

IMPLEMENTASI ALGORITMA NAÏVE BAYES DAN K-NEAREST NEIGHBOR UNTUK ANALISIS SENTIMEN TERHADAP LAYANAN INDIHOME BERDASARKAN OPINI TWITTER

1st Aida SolehahUniversitas Buana Perjuangan
Karawang, Indonesiaif19.aidasolehah@mhs.ubpkarawang.ac.id2nd Amril Mutoi SiregarUniversitas Buana Perjuangan
Karawang, Indonesiaamrilmutoi@ubpkarawang.ac.id3rd Adi Rizky PratamaUniversitas Buana Perjuangan
Karawang, Indonesiaadirizky@ubpkarawang.ac.id

Abstract— Dengan berkembang pesatnya teknologi, maka semakin besar juga penggunaan internet. Untuk menunjang kebutuhan internet agar semakin cepat dan nyaman digunakan maka diperlukan pula *provider* jaringan yang cepat untuk mengakses internet. Pada tahun 2018, Sebanyak 18,9% dari pengguna internet menyatakan bahwa media sosial adalah penggunaan utama internet mereka. Twitter adalah salah satu media sosial yang banyak digunakan untuk mencari hiburan, bisnis, ekonomi, politik, dan jenis informasi lainnya. Indihome merupakan *provider* yang banyak digunakan pengguna untuk mengakses internet. Di tahun 2022, IndiHome dinyatakan sebagai pemilik *Internet Service Provider (ISP)* yang paling banyak digunakan oleh pengguna ISP dibandingkan *provider* lainnya sehingga indihome dapat menarik perhatian masyarakat yang lebih besar. Penelitian ini akan melakukan analisis sentimen terhadap pelayanan indihome, data diambil dari opini-opini pengguna twitter. Data yang digunakan berjumlah 1008 data *tweet* dan data akan diklasifikasikan menjadi dua yaitu opini negatif dan positif. Algoritma *Naive Bayes* dan *K-Nearest Neighbor* digunakan untuk mengklasifikasikan data setelah dibagi menjadi 70% data latih dan 30% data uji. Algoritma *Naive Bayes* memiliki akurasi 81,8%, presisi 88,5%, dan *recall* 86,6%, sedangkan Algoritma *K-Nearest Neighbor* memiliki akurasi 59,0%, presisi 95,4%, dan *recall* 46,8% saat diuji menggunakan *Confusion Matrix*.

Kata kunci — *Internet Service Provider (ISP)*, Indihome, Analisis Sentimen, *Naive Bayes*, *K-Nearest Neighbor*.

I. PENDAHULUAN

Dengan berkembangnya teknologi, maka diperlukan pula *provider* jaringan yang cepat untuk mengakses internet. Sebesar 18,9% diantaranya mengungkapkan bahwa alasan utama mereka dalam menggunakan internet adalah untuk mengakses media sosial data tersebut diunggah oleh APJII tahun 2018. Twitter merupakan salah satu media sosial yang banyak digunakan untuk mencari berbagai informasi, hiburan, bisnis, ekonomi, politik dan lainnya. Banyaknya *tweet* yang disebar di twitter, *tweet-tweet* tersebut dapat berisi opini-opini pengguna terhadap suatu objek, sehingga banyak perusahaan yang menggunakan media sosial twitter untuk menyebarkan informasi maupun untuk mendapatkan *feedback* dari pengguna. Contohnya pada perusahaan jasa *Internet Service Provider* seperti Indihome.

