

PERBANDINGAN *NAÏVE BAYES* DAN *SUPPORT VECTOR MACHINE* UNTUK KLASIFIKASI ULASAN PELANGGAN INDIHOME

Aan Rohanah¹, Dwi Latifah Rianti², Betha Nurina Sari³
Teknik Informatika, Universitas Singaperbangsa Karawang^{1,2,3}
aan.rohanah17003@student.unsika.ac.id¹, dwi.latifah17001@student.unsika.ac.id²,
betha.nurina@staff.unsika.ac.id³

Submitted March 3, 2021; Revised June 8, 2021; Accepted June 12, 2021

Abstrak

IndiHome merupakan penyedia layanan internet dari PT. Telekomunikasi Indonesia, Tbk dengan jangkauan internet terluas di Indonesia. Kepuasan pelanggan menjadi salah satu hal yang harus diperhatikan dalam sebuah perusahaan, termasuk perusahaan IndiHome. Tingkat kepuasan layanan pelanggan IndiHome ini dapat dilihat dari ulasan pelanggan melalui media sosial Twitter. Penelitian ini membahas mengenai klasifikasi ulasan pelanggan IndiHome dengan mengaplikasikan tahapan penelitian CRISP-DM serta penerapan algoritme *Naïve Bayes Classifier* dan *Support Vector Machine Kernel Linear*. Data ulasan pelanggan diperoleh dari Twitter yang berjumlah 1000 *tweet* dengan menggunakan *tools* RapidMiner dan *library* R. Adapun tahapan *preprocessing* yang diterapkan yaitu *cleansing*, *case folding*, *tokenizing*, *convert word*, *stopword*, dan *stemming*. Hasil visualisasi data disajikan dalam bentuk *word cloud* yang dikategorikan berdasarkan opini positif dan negatif dari kata yang sering muncul. Hasil penelitian menunjukkan penerapan algoritme *Support Vector Machine Kernel Linear* lebih baik dibandingkan algoritme *Naïve Bayes Classifier* dengan nilai *accuracy* 82,11%, *precision* 76,44%, *recall* 88,01%, dan nilai *AUC* 0,909.

Kata Kunci : IndiHome, Kepuasan Pelanggan, *Naïve Bayes Classifier*, *Support Vector Machine*.

Abstract

IndiHome is an internet service provider from PT. Telekomunikasi Indonesia, Tbk with the widest internet coverage in Indonesia. Customer satisfaction is one of the things that must be considered in a company, including the IndiHome company. IndiHome's customer service satisfaction level can be seen from customer reviews via Twitter social media. This study discusses the classification of IndiHome customer reviews by applying the CRISP-DM research stages and the application of the Naïve Bayes Classifier algorithm and the Linear Support Vector Machine Kernel. Customer review data were obtained from Twitter, totaling 1000 tweets using the Rapid Miner and R library tools. The preprocessing stages applied were cleansing, case folding, tokenizing, word conversion, stopwords, and stemming. The results of data visualization are presented in the form of a word cloud which is categorized based on positive and negative opinions of words that often appear. The results showed that the application of the Support Vector Machine Kernel Linear algorithm is better than the Naïve Bayes Classifier algorithm with an accuracy value of 82.11%, 76.44% precision, 88.01% recall, and an AUC value of 0.909.

Key Words : *IndiHome, Customer Satisfaction, Naïve Bayes Classifier, Support Vector Machine.*

1. PENDAHULUAN

Teknologi yang semakin berkembang menjadikan internet sebagai salah satu media yang umum digunakan. Berdasarkan data hasil studi yang dikumpulkan oleh *Polling* Indonesia bersama Asosiasi Penyelenggaraan Jasa Internet Indonesia (APJII) menyebutkan bahwa jumlah

pengguna internet di Indonesia tumbuh 10,17 persen di tahun 2019 atau setara dengan kurang lebih 171 juta orang [1]. Penggunaan internet umumnya digunakan untuk mengakses media sosial. Adapun penggunaan media sosial yang sering diakses yaitu Facebook, Instagram, LinkedIn, dan Twitter [2].

