

Analisis Sentimen pada Komentar Sosial Media Instagram Layanan Kesehatan BPJS Menggunakan Naïve Bayes Classifier

Ahmad Karim¹, Sam Farisa C², Mustafa³

^{1,2,3} Teknik Informatika, Universitas Islam Sultan Agung Semarang

¹ahmadkarim@std.unissula.ac.id

Abstrak - Instansi Pelayanan Kesehatan Pemerintah seperti BPJS Kesehatan membutuhkan feedback dari masyarakat yang merupakan sumber informasi untuk meningkatkan kinerja. Umpan balik yang didapatkan dari masyarakat tidak hanya yang bersifat positif, melainkan juga bersifat negatif. Tujuan penelitian ini adalah menganalisis sentimen berdasarkan komentar masyarakat terhadap akun sosial media pemerintahan dengan mengklasifikasikan secara otomatis menggunakan algoritma naïve bayes classifier. Hasil implementasi dapat membantu pemilih akun menganalisis respon masyarakat terhadap suatu postingan berupa respon positive, negative atau netral. Metode klasifikasi berbasis pencocokan teks boyer moore berhasil diimplementasikan untuk melakukan klasifikasi teks komentar di postingan instagram. Metode yang diusulkan memiliki performa baik dengan akurasi mencapai 73 persen.

Kata Kunci: Instagram, Text Mining, Sentiment Analysis, Naïve Bayes

Abstract Government Health Service Agencies such as BPJS Health need feedback from the public which is a source of information to improve performance. Feedback obtained from the community is not only positive, but also negative. The purpose of this study is to analyze sentiment based on public comments on government social media accounts by classifying automatically using the naïve Bayes classifier algorithm. The results of the implementation can help account voters analyze the public's response to a post in the form of a positive, negative or neutral response. The Boyer Moore text matching-based classification method was successfully implemented to classify comment text on Instagram posts. The proposed method has good performance with accuracy reaching 73 percent.

Keywords: Instagram, Text Mining, Sentiment Analysis, Naïve Bayes

I. PENDAHULUAN

Media sosial sekarang ini tidak hanya digunakan sebagai sarana pertemanan, mencari teman, tapi sudah banyak digunakan untuk kegiatan lain namun juga digunakan untuk promosi produk, sarana jual beli sampai kampanye politik. Media sosial telah membantu masyarakat untuk mendapatkan informasi terbaru terkait peristiwa atau kejadian di lingkungan sekitar ataupun lingkungan yang lebih luas. Atas dasar hal itulah, banyak pemerintah daerah menggunakan media sosial sebagai satu layanan electronic government (E-Government) sebagai sarana dalam menyampaikan informasi kepada masyarakat [1].

Media sosial sebagai alat bantu melengkapi layanan E-Government menjadikan institusi publik lebih responsif dan transparan, mempromosikan kemitraan pemerintah lebih efisien, dan memberdayakan warga dengan membuat pengetahuan dan sumber daya lainnya lebih dapat diakses langsung. Media sosial digunakan instansi pemerintah sebagai alternatif pemberian pelayanan dan pembuatan kebijakan publik, pengumpulan berbagai masukan dari masyarakat, dan untuk lebih mendekatkan diri kepada masyarakat.

Semakin maraknya penggunaan teknologi smartphone menjadikan masyarakat dengan mudah menyampaikan opini melalui media sosial ke akun instansi pemerintah [2]. Banyaknya jumlah opini masyarakat di media sosial kepada akun-akun instansi pemerintah dapat dimanfaatkan pemerintah untuk memantau opini publik. Pemerintah selaku pelayan public dan pembuat kebijakan membutuhkan feedback dari masyarakat yang merupakan sumber informasi untuk meningkatkan kinerja. Umpan balik yang didapatkan dari masyarakat tidak hanya yang bersifat positif, melainkan juga bersifat negatif. Masyarakat seringkali memberikan keluhan-keluhan terhadap pelayanan yang telah diberikan atau memberikan usulan-usulan terhadap program kerja atau sekedar ingin mengetahui informasi yang ada saat ini Opini-opini yang terkumpul dalam media sosial dapat digali informasinya yang akan menghasilkan suatu kumpulan Sentimen terhadap kebijakan-kebijakan maupun informasi yang disampaikan instansi pemerintah [3]. Sentimen terhadap suatu informasi yang disampaikan instansi pemerintah dapat diperoleh dengan menggali jumlah like dan teks-teks komentar pada suatu postingan. Sentimen public atau masyarakat terhadap kebijakan atau sosialisasi dari pemerintah diperlukan agar dalam pembuatan keputusan benar-benar yang dibutuhkan oleh masyarakat sehingga dapat dijadikan sebagai landasan dalam

pembuatan keputusan. Selain itu dengan adanya sentimen masyarakat tersebut dapat dijadikan pendukung dalam pembuatan keputusan.

