

Analisis Sentimen Pengguna Twitter pada Layanan BPJS Kesehatan menggunakan

Algoritma Naïve Bayes Classifier

SKRIPSI

Untuk Memenuhi Sebagian Persyaratan

Mencapai Derajat Sarjana S1



Disusun oleh :

ASA CAHYO NUGROHO

2071100038

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI DAN KOMPUTER

UNIVERSITAS WIDYA DHARMA KLATEN

2024

HALAMAN PERSETUJUAN

**Analisis Sentimen Pengguna Twitter pada Layanan BPJS Kesehatan Menggunakan
Algoritma *Naïve Bayes Classifier***

Diajukan Oleh

Asa Cahyo Nugroho

2071100038

Telah disetujui oleh
Dosen Pembimbing I



Aryati Wuryandari, S.T., M.kom.

NIK. 690 208 292

Tanggal.....

Dosen Pembimbing II

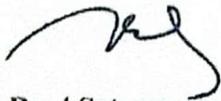


Muryanto, S.T., M.Cs.

NIK. 690 903 277

Tanggal.....

Mengetahui
Ketua Program Studi TI



Doni Setyawan, M.Cs.

NIK. 690 208 288

Tanggal.....

HALAMAN PENGESAHAN

Diterima dan disetujui oleh Dewan Penguji Skripsi Fakultas Teknologi dan Komputer
Universitas Widya Dharma Klatenm pada:

Hari : JUM'AT

Tanggal : 16 AGUSTUS 2024

Tempat : UNIVERSITAS WIDYA DHARMA

Susunan Dewan Penguji

Ketua



Aryati Wuryandari, S.T., M.Kom

NIK. 690 208 292

Sekretaris



Muryanto, S.T., M.Cs

NIK. 690 903 277

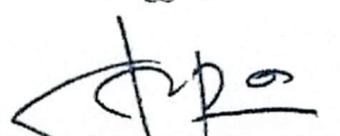
Anggota I



Syams Kurniawan Hidayat, S. T., M.Kom

NIK. 690 116 374

Anggota II



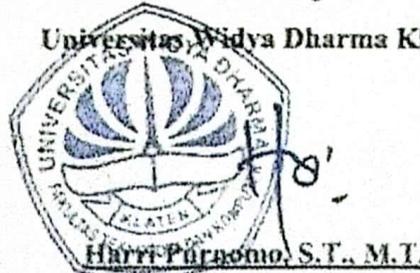
Hendro Joko Prasetyo, M.Kom

NIK. 690 903 27

Skripsi ini telah diterima dan disahkan sebagai salah satu syarat untuk memperoleh gelar
sarjana oleh:

Dekan Fakultas Teknologi dan Komputer

Universitas Widya Dharma Klaten



Hari Purnomo, S.T., M.T.

NIK. 690 196 499

SURAT PERNYATAAN KEASLIAN PENELITIAN

Yang bertanda tangan dibawah ini :

Nama : Asa Cahyo Nugroho
NIM : 2071100038
Program Studi : Teknik Informatika
Fakultas : Fakultas Teknologi dan Komputer
Pembimbing : 1. Aryati Wuryandari, S.T., M.Kom
2. Muryanto, S.T., M.Cs

Dengan ini menyatakan dengan sesungguhnya bahwa karya ilmiah Skripsi yang berjudul. **“Analisis Sentimen Pengguna Twitter pada Layanan BPJS Kesehatan Menggunakan Algoritma *Naïve Bayes Classifier*”** adalah benar-benar karya saya sendiri dan bebas dari plagiat. Hal-hal yang bukan merupakan karya saya dalam skripsi ini telah diberi tanda sitasi dan ditunjukkan dalam Daftar Pustaka.

Apabila dikemudian hari terbukti pernyataan saya tidak benar, saya bersedia menerima sanksi akademik berupa pembatalan ijazah dan pencabutan gelar yang saya peroleh dari skripsi.

Klaten, 13 Agustus 2024

Yang membuat Pernyataan



Asa Cahyo Nugroho

PERSEMBAHAN

Segala puji bagi Allah SWT yang telah memberikan limpahan rahmat, hidayah, serta inayah-Nya sehingga skripsi yang sederhana ini dapat terselesaikan dengan baik dan lancar. Sholawat serta salam senantiasa selalu tercurahkan kepada junjungan kita Baginda Nabiyullah Muhammad SAW semoga kelak kita mendapat syafaatnya di yaumul akhir aamiin. Untuk ini peneliti persembahkan skripsi ini dan rasa terimakasih kepada:

1. Kedua orang tua saya terkasih, serta adik saya tercinta yang telah memberikan kasih sayang, do'a serta motivasi baik secara moril maupun materil dan terima kasih selalu menasehati agar menjadi lebih baik.
2. Keluarga besar Fakultas Teknologi dan Komputer terutama kepada teman-teman saya di prodi Informatika terima kasih atas kenangan indah dan kebersamaanya selama ini.
3. Bapak Ibu dosen Fakultas Teknologi dan Komputer yang selalu membimbing saya dengan baik.
4. Para sahabat-sahabat yang selalu mendukung dan memberikan semangat agar saya bisa menyelesaikan skripsi ini.

