

Penerapan Naïve Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen dan Pemerintah

Applications of Naïve Bayes Classifier, K-Nearest Neighbor and Decision Tree to Analyze Sentiment on Netizen and Government Interaction

M. Khairul Anam¹, Bunga Nanti Pikir², Muhammad Bambang Firdaus³, Susi Erlinda⁴, Agustin⁵
^{1,2,4,5}STMIK Amik Riau, Indonesia
³Universitas Mulawarman, Indonesia

Informasi Artikel

Genesis Artikel:

Diterima, 25 Februari 2021
Direvisi, 17 April 2021
Disetujui, 13 November 2021

Kata Kunci:

Komentar Netizen
Analisis Sentimen
Naïve Bayes Classifier (NBC)
K-Nearest Neighbor (KNN)
Decision tree

Keywords:

Netizen Comments
Text Analysis
Sentiment Analysis
Naïve Bayes Classifier (NBC)
Decision tree

ABSTRAK

Pemerintah Pekanbaru saat ini sudah menerapkan teknologi dalam sistem pemerintahan, penerapannya saat ini masih mendapat keluhan dari masyarakat seperti layanan *publik command center* yang hanya sebagian masyarakat mengetahuinya dan penerapan cctv yang ada di Alat Pemberi Isyarat Lalu Lintas (APILL) yang belum berfungsi dengan baik. Penerapan teknologi lainnya oleh Pemerintah Pekanbaru dapat kita lihat dari keberadaan *portal-portal web* situs resmi Pemerintah. Sedangkan untuk melihat beragam komentar netizen dari *twitter*. *Twitter* menjadi tempat untuk mendapatkan data yang diungkapkan masyarakat melalui *tweets* yang *diposting* ke *timeline*. Analisa sentimen dilakukan untuk melihat pendapat atau kecenderungan opini netizen terhadap pemerintah Pekanbaru yang mengandung sentimen positif, negatif, dan netral. Data yang digunakan adalah tweet dengan jumlah dataset sebanyak 150 tweets. Data tersebut kemudian di analisa agar menjadi informasi. Analisa dilakukan menggunakan metode data *mining* yaitu *Naïve Bayes Classifier*, *K-Nearest Neighbor* (KNN), dan *Decision tree*. Penggunaan ketiga pendekatan ini berupaya untuk mengkategorikan hasil komentar netizen terkait penggunaan teknologi yang telah melalui proses analisis sentimen dan membandingkan keakuratan ketiga cara tersebut. Hasil akurasi yang didapatkan cukup beragam yaitu dari metode *Naïve Bayes* akurasi 100%, metode KKN akurasi 98,25%, dan metode *decision tree* akurasi 62,28%.

ABSTRACT

The Pekanbaru government has currently implemented technology in the government system, its application is currently still receiving complaints from the public such as the command center public service which only some people know about it and the application of cctv in the Traffic Signal Signaling Tool (APILL) which is not functioning properly. We can see the application of other technologies by the Pekanbaru Government by the existence of the Government's official website web portals. Meanwhile, to see the various comments from netizens from Twitter. Twitter is a place to get data that is disclosed by the public through tweets posted to the timeline. Sentiment analysis is carried out to see the opinions or tendencies of Netizen opinions towards the Pekanbaru government, which contain positive, negative and neutral sentiments. The data used are tweets with a total dataset of 150 tweets. The data are then analyzed for information. Data mines, such as the Naïve Bayes Classifier, K-Nearest Neighbor (KNN) and Decision Tree, were used to analyze. The employment of these three approaches seeks to categorize the outcomes of netizen comments connected to the use of technology that has gone through a sentiment analysis process and compare the accuracy of the three ways. The obtained accuracy results are very different: the 100 percent accuracy Naïve Bayes method, the 98,25 percent KKN method and a 62,28 percent accurate decision-making tree method.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Korespondensi:

M. Khairul Anam,
Program Studi Teknologi Informasi,
STMIK Amik Riau,
Email: khairulanam@sar.ac.id

1. PENDAHULUAN

Kota Pekanbaru merupakan ibu Kota Provinsi Riau yang terletak di pulau Sumatera [1]. Kota ini merupakan salah satu sentra ekonomi terbesar di bagian Timur Pulau Sumatera sebagai Kota dengan tingkat pertumbuhan, migrasi, dan urbanisasi yang tinggi [2]. Kota Pekanbaru memiliki kebijakan pemerintah untuk perkembangan Kota tersebut. Kebijakan pemerintah yang berjalan pada saat ini tidak luput dari komentar *netizen* di media sosial. Istilah *netizen* dibentuk dari dua kata yaitu internet dan citizen (warga). Jadi, bisa disimpulkan, *Netizen* adalah *user* (pengguna) internet aktif dalam berkomunikasi, mengeluarkan pendapat, dan berkolaborasi di media internet [3]. Pengguna internet melakukan aktivitas pada sebuah media *online* seperti media sosial dan portal web berita. Salah satu dampak positif penggunaan dari internet adalah penggunaan media sosial untuk berinteraksi antara pemimpin dengan warganya [4]. Media sosial menjadi salah satu tempat atau media untuk berekspresi dan berpendapat tentang berbagai macam topik. Sosial media saat ini membuat berbagai macam aspek kehidupan manusia, seperti komunikasi maupun interaksi, mengalami perubahan yang sebelumnya tidak pernah diduga. Dunia seolah-olah tidak memiliki batasan (*borderless*) tidak ada kerahasiaan yang bisa ditutupi [5]. Interaksi *netizen* atau partisipasi *online* di media sosial terhadap pemerintah Kota Pekanbaru mendapat berbagai macam komentar mulai dari pujian, kritik, saran, sindiran, bahkan ujaran kebencian.

Pemerintah Kota Pekanbaru saat ini memiliki pelayanan publik yang berbasis teknologi informasi. Perkembangan teknologi informasi menuntut pemerintah Kota Pekanbaru dalam menyediakan layanan yang lebih cepat, efektif, dan efisien, untuk itu dikembangkanlah sistem pemerintahan berbasis teknologi. Permasalahan yang terjadi dalam penerapan teknologi oleh Pemerintah Kota Pekanbaru saat ini masih banyak mendapat keluhan dari masyarakat seperti layanan publik *command center* yang hanya sebagian masyarakat mengetahui mengenai pelayanan publik ini dan penerapan cctv yang diterapkan di Alat Pemberi Isyarat Lalu Lintas (APILL) yang belum berfungsi dengan baik. Penerapan teknologi oleh pemerintah Kota Pekanbaru dapat kita lihat dari beberapa portal web situs serta portal web resmi lainnya. Sedangkan untuk melihat beragam komentar netizen dilihat dari media sosial twitter. Data pada Twitter berupa tweet yang akan diklasifikasikan menjadi 3 jenis komentar yaitu komentar positif, negatif dan netral. Jenis tweet didapatkan setelah melakukan *crawling* data dengan memanfaatkan API twitter yakni dengan menggunakan analisa sentimen dalam proses klasifikasi komentarnya. Analisa sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini [6]. Dalam penelitian ini analisa sentimen dilakukan untuk melihat pendapat atau kecenderungan opini netizen terhadap pemerintah Kota Pekanbaru yang mengandung sentimen positif, negatif dan netral mengenai penerapan teknologi, kemudian hasil sentimen dianalisa menggunakan metode *Naïve Bayes classifier*, *K-Nearest Neighbor* dan *Decision Tree*.