Penelitian tentang analisis Sentimen berdasarkan data teks dari media sosial banyak dilakukan. Salah satunya penelitian yang menggunakan metode machine learning SVM (Support Vector Machine) yang melakukan analisis sentimen pada kolom komentar Instagram yang berupaya untuk mengetahui sentimen dari setiap komentar dengan objek cyberbullying (perundungan di internet) [4]. Selain itu terdapat penelitian yang membandingkan SVM dengan Naïve bayes Classifier untuk menganalisis Sentimen tokoh publik di twitter [5]. Pada umumnya algoritma berbasis mesin pembelajaran seperti Neural Network, k-Nearest Neighbor, SVM dan Naïve bayes digunakan sebagai metode dalam klasifikasi teks [6].

Penelitian ini mengusulkan algoritma klasifikasi naïve bayes untuk menganalisis Sentimen pada akun instansi pemerintah terkait dengan layanan BPJS kesehatan pada media sosial Instagram. Instagram merupakan media sosial yang paling populer pada zaman sekarang. Pengguna yang dimulai dari anak-anak, remaja hingga orang dewasa turut mendongkrak popularitas Instagram. Sentimen analisis diusulkan dengan menganalisis teks pada kolom komentar yang dikemukakan masyarakat pada akun Instagram instansi pemerintah. Algoritma naïve bayes diusulkan karena merupakan salah satu teknik klasifikasi yang handal dalam mengklasifikasikan teks dan sudah banyak digunakan.

II. TINJAUAN PUSTAKA DAN DASAR TEORI

A. Tinjauan Pustaka

Sebelumnya terdapat penelitian yang dijadikan acuan dan referensi pada penelitian ini. Analisis Sentimen Dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter, [7]. Tokoh publik yang dinilai adalah tokoh yang dianggap layak dan memiliki kemampuan untuk dipilih menjadi pemimpin. Oleh karena itu, penelitian ini mencoba menganalisis tweet berbahasa Indonesia yang membicarakan tentang tokoh publik. Analisis dilakukan dengan melakukan klasifikasi tweet yang berisi sentimen masyarakat tentang tokoh tertentu. Metode klasifikasi yang digunakan dalam penelitian ini adalah Naive Bayes Classifier. Naive Bayes Classifier dikombinasikan dengan fitur untuk dapat mendeteksi negasi dan pembobotan menggunakan term frequency serta TF-IDF. Klasifikasi tweet pada penelitian ini diperoleh berdasarkan kombinasi antara kelas sentimen dan kelas kategori. Klasifikasi sentimen terdiri dari positif dan negatif sedangkan klasifikasi kategori terdiri dari kapabilitas, integritas, dan akseptabilitas. Hasil pengujian pada aplikasi yang dibangun dan pada tools RapidMiner memperlihatkan bahwa akurasi dengan term frequency memberikan hasil akurasi yang lebih baik daripada akurasi dengan fitur TF-IDF. Metode Support Vector Machine menghasilkan akurasi performansi yang lebih baik daripada metode Naive Bayes baik dalam klasifikasi sentimen maupun dalam klasifikasi kategori. Namun demikian, secara keseluruhan penggunaan metode Support Vector Machine dan Naive Bayes sama-sama memiliki performansi yang cukup baik untuk melakukan klasifikasi tweet.

Analisis Sentimen Hatespeech Pada Twitter Dengan Metode Naïve bayes Classifier Dan Support Vector Machine, [8]. Gerakan terbaru ada tagar Hatespeech (#HateSpeech), dilihat dari namanya sudah jelas yaitu ucapan kebencian. Penelitian ini menganalisis tagar tersebut, benarkan semua yang diberi tagar itu sentimennya kebencian. Proses klasifikasi pada penelitian ini menggunakan metode klasifikasi Naïve bayes Classifier (NBC) dan Support Vector Machine (SVM) dengan preprocessing data menggunakan tokenisasi, cleansing dan filtering. Data yang digunakan adalah tweet dalam bahasa Indonesia dengan tagar HateSpeech (#HateSpeech), dengan jumlah dataset sebanyak 522 tweet yang didistribusikan secara merata menjadi dua sentimen HateSpeech dan GoodSpeech. Hasil akurasi tertinggi didapatkan saat menggunakan metode klasifikasi Support Vector Machine (SVM) dengan tokenisasi unigram, stopword list Bahasa Indonesia dan emoticons, dengan nilai rata-rata akurasi mencapai 66,6%, nilai presisi 67,1%, nilai recall 66,7% nilai TP rate 66,7% dan nilai TN rate 75,8%.