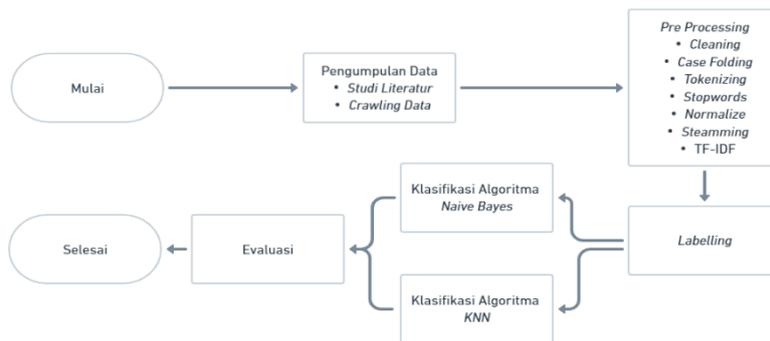
Indihome merupakan anak perusahaan dari PT.Telkom Akses yang beroperasi pada bidang pelayanan konstruksi dan pengelolaan *infrastruktur* jaringan pada produknya [1]. Diantara pengguna *fixed broadband*, 64,54% merupakan pengguna indihome yang dimana jumlah tersebut lebih banyak dari pada *provider* lain menurut APJII, 2022. Data tersebut menunjukkan indihome memiliki perhatian yang lebih dari masyarakat. Untuk mengetahui seberapa baik layanan *provider* indihome tersebut maka diperlukan opini-opini para pengguna indihome, opini-opini tersebut akan diambil dari media sosial twitter.

Analisis Sentimen merupakan proses penggalan secara otomatis diambil dari data tekstual untuk mendapatkan sebuah informasi pada suatu sentimen yang berada pada kalimat sehingga menghasilkan opini positif atau negatif [2]. Adapun beberapa penelitian terkait dalam bidang analisis sentimen terhadap suatu topik dengan memanfaatkan data twitter yang telah dilakukan. Salah satu penelitian di bidang ini adalah Analisis Sentimen Menggunakan Metode Klasifikasi *Support Vector Machine (SVM)* Terhadap Layanan Indihome Berdasarkan Twitter [3], Nilai akurasi pada penelitian ini sebesar 87%, dengan 86% hasil prediksi dengan data sebenarnya (*precision*), tingkat keberhasilan sistem sebesar 95% dalam memprediksi data (*recall*), tingkat error sebesar 13% untuk semua data prediksi, dan nilai rata-rata presisi dan daya ingat (*f1score*) sebesar 90%. Penelitian lainnya tentang Analisis Sentimen *Review Customer* Menggunakan Algoritma *Convolutional Neural Network* Terhadap Produk Indihome Dan *First Media* [4], penelitian ini menghasilkan akurasi tertinggi sebesar 98% untuk *provider* IndiHome dan 91% untuk *provider First Media*. Lalu penelitian terhadap Analisis Sentimen Pada Pelayanan Jaringan Internet Indihome Dengan Metode *Multinomial Naive Bayes* Masa Pandemi Covid-19 [5], hasil pada penelitian tersebut mendapatkan nilai akurasi sebesar 85%. Penelitian selanjutnya tentang Penelitian lainnya tentang Perbandingan *Naive Bayes* Dan *Support Vector Machine* Untuk Klasifikasi Ulasan Pelanggan [6], penelitian tersebut menghasilkan akurasi sebesar 82,11%. *precision* 76,44%, *recall* 88,01%.

Penelitian ini akan menerapkan metode klasifikasi dengan menggunakan Algoritma *Naive Bayes* dan *K-Nearest Neighbor (KNN)*. Metode *Naive Bayes* merupakan metode pengklasifikasian dalam pengembangan *text mining*, metode ini bekerja dengan menghitung nilai probabilitas dari sebuah data [7]. Algoritma *K-Nearest Neighbor* merupakan metode pendekatan untuk mencari solusi baik kasus lama maupun kasus baru berdasarkan kecocokan bobot antar fitur yang ada [2]. Sesuai pada saran-saran yang berada pada penelitian-penelitian sebelumnya yaitu agar ditambahkan algoritma lainnya sebagai perbandingan untuk mendapatkan hasil akurasi tertinggi. Penelitian ini akan melakukan analisis terhadap hasil *tweet* dengan kata kunci “indihome”. Data yang diperoleh dari hasil *crawling* pada twitter akan diklasifikasikan menjadi dua kategori sentimen yaitu positif dan negatif. Metode

Naïve Bayes dan *KNN* yang digunakan pada penelitian diharapkan dapat menghasilkan metode yang paling baik dalam melakukan klasifikasi pada data tersebut.

II. METODE PENELITIAN



Gambar 1 *Flowchart* Penelitian

A. Pengumpulan Data

Objek penelitian pada penelitian ini ialah opini masyarakat pada sosial media Twitter mengenai pelayanan Indihome. Data diambil dengan menggunakan proses *crawling* pada *python* yaitu sebanyak 1.008 data *tweet* data dari tanggal 01 Maret 2020-11 Desember 2022. Data yang di *crawling* terbagi menjadi 2 kelas yaitu kelas positif dan negatif.