Naïve Bayes, *K-Nearest Neighbor* (KNN) dan *Decision tree* merupakan algoritma data *mining* dengan metode klasifikasi. *Naïve Bayes* adalah klasifikasi menggunakan metode probabilitas dan statistik [7]. Keuntungan dari penggunaan *Naïve Bayes* adalah metode ini hanya membutuhkan jumlah data pelatihan (*Training Data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses klasifikasi. *Naïve Bayes* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan [8]. Penelitian *Naïve Bayes classifier* bertujuan untuk melakukan proses klasifikasi hasil komentar netizen mengenai penerapan teknologi yang telah melalui proses analisa sentimen. Kemudian metode selain *naïve bayes* yang digunakan adalah *decision Tree*. *Decision tree* merupakan suatu pendekatan yang sangat populer dan praktis dalam *machine learning* untuk menyelesaikan permasalahan klasifikasi [9]. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami [10]. Selanjutnya adalah metode *K-Nearest Neighbor* (KNN) yang sering juga digunakan untuk menganalisa sentimen. Metode KNN adalah proses untuk mengelompokkan data ke dalam kelas-kelas yang telah ditentukan sebelumnya berdasarkan jarak terdekat/tingkat kemiripan data tersebut dengan dataset/data latih yang ada [11].

Banyak metode yang digunakan untuk menerapkan analisa sentimen diantaranya seperti *Naïve Bayes Classifier*, *Support Vector Machine* (SVM), *Neural Network*, *K-Nearest Neighbor*, dan *K-Means*. Berdasarkan dari 4 penelitian sebelumnya mengenai metode klasifikasi untuk analisa sentimen [12–15]. Metode *Naïve Bayes classifier* mempunyai kecepatan dan akurasi yang tinggi ketika diaplikasikan dalam basis data yang besar dan data yang beragam serta kelebihan metode *Naïve Bayes classifier* adalah sederhana dan ketepatan akurasi yang cukup tinggi sebesar 97%, untuk itu penelitian ini menggunakan metode *Naïve Bayes classifier*. Penelitian [16] melakukan analisis sentimen pada ulasan produk menggunakan metode *Naïve Bayes* yang menghasilkan akurasi terbaik yaitu 77,78%. Kemudian penelitian [17] menghasilkan akurasi yang cukup tinggi yakni 98%. Selanjutnya penelitian [18] melakukan analisis sentimen pada *review* restoran di Singapura, menghasilkan akurasi sebesar 70%. Selain *naïve bayes* metode yang umum digunakan dan memiliki akurasi yang cukup tinggi adalah *decision tree*. Penelitian dilakukan dengan perbandingan beberapa metode dan hasilnya *decision tree* yang memiliki akurasi tertinggi yaitu 83,3% [19]. Peneliti lain juga melakukan perbandingan juga dengan beberapa metode seperti *decision tree*, *naïve bayes* dan KNN, dan akurasi tertinggi dari penelitian ini adalah *decision tree* dengan akurasi 96,83% [20], dan selanjutnya penelitian lain juga membandingkan tiga metode, hasilnya juga *decision tree* memiliki akurasi sempurna yakni 100% [21]. Selain *naïve bayes* dan *decision tree*, metode yang sering digunakan lainnya adalah *K-Nearest Neighbor* (KNN). Penelitian [22] melakukan analisis sentimen menggunakan KNN dan mendapatkan akurasi sebesar 79,21%. Penelitian [23] mendapatkan nilai akurasi yang cukup tinggi, yaitu 94,23% menggunakan KNN.

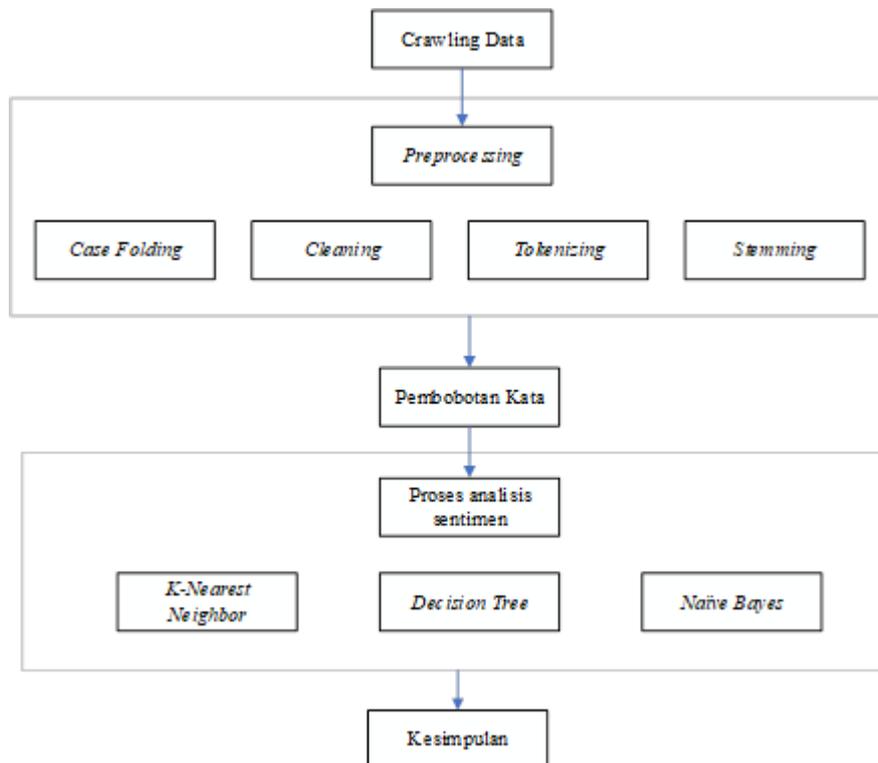
Dari beberapa penelitian, penelitian ini akan melakukan perbandingan sentimen analisis menggunakan metode *Naïve Bayes*, *K-Nearest Neighbor*, dan *decision tree* yang memiliki akurasi cukup tinggi. Penelitian ini merupakan lanjutan dari penelitian terdahulu [4] yang membahas kesiapan warga kota pekanbaru dalam menghadapi penerapan *smart city* di Kota Pekanbaru. Media sosial menjadi hal yang tidak terpisahkan dari sebuah proses dalam menyelenggarakan pemerintahan sekarang [24, 25]. Masyarakat memantau kinerja dan branding yang dilakukan oleh pemerintah melalui jenis media yang satu ini [4, 26]. Tak hanya itu, publik dan pelanggan juga lekat dengan opini di media sosial. Oleh karena itu, penting untuk mengetahui sentimen apa yang dibicarakan publik terhadap penyelenggaraan pemerintahan terlebih dalam implementasi teknologi informasi [27]. Data yang diperoleh untuk diolah pada penelitian ini tidak cukup banyak yaitu 115 tweet yang di *crawling* dari *twitter*. Data yang didapat sedikit jika dibandingkan