Eksperimen Sistem Klasifikasi Analisa Sentimen Twitter Pada Akun Resmi Pemerintah Kota Surabaya Berbasis Pembelajaran Mesin, [9]. Salah satu sumber informasi yang dibutuhkan oleh Pemerintah untuk dapat meningkatkan kinerjanya adalah umpan balik dari masyarakat. Saat ini media sosial digunakan oleh masyarakat sebagai alat untuk menyalurkan opini dan kondisi yang terjadi di sekitar. Pemerintah Kota Surabaya memiliki akun media sosial yaitu twitter bertujuan untuk meningkatkan interaksi kepada masyarakat Surabaya secara luas. Untuk dapat memantau seluruh tweet pada seluruh akun bukanlah sesuatu yang mudah dilakukan secara manual. Dalam penelitian ini dilakukan pengklasifikasian opini masyarakat pada media twitter menggunakan pembelajaran mesin. Pengklasifikasian dilakukan dengan menggunakan algoritma Naïve bayes dan Support Vector Machine (SVM). Data diambil dari twitter @e100ss dan @SapawargaSby selama periode 1 September 2015 sampai dengan 13 Oktober 2015, dimana akan dikelompokkan menjadi 3 kelas sentimen yaitu positif, negatif, dan netral. Model klasifikasi terbaik didapatkan dengan menggunakan algoritma Support Vector Machine (SVM) dengan hasil akurasi sebesar 79,81%. Setelah klasifikasi, dibuat web framework pembuatan visualisasi berupa wordcloud dan grafik streamgraph yang ditampilkan secara interaktif dengan aplikasi berbasis web yaitu R Shiny.

Text Mining Pada Media Sosial Twitter Studi Kasus: Masa Tenang Pilkada Dki 2017 Putaran 2, [10]. Penelitian ini menggunakan data sebanyak 20.000 data tweet diambil pada 15-19 April 2017 selama pelaksanaan Pilkada DKI Putaran 2. Kemudian data tersebut direduksi dengan menggunakan langkah preprocessing serta menghapus data dengan nilai TD-TDF yang rendah. Kemudian sentimen diberikan kepada data dengan menghitung jumlah kata positif dan negatif yang telah definisikan oleh peneliti berdasarkan observasi terhadap beberapa sampel data yang diambil secara acak. Penelitian ini menemukan bahwa terdapat ledakan tweet bersentimen negatif pada hari kedua masa tenang. Sedangkan ledakan

selanjutnya terjadi pada hari ketiga, namun pada sentimen positif. Temuan kami yang lain yaitu “ahok” selalu mendapatkan sentimen negatif lebih tinggi dan sentimen positif lebih rendah dari pada tweet “anies”. Hasil temuan ini ternyata sejalan dengan hasil pemilihan, yaitu kemenangan pada pasangan Anies-Sandi. Selain itu pada metode-metode unsupervised learning kami menemukan bahwa metode k-means tidak dapat memberikan hasil yang merata pada setiap kelompoknya. Sebaliknya, hasil luaran dari pemodelan topik (topic modeling – Latent Dirichlet Allocation) lebih merata. Selain itu hasil pengelompokan dari metode k-means dan topic modeling pada data tanggal 18 April 2017 memiliki nilai SSE (k-means) dan marginal likelihood (topic model) yang lebih baik dari pada data tanggal lainnya. Hal ini disebabkan oleh karena keragaman data yang terdapat pada tanggal 18 April 2017 tersebut lebih rendah

“Rancang Bangun Aplikasi Pengaduan Di Kabupaten Tegal Menggunakan Algoritma Bruteforce” [11]. Menurut peneliti Dinas-dinas pelayanan publik kabupaten Tegal dituntut secara profesional melayani masyarakat agar merasa nyaman dan aman. Terdapat beberapa unsur aman dan nyaman diantaranya dalam pengaksesan informasi dan pengaduan kepada pelayanan publik. Pengaduan merupakan salah satu sarana yang disediakan bagi masyarakat untuk menyampaikan masukan implementasi peran serta atau pengawasan yang berbentuk pengaduan, keluhan, sumbangan pikiran, gagasan, serta saran yang bersifat membangun. Pengaduan pada dinas-dinas pelayanan masyarakat di kabupaten Tegal masih menggunakan cara manual yang tidak efektif. Dalam upaya meningkatkan pelayanan dibutuhkan media yang dapat mempermudah masyarakat dalam melakukan pengaduan pada dinas-dinas pelayanan publik kabupaten Tegal. Tujuan dari tugas akhir ini adalah Merancang dan mengimplementasikan aplikasi pengaduan masyarakat berbasis sms gateway dan telegram. Berdasarkan hasil penelitian SMS Gateway Service Gammu dan Telegram berhasil diimplementasikan pada system berbasis website yang dibuat dengan bahasa pemrograman PHP. Sistem yang telah dibuat dapat membantu warga dan pemerintah kabupaten tegal dalam menangani permasalahan-permasalahan yang terjadi pada masyarakat. Permasalahan yang dialami berdasarkan aduan dapat ditangani sesuai dengan tugas dan wewenang karena aduan diteruskan dan ditujukan langsung pada dinas yang terkait pada isi SMS aduan.

B. Dasar Teori

1. Sentimen Analisis

Sentimen Analisis atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Sentimen Analysis dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif [13].

Sentimen Analysis dapat dibedakan berdasarkan sumber datanya, beberapa level yang sering digunakan dalam penelitian Sentimen Analisis adalah Sentimen Analisis pada level dokumen dan Sentimen Analisis pada level kalimat. Berdasarkan level sumber datanya Sentimen Analisis terbagi menjadi 2 kelompok besar yaitu :

- a) Coarse-grained Sentiment Analysis
- b) Fined-grained Sentiment Analysis.

