, *access token* dan *access key* yang diperoleh dari Twitter API. Setelah melalui tahapan autentifikasi baru data *tweets* dapat diambil dengan menentukan *key word* yang akan diambil.

Adapun *key word* yang digunakan adalah "@JNECare" dari tanggal 14 Februari 2018 sampai dengan 19 Februari 2018. Data yang telah diambil akan disimpan kedalam file dengan format .df. Dari hasil *crawling* diperoleh data *tweets* sebanyak 2.349 data. Contoh hasil data *tweets* yang diperoleh: "@JNECare Min paket ane kok gak sampe2 , trus input resinya online kok telat dan lamaa banget,, resi 081370001818018", dan "@JNECare min 8811621800615010 sudah sampai thank!!!".

Dari contoh data *tweets* yang diperoleh dapat dilihat bahwa banyak cara penulisan kata yang tidak benar. Perlu dilakukan proses *preprocessing* data agar data menjadi seragam dan lebih terstruktur.

3.3. Preprocessing

Tahapan *preprocessing* diawali dengan *case folding* yaitu penghapusan karakter yang tidak valid seperti Uniform Resource Locator (URL), *mention* (@), *hashtag* (#) dan penghapusan semua karakter selain huruf dan spasi. Setelah hanya tersisa huruf dan spasi, dilakukan perubahan semua karakter huruf menjadi kecil semua (*lowercase*).

Tahap kedua *preprocessing* terjadi perubahan yang seharusnya *tokenizing* tetapi sebelumnya diperlukan penormalan data terlebih dahulu. Data *tweets* yang diperoleh banyak memiliki kata yang tidak mengikuti kaedah Bahasa Indonesia. Karena media sosial, terutama Twitter memiliki batasan jumlah karakter setiap *tweet* hanya boleh 280 karakter. Karena itu banyak pengguna yang menyingkat kata-kata yang sering digunakan. Selain itu, banyak juga pengguna Twitter yang tidak menggunakan bahasa Indonesia yang baku dan banyak juga kata yang memiliki makna yang sama seperti "mohon" dan "tolong" yang harus diseragamkan. Dengan demikian, diperlukan proses normalisasi kata-kata yang dilakukan secara manual. Normalisasi kata dilakukan untuk setiap kata *tweets* satu persatu.

Setelah data normal dilakukan tahapan *tokenizing*. *Tokenizing* yaitu pemenggalan setiap kata berdasarkan karakter spasi dengan fungsi:

$dtm = \text{DocumentTermMatrix}(\text{myCorpus})$.

Terakhir, *stopword removing* adalah pembersihan data dari kata-kata umum. Beberapa kata yang masuk dalam kata umum adalah di, ke, kamu, saya, untuk, ini itu, serta, dan lain lain. Namun, ada beberapa kata umum yang tidak akan dihapus karena kata umum tersebut mempengaruhi opini pelanggan, seperti

belum, tidak dan sudah. Berikut contoh data tweet setelah melalui tahap praproses: “admin paket tidak sampai input resinya online telat lama resi”, dan “sudah sampai terimakasih”.

3.4. Pelabelan Data Tweets

Dari data tweets data yang diambil hanya atribut teks saja dan setiap tweet akan dilakukan pelabelan secara manual apakah tweet tersebut termasuk opini negatif, positif atau netral. Pelabelan dilakukan berdasarkan kata kunci terkait dengan kategori. Kategori negatif adalah kumpulan komentar yang berisikan komplain, sedangkan untuk kategori netral adalah kategori yang bertanya mengenai JNE tanpa ada pernyataan negatif atau positif, dan untuk kategori positif merupakan kumpulan komentar pelanggan yang memberikan komentar positif. Kategori negatif akan diberikan label -1, kategori positif 1 dan sedangkan kategori netral akan diberi label 0.

Setelah data melalui tahap *preprocessing*, beberapa data *tweets* tidak digunakan. Hal ini disebabkan oleh beberapa hal, salah satunya data tweet kosong karena hanya mengandung beberapa kata dan kata tersebut kata umum, hanya mengandung *emoticon*, *Retweets* (RT) atau URL tanpa diikuti keterangan sehingga pada tahap *preprocessing*, *tweets* menjadi kosong. Selain itu ada beberapa data *tweets* yang ganda sehingga dilakukan penghapusan. Data yang akan dipakai hanya 1725 data *tweets*. Jumlah data kelas positif sebanyak 90 data, data kelas negatif sebanyak 1058 data dan kelas netral 577 data.