dengan penelitian lain seperti [20, 28, 29] yang memiliki data diatas 900 data *tweet*. Penelitian-penelitian tersebut memiliki akurasi yang cukup tinggi yakni diatas 90%. Dengan adanya gap data tersebut, penelitian ini akan mencoba membuktikan terhadap metode yang digunakan dengan data yang relatif sedikit. Hasil akurasi yang didapat bisa menjadi acuan dalam menggunakan metode analisis sentimen dengan data yang kurang dari 200 data sampai diatas 1000.

Dari pemaparan diatas, penelitian ini bertujuan untuk memberikan rekomendasi metode yang digunakan jika mengolah data dengan jumlah sedikit. Selain itu penelitian ini dapat digunakan untuk melihat keluhan ataupun komentar *netizen* mengenai penerapan teknologi oleh pemerintahan kota Pekanbaru. Serta hasil sentimen dapat dimanfaatkan oleh pemerintah Kota Pekanbaru dalam melakukan evaluasi pemerintahan dalam penerapan teknologi kedepannya untuk mendukung penerapan smart city di kota Pekanbaru.

2. METODE PENELITIAN

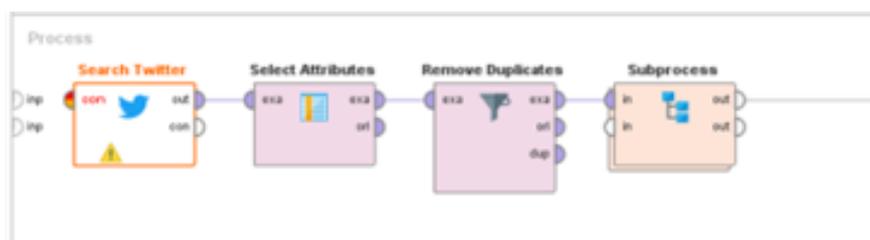
Tahapan proses penelitian yang dilakukan dalam penelitian ini digambarkan dalam sebuah alur metodologi penelitian pada Gambar 1.



Gambar 1. Metodologi penelitian

2.1. Crawling Data

Pada pengambilan data atau yang biasa disebut *crawling* pada penelitian ini menggunakan *rapidminer* yang ditunjukkan pada Gambar 2.



Gambar 2. Halaman Search Twitter

Gambar 2 merupakan tahap mengambil data langsung dari twitter dengan memanfaatkan API twitter dengan menggunakan tools *rapidminer*. Langkah pertama yang dilakukan dalam *crawling* data di twitter adalah *search* twitter. Kemudian untuk menghasilkan data yang normal perlu dilakukan teknik *cleaning* yaitu dengan menggunakan *select atribut*, *remove duplikat* serta subproses. Setelah mendapat data yang sudah di *cleaning* maka perlu kita lakukan tahap *preprocessing* untuk mendapatkan pembobotan kata.

2.2. Preprocessing

Pada proses perhitungan metode *Naïve Bayes* akan dilakukan dengan menggunakan *sample* 4 data, 3 data sebagai data latih dan 1 sebagai data uji. Proses secara menyeluruh yaitu menggunakan 100 data sentimen di twitter akan dilakukan pada bab selanjutnya. Berikut ini adalah contoh langkah-langkah penyelesaian data dengan data 4 dari sentimen di twitter.

Tabel 1. Sampel Data Tweet atau komentar

| No | Komentar | Kelas |
|----|--|---------|
| 1 | @sunggukkusuik buat jd <i>smartcity</i> dari pemerintahan kota maupun kabupatennya hrs melek tentang teknologi dan jg harus mempunyai ide ide kreatif yg ga ketinggalan zaman dan mengajak masyarakatnya berpartisipasi jd ada interaksi dari pemerintah dan masyarakatnya jg jd bs terlaksana #PkuSmartCity | Negatif |
| 2 | @sunggukkusuik menurut saya, pemerintah perlu mengadakan sosialisasi pengenalan <i>smart city</i> , sehingga utk menghadapi dan menerapkan dapat dimulai dari masyarakat yg otomatis akan membentuk konsep <i>smart city</i> dalam pembangunan kota #PkuSmartCity | Positif |
| 3 | @sunggukkusuik Karna terkadang, sebaik apapun tata kelola pemerintah jika tidak ada kesadaran yg tinggi dari masyarakat, rencana tsb tidak akan berjan baik & #PkuSmartCity pun tidak akan terbentuk. | Netral |

Tabel 1 merupakan sampel yang digunakan untuk *preprocessing*. Sampel tersebut diambil pada *hashtag* PkuSmartCity (#PkuSmartCity). *Hashtag* tersebut merupakan salah satu *hashtag* yang digunakan pada penelitian ini terkait dengan penerapan teknologi di kota Pekanbaru yang telah dilakukan proses pelabelan secara manual. Selanjutnya dilakukan proses *Preprocessing*.

1. Case Folding

Case folding merupakan proses untuk mengubah semua teks dokumen menjadi huruf kecil [30]. Pada Tabel 2 merupakan hasil *preprocessing* menggunakan *case folding*.

Tabel 2. Case Folding

| No | Komentar | Hasil Case Folding |
|----|--|--|
| 1 | @sunggukkusuik buat jd <i>smartcity</i> dari pemerintahan kota maupun kabupatennya hrs melek tentang teknologi dan jg harus mempunyai ide ide kreatif yg ga ketinggalan zaman dan mengajak masyarakatnya berpartisipasi jd ada interaksi dari pemerintah dan masyarakatnya jg jd bs terlaksana #PkuSmartCity | @sunggukkusuik buat jd <i>smartcity</i> dari pemerintahan kota maupun kabupatennya hrs melek tentang teknologi dan jg harus mempunyai ide ide kreatif yg ga ketinggalan zaman dan mengajak masyarakatnya berpartisipasi jd ada interaksi dari pemerintah dan masyarakatnya jg jd bs terlaksana #pkusmartcity |
| 2 | @sunggukkusuik menurut saya, pemerintah perlu mengadakan sosialisasi pengenalan <i>smart city</i> shg utk menghadapi dan menerapkan dapat dimulai dari masyarakat yg otomatis akan membentuk konsep <i>smart city</i> dalam pembangunan kota #PkuSmartCity | @sunggukkusuik menurut saya, pemerintah perlu mengadakan sosialisasi pengenalan <i>smart city</i> shg utk menghadapi dan menerapkan dapat dimulai dari masyarakat yg otomatis akan membentuk konsep <i>smart city</i> dalam pembangunan kota #pkusmartcity |
| 3 | @sunggukkusuik Karna terkadang, sebaik apapun tata kelola pemerintah jika tidak ada kesadaran yg tinggi dari masyarakat, rencana tsb tidak akan berjan baik & #PkuSmartCity pun tidak akan terbentuk. | @sunggukkusuik karna terkadang, sebaik apapun tata kelola pemerintah jika tidak ada kesadaran yg tinggi dari masyarakat, rencana tsb tidak akan berjan baik & #pkusmartcity pun tidak akan terbentuk. |

Dari Tabel 2 terdapat perbedaan dimana terlihat bahwa hasil *case folding* ini mengubah huruf kapital menjadi huruf kecil.