C. Naïve Bayes

Metode ini memandang frekuensi kemunculan fitur ke-j pada dokumen b_i (f_{ij}) yang ingin diklasifikasi sebagai jumlah kemunculan event ataupun fitur independen. Kemunculan fitur f_{ij} pada dokumen b akan dianggap tidak berhubungan ataupun independen terhadap kemunculan fitur yang lain [15]. Berikut merupakan persamaan untuk menghitung nilai probabilitas dengan menggunakan metode *Naïve Bayes*:

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

Penjelasan dari formula tersebut adalah sebagai berikut :

$P(H|E)$: Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis H terjadi jika diberikan bukti (*evidence*) E terjadi.

$P(E|H)$: Probabilitas sebuah bukti E terjadi akan memengaruhi hipotesis H.

$P(H)$: Probabilitas awal (priori) hipotesis H terjadi tanpa memandang bukti apapun.

$P(E)$: Probabilitas awal (priori) bukti E terjadi tanpa memandang hipotesis/bukti yang lain. $\frac{P(E)}{P(H)}$

Ide dasar dari aturan *Bayes* adalah bahwa hasil dari hipotesis atau peristiwa (H) dapat diperkirakan berdasarkan pada beberapa bukti (E) yang diamati. Ada beberapa hal penting dari aturan *Bayes* tersebut, yaitu Sebuah probabilitas awal/prior H atau $P(H)$ adalah probabilitas dari suatu hipotesis sebelum bukti diamati. Sebuah probabilitas akhir H atau $P(H|E)$ adalah probabilitas dari suatu hipotesis setelah bukti diamati

D. Demam *Dengue*

TF-IDF digunakan untuk menentukan nilai frekuensi sebuah kata di dalam sebuah dokumen atau artikel dan juga frekuensi di dalam banyak dokumen. Perhitungan ini menentukan seberapa relevan sebuah kata di dalam sebuah dokumen [6].

Penggunaan *term frequency* dan *document frequency* untuk memberikan peringkat pada dokumen dipelajari secara ekstensif, untuk model ruang vektor. Mengikuti pertimbangan dari model diskriminasi kata, mereka berpendapat bahwa kata yang muncul dalam dokumen harus diberi nilai sebanding dengan frekuensi kata dan terbalik sebanding dengan frekuensi dokumen. Pembobotan skema yang mengikuti pendekatan ini disebut $TF \times IDF$ (*term frequency* \times *inverse document frequency*).

Prosedur dalam implementasi TF-IDF terdapat perbedaan kecil di dalam semua aplikasinya, tetapi pendekatannya kurang lebih sama. TF (*Term Frequency*) merupakan frekuensi sebuah kata terdapat di dalam sebuah dokumen. Nilai dari TF didapat menggunakan persamaan.

$$TF(t) = f_{t,d} / \sum t, d \quad (2)$$

dimana $f_{t,d}$ merupakan frekuensi sebuah kata (t) muncul di dalam dokumen d, sedangkan $\sum t, d$ merupakan total keseluruhan kata yang terdapat di dalam dokumen d.

Kemudian untuk menghitung nilai IDF (*Inverse Document Frequency*) dari sebuah kata di dalam kumpulan dokumen menggunakan persamaan:

$$IDF(t) = \log\left(\frac{|D|}{f_{t,d}}\right) \quad (3)$$

$|D|$ merupakan jumlah dokumen yang ada di dalam koleksi, sedangkan $f_{t,D}$ merupakan jumlah dokumen dimana t muncul di dalam D. Dalam koleksi dokumen D, sebuah kata t dan dokumen individu $d \in D$, dapat dihitung nilai TF-IDF menggunakan rumus:

$$TF - IDF(t) = TF(t) * IDF(t)$$

Diasumsikan bahwa $|D| \sim f_{t,D}$, ukuran dari kumpulan dokumen hampir sama dengan frekuensi t di dalam D. Jika $1 < \log(|D|/f_{t,D}) < c$ untuk sebuah konstanta c dengan nilai yang kecil, maka wd akan lebih kecil daripada $f_{t,d}$ tetapi tetap bernilai positif. Hal berarti bahwa w mempunyai relasi yang biasa dengan seluruh dokumen tetapi tetap menyimpan beberapa informasi yang penting.

III. METODE PENELITIAN

A. Metode Pengumpulan Data

Metode pengumpulan data adalah teknik atau cara yang dilakukan oleh peneliti untuk mengumpulkan data. Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian. Adapun dalam pengumpulan data dilakukan dengan dua metode yaitu metode primer dan metode sekunder.

Metode primer dilakukan dengan cara pengumpulan data secara langsung pada situs instagram akun BPJS Kesehatan dengan mengambil komentar-komentar pada postingan. Sedangkan metode sekunder dilakukan dengan cara mengumpulkan dokumen, informasi dan browsing terkait informasi sosial media instagram dan metode naïve bayes untuk klasifikasi

B. Metode Pengembangan Sistem

a. Definisi Kebutuhan/ Analisis

Ada 2 aktifitas analisis sebagai berikut

i. Analisis kebutuhan

Analisis Kebutuhan yaitu menghasilkan sebagian besar kebutuhan. .

ii. Definisi Kebutuhan

Definisi kebutuhan menghasilkan dokumen kebutuhan

Tahap analisis kebutuhan harus ada informasi tentang fungsi, sifat dan tujuan, kendala yang ada dalam sistem.