Twitter merupakan salah satu jejaring sosial serta alat komunikasi dengan pengguna paling banyak. Penggunaan Twitter sebagai jejaring sosial memiliki karakteristik dan format yang unik. Pada penggunaannya Twitter dapat mengirim pesan serta membaca pesan dengan 140 karakter sebagai batas maksimal dari penulisan [3]. Selain keunikan Twitter dan mempunyai jumlah pengguna yang banyak, memungkinkan Twitter memiliki data yang melimpah dan berpotensi besar dalam mengekspresikan tentang berbagai topik. Sehingga sangat efisien dalam penambahan data salah satunya tentang kualitas pelayanan internet [4].

PT Telekomunikasi Indonesia, Tbk (PT Telkom) salah satu perusahaan penyelenggara jejaring telekomunikasi di Indonesia dengan produk unggulan IndiHome. IndiHome saat ini memiliki pelanggan dengan sebanyak 8,7% pengguna [5]. Kompetisi yang cukup tajam di sektor telekomunikasi ini membuat banyak perusahaan telekomunikasi bersaing untuk memberikan pelayanan terbaik kepada pelanggan. Oleh karena itu, tingkat kepuasan layanan pelanggan IndiHome juga perlu menjadi salah satu hal yang diperhatikan. Karena, berdasarkan hasil penelitian yang dilakukan Tineges (2020), IndiHome masih memiliki tingkat kepuasan pelanggan yang cukup rendah. Kepuasan pelanggan yang masih rendah ini perlu diperhatikan untuk segera ditingkatkan oleh perusahaan dari segi kualitas pelayanan. Salah satu cara yang dapat dilakukan adalah dengan memanfaatkan data dari pelanggan yang dapat diekstrak menjadi informasi dan pengetahuan. Pada proses pengolahan informasi menjadi pengetahuan, perlu adanya pengolahan dari data yang tersebar oleh pelanggan. Data yang dikumpulkan biasanya diungkapkan dalam kritikan atau keluhan pelanggan yang dapat dilihat dari sentimen opini pengguna pada media sosial

Twitter. Melalui Twitter pengguna dapat berdiskusi dalam memberikan kritik, saran, ataupun kepuasan mereka terhadap layanan IndiHome yang telah diberikan. Sehingga pandangan masyarakat dapat dijadikan sumber informasi untuk analisis sentimen [6].

Analisis sentimen merupakan suatu metode dalam memahami opini berdasarkan kategori polaritasnya, apakah itu positif atau negatif [6]. Saat ini analisis sentimen telah banyak digunakan untuk penelitian dengan tujuan memberikan informasi mengenai penilaian atau penyaringan opini. Seperti pada penelitian yang telah dilakukan oleh Shania Rizkia (2019) mengenai analisis kepuasan pelanggan terhadap Internet Provider IndiHome di Twitter dengan menggunakan metode *Decision Tree* yang diklasifikasikan berdasarkan opini positif, negatif, maupun netral. Pada penerapan metode *Decision Tree* masih belum maksimal dikarenakan jumlah data yang diambil masih kurang besar dan penerapan fitur yang masih sedikit sehingga menghasilkan akurasi 80,1% [5]. Penelitian sebelumnya yang telah dilakukan oleh Mardiansyah (2018) mengenai analisis algoritme *Naïve Bayes Classifier* pada aplikasi berbasis *web* dengan objek penelitian perusahaan PT. Telkom Indonesia menghasilkan penelitian yang lebih baik dikarenakan algoritme tersebut tidak terpengaruh oleh atribut yang lain sehingga untuk proses klasifikasi sentimen menghasilkan akurasi mencapai 83.% dengan pengujian data 100-200 *tweet* [7]. Pada penelitian Tineges (2020) dengan analisis sentimen terhadap layanan IndiHome menggunakan metode klasifikasi *Support Vector Machine* (SVM) mampu memberikan hasil akurasi sebesar 87% dengan evaluasi yang digunakan yaitu *Confusion Matrix* [6]. Beberapa penelitian di atas membuktikan bahwa algoritme *Naïve Bayes Classifier* dan *Support Vector Machine* mampu bekerja lebih baik

dibandingkan algoritme klasifikasi yang lainnya seperti *Decision Tree*.