2. Cleaning

Cleaning merupakan salah satu metode untuk menghilangkan noise data yang tidak konsisten [31]. Pada tabel 3 merupakan hasil *cleaning* yang telah dilakukan.

Tabel 3. Cleaning

| No | Komentar | Hasil Cleaning |
|----|--|--|
| 1 | @sunggukkusuik buat jd <i>smartcity</i> dari pemerintahan kota maupun kabupatennya hrs melek tentang teknologi dan jg harus mempunyai ide ide kreatif yg ga ketinggalan zaman dan mengajak masyarakatnya berpartisipasi jd ada interaksi dari pemerintah dan masyarakatnya jg jd bs terlaksana #PkuSmartCity | sunggukkusuik buat jd <i>smartcity</i> dari pemerintahan kota maupun kabupatennya hrs melek tentang teknologi dan jg harus mempunyai ide ide kreatif yg ga ketinggalan zaman dan mengajak masyarakatnya berpartisipasi jd ada interaksi dari pemerintah dan masyarakatnya jg jd bs terlaksana pkusmartcity |
| 2 | @sunggukkusuik menurut saya, pemerintah perlu mengadakan sosialisasi pengenalan <i>smart city</i> , shg utk menghadapi dan menerapkan dapat dimulai dari masyarakat yg otomatis akan membentuk konsep <i>smart city</i> dalam pembangunan kota #PkuSmartCity | sunggukkusuik menurut saya pemerintah perlu mengadakan sosialisasi pengenalan <i>smart city</i> shg utk menghadapi dan menerapkan dapat dimulai dari masyarakat yg otomatis akan membentuk konsep <i>smart city</i> dalam pembangunan kota pkusmartcity |
| 3 | @sunggukkusuik Karna terkadang, sebaik apapun tata kelola pemerintah jika tidak ada kesadaran yg tinggi dari masyarakat, rencana tsb tidak akan berjan baik & #PkuSmartCity pun tidak akan terbentuk. | sunggukkusuik karna terkadang sebaik apapun tata kelola pemerintah jika tidak ada kesadaran yg tinggi dari masyarakat, rencana tsb tidak akan berjan baik pkusmartcity pun tidak akan terbentuk |

3. Tokenizing

Tokenizing merupakan salah satu proses pembagian teks atau kata yang berasal dari kalimat atau sebuah paragraf menjadi bagian-bagian tertentu. *Tokenizing* seringkali dipakai dalam ilmu *linguistic* dan hasil tokenisasi berguna untuk analisis teks lebih lanjut [32]. Tabel 4 merupakan hasil *tokenizing* yang telah dilakukan.

Tabel 4. *Tokenizing*

| Kata | TF | Kata | TF | Kata | TF |
|----------------|----|---------------|----|---------------|----|
| akan | 3 | konsep | 1 | Perlu | 1 |
| apapun | 1 | kota | 2 | pkusmartcity | 3 |
| baik | 1 | kreatif | 1 | Rencana | 1 |
| berjalan | 1 | masyarakat | 2 | Saya | 1 |
| berpartisipasi | 1 | masyarakatnya | 2 | Sebaik | 1 |
| buat | 1 | maupun | 1 | Smart | 2 |
| city | 2 | melek | 1 | smartcity | 1 |
| dalam | 1 | membentuk | 1 | sosialisasi | 1 |
| dapat | 1 | mempunyai | 1 | sungguikkusuk | 3 |
| dari | 4 | menerapkan | 1 | tata | 1 |
| dimulai | 1 | mengadakan | 1 | teknologi | 1 |
| harus | 1 | mengajak | 1 | tentang | 1 |
| interaksi | 1 | menghadapi | 1 | terbentuk | 1 |
| jika | 1 | menurut | 1 | terkadang | 1 |
| kabupatennya | 1 | otomatis | 1 | terlaksana | 1 |
| karna | 1 | pembangunan | 1 | tidak | 3 |
| kelola | 1 | pemerintah | 3 | tinggi | 1 |
| kesadaran | 1 | pemerintahan | 1 | zaman | 1 |
| ketinggalan | 1 | pengenalan | 1 | | |

4. Stemming

Stemming merupakan proses untuk mencari *stem* (kata dasar) dari kata hasil *stopword removal (filtering)*. Terdapat dua aturan dalam melakukan *stemming* yaitu dengan pendekatan kamus dan pendekatan aturan [33]. Pada penelitian proses *stemming* menggunakan *stemming nazief-andriani*. Algoritma ini merupakan salah satu cara untuk melakukan *stemming* bahasa Indonesia [34]. Tabel 5 merupakan hasil dari *stemming* yang telah dilakukan.

Tabel 5. *Stemming*

| Kata | TF | Kata | TF | Kata | TF |
|-----------|----|--------------|----|---------------|----|
| ada | 1 | kadang | 1 | rencana | 1 |
| ajak | 1 | karna | 1 | sadar | 1 |
| akan | 3 | kelola | 1 | saya | 1 |
| apapun | 1 | kenal | 1 | sebaik | 1 |
| baik | 1 | konsep | 1 | smart | 2 |
| bangun | 1 | kota | 2 | smartcity | 1 |
| bentuk | 2 | kreatif | 1 | sosialisasi | 1 |
| buat | 1 | laksana | 1 | sungguikkusuk | 3 |
| city | 2 | masyarakat | 4 | tata | 1 |
| dalam | 1 | maupun | 1 | teknologi | 1 |
| dapat | 1 | melek | 1 | tentang | 1 |
| dari | 4 | mulai | 1 | terap | 1 |
| hadapi | 1 | otomatis | 1 | tidak | 3 |
| harus | 1 | partisipasi | 1 | tinggal | 1 |
| interaksi | 1 | pemerintah | 4 | tinggi | 1 |
| jalan | 1 | perlu | 1 | turut | 1 |
| jika | 1 | pkusmartcity | 3 | zaman | 1 |
| kabupaten | 1 | punya | 1 | | |

2.3. Pembobotan Kata

Pembobotan Kata bertujuan untuk memberikan nilai/bobot pada term yang terdapat pada suatu dokumen [35]. Untuk melakukan pembobotan pada penelitian ini menggunakan *Term Frekuensi*. *Term frequency (TF)* adalah metode sederhana dari pembobotan kata. *Term frequency* memperhatikan kemunculan *term* di dalam dokumen [36]. Tabel 6 merupakan hasil pembobotan yang telah dilakukan.