Tujuan tahapan sistem untuk mengetahui bagaimana sistem akan dikembangkan.

b. Desain sistem dan software

Tahapan desain sistem dan software yaitu

i. Desain sistem

Desain Sistem adalah menghasilkan konfigurasi sistem keseluruhan dan penjelasan

ii. Desain software

Desain software adalah proses mempunyai fokus pada atribut pemrograman yaitu strukturdata, arsitektur, software, spesifikasi interface dan algoritma yang digunakan.

Tujuan tahapan desain software untuk menerjemahkan kebutuhan menjadi representasi software sebelum dilakukan pemrograman/pengkodean.

iii. Implementasi dan Testing Unit

Tahap desain yang dibuat diterjemahkan dalam bentuk kode program yang dapat dieksekusi/dimengerti oleh mesin. setelah tahap penerjemahan dalam koding program tahap selanjutnya adalah pengujian untuk mencari eror dalam penulisan kode dihasilkan unit program yang valid.

c. Integrasi dan Sistem

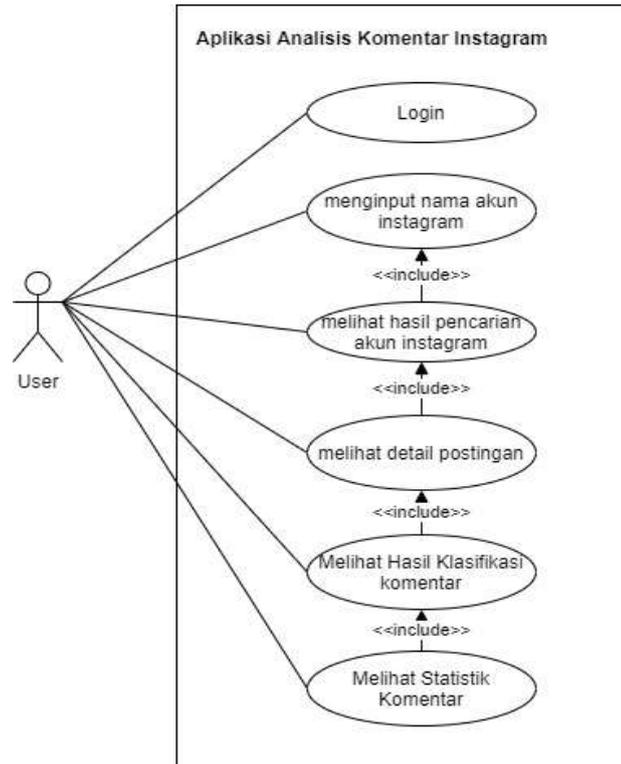
Tahapan yang diterapkan jika sistem terbagi dalam subsub sistem yang lebih kecil.dalam tahapan tersebut unit program yang telah dibuat dan valid akan di integrasikan dengan unit program lainnya kemudian dilakukan pengujian secara keseluruhan sehingga siap digunakan oleh user/dipublikasikan.

d. Operasi Dan Perawatan.

Tujuan perawatan sistem yaitu agar sistem yang telah dikembangkan dapat mengakomodasikan perubahan-perubahan yang terjadi pada lingkungan sistem,sehingga kegiatan operasional dapat berjalan dengan baik, ditahap perawatan fase awal pengembangan sistem diterapkan kembali

C. Use Case Diagram

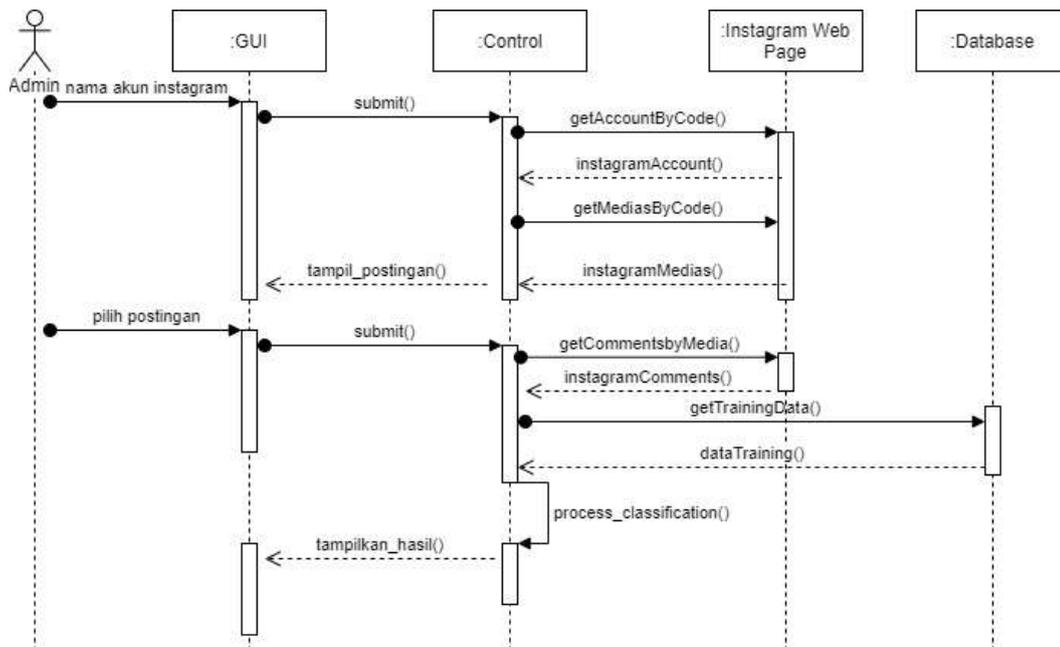
Gambar 1 ditunjukkan usecase diagram aplikasi. Pada usecase diagram terdapat komponen user dan usecase. Usecase diagram pada aplikasi yang diusulkan terdiri dari login, menginput nama akun, melihat hasil pencarian dan melihat detail postingan.