Berdasarkan permasalahan tersebut untuk mengetahui penerapan algoritme terbaik maka akan dilakukan perbandingan algoritme dengan menerapkan algoritme *Naive Bayes Classifier* dan *Support Vector Machine*, terhadap pengaruh sentimen kepuasan pelanggan IndiHome di Twitter dengan klasifikasi berdasarkan sentimen positif dan negatif. Adapun manfaat dari hasil analisis sentimen pelanggan IndiHome ini dapat digunakan sebagai bahan evaluasi atau *feedback* untuk IndiHome dalam melakukan peningkatan pelayanan agar dapat memberikan kepuasan terhadap pelanggan yang lebih baik lagi.

2. METODE PENELITIAN

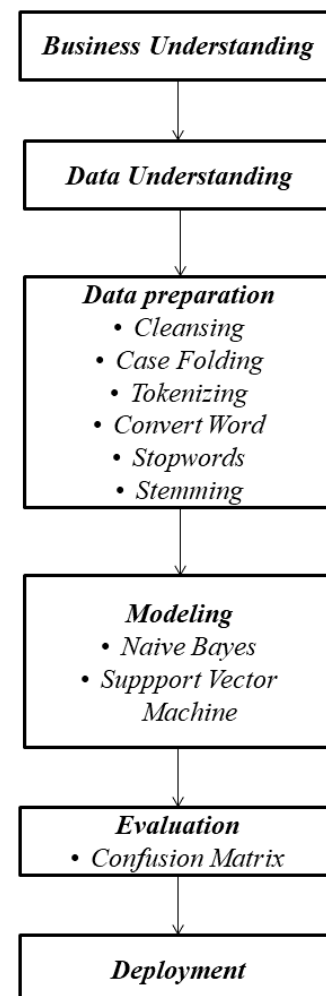
Teknik Pengumpulan Data

Data diperoleh dari proses *crawling* yaitu proses pengambilan kalimat opini dari media sosial Twitter melalui *tools* RapidMiner. Data diambil dengan menggunakan operator *Search* Twitter. Kalimat opini dicari menggunakan kata kunci @IndiHome dan diambil sebanyak 1000 *tweet*. Pada tahapan pengumpulan data, hasil data *crawling* merupakan data yang akan diolah untuk tahapan selanjutnya sesuai dengan kebutuhan penelitian yang akan digunakan.

Teknik Analisis dan Pemrosesan Data

Proses analisis data akan dilakukan dengan menggunakan *tools* RapidMiner. Penelitian ini juga akan menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM) untuk analisis dan pemrosesan data [8]. CRISP-DM merupakan metode yang digagas oleh beberapa perusahaan besar Daimler-Benz, penyedia asuransi OHRA, perusahaan pembuat perangkat lunak statistika SPSS.INC dan produsen perangkat keras perangkat lunak NCR Corp yang tujuannya

untuk memformalkan dan membakukan pendekatan terhadap *data mining*. Tahapan CRISP-DM ini dimulai dari *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*. CRISP-DM digunakan untuk penggabungan kerjasama yang bersifat konseptual yang dapat diterapkan untuk jenis data tertentu [9]. Adapun proses yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Proses CRISP-DM

1. Business Understanding

Tahap ini merupakan tahap pemahaman bisnis dengan memahami masalah apa yang akan dipecahkan. Hal ini perlu dilakukan dengan kajian dan pencarian solusi secara berulang kali.

2. *Data Understanding*

Merupakan tahap memahami struktur data yang tersedia agar dapat membantu memecahkan masalah penelitian. Pada tahap ini juga akan dilakukan pelabelan pada data yang diperoleh yaitu positif atau negatif.