B. Pre-Processing

Pre-processing merupakan tahap sebelum klasifikasi yaitu untuk mengganti kata yang tidak beraturan menjadi lebih teratur sehingga dapat mempermudah pemrosesan dataset kata tersebut [6]. *Pre-processing* ini dilakukan menggunakan google colab dengan library *nlk* dan *sastrawi*. Pada tahap ini akan dilakukan proses *Cleaning*, *Case Folding*, *Tokenizing*, *Filtering*, *Stemming* dan *TF-IDF*.

C. Labelling Data

Setelah *Pre-processing* dan Pembobotan *TF-IDF* tahap selanjutnya *Labelling* atau Pelabelan dengan mengklasifikasikan setiap data menggunakan label yang berbeda, seperti kelas positif dan negatif. Label positif ditandai dengan angka 1 sedangkan negatif ditandai dengan -1 [3].

Tabel 1 Contoh *Labelling* Data

Datetime	Text	Label
2022-10-29	nonton netflix paling seru di barengi jaringan internet indihome yang super super luar biasa	1
2022-10-29	sekarang kalau mau tanya terkait layanan indihome gampang banget. Bisa telepon 147 sampai dm di sosmed	1
2022-10-29	indihome keren promonya good akses internet cepat dan stabil wow	1
2022-10-29	indihome trouble	-1
2022-10-29	kualitas dan harga kalian yang gak masuk akal	-1

D. Implementasi Algoritma *Naïve Bayes* dan *K-Nearest Neighbor*

Algoritma *Naive Bayes* dan algoritma *K-Nearest Neighbor* dapat digunakan untuk memproses dan mengelompokkan sebuah data setelah melalui *pre-processing* dan pelabelan data. Klasifikasi merupakan metode untuk mengelompokkan dan menemukan *itemset* sesuai pada nilai bobot yang terdapat pada setiap *itemset*. Algoritma *Naive Bayes* didasarkan pada *teoroma bayes*, algoritma ini sangat baik dalam mengolah data besar [8]. Algoritma *K-Nearest Neighbor* algoritma untuk mengklasifikasikan sesuatu berdasarkan seberapa dekat mereka satu sama lain [9]. Data dalam penelitian ini akan klasifikasi menjadi dua bagian, yaitu positif (1) dan negatif (-1). Berikut merupakan metode perhitungan algoritma *Naive Bayes*.

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

X: Data pada suatu kelas yang tidak diketahui

H : Kelas yang spesifik merupakan data hipotesis X

P(H)|(X) : Peluang pada hipotesis H berdasar kondisi

P(H) : Peluang pada hipotesis H

P(X|H) : Peluang pada X berdasarkan kondisi pada hipotesis

P(X) : Peluang pada X

Berikut merupakan metode perhitungan algoritma *K-Nearest Neighbor* dengan *Euclidean distance*.

$$d = \sum_{i=1}^n (a_i - b_i)^2$$

- d : Jarak
- a : Data uji
- b : Spesimen pada data
- i : Variabel pada data
- n : Dimensi pada data

E. Evaluasi Algoritma *Naïve Bayes* dan *K-Nearest Neighbor*

Tahap selanjutnya adalah tahap evaluasi menggunakan *Confusion Matrix* untuk menghitung akurasi setelah melakukan proses klasifikasi dengan *Naive Bayes* dan *K-Nearest Neighbor*. Langkah-langkah pengujian *Confusion Matrix* untuk menentukan *accuracy*, *precision*, dan *recall* adalah sebagai berikut.