MOTTO

"Menuntut ilmu itu wajib atas setiap Muslim" (HR. Ibnu Majah no. 224, dari sahabat Anas bin Malik radhiyallahu 'anhu, dishahihkan Al Albani dalam Shahiih al-Jaami'ish Shaghiir no. 3913).

KATA PENGANTAR

Alhamdulillah, segala Puji dan Syukur kita curahkan kepada Allah Subhanahu Wata'ala. Dzat yang hanya kepada-Nya lah kita memohon pertolongan. Alhamdulillah atas segala pertolongan, rahmat, kasih sayang, dan hidayah-Nya peneliti dapat menyelesaikan skripsi yang berjudul "Implementasi Algoritma Apriori Pada Sistem Informasi Untuk Mencari Pola Pada Transaksi Pariwisata". Shalawat serta salam selalu kita curahkan kepada junjungan Nabi Agung kita, Suri Tauladan kita, Baginda Nabi Muhammad Shallallahu Alaihi Wasallam yang senantiasa menjadi sumber inspirasi dan teladan terbaik bagi seluruh umat manusia yang kita tunggu syafaatnya di Yaumul Akhir nanti.

Peneliti menyadari ada berbagai pihak yang memberikan dukungan dan bantuan selama menyelesaikan studi dan tugas akhir ini. Oleh karena itu, sudah sepantasnya peneliti dengan rasa penuh hormat mengucapkan terimakasih dan mendoakan semoga Allah memberikan balasan terbaik kepada:

1. Bapak Prof. Dr. H Triyono, M.Pd. selaku Rektor Universitas Widya Dharma Klaten.
2. Bapak Harri Purnomo, ST. MT. selaku Dekan Fakultas Teknologi dan Komputer Universitas Widya Dharma Klaten.
3. Bapak Doni Setyawan, ST, M.Cs. selaku Ketua Program Studi Teknik Informatika Universitas Widya Dharma Klaten.
4. Ibu Aryati Wuryandari, ST, M.Kom. selaku Pembimbing II saya, yang juga membimbing saya dengan penuh kesabaran sehingga saya dapat menyusun tugas akhir ini

5. Bapak Muryanto, ST, M.Cs. selaku dosen pembimbing I saya, yang telah membimbing saya dengan penuh kesabaran sehingga saya dapat menyusun tugas akhir ini.
6. Seluruh dosen Fakultas Teknologi dan Komputer Universitas Widya Dharma Klaten yang telah memberikan bekal ilmu pengetahuannya.
7. Bapak Ibu selaku orang tua saya yang selalu mendoakan dan memberikan semangat tanpa henti.
8. Sahabat-sahabatku seperjuangan dari TI A dan TI B angkatan 2024 dan seluruh teman-teman saya yang tidak bisa saya sebutkan satu per satu baik dari lingkungan kampus Universitas Widya Dharma Klaten maupun luar kampus yang telah memberikan saya rasa semangat, dukungan moril untuk menyelesaikan tugas akhir ini.

Dan akhir kata, peneliti menyadari bahwa tidak ada yang sempurna di dunia ini, peneliti masih melakukan kesalahan dalam penyusunan skripsi ini. Oleh karena itu, peneliti meminta maaf yang sebesar-besarnya atas kesalahan yang telah diperbuat.

Peneliti berharap semoga skripsi ini dapat memberikan manfaat bagi pembaca dan dapat dijadikan referensi demi pengembangan ke arah yang lebih baik lagi. Kebenaran datangnya dari Allah dan kesalahan datangnya dari diri peneliti. Semoga Allah SWT senantiasa melimpahkan Rahmat dan Ridho-Nya kepada kita semua.

Klaten, 6 Agustus 2024

Peneliti,

Asa Cahyo Nugroho

DAFTAR ISI

HALAMAN PERSETUJUAN.....	i
HALAMAN PENGESAHAN	ii
SURAT PERNYATAAN KEASLIAN PENELITIAN	iii
PERSEMBAHAN.....	iv
MOTTO	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN.....	xiii
ABSTRAK.....	xiv
BAB I PENDAHULUAN.....	1
A. Latar Belakang.....	1
1. Alasan Pemilihan Judul.....	3
2. Rumusan Masalah	4
3. Batasan Masalah.....	4
4. Keaslian Penelitian.....	4
5. Manfaat Penelitian.....	5
B. Tujuan Penelitian	5
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI	7
A. Tinjauan Pustaka.....	7
1. Konteks Penelitian.....	7
2. Penelitian Terdahulu	9
B. Landasan Teori.....	14
1. Analisis Sentimen.....	14
2. Natural Language Processing (NLP).....	16
3. Media Sosial.....	18
4. Twitter	19
5. API Twitter.....	21
6. Text Mining.....	21
7. Text Preprocessing	23
8. Machine Learning	25