Tabel 6. Pembobotan Kata

| Term (t) | TF | | | |
|--------------|----|---------|---------|--------|
| | TF | Negatif | Positif | Netral |
| ada | 1 | 1 | 0 | 0 |
| ajak | 1 | 1 | 0 | 0 |
| akan | 3 | 0 | 1 | 2 |
| apapun | 1 | 0 | 0 | 1 |
| baik | 2 | 0 | 0 | 2 |
| bangun | 1 | 0 | 1 | 0 |
| bentuk | 2 | 0 | 1 | 1 |
| buat | 1 | 1 | 0 | 0 |
| city | 2 | 0 | 2 | 0 |
| dalam | 1 | 0 | 1 | 0 |
| dapat | 1 | 0 | 1 | 0 |
| dari | 4 | 2 | 1 | 1 |
| hadapi | 1 | 0 | 1 | 0 |
| harus | 1 | 1 | 0 | 0 |
| interaksi | 1 | 1 | 0 | 0 |
| jalan | 1 | 0 | 0 | 1 |
| jika | 1 | 0 | 0 | 1 |
| kabupaten | 1 | 1 | 0 | 0 |
| kadang | 1 | 0 | 0 | 1 |
| karna | 1 | 0 | 0 | 1 |
| kelola | 1 | 0 | 1 | 0 |
| kenal | 1 | 1 | 0 | 0 |
| konsep | 1 | 0 | 0 | 1 |
| kota | 2 | 1 | 1 | 0 |
| kreatif | 1 | 1 | 0 | 0 |
| laksana | 1 | 1 | 0 | 0 |
| masyarakat | 4 | 2 | 1 | 1 |
| maupun | 1 | 1 | 0 | 0 |
| melek | 1 | 1 | 0 | 0 |
| mulai | 1 | 0 | 1 | 0 |
| otomatis | 1 | 0 | 1 | 0 |
| partisipasi | 1 | 1 | 0 | 0 |
| pemerintah | 4 | 2 | 1 | 1 |
| perlu | 1 | 0 | 1 | 0 |
| pkusmartcity | 3 | 1 | 1 | 1 |
| Total | | 20 | 17 | 13 |

2.4. Analisis Sentimen

Proses Analisis sentimen yaitu membandingkan tiga metode yaitu *Naïve Bayes*, KNN, dan *decision tree*. Ketiga metode tersebut menghasilkan akurasi yang berbeda. Langkah pertama untuk menganalisa adalah melakukan *preprocessing* seperti yang dilakukan pada Tabel 6, kemudian memasukan satu persatu metode yang digunakan pada operator *cross validation* pada *rapidminer*.

2.5. Kesimpulan

Pada tahap ini, menyimpulkan hasil dari penelitian yang dilakukan mulai dari tahap *crawling* data sampai didapatnya hasil perhitungan melalui sistem maupun manual.

3. HASIL DAN ANALISIS

Setelah dilakukan tahap *preprocessing*, tahap selanjutnya yaitu melakukan proses klasifikasi secara keseluruhan. Pada proses yang dilakukan akan membandingkan dengan penelitian-penelitian yang memiliki data lebih dari 900 tweet dan memiliki akurasi diatas 90%. Proses yang dilakukan hampir sama tahapannya, tetapi data yang digunakan lebih sedikit dengan peneliti sebelumnya seperti [20, 28, 29]. Berikut ini adalah tahapan penelitian yang dilakukan:

3.1. Proses Klasifikasi

Tahap pertama yaitu memasukan data *training* dan *testing* yang telah dilakukan *preprocessing* sebelumnya. Data yang didapatkan dari twitter terkait dengan interaksi antara netizen dengan pemerintah sekitar 115 data, data ini telah *filter* dari beberapa data tweet, karena banyak tweet yang mengandung pornografi. Dari 115 data, 70% untuk data *training* dan 30% lagi untuk data *testing*. Tabel 7 merupakan sampel yang digunakan.

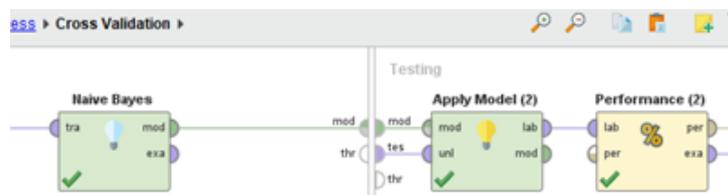
Tabel 7. Sampel data yang digunakan

| Text | Sentiment |
|---|-----------|
| Bagaimana penerapan <i>smartcity</i> yang diterapkan oleh pemerintahan pekanbaru masih under | negatif |
| bagaimana nih menurut kalian warga kota pekanbaru mengenai penggunaan teknologi terhadap pelayanan pemerintahan kota Pekanbaru | netral |
| sticker uji kir nya pun sekarang sudah smart ada qr kodenya Jika di scan akan menunjukkan identitas kendaraan dan pemilik kendaraan itu sendiri | positif |
| menurut aku sih penerapan cctv di setiap lampu merah belum terealisasi dengan baik sepenuhnya karena masih ada cctv dilampu merah belum berfungsi dengan baik mungkin karena tidak ada biaya perawatan cctv dari pemerintah makanya cctv dilampu merah kurang berfungsi | negatif |
| halo Pekanbaru gimana pendapat kalian mengenai penerapan teknologi yg dibuat oleh pemerintah kota pekanbaru penerapan teknologi sprti cctv disetiap lampu merah sudah terealisasi dengan benar atau layanan yg di berikan pemerintah sudah efektif | netral |
| lebih aman dan simple aja sih karna kan ngga harus bawa buku kir lagi, kita udah bisa nunjukin bukti lulus uji elektronik dengan <i>smart card</i> yang kita punya bukti lulus uji elektronik ini juga keihatan lebih modern | positif |
| hallo warga pekanbaru, bagaimana pendapatmu tentang peluncuran sistem baru yakni bukti lulus uji elektronik | netral |
| menurutku lebih bagus sih yaa dapat membantu pemerintah juga dalam menerapkan <i>smart city</i> | positif |
| layanan pemerintah juga belum efektif sampai saat ini karna banyak aja kendala setiap masyarakat mau ngurus ini-itu kuarang dilayanin sama petugasnyalah dan lain sebagainya semoga aja kedepannya pemerintah lebih memperhatikan keluhan warganyaa | negatif |
| hallo man teman warga pekanbaru ada yg tau tntng visi kota pekanbaru yg tinggal di pekanbaru harus tau dong ya | netral |
| semoga dengan adanya konsep <i>smart mobility</i> bisa berdampak tidak ada kemacetan di kota pekanbaru ya | positif |
| wah ini termasuk impian ya semoga saja mulai dengan konsep <i>smart city</i> tentunya untuk kedepannya semoga lebih baik menurutmu sudah siapkah masyarakat pekanbaru sendiri menghadapinya | positif |