1. Accuracy

Accuracy, yakni untuk mengetahui jumlah data yang di klasifikasikan secara benar. Berikut rumus perhitungan *Accuracy*

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100$$

2. Precision

Precision, digunakan untuk mengetahui tingkat ketepatan prediksi dari suatu sistem, dengan menghitung berdasarkan jumlah prediksi positif dari total data yang diprediksi sistem. Berikut rumus perhitungan *Precision*

$$Precision = \frac{TP}{TP+FP}$$

3. Recall

Recall, digunakan untuk mengenali tingkat keberhasilan dari suatu kelas. Berikut rumus perhitungan *Recall*

$$Recall = \frac{TP}{TP+FN}$$

Keterangan :

- A. TP adalah *True Positif*, yaitu jumlah data positif yang diklasifikasikan dengan benar oleh sistem.
- B. FP adalah *False Positif*, yaitu jumlah data positif yang diklasifikasikan dengan salah oleh sistem
- C. FN adalah *False Negatif*, yaitu jumlah data negatif namun diklasifikasikan salah oleh sistem.
- D. TN adalah *True Negatif*, yaitu jumlah data negatif yang diklasifikasikan dengan benar oleh sistem

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pengumpulan data (*Crawling*) pada penelitian ini menggunakan library *textblob* dan *snsrape* menggunakan Bahasa python pada *Google Colab*. Data yang dikumpulkan yaitu berupa *tweet* atau opini publik berbahasa Indonesia, dalam proses pengambilan data

	Datetime	Tweet Id	Text	Username	Label
0	2020-03-29 23:56:28+00:00	1244413146749472769	@putranto1125 @TelkomCare @IndiHome Gangerti l...	bchtirfi	1
1	2020-03-29 23:44:26+00:00	1244410115127570432	@IndiHome perimis qaqa, kenapa akhir2 ini jari...	orinkrisna	1
2	2020-03-29 22:53:24+00:00	1244397273506656256	@IndiHome Mohon bantuannya, internet saya LOS....	mikhaahkim	1
3	2020-03-29 21:10:59+00:00	1244371499550556162	Tolong dong Indihome, wifi dirumah ku aktifin ...	brightfat	1
4	2020-03-29 20:02:34+00:00	1244354281982980105	@jordi154 Ok Kak Jordi, sudah di respon DM. -...	IndiHome	1
...
1004	2022-11-29 12:51:06+00:00	1597573869690200064	@moanamenfess Poor signal cenah....indihome ru...	rejstancetxt	-1
1005	2022-11-29 12:26:06+00:00	1597567576980455425	hari ini namatin reparasi indihome, bookings s...	notsannyforreal	-1
1006	2022-12-11 10:48:34+00:00	1601891687592128515	Yg hobby maen game tapi saat maen internet nya ...	IrfanJa44129301	-1
1007	2022-12-11 10:45:12+00:00	1601890840145805316	Yg hobby maen game tapi saat maen internet nya...	UZulviana	-1
1008	2022-12-11 06:42:09+00:00	1601829674287448064	Mumpung di kantor sudah berlangganan wifi Indl...	ArdouAjawalla	-1

twitter dengan *Google colab* hanya memasukan *query* yang mengandung kata kunci “Indihome”. Setelah itu data dipecah menjadi dua kategori dengan mengembalikan polaritas dan subjektivitas kalimat yang dibuat oleh *textblob*, *library* tersebut dapat memberikan skor individu pada semua kata, kemudian hasil sentimen tersebut dihitung dan menghasilkan polaritas negatif dan

Gambar 2 Hasil *Crawling* Data

positif. Data yang terkumpul sebanyak 1.008 tweet data yang diambil sejak meningkatnya pengguna *provider* indihome dari 01 Maret 2020 hingga 11 Desember 2022.

B. Pre-processing Data

Dengan menggunakan bahasa pemrograman *Python* dan *Google Colab*, proses *Pre-processing* data dapat digunakan untuk membersihkan data teks atau mengubah kata yang tidak beraturan menjadi kata yang beraturan.

1. Cleaning dan Casefolding

Proses *Cleaning* melibatkan penghapusan karakter non-abjad, seperti tanda baca, simbol seperti tanda "@" di nama pengguna, tagar (#), *emoticon*, dan url situs web, serta mengurangi kesalahan.

Tabel 2 *Cleaning* dan *Casefolding*

Sebelum	Sesudah
woi indihome! sering banget sih trouble astagaa !!! buruan benerin koneksinya !!! @indihome @indihomecare #indihomelemot #lemotparah	woi indihome sering banget sih trouble astagaa buruan benerin koneksinya

2. Tokenizing

Tahap *Tokenizing* adalah proses pemecahan string (kalimat dipecah menjadi perkata) teks setelah prosedur *Case Folding* selesai (kalimat dipecah menjadi kata-kata).