9.	Klasifikasi.....	27
10.	Naïve Bayes	30
11.	Term Frequency Inverse Document Frequency (TF-IDF)	31
12.	K-Fold Cross-Validation	33
13.	Python.....	34
14.	Database.....	35
15.	Diagram Konteks	35
16.	Entity Relationship Diagram	36
17.	Data Flow Diagram	37
18.	Tweet-Harvest	38
19.	Flowchart.....	38
BAB III METODOLOGI PENELITIAN		40
A.	Alat dan bahan Penelitian	40
1.	Alat Penelitian	40
2.	Bahan penelitian	44
B.	Jalannya Penelitian.....	44
1.	Perancangan Sistem.....	44
2.	Perancangan Database.....	48
3.	Pengumpulan data	52
4.	Persiapan data.....	53
5.	Ekstraksi Fitur.....	65
6.	Pemodelan data	67
7.	Evaluasi Model.....	74
8.	Visualisasi data.....	76
C.	Hambatan Penelitian	76
BAB IV HASIL DAN PEMBAHASAN		79
A.	Hasil Pengumpulan data	79
1.	Hasil Crawling Data	79
2.	Hasil Pre-processing Data	79
3.	Hasil Visualisasi Dataset.....	83
4.	Pelabelan Dataset	84
5.	Ekstraksi Fitur	86
6.	Klasifikasi Model	87
B.	Pembahasan.....	88
C.	Kelemahan Sistem	91

BAB V KESIMPULAN DAN SARAN	92
A. Kesimpulan	92
B. Saran	93
DAFTAR PUSTAKA	94

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	12
Tabel 2. 2 Contoh Case Folding	23
Tabel 2. 3 contoh Cleansing atau Removing	24
Tabel 2. 4 Contoh Filtering	25
Tabel 2. 5 Contoh Tokenizing	25
Tabel 2. 6 Contoh Stemming	25
Tabel 2. 7 Confusion Matrix	29
Tabel 2. 8 simbol diagram konteks	36
Tabel 2. 9 Entity Relationship Diagram	36
Tabel 2. 10 Simbol Data Flow Diagram	37
Tabel 2. 11 Simbol Flowchart.....	39
Tabel 3. 1 Database Logis Tabel Crawling.....	50
Tabel 3. 2 Database Logis Tabel Labeling	51
Tabel 3. 3 Tahap pelabelan data	52
Tabel 3. 4 hasil dari penerapan proses case folding.....	55
Tabel 3. 5 hasil penerapan proses cleansing dan removing	57
Tabel 3. 6 hasil proses penerapan normalization	59
Tabel 3. 7 hasil proses penerapan stopwords removal.....	61
Tabel 3. 8 hasil penerapan proses tokenizing	63
Tabel 3. 9 hasil proses penerapan stemming	65
Tabel 3. 10 Data/Dokumen	66
Tabel 3. 11 hasil perhitungan tf-idf.....	66
Tabel 3. 12 data latih.....	67
Tabel 3. 13 data uji.....	72
Tabel 3. 14 data uji setelah preprocessing	72
Tabel 3. 15 Data latih dan data uji	72
Tabel 3. 16 Hasil Confusion Matrix	75

DAFTAR GAMBAR

Gambar 2. 1 cara kerja analisis sentiment (Lyman, 2023)	15
Gambar 2. 2 proses text mini (Fahlevi, 2021).	22
Gambar 2. 3 Cara kerja Machine Learning.....	27
Gambar 2. 4 proses klasifikasi(Asdar, 2022).....	28
Gambar 2. 5 Contoh Simulasi K-Fold Cross-Validation	34
Gambar 3. 1 Diagram Konteks Sistem Analisis Sentimen	45
Gambar 3. 2 Data Flow Diagram Level 1	46
Gambar 3. 3 Data Flow Diagram Level 2 Proses 2 Pre-processing.....	47
Gambar 3. 4 Data Flow Diagram Level 2 Proses 3 Labeling	48
Gambar 3. 5 Entity Relationship Diagram.....	49
Gambar 3. 6 Rancangan Database Fisik	51
Gambar 3. 7 Flowchart alur case folding.....	54
Gambar 3. 8 Kode Pemrograman proses case folding	54
Gambar 3. 9 Flowchart proses Cleansing atau removing	56
Gambar 3. 10 Kode Program proses cleansing atau removing	57
Gambar 3. 11 flowchart proses normalization	58
Gambar 3. 12 kode program proses normalization	59
Gambar 3. 13 Flowchart proses stopword removal	60
Gambar 3. 14 kode program proses stopwords removal.....	61
Gambar 3. 15 Flowchart proses tokenization	62
Gambar 3. 16 kode program proses tokenization	63
Gambar 3. 17 Flowchart proses stemming	64
Gambar 3. 18 kode program proses stemming	65
Gambar 4. 1 hasil crawling data	79
Gambar 4. 2 Hasil Cleaning.....	80
Gambar 4. 3 Hasil Case Folding	80
Gambar 4. 4 Kamus normalisasi.....	81
Gambar 4. 5 Hasil Normalisasi.....	81
Gambar 4. 6 Hasil Tokenizing	82
Gambar 4. 7 Hasil Stopword Removal	82
Gambar 4. 8 hasil stemming	83
Gambar 4. 9 Wordcloud Dataset.....	83
Gambar 4. 10 Hasil pelabelan manual dan otomatis.....	84
Gambar 4. 11 Visualisasi hasil pelabelan	85
Gambar 4. 12 Visualisasi Positif.....	85
Gambar 4. 13 visualisasi negatif.....	86
Gambar 4. 14 hasil TF-IDF.....	86
Gambar 4. 15 Hasil Split Data	87
Gambar 4. 16 Hasil Klasifikasi Naïve Bayes.....	88
Gambar 4. 17 Hasil Confusion Matrix.....	89
Gambar 4. 18 K-Fold Cross-Validation.....	90