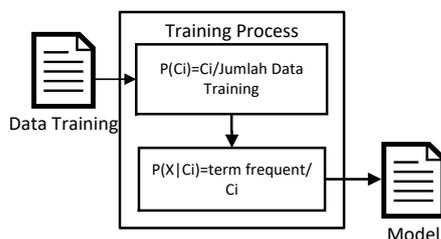
Data dibagi menjadi dua bagian, yaitu *data training* dan *data testing*. Untuk *data training* menggunakan 1208 data *tweets* diambil dari 70% dari setiap kategori. Sedangkan untuk *data testing* 30% dari setiap kategori atau menggunakan 517 data *tweets*. Jumlah data *training* dan *data testing* untuk masing-masing kategori dapat dilihat pada Tabel 1.

Tabel 1. Pembagian data *training* dan *testing*

	Negatif	Netral	Positif
Data Training	740	405	63
Data Testing	318	172	27

3.5. Training

Proses *training* akan menggunakan algoritma Naïve Bayes yang sudah ada pada Rstudio. Adapun tahapan dari proses *training* menggunakan algoritma Naïve Bayes dapat dilihat pada Gambar 4.



Gambar 4. Tahapan Proses *Training*

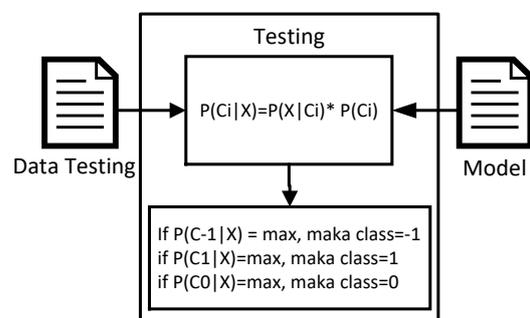
Dari data *training* akan dihitung jumlah data setiap kategori (C_i). Setelah diperoleh, jumlah setiap kategori (C_i) dibagi dengan jumlah keseluruhan data atau yang sering disebut dengan nilai prior setiap kelas ($P(C_i)$). Nilai prior adalah nilai peluang atau probabilitas data akan diklasifikasikan ke dalam kelas tersebut. Kemudian, hitung nilai peluang kemunculan dari setiap kata untuk setiap kategori ($P(X|C_i)$). Jika semua peluang kata telah dihitung, simpan semua nilai probabilitas ke dalam model.

Pada Rstudio untuk mendapatkan nilai probabilitas $P(C_i)$ dan $P(X|C_i)$ cukup menggunakan fungsi Naïve bayes yang ada di *library* “e1071”. Berikut contoh proses *training* dengan Naïve Bayes yang akan disimpan kedalam variabel “classifier”.

```
classifier <- naiveBayes(trainNB, df.train$class, laplace = 1)
```

3.6. Testing

Hasil dari tahap *training* akan menghasilkan model yang akan digunakan untuk klasifikasi data *testing*. Dimana klasifikasi akan memiliki beberapa tahapan yang dapat dilihat pada Gambar 5.



Gambar 5. Tahapan Proses *Testing*

Dari data *testing* akan dilihat kata-kata apa saja yang akan muncul setiap tweetnya. Untuk menentukan *tweets* tersebut apakah masuk kategori negatif, netral atau positif dibutuhkan nilai probabilitas setiap kata ($P(X|C_i)$) untuk masing kategori. Kemudian, hitung $P(C_i|X)$ dengan cara mengalikan $P(X|C_i)$ semua kata dan prior kelas ($P(C_i)$). Nilai $P(X|C_i)$ dan $P(C_i)$ diperoleh dari model hasil proses *training*. Hitung $P(C_i|X)$ untuk setiap kelas. Kelas yang memiliki nilai $P(C_i|X)$ paling tinggi maka *tweets* tersebut akan masuk kedalam kelas tersebut. Apabila semua *tweets* telah ditentukan kelasnya maka dapat dibandingkan dengan kelas sebenarnya apakah sama atau tidak.