Dalam melakukan proses klasifikasi penelitian ini menggunakan *cross validation*. Teknik ini digunakan untuk pembentukan model yang fit dari data *training* sebaik mungkin [32]. Pada perangkat lunak *Rapidminer* dapat dilakukan dengan operator *Cross-Validation* dengan parameter *number of validations* (jumlah validasi) dan *sampling type* (tipe pengambilan sampel). Jumlah validasi x yang diujicobakan pada penelitian ini sebanyak 1-10 kali. Hal ini bertujuan untuk melihat nilai performa terbaik yang diperoleh masing-masing algoritma klasifikasi. Selain menggunakan operator *Cross Validation*, operator yang digunakan adalah *Apply* model yang digunakan untuk menampung data hasil dari data *training* dan data *testing*. Kemudian *Naïve Bayes*, KNN, dan *decision tree* yang digunakan untuk perhitungan metode yang digunakan pada *rapidminer*. Selanjutnya yaitu *Performance* yang digunakan untuk melihat hasil akurasi data dalam bentuk persentase dan dalam bentuk *confidence*.

3.2. Hasil Klasifikasi

Proses analisis sentimen yang pertama dilakukan yaitu menggunakan metode *Naïve Bayes*. *Naïve Bayes* digunakan karena metode ini memiliki akurasi yang cukup tinggi jika dibandingkan dengan metode klasifikasi seperti KNN, C45 dan lain sebagainya untuk analisis sentimen [37–41]. Pada Gambar 3 merupakan proses desain yang dilakukan pada *rapidminer* menggunakan metode *Naïve Bayes*.

Gambar 3. Proses Desain Metode *Naïve Bayes*

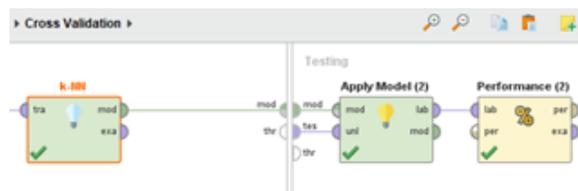
Proses analisa tersebut menghasilkan akurasi yang cukup tinggi yaitu 100%, seperti pada Gambar 4 yang merupakan hasil analisa menggunakan metode *Naïve Bayes*.

accuracy: 100.00%

| | true negative | true neutral | true positive | class pre |
|----------------|---------------|--------------|---------------|-----------|
| pred. negative | 11 | 0 | 0 | 100.00% |
| pred. neutral | 0 | 71 | 0 | 100.00% |
| pred. positive | 0 | 0 | 32 | 100.00% |
| class recall | 100.00% | 100.00% | 100.00% | |

Gambar 4. Hasil Analisa Sentimen Menggunakan *Naïve Bayes*

Naïve Bayes merupakan salah satu metode terbaik untuk melakukan klasifikasi berdasarkan beberapa penelitian yang selalu mendapatkan nilai akurasi diatas 90% [29, 42]. Hal ini juga sejalan dengan penelitian yang kami lakukan, bahwa metode *Naïve Bayes* juga masih konsisten diatas 90% yaitu dengan akurasi 100%. Selain itu untuk data dibawah 200 sangat direkomendasikan menggunakan metode Nave Bayes ini karena memiliki tingkat akurasi yang sangat baik. Berikutnya adalah pengujian menggunakan metode lainnya yaitu menggunakan KNN. Gambar 5 merupakan desain proses menggunakan *rapidminer*.



Gambar 5. Proses Desain Metode KNN

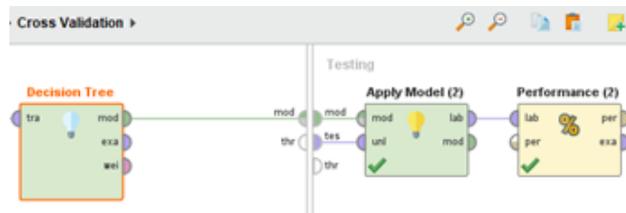
Proses desain pada Gambar 5 didapat akurasi 98,25%. Meskipun hasil akurasi masih dibawah *Naïve Bayes*, tetapi hasil 98,25% sudah cukup tinggi. Hal ini membutuhkan bahwa penggunaan metode KNN bisa menghasilkan akurasi yang cukup tinggi dengan data dibawah 200 data atau tweet. Berikutnya pada gambar 8 merupakan hasil yang didapatkan pada *rapidminer*.

accuracy: 98.25%

| | true negative | true neutral | true positive | class pre |
|----------------|---------------|--------------|---------------|-----------|
| pred. negative | 11 | 1 | 1 | 84.62% |
| pred. neutral | 0 | 70 | 0 | 100.00% |
| pred. positive | 0 | 0 | 31 | 100.00% |
| class recall | 100.00% | 98.59% | 96.88% | |

Gambar 6. Hasil Analisa Sentimen Menggunakan KNN

Kemudian metode yang ketiga adalah menggunakan *decision tree*. *Decision tree* juga merupakan salah satu metode terbaik untuk melakukan klasifikasi. Gambar 7 merupakan desain prosesnya.



Gambar 7. Proses Desain Metode Decision Tree

Desain pada Gambar 7 menghasilkan akurasi sebesar 62,28%. Gambar 10 merupakan hasil yang didapatkan dari *rapidminer*.

accuracy: 62.28%

| | true negative | true neutral | true positive | class pre |
|----------------|---------------|--------------|---------------|-----------|
| pred. negative | 0 | 0 | 0 | 0.00% |
| pred. neutral | 11 | 71 | 32 | 62.28% |
| pred. positive | 0 | 0 | 0 | 0.00% |
| class recall | 0.00% | 100.00% | 0.00% | |

Gambar 8. Hasil Analisa Sentimen Menggunakan *Decision Tree*

Hasil pada Gambar 8 merupakan hasil terendah pada penelitian ini. Namun hasil ini juga sudah cukup tinggi karena di atas 50% tingkat akurasinya. Namun untuk data yang kurang dari 200, metode ini kurang direkomendasikan karena akurasi yang dihasilkan kurang dari 90

Dari ketiga metode yang digunakan memang ada penurunan akurasi yang dihasilkan, hal ini terjadi dikarenakan data yang didapat dari proses *mining* pada twitter kurang dari 200 data. Hal ini menyebabkan pada beberapa metode tidak menghasilkan akurasi yang cukup baik dibandingkan dengan yang lainnya. Akurasi tertinggi didapat dari menggunakan metode *naïve bayes*, yaitu sebesar 100%. Akurasi selanjutnya mengalami sedikit penurunan saat menggunakan metode KNN, akurasi yang didapatkan adalah sebesar 98,25%. Walaupun mengalami penurunan, akurasi ini cukup tinggi jika dibandingkan dengan penelitian terdahulu yang masih dibawah 95% [19, 43]. Kemudian saat menggunakan *decision tree*, mengalami penurunan yang cukup signifikan yakni *decision tree* menghasilkan akurasi sebesar 62,28%.