Tabel 3 *Tokenizing*

Sebelum	Sesudah
woi indihome sering banget sih trouble astagaa buruan benerin koneksinya	['woi', 'indihome', 'sering', 'banget', 'sih', 'trouble', 'astagaa', 'buruan', 'benerin', 'koneksinya']

3. Stopwords

Penghapusan kata-kata yang tidak berarti dikenal sebagai "*Stopword*" atau "*filtering*", Itu berarti membuang kata-kata yang tidak berarti apa-apa dan tidak memengaruhi analisis sentimen. *List stopwords* pada Bahasa Indonesia dapat menggunakan *library NLTK*, selain itu dapat menambahkan secara manual list tersebut pada *list_stopword* agar teks menjadi lebih bersih dan terstruktur.

Tabel 4 *Stopwords*

Sebelum	Sesudah
['woi', 'indihome', 'sering', 'banget', 'sih', 'trouble', 'astagaa', 'buruan', 'benerin', 'koneksinya']	['indihome', 'sering', 'banget', 'sih', 'trouble', 'astagaa', 'buruan', 'benerin', 'koneksinya']

4. Normalize

Kata-kata yang tidak lengkap atau yang ditulis salah (*typo*) diubah melalui proses *Normalize* menjadi kata-kata yang benar dan dapat dimengerti.

Tabel 5 *Normalize*

Sebelum	Sesudah
['indihome', 'sering', 'banget', 'sih', 'trouble', 'astagaa', 'buruan', 'benerin', 'koneksinya']	['indihome', 'sering', 'sangat', 'sih', 'masalah', 'astaga', 'buruan', 'benerin', 'koneksinya']

5. Stemming

Proses *Stemming* merupakan tahap mengubah kata menjadi kata dasar seperti mengubah kata hasil *Filtering* dengan cara menghilangkan imbuhan (awalan dan akhiran) menjadi ke bentuk dasar pada kata.

Tabel 6 *Stemming*

Sebelum	Sesudah
['indihome', 'sering', 'sangat', 'sih', 'masalah', 'astaga', 'buruan', 'benerin', 'koneksinya']	['indihome', 'sering', 'sangat', 'sih', 'masalah', 'astaga', 'buru', 'benerin', 'koneksi']

6. Pembobotan Kata

Pembobotan TF-IDF (*Term Frequency-Inverse Document Frequency*) digunakan untuk mengubah data tekstual menjadi data numerik sehingga setiap kata atau fitur kalimat dapat diberi bobot [10].

Tabel 7 Hasil TF-IDF

Term	TF	TF-IDF
kak	0.037037037037037035	0.07646980579385154

mohon	0.037037037037037035	0.1369570072402003
maaf	0.037037037037037035	0.11863492420923337
keluh	0.037037037037037035	0.15843176632364261
tiket	0.037037037037037035	0.14522158321480066

C. Labelling Data

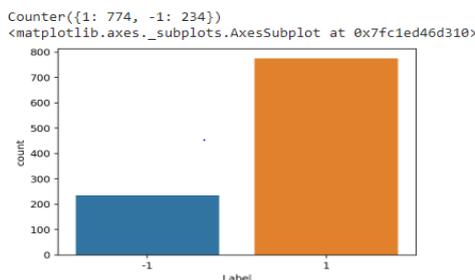
Proses *Labeling* yaitu proses pelabelan data setelah tahapan *Text Pre-processing, labeling* pada penelitian ini di bagi menjadi 2 kategori sentimen berdasarkan kalimatnya yaitu sentimen *Negatif* dan sentimen *Positif*. Proses *labelling* dilakukan dengan cara manual oleh Pionika Linda Sari (Lulusan Sarjana Sastra Indonesia) dengan format excel. Pelabelan positif ditandai dengan angka 1, sedangkan pelabelan negative ditandai dengan angka -1.