DAFTAR LAMPIRAN

Lampiran 1 Kode Program.....	98
------------------------------	----

ABSTRAK

ASA CAHYO NUGROHO NIM. 2071100038. Program Studi Teknik informatika, Fakultas Teknik dan Komputer, Univeritas Widya Dharma Klaten. Judul Skripsi : **Analisis Sentimen Pengguna Twitter pada Layanan BPJS Kesehatan menggunakan Algoritma Naïve Bayes Classifier.**

Twitter merupakan media sosial memiliki pengguna dengan jumlah yang cukup besar di Indonesia. Twitter menduduki peringkat ke 6 dengan jumlah akses di Indonesia dengan pengguna aktif di angka 100,9 juta. Pada penelitian ini twitter digunakan untuk mengumpulkan data tweet sesuai dengan tema yang diambil adalah data yang mengandung kata kunci terkait BPJS Kesehatan yang digunakan akan di gunakan untuk proses klasifikasi dengan membagi kelas menjadi 2 sentimen yaitu sentimen positif dan negatif.

Penelitian ini menggunakan metode *data crawling* dengan menggunakan *tool tweet-harvest* dan *Twitter API* untuk proses pengumpulan data yang akan digunakan untuk tahap pre-processing untuk mencari data yang lebih terstruktur. Setelah itu dilakukan pelabelan dengan cara manual dan otomatis, setelah itu data digunakan untuk tahap ekstraksi fitur menggunakan *TF-IDF*. Setelah itu dilakukan split data untuk melakukan proses klasifikasi sentimen dengan menggunakan algoritma *naïve bayes*.

Untuk hasil evaluasi model menggunakan confusion matrik dan *k-fold cross-validation* dan menghasilkan akurasi dari model *naïve bayes* sebesar 79,59%. Dan hasil dari *f-fold cross-validation* dengan mencari *k (fold) = 5* memiliki akurasi sebesar 78,72% dengan menghasilkan nilai yang konsisten untuk setiap proses *fold* nya. Dengan ini model dari *naïve bayes* memiliki kemampuan yang baik pada penelitian ini dengan mengambil topik BPJS Kesehatan.

Kata Kunci: Analisis Sentimen, Twitter, BPJS Kesehatan, *Naïve Bayes*

BAB I

PENDAHULUAN

A. Latar Belakang

Di era globalisasi seperti sekarang penggunaan media sosial mengalami peningkatan yang sangat signifikan di kalangan masyarakat dunia. Perkembangan media sosial ini menyoroti transformasi yang mendalam dalam interaksi manusia, komunikasi, dan budaya di kalangan masyarakat. Media sosial telah menjadi kekuatan yang mempengaruhi aspek kehidupan manusia di era digital saat ini, seperti cara kita berkomunikasi dan berinteraksi hingga memperoleh informasi, membentuk opini, dan identitas online. Dengan laju pertumbuhan media sosial yang sangat pesat dan penetrasi yang luas di seluruh dunia, pemaham tentang evolusi, dampak dan tantangan penggunaan media sosial. Dengan demikian para akademisi, praktisi dan pengambil keputusan menjadi terbantu untuk mengeksplorasi sejarah, struktur, dampak, serta tantangan yang di hadapi dalam perkembangan media sosial di era digital saat ini. Dampak Perkembangan dari media sosial ini ditandai dengan munculnya platform-platform besar seperti, Facebook, Twitter, Instagram, Youtube, Google, Path, dan WhatApps.