Hasil yang diperoleh oleh *Naïve Bayes* tidak lepas dari pembobotan kata yang dilakukan pada penelitian ini, pada [17] mengatakan bahwa adanya peningkatan akurasi pada metode *Naïve Bayes* ketika dilakukan pembobotan kata terlebih dahulu. Alasan lainnya adalah data yang digunakan untuk proses klasifikasi sudah melalui tahap *preprocessing*, sehingga datanya menjadi lebih bersih sehingga bisa menaikkan akurasi pada metode *Naïve Bayes* [18].

4. KESIMPULAN

Setelah menyelesaikan serangkaian tahapan terhadap analisa data sentimen pada twitter dengan objek tanda pagar #pemerintahpku, #teknologiku, #smartcitypekanbaru, #pekanbaru, #infopku mengenai penerapan teknologi oleh pemerintah kota pekanbaru, maka dapat diambil beberapa kesimpulan diantaranya adalah (1) Data yang diperoleh dari tweet yang telah diolah menggunakan *sentiment* analisis dengan metode *Naïve Bayes* menghasilkan akurasi sebesar 100%, metode KKN menghasilkan akurasi 98,25%, dan metode *decision tree* menghasilkan akurasi sebesar 62,28%; (2) Data yang diperoleh juga menunjukkan bahwa netizen atau masyarakat belum sepenuhnya peduli dengan penerapan teknologi yang diterapkan pemerintah kota Pekanbaru dilihat dari hasil *sentimen* yang telah dianalisa yang bersifat netral yang mendominasi. Saran pada penelitian selanjutnya yaitu melakukan perbandingan terhadap metode lain seperti SVM yang dikenal juga sebagai metode klasifikasi terbaik. Selain itu data yang di jadikan *training* maupun *testing* disarankan berjumlah lebih dari 1000 data, agar proses klasifikasi menjadi lebih baik. Kemudian untuk menambah akurasi pada setiap metode yang mendapatkan nilai akurasi rendah disarankan menggunakan *feature selection* agar bisa meningkatkan akurasinya.

REFERENSI

- [1] Pirawa, "Profil Kota Pekanbaru Provinsi Riau," 2018. [Online]. Available: <https://www.riaumagz.com/2018/03/profil-kota-pekanbaru-provinsi-riau.html>
- [2] Ruswana, "Titik Promosi Strategis di Kota Pekanbaru," 2020. [Online]. Available: <https://swastikaadvertising.com/titik-promosi-strategis-di-kota-pekanbaru>
- [3] I. Gamayanto, F. E. Nilawati, and Suharnawi, "Pengembangan dan Implementasi dari Wise Netizen (E-Comment) di Indonesia," *Techno.Com*, vol. 16, no. 1, pp. 80–95, feb 2017.
- [4] M. K. Anam, "Analisis Respons Netizen Terhadap Berita Politik Di Media Online," *JURNAL ILMIAH ILMU KOMPUTER*, vol. 3, no. 1, pp. 14–21, apr 2017.
- [5] M. Mulawarman and A. D. Nurfitri, "Perilaku Pengguna Media Sosial beserta Implikasinya Ditinjau dari Perspektif Psikologi Sosial Terapan," *Buletin Psikologi*, vol. 25, no. 1, pp. 36–44, jun 2017.
- [6] I. Rozi, S. Pramono, and E. Dahlan, "Implementasi Opinion Mining (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik Pada Perguruan Tinggi," *Jurnal EECCIS*, vol. 6, no. 1, pp. pp. 37–43, 2012.
- [7] A. Suryanto, I. Alfarobi, T. A. Tutupoly, R. Fauziahti, T. Industri, S. I. Akuntansi, T. Informatika, S. Informasi, N. Bayes, and O. Weights, "Optimasi Naïve Bayes Menggunakan Optimize Weights dan Stratified Pada Data Kredit Operasi," *Jurnal Mantik*, vol. 3, no. 1, pp. 211–219, 2019.
- [8] A. Pattekari, S.A.; Parveen, "Prediction system for heart disease using Naïve Bayes," *International Journal of Advanced Computer and Mathematical Sciences*, vol. 3, no. 3, pp. 290–294, 2012.
- [9] G. Wahyuningtyas, I. Mukhlash, and Soetrisno, "Aplikasi Data Mining untuk Penilaian Kredit Menggunakan Metode Fuzzy Decision Tree," *Jurnal Sains Dan Seni Pomits*, vol. 2, no. 1, pp. 1–6, 2014.
- [10] Y. Sunoto and B. Wasito, "Analisis Testimonial Wisatawan Menggunakan Text Mining Dengan Metode Naïve Bayes Dan Decision Tree , Studi Kasus Pada Hotel Hotel Di Jakarta," *Jurnal Informatika dan Bisnis ANALISIS*, vol. 3, no. 2, pp. 39–49, 2014.
- [11] A. Deviyanto and M. D. R. Wahyudi, "Penerapan Analisis Sentimen pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 3, no. 1, pp. 1–13, dec 2018.