3. *Data Preparation*

Data yang telah diberi label, kemudian akan dilakukan *preprocessing* sebagai berikut[10]:

- a. *Cleansing*, yaitu pembersihan atribut yang tidak berpengaruh terhadap hasil klasifikasi, seperti tanda baca.
- b. *Case folding*, yaitu mengubah semua huruf menjadi huruf kecil.
- c. *Tokenizing*, yaitu pemotongan kalimat menjadi kata tunggal.
- d. *Convert word*, yaitu memperbaiki kesalahan pada penulisan kata.
- e. *Stopwords*, yaitu penghilangan kata yang tidak berpengaruh terhadap hasil klasifikasi.
- f. *Stemming*, yaitu tahap mengubah kata imbuhan menjadi kata dasarnya.

4. *Modeling*

Modeling atau pemodelan merupakan tahap utama yaitu penerapan teknik *data mining* yang digunakan. Pada penelitian ini menggunakan teknik klasifikasi dengan algoritme *Naive Bayes Classifier* dan *Support Vector Machine*.

Pembagian data pada penelitian ini menggunakan *cross validation* dengan k $fold=10$. Nilai k $fold=10$ pada *cross validation* direkomendasikan sebagai pemilihan model terbaik karena cenderung memberikan estimasi akurasi yang lebih baik dibandingkan dengan nilai k yang lain. Data akan dibagi menjadi 10 *fold* berukuran sama, sehingga menghasilkan 10 *subset* data yang dapat digunakan untuk mengevaluasi kinerja algoritme [11].

5. *Evaluation*

Tahap ini bertujuan untuk menilai hasil *data mining* yang telah dilakukan. Adapun dalam penelitian ini akan membandingkan

hasil dari algoritme *Naive Bayes Classifier* dan *Support Vector Machine* dengan *confusion matrix*.

6. *Deployment*

Merupakan tahapan akhir untuk mengetahui hasil dari penelitian yang telah dilakukan serta jika memungkinkan hasil dari penelitian yang telah dilakukan dapat dikembangkan untuk penelitian selanjutnya. Adapun hasil *deployment* yang dilakukan pada penelitian ini tidak diimplementasikan dalam bentuk program.

3. HASIL DAN PEMBAHASAN

1. *Business Understanding*

IndiHome merupakan salah satu penyedia layanan internet yang saat ini memiliki jangkauan terluas dibandingkan pesaingnya yang lain. Oleh karena itu, IndiHome mampu menarik banyak minat masyarakat di Indonesia untuk berlangganan. Salah satu hal yang perlu diperhatikan oleh sebuah perusahaan adalah kepuasan pelanggan terhadap kualitas barang atau jasa yang diberikan. Maka berdasarkan hal tersebut, perlu adanya analisis kepuasan pelanggan IndiHome yang dapat meningkatkan kualitas pelayanannya.

2. *Data Understanding*

Jumlah data ulasan pelanggan IndiHome yang berhasil dikumpulkan adalah 1000 *tweet*. Data tersebut merupakan sampel dari ulasan pelanggan pada akhir tahun 2020 yang diambil dari tanggal 24 sampai 31 Desember 2020.

Selanjutnya dilakukan pelabelan secara manual pada data dengan mengkategorikan ulasan menjadi opini positif atau negatif. Opini positif berjumlah 568 dan opini negatif 432.

3. *Data Preparation*

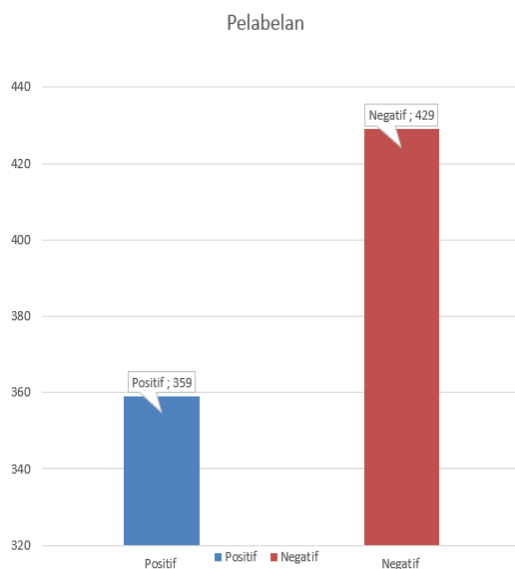
Pada tahap ini dilakukan seleksi data dengan menghilangkan data duplikat yang memiliki ulasan yang sama, seperti pada

Gambar 2.