Gambar 1. Use Case Diagram

a. Sequence Diagram

Sequence diagram menjelaskan secara detail fungsi-fungsi yang digunakan dalam aplikasi. Sequence diagram diatas menjelaskan proses klasifikasi yang dimulai dari user menginput nama akun instagram yang ingin dicari. Kemudian proses dimulai dengan memanggil fungsi `getAccountByCode` dan `getMediaByCode` yang mencari data ke instagram dengan cara scrapping. Pemanggilan fungsi tersebut menghasilkan data akun dan postingannya. User lalu memilih postingan yang ingin dilihat. Kemudian aplikasi akan mengambil detail postingan beserta komentar menggunakan fungsi `getCommentsByMedia`. Komentar yang diambil akan diklasifikasi menggunakan algoritma naïve bayes berdasarkan preprocessing TFIDF. Hasil akhir ditampilkan pada halaman detail postingan



Gambar 2. Class Diagram

IV. HASIL DAN PEMBAHASAN

A. Pengujian Sistem

Pengujian pada dasarnya adalah menemukan serta menghilangkan *bug* (kesalahan-kesalahan) yang ada di dalam sistem/perangkat lunak. Metode yang digunakan dalam pengujian validasi pada penelitian ini adalah metode *black box*. Pada pengujian *black box* tidak perlu tahu apa yang sesungguhnya terjadi dalam sistem/perangkat lunak. Di dalam *black box*, item-item yang diuji dianggap gelap karena logikanya tidak diketahui, yang diketahui hanya apa yang masuk dan apa yang keluar dari kotak hitam. Yang diuji adalah masukan serta keluarannya artinya dengan berbagai masukan yang diberikan, apakah sistem/perangkat lunak memberikan keluaran seperti yang diharapkan. *Black box* dapat menemukan kesalahan dalam kategori berikut:

1. Fungsi-fungsi yang tidak benar atau hilang.
2. Kesalahan antar muka sistem.
3. Kesalahan dalam struktur data atau akses basis data eksternal.
4. Inisialisasi dan kesalahan terminasi.
5. Validitas fungsional.
6. Kesensitifan sistem terhadap nilai input tertentu.
7. Batasan dari suatu data.

B. Hasil Implementasi Sistem



Gambar 3. Login

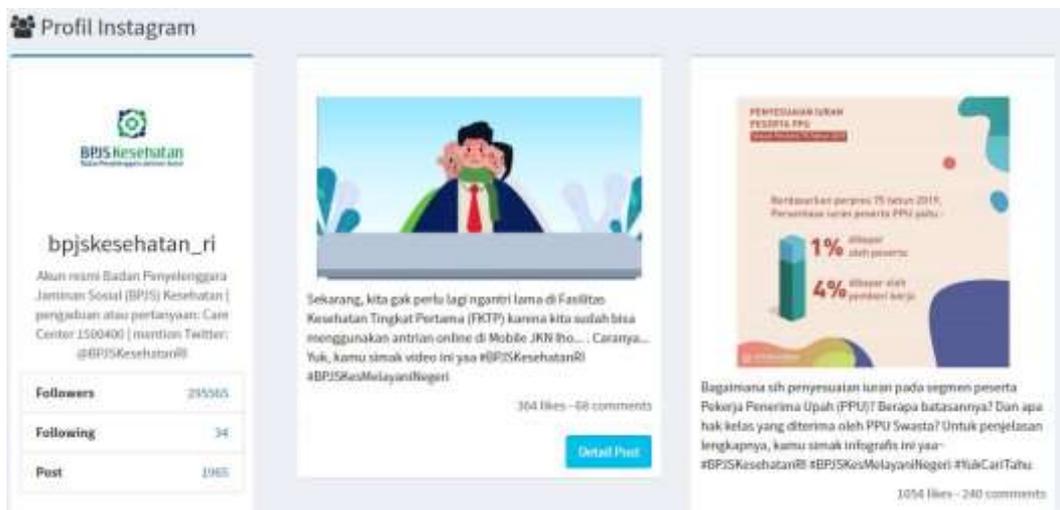
Halaman login digunakan user untuk masuk kedalam aplikasi. Jika memasukan email dan password dengan benar maka aplikasi akan mengarahkan ke halaman beranda admin dan dapat melakukan pengolahan data.