Untuk melakukan proses *testing* dapat dipermudah dengan adanya fungsi “predict” cukup menginput model yang telah disimpan dalam variabel “classifier” dan data *testing* saja.

```
pred <- predict(classifier, newdata=testNB)
```

Data akan diklasifikasikan berdasarkan nilai probabilitas. Jika nilai probabilitas untuk kelas positif yang paling tinggi maka *tweets* tersebut akan masuk kelas positif begitu juga sebaliknya.

3.7. Analisis Hasil

Probabilitas yang diperoleh dari proses *training* akan diuji dengan *data testing* untuk mengukur keakuratannya. Dari hasil *testing* akan dibandingkan dengan label sebenarnya dan akan menghasilkan *confusion matrix*. Dari *confusion matrix* dapat dihitung nilai akurasi, *precision* dan *recall* (Hossin dan Sulaiman, 2015). Akurasi digunakan untuk melihat gambaran seberapa akurat sistem mengklasifikasikan dengan tepat. Akurasi dapat dihitung dengan menggunakan Persamaan 2.

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

Akurasi didapat dari jumlah data yang berhasil dikategorikan sesuai kategorinya (TP) ditambah jumlah data negatif benar (TN) dibagi dengan jumlah semua data. *Precision* digunakan untuk mengukur pola positif yang berhasil diprediksi dengan benar (TP) dari total pola prediksi dalam kelas positif baik itu data yang berhasil diprediksi dengan benar (TP) dengan data yang berhasil diprediksi bukan kelas positif (FP). Untuk menghitung *precision* menggunakan Persamaan 3.

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

Sedangkan *recall*, untuk melihat ukuran keberhasilan dengan membagi data yang berhasil diprediksi dengan benar (TP) dibagi dengan jumlah dari data yang berhasil diprediksi dengan benar (TP) dan data yang berhasil diprediksi negatif salah (FN). Untuk menghitung nilai *recall* dapat dilihat pada Persamaan 4.

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

4. Hasil dan Pembahasan

Setelah data melalui tahap *preprocessing* maka baru dapat dilakukan *data training* dan *data testing*. Dimana *data training* dan *testing* dilakukan dengan menggunakan algoritma Naïve Bayes. Dari hasil *training* akan menghasilkan nilai priori probabilitas untuk setiap kategori. Nilai a-priori probabilitas yang dihasilkan untuk kategori negatif adalah 0.591, kategori netral sebesar 0,359 dan kategori positif sebesar 0,05. Dari *data set* yang digunakan sudah terlihat bahwa data yang paling banyak digunakan adalah kategori negatif sehingga nilai probabilitas yang paling besar juga kategori negatif.

Selain nilai perior, juga diperoleh nilai probabilitas setiap kata yang muncul untuk setiap kategori. Contoh nilai probabilitas kata dapat dilihat pada Tabel 2.

Tabel 2. Contoh nilai P(X|C_i)

Kata	Negatif	Netral	Positif
Komplain	0.02156	0.00737	0.01538
Cepat	0.00539	0.00491	0.03077
Informasi	0.02291	0.04422	0.01538

Setelah diperoleh nilai probabilitas, maka tahap pengujian siap dilakukan. Data *tweets* yang diujikan akan dilihat probabilitas untuk setiap kata dan dihitung jumlah P(X|C_i). Dari hasil pengujian 517 data dimana *tweets* kategori negatif sebanyak 318 data, kategori netral sebanyak 172 data dan sebanyak 27 data *tweets* positif diperoleh hasil yang dapat dilihat pada Tabel 3.

Tabel 3. Daftar hasil pengujian pertama

	Negatif	Netral	Positif
Negatif	256	39	4
Netral	60	133	11
Positif	2	0	12

Berdasarkan data pada Tabel 3, maka dapat diubah menjadi *confusion matrix* yang akan digunakan untuk menghitung akurasi, *precision*, dan *recall* dimana hasilnya untuk setiap kelas dapat dilihat pada Tabel 4.

Tabel 4. Perbandingan akurasi, *precision* dan *recall*

Kelas	Akurasi	<i>Precision</i>	<i>Recall</i>
Negatif	0,80	0,86	0,80
Netral	0,79	0,65	0,77
Positif	0,97	0,85	0,44
Rata-rata	0,85	0,79	0,67

Dari Tabel 4 diperoleh bahwa kategori yang memiliki akurasi tertinggi adalah kategori positif tetapi nilai *recall* paling kecil. Hal ini membuktikan bahwa dari model hanya mampu mengenali mana data yang bukan kelas positif tetapi belum mampu mengidentifikasi data kelas positif dengan baik. Diduga, data yang digunakan pada proses *training* terlalu sedikit karena sedikitnya pelanggan JNE yang memberikan komentar positif.