A	
431	@IndiHome Jangan lupa nonton film ini di indihome tv#thnbaruathome
432	@IndiHome Jangan lupa nonton film ini di indihome tv#thnbaruathome
433	@IndiHome Jangan lupa nonton film ini di indihome tv#thnbaruathome
434	@IndiHome Jangan lupa nonton film ini di indihome tv#thnbaruathome
435	@IndiHome Jangan lupa nonton film ini di indihome tv#thnbaruathome
436	@IndiHome streaming valorant di twitch ping nya gede terus. sebelumnya jarang gangguan

Gambar 2. Data Duplikat

Adapun hasil dari proses penghilangan *tweet* duplikat dan *tweet* yang tidak relevan, berhasil mengurangi data menjadi 788 *tweet* dengan opini positif berjumlah 359 dan opini negatif berjumlah 429 seperti yang ditunjukkan pada Gambar 3.



Gambar 3. Pelabelan Data

Selain itu dilakukan juga penerapan seleksi fitur untuk pembobotan dengan TF-IDF. TF-IDF diterapkan karena data yang diperoleh merupakan data nominal dan jika ingin dilakukan *preprocessing* untuk klasifikasi harus berupa data numerik. Setelah data diubah menjadi numerik, kemudian data diolah melalui tahap *preprocessing text mining* dengan tahapan sebagai berikut:

a. *Cleansing*

Tahap *cleansing* dilakukan pembersihan *tweet* dari karakter yang tidak berpengaruh seperti *mention* yang diawali dengan simbol @, *hashtag* yang diawali dengan

simbol #, URL yang diawali dengan *http*, dan karakter simbol “-\$%&'()*+,-./:;<=>?@[\|_`{}~”. Selain itu tahapan *cleansing* ini digunakan untuk mengatasi data yang duplikat, dan atribut yang tidak relevan terhadap hasil. Penghilangan data duplikat dilakukan agar tahap klasifikasi lebih cepat dan akurasi yang dihasilkan lebih baik.

Tabel 1. Tahap *Cleansing*

Sebelum	Sesudah
@IndiHome @IndiHome Min kenapa saya udah ajukan henti langganan tapi belum dijawab lama bgt. Saya mau pindah sebentar lg ini mau ke jepang.	Min kenapa saya udah ajukan henti langganan tapi belum dijawab lama bgt. Saya mau pindah sebentar lg ini mau ke jepang

b. *Case folding*

Tahap *case folding* yaitu tahapan menyamakan kata. Tahapan ini dilakukan dengan mengubah huruf besar menjadi *lowercase* (huruf kecil).

Tabel 2. Tahap *Case Folding*

Sebelum	Sesudah
Min kenapa saya udah ajukan henti langganan tapi belum dijawab lama bgt. Saya mau pindah sebentar lg ini mau ke jepang	min kenapa saya udah ajukan henti langganan tapi belum dijawab lama bgt. saya mau pindah sebentar lg ini mau ke jepang

c. *Tokenizing*

Tahap *tokenizing* merupakan tahapan pemotongan *string* untuk memisahkan setiap kata menjadi kata tunggal.

Tabel 3. Tahap *Tokenizing*

Sebelum	Sesudah
min kenapa saya udah ajukan henti langganan tapi belum dijawab lama bgt. Saya mau pindah sebentar lg ini mau ke jepang	“min”, “kenapa”, “saya”, “udah”, “ajukan”, “henti”, “langganan”, “tapi”, “belum”, “dijawab”, “lama”, “bgt”, “saya”, “mau”, “pindah”, “sementar”, “lg”, “ini”, “mau”, “ke”, “jepang”

d. *Convert Word*

Tahapan ini akan memperbaiki kesalahan dalam penulisan kata dengan menggantinya menjadi kata yang sesuai agar dapat lebih mudah untuk dipahami.