Pertumbuhan media sosial di Indonesia mengalami peningkatan yang sangat signifikan, dengan munculnya platform media sosial yang semakin canggih dan kompatibel seperti sekarang ini. Pengguna media sosial di Indonesia yang aktif tercatat dari Januari tahun 2024 berdasarkan data dari sumber datareportal.com (2024), berjumlah 139,4 juta pengguna aktif. Jumlah pengguna aktif media sosial yang tercatat sampai dengan Januari tahun 2024 setara dengan 49.9% dari total populasi terbaru di Indonesia. Sumber dari datareportal.com (2024) mengenai pengguna aktif media sosial di Indonesia terdiri atas 46,5% pengguna media sosial di Indonesia adalah perempuan dan 53,5% pengguna media sosial aktif adalah laki-laki. Dan beberapa platform yang digunakan oleh

pengguna media sosial yang aktif di Indonesia berdasarkan dari sumber data portal (2024), pengguna facebook tercatat pada Januari tahun 2024 mencapai angka 117,6 juta, untuk penggunaan Youtube di Indonesia memiliki 139,0 juta pengguna, untuk pengguna Twitter di Indonesia memiliki angka 100,9 juta.

Para pengguna sosial media di Indonesia mulai menggunakan media sosial nya untuk berbisnis, dengan munculnya akun-akun yang memasarkan produk-produk seperti elektronik, kecantikan, obat-obatan dan buku-buku untuk pendidikan. Platform media sosial digunakan sebagai alat bantu untuk menyampaikan informasi yang lebih responsif dan transparan. Selain itu juga sebagai sarana promosi kemitraan instansi pemerintahan agar lebih efisien, dan memberdayakan warga dengan membuat pengetahuan dan sumber daya lainnya agar lebih mudah diakses. Instansi pemerintahan saat ini memanfaatkan media sosial sebagai alternatif pemberian layanan dan membuat kebijakan-kebijakan untuk publik, serta pengumpulan keluhan-keluhan dan masukan yang di berikan oleh masyarakat serta lebih mendekatkan diri kepada masyarakat lewat layanan media sosial tersebut. Semakin mudahnya akses layanan media sosial maka masyarakat lebih pintar dalam memberikan sebuah opini-opini misalnya opini yang ditujukan ke instansi pemerintahan. Opini yang diberikan oleh masyarakat yang ditujukan untuk akun instansi pemerintahan dapat di manfaatkan untuk meninjau opini publik yang berguna untuk instansi pemerintahan.

Penelitian ini mengusulkan penggunaan algoritma berbasis mesin pembelajaran naïve bayes classifier untuk menyelesaikan penganalisaan sentimen pada BPJS Kesehatan. Banyak masyarakat yang beropini terhadap layanan dan pelayanan dari BPJS Kesehatan. Opini-opini yang telah diberikan masyarakat untuk layanan BPJS Kesehatan memiliki beragam sifat, ada yang memberikan opini positif, ada yang memberikan opini negatif. Opini-opini dari masyarakat tersebut terkumpul dan dapat digali lagi untuk

menghasilkan suatu kumpulan sentimen ini biasanya diperoleh melalui jumlah like dan tweet yang terpadat dalam suatu postingan dari layanan media sosial instansi pemerintah. Sentimen tersebut berguna sebagai evaluasi untuk pemerintahan untuk membuat kebijakan-kebijakan yang dibutuhkan sesuai hasil dari sentimen masyarakat yang sudah di olah dan di kumpulkan.

Oleh karena itu peneliti melakukan penelitian tentang sentimen analisis untuk menganalisis teks opini yang dikemukakan masyarakat yang tertuang pada kolom tweet akun layanan BPJS Kesehatan pada platform media sosial yakni Twitter dengan menggunakan Algoritma naïve bayes.

1. Alasan Pemilihan Judul

Ada beberapa alasan sehingga penulis memilih judul analisis sentimen tweet pengguna Twitter pada layanan BPJS Kesehatan menggunakan algoritma naïve bayes classifier antara lain :

- a. Tujuan penulis yaitu menganalisis sentiment tweet pengguna platform media sosial terhadap akun layanan milik instansi pemerintah BPJS Kesehatan.
- b. Naive Bayes Classifier sebagai metode analisis sentimen menunjukkan penerapan teknologi dalam menganalisis data sosial media. Ini merupakan topik yang menarik karena berkaitan dengan bidang kecerdasan buatan dan pemrosesan bahasa alami.
- c. BPJS Kesehatan adalah lembaga penting dalam konteks kesehatan di Indonesia. Analisis sentimen terhadap layanan mereka bisa memberikan wawasan berharga tentang persepsi masyarakat terhadap lembaga tersebut. Dengan memilih topik ini, peneliti mungkin ingin menangkap isu-isu aktual yang berkaitan dengan layanan publik dan respons masyarakat terhadapnya
- d. Twitter adalah platform yang sangat populer di kalangan pengguna internet, dan tweet di sana sering mencerminkan opini dan persepsi yang beragam. Dengan fokus

pada tweet yang terdapat pada twitter, penelitian ini mungkin mencoba untuk menggali sentimen masyarakat yang terkait dengan layanan BPJS Kesehatan di platform ini.