Gambar 4. Beranda

Halaman beranda berisi form pencarian dan data akun yang pernah dicari. Untuk memulai pencarian user menginput nama akun pada form yang telah disediakan. Lalu dapat melakukan submit data dengan menekan tombol enter.

Hasil dari menginput username dari akun instagram yang akan dicari system menampilkan halaman profil. Halaman profil ditunjukkan pada gambar5.



Gambar 5. Feed Instagram

Setelah admin melakukan submit system akan melakukan pencarian data ke instagram berdasarkan nama akun yang diinput. Apabila data ditemukan akan ditampilkan informasi-informasi terkait dengan akun seperti nama akun, caption, follower, follow, jumlah post dan postingan-postingan. Untuk melihat detail postingan user dapat menekan tombol detail post.

System akan menampilkan halaman detail postingan yang dipilih user. Halaman detail posting ditunjukkan pada gambar 6.



Gambar 6. Hasil Klasifikasi

Setelah memilih postingan yang ingin dilihat system akan menampilkan detail postingan. Halaman detail postingan berisi detail informasi terkait posting yang terdiri dari foto dan caption, jumlah like dan komentar. Data komentar dari postingan juga ditampilkan dalam bentuk tabel beserta hasil klasifikasinya. Hasil klasifikasi dirangkum dan ditampilkan dalam bentuk grafik pie.

B. Pengujian

Pengujian fungsional dilakukan untuk memastikan tanggapan/respons atas suatu event atau masukan pada fungsi yang dibuat akan menjelaskan proses yang tepat dan menghasilkan keluaran/output sesuai dengan rancangan yang telah dibuat.

Tabel 1. Tabel Pengujian Blackbox

Input / Event	Output	Hasil
Membuka aplikasi	Halaman Login	Sesuai / OK
Input Email dan Password (Benar)	Halaman beranda aplikasi	Sesuai/ OK
Input Email dan Password (Salah)	Menampilkan halaman login	Sesuai/ OK
Input nama akun	Menampilkan halaman profil	Sesuai/ OK
Klik detail posting	Menampilkan detail posting	Sesuai/ OK
Tambah keyword	Menambahkan keyword kategori	Sesuai/ OK
Klik halaman account	Menampilkan halaman dashboard account	Sesuai/ OK

Pengujian klasifikasi dilakukan untuk mengetahui performa klasifikasi menggunakan naïve bayes yang diusulkan. Performa yang akan dihitung merupakan akurasi sebagai tolak ukur metode klasifikasi. Sampel pengujian ditampilkan pada tabel 2.

Tabel 2.. Pengujian Klasifikasi

Akun	Komentar	Output	Asli	Keterangan
rivai_simbolon	selamat malam @bpjskesehatan_ri masa setiap bulan saya harus seperti ini sih. belum terpotong saldonya lho min.. tolong dibantu pendeбетannya ya.. makasih mba2 cantik / mas2 guanteng	Positive	Negative	FALSE
cak.uus	Slmt siang, bpjskesehatan. Apakah ada perbdaan laynan yg didapat dari pesrta bpjs dengan yg non bpjs ketika berobat di puskesmas	Netral	Netral	TRUE
snaimaaa	sya kok merasa salah satu rakyat yang tdk terbantu adanya bpjs ya min padahal bpjs sya pbi(bantuan pemerintah) tapi tdk bermanfaat karna stiap mau ke puskesmas merka tdk menerima krn faskes di kartu masih alamat kampung halaman sdgkan sy tinggal di kota	Negative	Negative	TRUE
...
bozz.kyle	Just like bpjs making the bill higher 2 times	Netral	Negative	FALSE
sodikun343	Selamat bertugas pa gub kemaren kita poto bersama d sentul.skrng sudah bertugas kmbali selamat jateng gayeng	Positive	Positive	TRUE

$$akurasi = \frac{Jumlah\ Relevan}{Jumlah\ Data} \times 100$$

$$akurasi = \frac{31}{50} \times 100 = 62 \%$$

V. KESIMPULAN

A. Kesimpulan

Berdasarkan hasil penelitian pada akhir laporan penulis dapat mempermudah masyarakat untuk mengidentifikasi sebuah opini atau pendapat tentang akun instagram BPJS Kesehatan, algoritma *Naive Bayes* dapat digunakan untuk melakukan klasifikasi komentar di sebuah postingan instagram BPJS kesehatan yang diambil dari hasil nilai probabilitas *Naive Bayes* terbesar. Berdasarkan hasil pengujian blackbox, system yang dibuat tidak memiliki error dan berjalan sesuai dengan inputnya. Penelitian ini menggunakan data training sebanyak 250 komentar dengan data testing 50 komentar. Dari data testing 50 komentar sebesar 31 menghasilkan klasifikasi yang benar, dan 19 komentar menghasilkan klasifikasi yang salah. Hasil pengujian klasifikasi menunjukkan performa yang cukup dengan hasil akurasi sebesar 62 persen.