Tabel 4. Tahap Convert Word

Sebelum	Sesudah
“min”, “kenapa”, “saya”, “udah”, “ajukan”, “henti”, “langganan”, “tapi”, “belum”, “dijawab”, “lama”, “bgt”, “saya”, “mau”, “pindah”, “sebenstar”, “lg”, “ini”, “mau”, “ke”, “jepang”	“admin”, “kenapa”, “saya”, “sudah”, “ajukan”, “henti”, “langganan”, “tapi”, “belum”, “dijawab”, “lama”, “banget”, “saya”, “mau”, “pindah”, “sebenstar”, “lagi”, “ini”, “mau”, “ke”, “jepang”

e. Stopwords

Tahap *stopword* merupakan teknik untuk menghilangkan kata yang tidak memiliki makna atau tidak berpengaruh terhadap hasil klasifikasi.

Tabel 5. Tahap Stopword

Sebelum	Sesudah
“admin”, “kenapa”, “saya”, “ajukan”, “henti”, “langganan”, “tapi”, “belum”, “dijawab”, “lama”, “banget”, “saya”, “mau”, “pindah”, “sebenstar”, “lagi”, “ini”, “mau”, “ke”, “jepang”	“admin”, “ajukan”, “henti”, “langganan”, “lama”, “pindah”, “banget”, “sebenstar”

f. Stemming

Tahap *stemming* berfungsi untuk mengubah kata imbuhan seperti “ber-, di-, ke-, me-, -nya, -kan” menjadi kata dasar. Tahapan ini juga mengubah kata turunan menjadi kata baku.

Tabel 6. Tahap Stemming

Sebelum	Sesudah
“admin”, “ajukan”, “henti”, “dijawab”, “banget”, “sebenstar”	“admin”, “aju”, “henti”, “langgan”, “jawab”, “lama”, “banget”, “pindah”, “bentar”

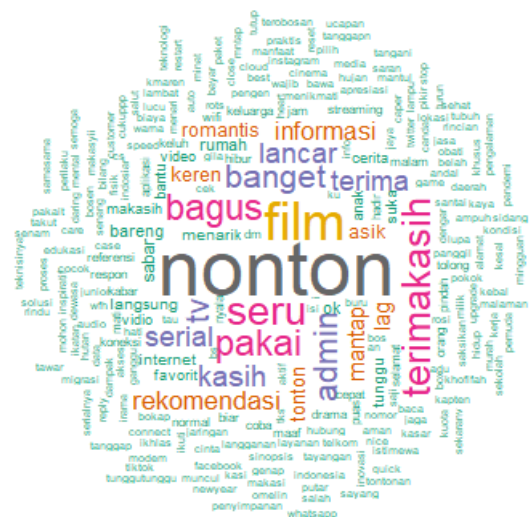
4. Modelling

Tahap *modeling* dilakukan dengan menerapkan dua algoritme klasifikasi yaitu *Naïve Bayes Classifier* dan *Support Vector Machine* kernel linear. Pembagian data dilakukan dengan menggunakan *cross validation* dengan nilai *k fold*=10. Tahap *modeling* juga dilakukan dengan membuat *word cloud* untuk ulasan

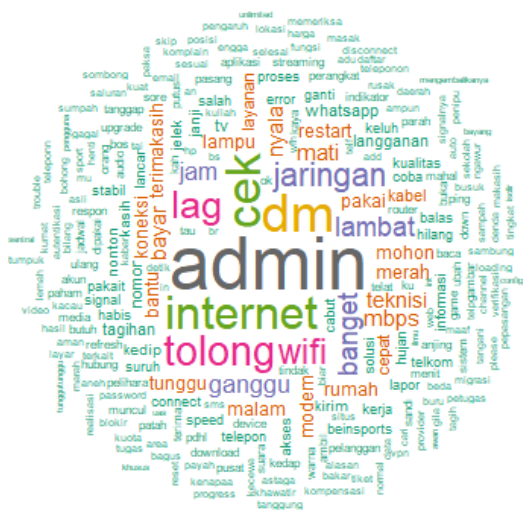
pengguna IndiHome yang bernilai positif atau negatif dengan memanfaatkan operator *execute R* pada *tools RapidMiner*.