Untuk klasifikasi data netral juga memiliki beberapa kesalahan hal ini diduga disebabkan data netral yang memiliki banyak kata yang ambigu. Dalam satu *tweets* mungkin mengandung pertanyaan dari pelanggan menanyakan paket mereka dan juga mengandung opini negatif. Sehingga terdapat kemiripan dengan kategori negatif yang komplain paket mereka belum sampai. Selain itu, data *tweet* banyak juga terlalu singkat sehingga menyebabkan kesulitan untuk pengkategorian.

5. Kesimpulan

Dari hasil penelitian ini diperoleh bahwa Naïve Bayes dapat diimplementasikan untuk mengklasifikasi data *tweets* cukup baik. Hal ini dibuktikan dengan nilai akurasi kategori negatif adalah 80% , kategori negatif 79% dan kategori positif 97%.

Data yang digunakan dalam penelitian ini sebagian besar adalah komentar netral dan negatif. Hal ini disebabkan *tweets* pelanggan sebagian besar bertujuan menanyakan paket mereka yang bermasalah. Sedangkan pelanggan yang tidak mengalami kendala tidak memberikan komentar. Untuk mengurangi komentar negatif yang ada di Twitter JNE diharapkan dapat melakukan perbaikan pada sistem update status paket pelanggan karena sebagian besar *tweets* pelanggan JNE menanyakan status paket mereka. Dari penelitian ini diharapkan dilakukan penelitian yang lebih dalam dan dengan menggunakan metode yang lain.

Daftar Pustaka

- Ahmed, I., Guan, D., Chung, T.C., 2014, SMS Classification Based on Naïve Bayes Classifier and Apriori Algorithm Frequent Itemset. *International Journal of Machine Learning and Computing* 4 (2), 183-187.
- Feldman, R., Sanger, J., 2007, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press, New York.
- Han, J., Kamber, M., Pei, J., 2006, *Data Mining Concept and Techniques*^{3rd}. Elsevier Inc, USA.
- Hidayatullah, A.F., Azhari, S.N., 2014, Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter. *Seminar Nasional Informatika*, Yogyakarta, 12 Agustus, 115-122.
- Hossin, M., Sulaiman, M.N., 2015, A Review on Evaluation Metrics for Data Classification Evaluations, *International Journal of Data Mining & Knowledge Management Process* 5 (2), 1-11.
- Indraloka, D.S., Santosa, B., 2017. Penerapan Text Mining untuk Melakukan Clustering Data Tweet Shopee Indonesia. *Jurnal Sains dan Seni Pomits* 6 (2), A-51-A56.
- Jumeilah, F.S., 2017, Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian. *Rekayasa Sistem dan Teknologi Informasi* 1(1), 19-25.
- Kini, M., S.H., Devi, Desai, P.G., Chiplunkar, N., 2015, Text mining Approach to Classify Technical Research Document using Naïve Bayes. *International Journal of Advanced Research in Computer and Communication Engineering* 4(7), 386-391.
- Kiruthika, M., Woon, S., Giri, P., 2016, Sentiment Analysis of Twitter Data. *International Journal of Innovations in Eengineering and Technology* 6 (4), 264-273.
- Li-guo, D., Peng, D., Ai-ping, L., 2014, A New Naïve Bayes Text Classification Algorithm. *TELKOMNIKA Indonesian Journal of Electrical Engineering* 12(2), 947-952.
- Pang, B., Lee, L., 2008, Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval* 2 (1), 1-135.
- Rani, V.V., Rani, K.S., 2016, Twitter Streaming and Analysis through R, *Indian Journal of Science and Technology* 9(45), 1-6.
- Saleh, A., 2015, Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Creative Information Technology Journal* 2(3), 207-217.
- Ting, S.L., Ip, W.H., Albert H.C. Tsang, 2011, Is Naïve Bayes a Good Classifier for Document Classification?. *International Journal of Software Engineering and Its Applications* 5(3), 37-46.