B. Saran

Berdasarkan kesimpulan dan analisis laporan, saran dari peneliti untuk penelitian lebih lanjut yaitu Penelitian lebih lanjut dapat menerapkan dan melakukan optimasi terhadap metode klasifikasi. Dalam proses klasifikasi teks terdapat banyak komponen yang dapat dioptimalkan seperti pra pengolahan teks dapat dilakukan stemming. Selain itu dapat menambahkan dataset training yang lebih banyak sehingga memperluas jangkauan klasifikasi. Dikembangkan penelitian yang lebih mendalam dan variasi atau kombinasi algoritma klasifikasi berbasis machine learning

DAFTAR PUSTAKA

- [1] W. Suciska, "OPTIMALISASI PENERAPAN E-GOVERNMENT MELALUI MEDIA SOSIAL DALAM MEWUJUDKAN GOOD GOVERNANCE," in *PROSIDING SEMINAR NASIONAL KOMUNIKASI 2016*, 2016.
- [2] A. Novantirani, "Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine," in *eProceedings of Engineering 2.1*, 2015.
- [3] N. Faradillah, "EKSPERIMEN SISTEM KLASIFIKASI ANALISA SENTIMEN TWITTER PADA AKUN RESMI PEMERINTAH KOTA SURABAYA BERBASIS PEMBELAJARAN MESIN," in *Seminar Nasional Sistem Informasi Indonesia*, 2016.
- [4] W. A. Luqyana, "Analisis Sentimen Cyberbullying Pada Komentar Instagram Dengan Metode Klasifikasi Support Vector Machine," 2018.
- [5] A. . F. Hidayatullah and A. SN, "ANALISIS SENTIMEN DAN KLASIFIKASI KATEGORI TERHADAP TOKOH PUBLIK PADA TWITTER," 2014.
- [6] M. Qasem, R. Thulasiram and P. Thulasiram, "Twitter Sentiment Classification Using Machine Learning Techniques for Stock Markets," in *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2015.
- [7] A. F. Hidayatullah, "Analisis Sentimen Dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter," *Seminar Nasional Informatika 2014*, 2014.
- [8] A. G. Buntoro, "Analisis Sentimen Hatespeech Pada Twitter Dengan Metode Naive Bayes Classifier Dan Support Vector Machine," *Jurnal Dinamika Informatika*, 2016.
- [9] N. Faradillah, R. Kusumawardani and I. Hafidz, "EKSPERIMEN SISTEM KLASIFIKASI ANALISA SENTIMEN TWITTER PADA AKUN RESMI PEMERINTAH KOTA SURABAYA BERBASIS PEMBELAJARAN MESIN," *Seminar Nasional Sistem Informasi Indonesia*, 2016.
- [10] A. F. Hadi, "Text Mining Pada Media Sosial Twitter Studi Kasus: Masa Tenang Pilkada Dki 2017 Putaran 2," *Seminar Nasional Matematika dan Aplikasinya*, 2017.
- [11] S. I. Setiawan, "Rancang Bangun Aplikasi Pengaduan Di Kabupaten Tegal Menggunakan Algoritma Bruteforce," *Skripsi Teknik Informatika UDINUS*, 2017.
- [12] A. Prasetya, "Penerapan Algoritma Boyer-Moore dan Algoritma Rabin- Karp dalam Mendeteksi Aksi Plagiatisme," *Makalah IF3051 Strategi Algoritma*, 2011.
- [13] B. Liu, *Sentiment analysis and opinion mining*, Claypool Publisher, 2012.
- [14] S. M. Weiss, *Text mining: predictive methods for analyzing unstructured information.*, Springer Science & Business Media, 2010.

- [15] X. Ji, S. A. Chun and J. Geller, "Monitoring Public Health Concerns Using Twitter Sentiment Classifications," in *IEEE International Conference on Healthcare Informatics*, 2013.
- [16] A. Nurmandi, Manajemen pelayanan publik, Sinergi Publishing, 2010.
- [17] R. S. Pressman, Software Engineering, Somerville, 2005.
- [18] a. kadir, Konsep Elemen Dasar Sistem., 2003.
- [19] J. HM, Analisis dan Desain Sistem Informasi, Yogyakarta: ANDI, 2006.
- [20] Jogiyanto, Analisis dan Desain Sistem Informasi pendekatan Terstruktur Teori dan Praktik Aplikasi Bisnis, dan E-Commerce, Yogyakarta: Andi Offset, 2005.
- [21] P. Group, "PHP: Hypertext preprocessor," 2006. [Online]. Available: <http://php.net/> .
- [22] S. .. A.S Rossa, Rekayasa Perangkat Lunak, Bandung: Modula, 2011.
- [23] A. Sunyoto, Ajax Membangun Web dengan Teknologi Asynchronous JavaScript&XML, Penerbit Andi, 2007.
- [24] A. Minandar, "Aplikasi Algoritma Pencarian String Boyer-Moore pada Pencocokan DNA," 2010.