- [12] Y. Cahyono, "Analisis Sentiment pada Sosial Media Twitter Menggunakan Nave Bayes Classifier dengan Feature Selection Particle Swarm Optimization dan Term Frequency," *Jurnal Informatika Universitas Pamulang*, vol. 2, no. 1, pp. 14–19, mar 2017.
- [13] Y. S. Mahardhika and E. Zuliarso, "Analisis Sentimen Terhadap Pemerintahan Joko Widodo Pada Media Sosial Twitter Menggunakan Algoritma Naives Bayes," in *Prosiding SINTAK 2018*, 2018, pp. 409–413.
- [14] F. Nurhuda and S. W. Sihwi, "Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier," *ITSmart: Jurnal Ilmiah Teknologi dan Informasi*, vol. 2, no. 2, pp. 35–42, 2013.
- [15] S. W. U. Vitandy, A. A. Supianto, and F. A. Bachtiar, "Analisis Sentimen Evaluasi Kinerja Dosen menggunakan Term Frequency- Inverse Document Frequency dan Naive Bayes Classifier," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 6, pp. 6080–6088, 2019.
- [16] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 4, no. 2, pp. 113–118, dec 2018.
- [17] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd. Id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 2, no. 2, pp. 681–686, 2019.
- [18] V. A. Permadi, "Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran di Singapura," *Jurnal Buana Informatika*, vol. 11, no. 2, pp. 140–151, oct 2020.
- [19] M. Syarifuddin, "Analisis Sentimen Opini Publik Terhadap Efek Psbb Pada Twitter Dengan Algoritma Decision Tree-Knn-Naive Bayes," *INTI Nusa Mandiri*, vol. 15, no. 1, pp. 87–94, 2020.
- [20] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *Jurnal Informatika Universitas Pamulang*, vol. 5, no. 4, pp. 646–651, dec 2021.
- [21] N. T. Romadloni, I. Santoso, and S. Budilaksono, "Perbandingan Metode Naive Bayes , Knn Dan Decision Tree Terhadap Analisis Sentimen Transportasi Krl," *Jurnal IKRA-ITH Informatika*, vol. 3, no. 2, pp. 1–9, 2019.
- [22] A. Salam, J. Zeniarja, and R. S. U. Khasanah, "Analisis Sentimen Data Komentar Sosial Media Facebook dengan K-Nearest Neighbor (Studi Kasus pada Akun Jasa)," in *Prosiding SINTAK*, 2018, pp. 480–486.
- [23] R. R. A. Siregar, Z. U. Siregar, and R. Arianto, "KLASIFIKASI SENTIMENT ANALYSIS PADA KOMENTAR PESERTA DIKLAT MENGGUNAKAN METODE K-NEAREST NEIGHBOR," *KILAT*, vol. 8, no. 1, pp. 81–92, may 2019.
- [24] H. Juanda, "Media Sosial Sebagai Penyebarluasan Informasi Pemerintah Aceh," *Jurnal Peurawi*, vol. 1, no. 1, pp. 1–22, 2017.
- [25] E. Novianti, A. Ruchiyat Nugraha, L. Komalasari, K. Komariah, and S. Rejeki, "Pemanfaatan Media Sosial Dalam Penyebaran Informasi Program Pemerintah (Studi Kasus Sekretariat Daerah Kabupaten Pangandaran)," *AL MUNIR : Jurnal Komunikasi dan Penyiaran Islam*, vol. 11, no. 1, pp. 48–59, 2020.
- [26] Yovinus, "Peran Komunikasi Publik Media Sosial dalam Implementasi Kebijakan Keterbukaan Informasi Publik Di Kota Bandung," *Jurnal Academia Praja*, vol. 1, no. 1, pp. 185–211, 2018.
- [27] C. M. Zellatifanny, "Respon Pengguna Twitter terhadap Regulasi Pengendalian Akses Ponsel Ilegal melalui Validasi IMEI (Twitter User's Response to Regulation of Contraband Cell Phone Access Control through IMEI Validation)," *JURNAL IPTEKKOM : Jurnal Ilmu Pengetahuan & Teknologi Informasi*, vol. 21, no. 2, p. 139, 2019.
- [28] R. Cahyani, I. S. Rozas, and N. Yalina, "Analisis Sentimen Pada Media Sosial Twitter Terhadap Tokoh Publik Peserta Pilpres 2019," *MATICS: Jurnal Ilmu Komputer dan Teknologi Informasi*, vol. 12, no. 1, pp. 46–54, 2020.
- [29] E. Fitri, "Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine," *Jurnal Transformatika*, vol. 18, no. 1, pp. 71–80, jul 2020.
- [30] P. M. Prihatini, "Implementasi Ekstraksi Fitur Pada Pengolahan Dokumen Berbahasa Indonesia," *Jurnal Matrix*, vol. 6, no. 3, pp. 174–178, 2016.
- [31] R. R. Rerung, "Penerapan Data Mining dengan Memanfaatkan Metode Association Rule untuk Promosi Produk," *Jurnal Teknologi Rekayasa*, vol. 3, no. 1, pp. 89–98, jun 2018.
- [32] A. N. Ulfah and M. K. Anam, "Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM)," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 7, no. 1, pp. 1–10, apr 2020.
- [33] W. A. Luqyana, I. Cholissodin, and R. S. Perdana, "Analisis Sentimen Cyberbullying Pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya*, vol. 2, no. 11, 2018.

- [34] M. Adriani, J. Asian, B. Nazief, S. M. Tahaghoghi, and H. E. Williams, "Stemming Indonesian: A confix-stripping approach," *ACM Transactions on Asian Language Information Processing*, vol. 6, no. 4, pp. 1–33, dec 2007.
- [35] A. Deolika, K. Kusriani, and E. T. Luthfi, "ANALISIS PEMBOBOTAN KATA PADA KLASIFIKASI TEXT MINING," *JURNAL TEKNOLOGI INFORMASI*, vol. 3, no. 2, pp. 179–184, dec 2019.
- [36] A. T. Ni'mah and A. Z. Arifin, "Perbandingan Metode Term Weighting terhadap Hasil Klasifikasi Teks pada Dataset Terjemahan Kitab Hadis," *Rekayasa*, vol. 13, no. 2, pp. 172–180, aug 2020.
- [37] R. R. Waliyansyah and C. Fitriyah, "Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naive Bayes dan k-Nearest Neighbor (k-NN)," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 5, no. 2, pp. 157–163, aug 2019.
- [38] M. A. Maricar and Dian Pramana, "Perbandingan Akurasi Naive Bayes dan K-Nearest Neighbor pada Klasifikasi untuk Mera-malkan Status Pekerjaan Alumni ITB STIKOM Bali," *Jurnal Sistem dan Informatika (JSI)*, vol. 14, no. 1, pp. 16–22, nov 2019.
- [39] R. N. Devita, H. W. Herwanto, and A. P. Wibawa, "Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa Indonesia," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 4, pp. 427–434, oct 2018.
- [40] N. A. Setifani, D. N. Fitriana, and A. Yusuf, "Perbandingan Algoritma Naive Bayes, Svm, Dan Decision Tree Untuk Klasifikasi Sms Spam," *JUSIM (Jurnal Sistem Informasi Musirawas)*, vol. 5, no. 02, pp. 153–160, 2020.
- [41] Y. Asri, "Analisa Perbandingan Keputusan Metode Klasifikasi Decision Tree dan Naive Bayes Dalam Penentuan Diagnosa Hipertensi," *Jurnal Kajian Ilmu dan Teknologi (KILAT)*, vol. 4, no. 1, pp. 41–46, 2015.
- [42] A. M. Raya, F. Nurbaiti, and D. Sofia, "Klasifikasi Sentimen Masyarakat Terhadap Kenaikan Harga Tiket Pesawat Pada Twitter Menggunakan Naive Bayes," *Instek*, vol. 4, no. 2, pp. 239–246, 2019.
- [43] S. Saidah and J. Mayary, "Analisis Sentimen Pengguna Twitter Terhadap Dompot Elektronik Dengan Metode Lexicon Based Dan K Nearest Neighbor," *Jurnal Ilmiah Informatika Komputer*, vol. 25, no. 1, pp. 1–17, 2020.