2. Rumusan Masalah

Berdasarkan latar belakang yang telah di atas, maka rumusan masalah dalam tugas akhir ini adalah bagaimana menganalisis sentimen masyarakat melalui tweet berbentuk teks pada platform media sosial akun milik instansi pemerintah BPJS Kesehatan menggunakan algoritma naïve bayes classifier.

3. Batasan Masalah

Untuk menghindari penyimpangan judul dan tujuan, maka penulis membuat batasan permasalahan pada penelitian ini, adapun batasan masalahnya adalah :

- a. Analisis sentimen hanya dilakukan pada tweet yang ditulis dalam Bahasa Indonesia.
- b. sumber data hanya pada tweet yang terkait dengan BPJS Kesehatan di Twitter.
- c. Teknik yang digunakan dalam penelitian ini adalah Klasifikasi, menggunakan naïve bayes classifier.
- d. Analisis hanya dilakukan pada dua kategori sentimen, positif dan negatif.

4. Keaslian Penelitian

Keaslian penelitian ini berdasarkan beberapa penelitian sebelumnya yang mempunyai karakteristik yang hampir sama, meskipun berbeda dalam hal pemilihan subjek, objek, waktu, dan tempat penelitian. Penelitian ini di harapkan memberikan ilmu yang bermanfaat bagi penulis dan dapat memberikan kontribusi yang berharga

dalam pemahaman terhadap persepsi masyarakat terhadap layanan tersebut melalui analisis sentimen.

5. Manfaat Penelitian

Adapun manfaat penelitian yang dapat diperoleh dari penelitian ini adalah :

a. Bagi Peneliti

Mengukur kemampuan dan wawasan serta menerapkan secara langsung ilmu yang telah dipelajari.

b. Bagi Pembaca

1. Mendapat pengetahuan cabang ilmu natural language processing yang merupakan salah satu *artificial intelligence*.
2. Mendapat inspirasi sebagai referensi penelitian yang dapat dikembangkan lebih lanjut.

c. Bagi Akademik

1. Sebagai tolak ukur sejauh mana pemahaman dan penguasaan materi terhadap teori yang diajarkan
2. Sebagai bahan referensi bagi mereka yang mengadakan penelitian untuk di kembangkan lebih lanjut dengan permasalahan yang berbeda

B. Tujuan Penelitian

Tujuan dilakukannya penelitian ini adalah untuk menggali dan memahami sentimen pengguna Twitter terhadap layanan BPJS Kesehatan menggunakan pendekatan algoritma Naive Bayes Classifier.

Secara Khusus dilakukannya penelitian ini adalah :

- a. Bertujuan untuk menentukan tweet yang diberikan oleh pengguna memiliki sentimen positif, negatif terhadap layanan pada BPJS Kesehatan

- b. Bertujuan untuk mengidentifikasi sentimen terhadap layanan BPJS Kesehatan dari kurun waktu tertentu untuk mengetahui pola dan tren yang sedang terjadi.
- c. Bertujuan untuk memberikan informasi bagi BPJS kesehatan untuk memperbaiki dan meningkatkan kualitas pelayanan berdasarkan hasil dari sentimen yang diberikan oleh pengguna
- d. Untuk mengidentifikasi masalah yang terjadi terhadap layanan BPJS Kesehatan untuk mengambil strategi dan kebijakan kedepannya agar lebih tepat dan efektif

BAB V

KESIMPULAN DAN SARAN

A. Kesimpulan

Setelah melakukan penelitian tentang analisis sentimen tweet pengguna twitter pada layanan BPJS Kesehatan menggunakan algoritma naïve bayes classifier maka dapat disimpulkan :

1. Penelitian melakukan crawling data pada tweet dengan keyword BPJS kesehatan RI di aplikasi Twitter mendapatkan dataset sebanyak 2506 data. Setelah melewati tahap *preprocessing* dataset berkurang menjadi 2425 data karena sudah melewati tahapan seperti cleaning, normalisasi, case folding , stopword removal, tokenizing, stemming dan pelabelan data yang dibagi menjadi dua kelas yaitu positif dan negatif. Dan penerapan fitur *TF-IDF* untuk mengetahui bobot kata pada data.
2. Hasil metode *naïve bayes* untuk model multinomial dengan jumlah 1364 sentimen dan 1061 sentimen negatif. Nilai akurasi sebesar 79,59 %, dengan margin error sebesar 1,67 %, presisi prediksi positif sebesar 80%, presisi prediksi negatif sebesar 78%, untuk recall data positif sebesar 77% dan negatif memiliki angka 77%. Untuk evaluasi model sendiri menggunakan *confusion matrix* dengan hasil *True positive* = 211 data , *True negative* = 175 data, *False positive* = 51 data, dan *False negative* = 48 data. Untuk evaluasi model teknik *cross validation* yaitu *kl-fold cross-validation* mencari $k = 5$ dan memperoleh hasil untuk $k = 1$ mendapatkan nilai akurasi 79.59%, $k = 2$ mendapatkan nilai akurasi 78.35%, $k = 3$ mendapatkan nilai akurasi 78.56%, $k = 4$ mendapatkan nilai akurasi 80.00%, dan $k = 5$ mendapatkan nilai akurasi 77.11%, maka rata-rata akurasi yang di dapatkan sebesar 78,72%.