Tabel 8 Data Hasil Labelling

Text	Label
indihome sering sangat sih masalah astaga buru benerin koneksi	-1
mumpung kantor sudah langgan wifi indihome kerja juga kurang sempat nih streaming film dulu game of thrones walaupun film sudah lama sih	1
sudah mahal sering ganggu internet lambat kalau bayar gapernah telat telat hari langsung ditelfon siapa lagi kalau bukan banget kalau ganggu mau ganti modem masalah cuma ada indihome doang rumah saya nangis sangat	-1
sebal banget sama indihome ganggu terus dasar	-1
...	...
indihome per bulan enak tidak lambat cinta sangat belum pake first media enggak sangat de wifi putus nyambung	1

D. Klasifikasi Algoritma Naïve Bayes dan K-Nearest Neighbor

Pada proses klasifikasi, data akan dikelompokkan menjadi 2 bagian yaitu data positif dan data negatif. Data yang sudah melalui tahap pelabelan berjumlah 1008 data *tweet*, setelah diklasifikasikan menjadi 774 *tweet* yang masuk dalam kelas positif, dan 234 *tweet* kelas negatif. Setelah itu data memasuki tahap proses *Splitting Data* (Pembagian Data) dataset di bagi menjadi 70% data latih dan 30% data uji data.



Gambar 3 Hasil Klasifikasi

Tahap selanjutnya yaitu melakukan proses klasifikasi dengan menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbor*. Pada proses klasifikasi data yang digunakan untuk penelitian ini berjumlah 1008 data dengan 705 data *training* dan 303 data *testing*. Proses klasifikasi *Naive Bayes* dilakukan dengan menghitung probabilitas kata dalam kalimat yang memiliki banyak peluang untuk setiap kelas, dilakukan perhitungan nilai akurasi dari model klasifikasi *naive bayes* yang dibuat oleh *sklearn.metrics*, didapatkan hasil dari perhitungan *matrix* dengan *code python* yaitu sebesar 97.0% untuk data *training* dan 81.8% untuk data *testing*. Proses klasifikasi *K-Nearest Neighbor* diimplementasikan dengan menghitung sebuah jarak ketetangaan sebuah data dengan data lainnya. Dilakukan perhitungan nilai akurasi dari model klasifikasi *K-Nearest Neighbor* yang telah dibuat dengan menggunakan *sklearn.metrics*, didapatkan hasil dari perhitungan *matrix* dengan *code python* yaitu sebesar 84.8% untuk data *training* dan 59.0% untuk data *testing*.

E. Evaluasi

Proses evaluasi dengan menggunakan *Confusion Matrix* merupakan tahap implementasi selanjutnya dengan maksud untuk mengetahui kinerja algoritma. Proses bekerja *Confusion Matrix* yaitu untuk membandingkan hasil pengujian klasifikasi dari sebuah sistem dengan hasil pengujian yang sebetulnya. Pengujian pada penelitian ini dilakukan pada objek sebuah opini pengguna twitter terhadap layanan Indihome menggunakan data *testing* yang berjumlah 303 data.

Tabel 9 Hasil *Confusion Matrix*

	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
Algoritma Naïve Bayes	81.8%	88.5%	86.6%
Algoritma K-Nearest Neighbor	59.0%	95.4%	46.8%

Berdasarkan hasil perhitungan *Confusion Matrix* diatas dapat diperoleh hasil klasifikasi dari kedua algoritma yaitu algoritma *Naïve Bayes* dan *K-Nearest Neighbor* menggunakan evaluasi *Confusion Matrix* pada analisis sentimen pelayanan indihome menurut opini pengguna twitter memperoleh hasil bahwa algoritma *Naïve Bayes* memiliki hasil yang lebih akurat daripada *K-Nearest Neighbor*. Algoritma *Naïve Bayes* menghasilkan akurasi 81.8% sedangkan *K-Nearest Neighbor* menghasilkan akurasi sebesar 59.0%.