5. Evaluation

Hasil ulasan pelanggan IndiHome ini divisualisasikan menggunakan *word cloud* yang di dalamnya dapat dilihat berdasarkan kata yang paling sering muncul dari ulasan pelanggan. *Word cloud* pada penelitian ini dibuat berdasarkan kata yang terdapat pada opini positif atau negatif. Adapun hasil *word cloud* ulasan positif pelanggan IndiHome disajikan pada Gambar 4, sedangkan untuk ulasan negatif dapat dilihat pada Gambar 5.



Gambar 4. Word cloud Ulasan Positif



Gambar 5. Word cloud Ulasan Negatif

Berdasarkan hasil *word cloud* di atas, dapat dilihat bahwa untuk ulasan pelanggan positif kata “nonton” merupakan kata yang sering muncul. Hal ini dikarenakan banyak pelanggan IndiHome yang menggunakan layanan internet untuk menonton film pada akhir tahun. Sedangkan untuk *word cloud* ulasan pelanggan negatif, kata “admin” menjadi sering muncul karena banyak pelanggan yang menyampaikan keluhannya kepada admin IndiHome di media sosial Twitter.

Selain visualisasi *wordcloud* dari hasil ulasan pelanggan IndiHome, dilakukan juga perbandingan kinerja algoritme. Evaluasi ini dilakukan menggunakan perbandingan nilai AUC dan *confusion matrix* yang terdiri dari *accuracy*, *precision*, dan *recall*.

Berikut ini merupakan perbandingan hasil dari penerapan algoritme *Naïve Bayes Classifier* dan *Support Vector Machine* dengan menerapkan *cross validation* dengan nilai *k fold* = 10.

Tabel 7. Evaluation Naïve Bayes Classifier

Accuracy	Precision	Recall	AUC
77,03%	69,35%	89,40%	0,619

Tabel 8. Evaluation SVM (Kernel Linear)

Accuracy	Precision	Recall	AUC
82,11%	76,44%	88,01%	0,909

Berdasarkan Tabel 7 dan Tabel 8 dapat dilihat bahwa nilai akurasi 82,11% merupakan akurasi tertinggi yang didapatkan oleh algoritme *Support Vector Machine*. Hal lain yang dapat dilihat yaitu tingkat keakuratan perkiraan pencarian atau *precision* algoritme *Support Vector Machine* juga lebih tinggi yaitu 76,44%. Sedangkan untuk tingkat keakuratan pencarian atau *recall*, algoritme *Naïve Bayes* menghasilkan *recall* yang lebih baik dengan nilai 89,40%. Akan tetapi, walaupun *Naïve Bayes* memiliki nilai *recall* yang lebih baik, *Naïve Bayes* memiliki nilai AUC yang rendah yaitu 0,619.

6. Deployment

Hasil dari penelitian ini dapat digunakan sebagai bahan evaluasi atau *feedback* untuk IndiHome dalam meningkatkan pelayanan agar dapat memberikan kepuasan terhadap pelanggan yang lebih baik lagi.

Model klasifikasi ulasan pelanggan yang telah dihasilkan juga dapat dikembangkan menjadi aplikasi *Machine Learning* agar dapat mengklasifikasikan ulasan secara otomatis. Aplikasi lain yang dapat dibuat dari hasil penelitian ini juga yaitu aplikasi untuk tahapan *preprocessing* dalam *text mining* agar dapat memudahkan proses pembuatan sebuah model klasifikasi.

4. SIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka dapat ditarik kesimpulan bahwa algoritme *Naïve Bayes* dan *Support Vector Machine* dapat digunakan untuk klasifikasi ulasan pelanggan Indihome dari media sosial Twitter. Adapun beberapa *preprocessing* yang digunakan pada penelitian ini yaitu *cleansing*, *case folding*, *tokenizing*, *convert word*, *stopword*, dan *stemming*.

Hasil penelitian didapatkan bahwa IndiHome mendapatkan hasil opini positif sebesar 45,56% sedangkan opini negatif

sebesar 54,44%. Kata yang sering muncul pada ulasan positif adalah “nonton”, sedangkan pada ulasan negatif yaitu kata “admin”. Selain itu, didapatkan juga algoritme terbaik untuk mengklasifikasikan ulasan pengguna IndiHome pada media sosial Twitter adalah *Support Vector Machine* lebih baik dengan nilai *accuracy* 82,11%, *Precision* 76,44%, *recall* 88,01%, dan nilai AUC 0,909.

DAFTAR PUSTAKA

- [1] B. Harto and R. Komalasari, “Optimalisasi Platform Online Internet Marketing untuk Sme Little Rose Bandung,” *Empower. Community*, vol. 1, no. 1, p. 1, 2020, doi: 10.31543/ecj.v1i1.357.
- [2] A. W. Irawan, A. Yusufianto, D. Agustina, and R. Dean, *Laporan Survei Internet Apjii 2019-2020 (Q2)*, vol. 2020. 2020.
- [3] R. S. Putra, “Analisis Sentimen Twitter dengan Klasifikasi Naïve Bayes menggunakan Seleksi Fitur Mutual Information dan Inverse Document Frequency,” Skripsi, Departemen Ilmu Komputer, Institut Pertanian Bogor, Bogor, Indonesia, 2017.
- [4] N. H. Saputra, “Analisis Sentimen Pemilihan Gubernur Jawa Barat 2018 pada Data Twitter menggunakan Naïve Bayes,” Skripsi, Departemen Ilmu Komputer, Institut Pertanian Bogor, Bogor, Indonesia, 2020.
- [5] M. P. Shakina Rizkia, Erwin Budi Setiawan S.Si., M.T, Diyas Pusandari S.S., “Analisis Sentimen Kepuasan Pelanggan Terhadap Internet Provider Indihome di Twitter Menggunakan Metode Decision Tree dan Pembobotan TF-IDF,” *e-Proceeding Eng.*, vol. 6, no. 2, pp. 9683–9693, 2019.
- [6] R. Tineges, A. Triayudi, and I. D. Sholihati, “Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM),” *J. Media Inform. Budidarma*, vol. 4, no. 3, pp. 650–658, 2020, doi: 10.30865/mib.v4i3.2181.
- [7] A. O. Murdiansyah and I. Siswanto, “Algoritma Naive Baiyes Classsifier pada Aplikasi Data Mining Berbasis Web,” vol. 1, no. 1, pp. 284–290, 2018.
- [8] C. Claudia and A. Budi, “Analisa dan Klasifikasi Sentiment Opini Penonton pada Website IMDB.com dengan Algoritma Support Vector Machine,” *J. Inform. dan Bisnis*, vol. 6, no. 2, pp. 18–28, 2017.
- [9] M. North, *Data Mining for the Messes*. Infinitte, 2016.
- [10] I. P. Windasari, F. N. Uzzi, and K. I. Satoto, “Sentiment analysis on Twitter posts: An analysis of positive or negative opinion on GoJek,” *Proc. - 2017 4th Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE 2017*, pp. 266–269, 2017, doi: 10.1109/ICITACEE.2017.8257715.
- [11] P. R. Sihombing and O. P. Hendarsin, “Perbandingan Metode Artificial Neural Network (ANN) dan Support Vector Machine (SVM) untuk Klasifikasi Kinerja Perusahaan Daerah Air Minum (PDAM) di Indonesia,” *J. Ilmu Komputer.*, vol. 8, no. 1, pp. 9–20, 2020, doi: 10.24843/jik.2020.v13.i01.p02.