B. Saran

Berdasarkan hasil penelitian tentang analisis sentimen tweet pengguna twitter pada layanan BPJS kesehatan, saran yang dapat diberikan oleh penulis untuk penelitian selanjutnya adalah sebagai berikut:

1. Menambahkan kelas sentimen netral untuk dataset yang akan dilakukan penelitian selanjutnya.
2. Menggunakan model *machine learning* lain yang memiliki kualitas yang lebih baik untuk proses pengujian data, serta melakukan perbandingan model lain untuk mencari performa terbaik dari model.
3. Menambah jumlah dataset dan mengambil dataset lain dari berbagai sosial media seperti facebook dan instagram
4. Penelitian selanjutnya dapat menggunakan lain seperti bahasa inggris karena fitur yang tersedia sangat baik untuk dilakukan penelitian.

DAFTAR PUSTAKA

- Aam, S. r., & marlina, L. (2020). Analisis Sentimen terkait Sertifikasi Halal. *Journal of Economics and Business Aseanomics*, 69-85.
- Alfarandy, M. H. (2024). *Text Preprocessing in Natural Language Processing*. (PT . Bisa Artifisial Indonesia) Retrieved from <https://bisa.ai/portofolio/detail/MzA5OA>
- Amazon, A. (2023). *Apa itu Pemrosesan Bahasa Alami (NLP)?* (Amazon) Retrieved from <https://aws.amazon.com/id/what-is/nlp/>
- Anggraeny, M. S. (2020). *Confusion Matrix*. (Bina Nusantara University) Retrieved from <https://socs.binus.ac.id/2020/11/01/confusion-matrix/>
- Asdar, A. R. (2022, October 19). *Classification Method: Basic Concepts*. (Kompasiana Beyond Blogging) Retrieved from <https://www.kompasiana.com/aulyarahmanasdar/634f9d7408a8b548ca2475c2/classfication-method-basic-concepts?page=all#section1>
- Asri, Y., NitaSuliyanti, W., Kuswardani, D., & Fajri, M. (2022). Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naive Bayes dalam menganalisis sentimen data ulasan PLN Mobile. *Jurnal Pengkajian dan Penerapan Teknik Informatika*, 264-275.
- Azis, H., Purnawansyah, Fattah, F., & Putri, I. P. (2020). Performa Klasifikasi K-NN dan Cross-validation pada Data Pasien Pengidap Penyakit Jantung. *ILKOM Jurnal Ilmiah*, XII(2), 81-86.
- Baehaqi, F., & Cahyono, N. (2024). Analisis Sentimen Terhadap Cyberbullying Pada Komentar di Instagram Menggunakan Algoritma Naïve Bayes. *Indonesian Journal of Computer Science*, 1051-1063.
- Batu Bara, E. A., Nasution, K. A., Ginting, R. Z., & Kartini. (2022). Penelitian tentang Twitter. *Jurnal Edukasi Nonformal*, III(2), 167-172.
- Chandradev, V., Suarjaya, I. M., & Bayupati, I. P. (2023). Analisis Sentimen Review Hotel Menggunakan Metode Deep Learning BERT. *Jurnal Buana Informatika*, 107-116.
- Debby, A., & Auliya, R. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi*, 50-58.
- Efrain, D. A., & Ermatita. (2023). Analisis Sentimen Pada Sosial Media Instagram Menggunakan Algoritma Naive Bayes (Studi Kasus : Timnas Futsal Indonesia). *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya*, 498-509.
- Facebook. (2024). *Platform Instagram Graph API*. (Meta) Retrieved from https://developers.facebook.com/docs/instagram-api?locale=id_ID
- Fahlevi, A. (2021, September 30). *Proses Data Mining KDD*. (BINUS University) Retrieved from <https://sis.binus.ac.id/2021/09/30/proses-data-mining-kdd/>
- Fani, S. M., Santoso, R., & Suparti. (2021). Penerapan Text Mining untuk Melakukan Clustering Data Tweet Akun Blibli pada Media Sosial Twitter Menggunakan K-Means Clustering. *Jurnal Gaussian* , X(4), 583-593.