IV. KESIMPULAN DAN SARAN

Penelitian ini menghasilkan nilai sentimen yang dihasilkan dari *tweet* pelayanan indihome menurut opini pengguna twitter yaitu bersentimen positif. *Algoritma Naive Bayes* dan *K-Nearest Neighbor* kemudian akan digunakan pada penelitian ini dengan membagi data menjadi 70% data *training* dan 30% data *testing*. kemudian dilanjutkan ke tahap evaluasi yaitu pengujian model klasifikasi dengan menggunakan *Confusion Matrix*. Pada penelitian ini evaluasi pengujian model klasifikasi algoritma *Naive Bayes* dan *K-Nearest Neighbor* dengan menggunakan *Confusion Matrix*. Dari pengujian tersebut didapatkan hasil akurasi yang paling baik pada data tersebut yaitu pada kinerja algoritma *Naive Bayes* dengan nilai *Accuracy* sebesar 81.8%, *Precision* 88.5%, dan *Recall* 86.6%. Adapun saran yang dapat dilakukan pada penelitian selanjutnya yaitu dapat diuji dan dibandingkan dengan algoritma atau metode lainnya. Serta data *tweet* pada bulan selanjutnya dapat terus diupdate untuk mengetahui opini masyarakat yang lainnya terhadap pelayanan indihome. Untuk kamus pada sentiment juga perlu terus ditambahkan agar kosa kata yang digunakan semakin luas.

PENGAKUAN

Naskah Ilmiah ini adalah sebagian dari penelitian Tugas Akhir milik Aida Solehah dengan judul Analisis Sentimen Terhadap Layanan Indihome Menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor yang dibimbing oleh Bapak Dr. Amril Mutoi Siregar dan Bapak Adi Rizky Pratama.

DAFTAR PUSTAKA

- [1] Mahira, P. Hadi, and H. Nastiti, "Pengaruh Kualitas Produk Dan Kualitas Pelayanan Terhadap Kepuasan Pelanggan Indihome," *Pros. Konf. Ris. Nas. Ekon. Manajemen, dan Akunt.*, vol. 2, no. 1, pp. 1267–1283, 2021.
- [2] A. Tangu Mar, E. Sedyono, and H. Purnomo, "Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Metode Pembelajaran Dalam Jaringan (DARING) Di Universitas Kristen Wira Wacana Sumba," *Jointer - J. Informatics Eng.*, vol. 2, no. 01, pp. 24–31, 2021, doi: 10.53682/jointer.v2i01.30.
- [3] 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, p. 650, 2020, doi: 10.30865/mib.v4i3.2181.
- [4] S. H. Badjrie, O. N. Pratiwi, and H. D. Anggana, "Analisis Sentimen Review Customer Terhadap Produk Indihome Dan First Media Menggunakan Algoritma Convolutional Neural Network Review Analysis Sentiment Customer Product Indihome And First Media Using Convolutional Neural Network," *eProceedings ...*, vol. 8, no. 5, pp. 9049–9061, 2021, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15877>.
- [5] A. Sentimen, "Jurnal Smart Teknologi Analisis Sentimen Pada Pelayanan Jaringan Internet Indihome Dengan Metode Multinomial Naïve Bayes Masa Pandemi Covid-19 Sentiment Analysis on Indihome Internet Network Services Using The Multinomial Naïve Bayes Method During The Cov," vol. 3, no. 6, pp. 612–623, 2022.
- [6] T. Kurniawan, "Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Media Mainstream Menggunakan Naïve Bayes Classifier Dan Support Vector Machine Media Mainstream Menggunakan Naïve Machine," *IT J.*, vol. 23, p. 1, 2017.
- [7] A. V. Sudiantoro *et al.*, "Analisis Sentimen Twitter Menggunakan Text Mining Dengan," vol. 10, no. 2, pp. 398–401, 2018.
- [8] A. M. Siregar, "Klasifikasi Untuk Prediksi Cuaca Menggunakan Esemble Learning," *Petir*, vol. 13, no. 2, pp. 138–147, 2020, doi: 10.33322/petir.v13i2.998.
- [9] A. Rizky pratama, "Klasifikasi Daging Sapi Berdasarkan Ciri Warna Dengan Metode Otsu dan K-Nearest Neighbor," *Techno Xplore J. Ilmu Komput. dan Teknol. Inf.*, vol. 6, no. 1, pp. 9–18, 2021, doi: 10.36805/technoxplore.v6i1.1239.
- [10] J. A. Septian, T. M. Fachrudin, and A. Nugroho, "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor," *J. Intell. Syst. Comput.*, vol. 1, no. 1, pp. 43–49, 2019, doi: 10.52985/insyst.v1i1.36.