- Firdaus, A., & Firdaus, W. I. (2021). Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan). *Jurnal JUPITER*, XIII(1), 66-78.
- Herlambang, B. A., & Setyawati, V. A. (2015). Perancangan Data Flow Diagram Sistem Pakar Penentuan Kebutuhan Gizi Bagi Individu Normal Berbasis Web. *Jurnal Informatika UPGRIS*, I, 78-85.
- IBM. (2023). *What is natural language processing (NLP)?* (IBM's ESG Framework) Retrieved from <https://www.ibm.com/topics/natural-language-processing>
- Intern , D. (2020, August 19). *Apa itu Machine Learning? Beserta Pengertian dan Cara Kerjanya.* (dicoding) Retrieved from <https://www.dicoding.com/blog/machine-learning-adalah/>
- Intern, D. (2023, May 31). *Python: Pengertian, Contoh Penggunaan, dan Manfaat Mempelajarinya.* (dicoding) Retrieved from <https://www.dicoding.com/blog/python-pengertian-contoh-penggunaan-dan-manfaat-mempelajarinya/>
- Karim, A. (2021). *Analisis Sentimen pada Komentar Sosial Media Instagram Layanan Kesehatan BPJS Menggunakan Naive Bayes.*
- Kurniawan, D., Purnomo, H., & Iriani, A. (2023). Analisis Sentimen Produk Herbal Jamu pada Media Sosial Instagram. *Jurnal Sistem Informasi*, 3202-3210.
- Lyman, C. (2023, February 24). *Apa itu Analisis Sentimen di AI dan Cara Kerjanya?* (PT. Pintu Kemana Saja) Retrieved from <https://pintu.co.id/blog/analisis-sentimen>
- Nugraha, K. A. (2021). Analisis Sentimen Berbasis Emoticon pada Komentar Instagram Bahasa Indonesia Menggunakan Naïve Bayes. *Jurnal Teknik Informatika dan Sistem Informasi*, 715-721.
- Populix. (2021, 09 18). Media Sosial. *Contoh hingga Manfaatnya bagi Pebisnis.*
- Rais, T. I. (2022). Analisis terhadap komentar Video Youtube Raiden Shogun – judgement of euthymia menggunakan metode Majority Voting.
- Revou. (2024). *Algoritma Naive Bayes.* (PT. Revolusi Cita Edukasi) Retrieved from <https://revou.co/kosakata/algoritma-naive-bayes>
- Rifaldi, D., Fadlil, A., & Herman. (2023). Teknik Preprocessing pada Text Mining Menggunakan Data Tweet 'Mental Health'. *Jurnal Pendidikan Teknologi Informasi*, III(2), 161-171.
- Salsabilla, N. A. (2022). Analisis Sentimen pada Media Sosial Twitter terhadap tokoh Gus Dur menggunakan metode Naïve Bayes dan Support Vector Machine (SVM).
- Satria, H. (2023). *Cara Mendapatkan Data (Crawl) Twitter X.* Retrieved from Helmi Satria Writings: Stories and Learnings: <https://helmisatria.com/blog>
- Septiani, D., & Isabela, I. (2022). Analisis Term Frequency Inverse Document Frequency (TF-IDF) dalam Temu Kembali pada Dokumen Teks. *Jurnal Sistem dan Teknologi Informasi Indonesia*, I(2), 81-88.

- Setyawan, R. (2021, october 31). *Apa itu Data Mining dan Bagaimana Metodenya?* (Dicoding) Retrieved from <https://www.dicoding.com/blog/apa-itu-data-mining/>
- Sihotang, R., Saputro, H., & Novari, S. (2021). Sistem Informasi penggajian LKPEnglish Academy menggunakan Embarcadero XE2 Berbasis Client Server. *Jurnal Teknik Informatika Mahakarya (JTIM)*, IV(1), 28-36.
- Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, XVII(1), 168-192.
- Thomas, S., Yuliana, & P, N. (2021). Studi Analisis Metode Analisis Sentimen pada YouTube. *Journal of Information Technology*, 1-7.
- Youtap. (2022, February 11). *Maksimalkan 5 Fitur Instagram Berikut untuk Bantu Tingkatkan Penjualan.* (PT. Mitra Digital Sukses) Retrieved from <https://www.youtap.id/blog/5-fitur-instagram-yang-perlu-dimaksimalkan>
- Yulianeu, A., & Oktamala, R. (2022). Sistem Informasi Geografis Trayek Angkutan Umum di Kota Tasikmalaya Berbasis Web. *Jurnal Teknik Informatika*, X(2), 125-134.
- Yunita, H., & Dina. (2021). Aplikasi Pelayanan Kesehatan pada Puskesmas. *Jurnal Informatika dan Teknologi Komputer*, I(1), 1-13.
- Yusuf, F., Rahman, H., Rahmi, S., & Lismayani, A. (2023). Pemanfaatan Media Sosial sebagai Sarana Komunikasi, Informasi, dan Dokumentasi : Pendidikan di Majelis Taklim Annur Sejahtera. *Jurnal Hasil-Hasil Pengabdian dan Pemberdayaan Masyarakat*, II, 1-8.