

Analisis Sentimen Aplikasi Identitas Kependudukan Digital (IKD) Menggunakan Metode Naïve Bayes

Rahmat Hidayat¹, Rohim Nur Rahman², Muhammad Reifin Perdana³, Arbansyah⁴

¹⁻⁴ Prodi Teknik Informatika, Fakultas Sains dan Teknologi,
Universitas Muhammadiyah Kalimantan Timur

Korespondensi penulis: 2011102441192@umkt.ac.id

Abstract. *Digital Population Identity (IKD) is a digital-based location data innovation through a mobile application with a photo or QR Code. The government's objective is to reduce the physical prints of KTP as well as the use of blank KTP-el in the hope of administrative efficiency. ICT is integrated with health services, education, banking, and taxation, facilitating public access in the era of technological development. However, in remote areas, limited internet access and minimal socialization raise concerns about the security of digital identity data being considered. Social media, especially YouTube, is a channel platform used by the public to convey opinions, opinions and comments about ICTs. So that's why sentimental analysis is needed using the Naive Bayes algorithm to help understand public opinion. The tests were conducted using Orange on 1,561 data showing accuracy, precision, recall, and F1 above 90%. The results of this analysis can serve as a guide for staff in interacting with the community for the implementation of Digital KTP through IKD, as well as improving services regarding the applications provided.*

Keywords: *Sentiment Analysis, Digital Population Identity, KTP, Naïve Bayes*

Abstrak. Identitas Kependudukan Digital (IKD) adalah inovasi data kependudukan berbasis digital melalui aplikasi mobile dengan foto atau QR Code. Tujuan pemerintah adalah mengurangi cetakan fisik KTP serta penggunaan blanko KTP-el dengan harapan efisiensi administrasi. IKD terintegrasi dengan layanan kesehatan, pendidikan, perbankan, dan pajak, memudahkan akses masyarakat pada era perkembangan teknologi. Namun, di daerah terpencil, akses internet terbatas dan minimnya sosialisasi menimbulkan kekhawatiran terhadap keamanan data identitas digital yang menjadi pertimbangan. Media sosial, terutama YouTube, menjadi platform saluran yang digunakan masyarakat dalam menyampaikan pendapat, opini, dan komentar mengenai IKD. Maka dari itu analisis sentimen dibutuhkan dengan menggunakan algoritma Naive Bayes dalam membantu memahami pandangan publik. Pengujian dilakukan menggunakan Orange pada 1.561 data menunjukkan akurasi, presisi, recall, dan F1 di atas 90%. Hasil analisis ini dapat menjadi panduan bagi pegawai dalam berinteraksi dengan masyarakat untuk penerapan KTP Digital melalui IKD, serta meningkatkan pelayanan mengenai aplikasi yang disediakan.

Kata kunci: Analisis Sentimen, Identitas Kependudukan Digital, KTP, Naïve Bayes

LATAR BELAKANG

Pelayanan publik saat ini mengalami perkembangan yang pesat tidak hanya dilayani secara manual namun ke pelayanan online (Prasetyo, Zakaria, and Wiliantoro 2022). Sejalan dengan kemajuan teknologi, pemerintah menyediakan pelayanan yang cepat dan efisien kepada seluruh warga negara, yang pada gilirannya meningkatkan pengelolaan data. Untuk mempercepat proses pendokumentasian kependudukan, pemerintah menciptakan inovasi baru yang bertujuan meningkatkan pengurusan dokumen setiap tahunnya. Salah satu inovasi yang dikembangkan untuk digitalisasi dokumen kependudukan adalah Identitas Kependudukan Digital (IKD) (Wahyuningsih 2023).

Identitas Kependudukan Digital (IKD) merupakan data kependudukan berbasis digital yang diperkenalkan sebagai aplikasi perangkat seluler dalam bentuk foto atau kode QR. Pemerintah berharap melalui penerapan Identitas Kependudukan Digital (IKD), dokumen kependudukan tidak perlu lagi dicetak atau disimpan dalam bentuk fisik dokumen Kartu Tanda Penduduk (KTP) untuk keperluan administrasi, dan juga menghemat anggaran untuk menggunakan identitas elektronik. KTP kosong. Aplikasi tersebut menghubungkan dan mengintegrasikan layanan kesehatan, pendidikan, pelayanan sosial, perbankan, perpajakan, payment gateway, dan lain-lain dalam satu aktivasi yang termasuk dalam fungsi IKD, sehingga memudahkan masyarakat dalam menggunakannya di era perkembangan teknologi.

Meskipun inovasi ini diharapkan dapat meningkatkan efisiensi dan kualitas pengelolaan data kependudukan, terdapat perdebatan di masyarakat terkait implementasi dan keamanan Identitas Kependudukan Digital (IKD). Beberapa masyarakat, terutama di daerah pelosok, merasa khawatir karena keterbatasan akses internet dan minimnya sosialisasi mengenai aplikasi ini. Keraguan juga timbul terkait keamanan data identitas dalam bentuk digital. Dalam konteks ini, media sosial, khususnya Twitter, menjadi saluran yang digunakan masyarakat untuk menyampaikan pendapat, opini, dan komentar terkait Identitas Kependudukan Digital (IKD).

Analisis sentimen merupakan bagian dari text mining yang bertujuan untuk mengklasifikasi dokumen teks berupa opini sehingga menghasilkan suatu informasi sentimen yang dapat bermakna positif maupun negatif (Ardianto et al. 2020). Metode analisis sentimen dapat digunakan untuk menganalisa pendapat maupun emosi seseorang dalam menyikapi sesuatu yang menyangkut topik tertentu (Cendana and Permana 2019). misalnya sentimen masyarakat terhadap pariwisata Lombok sehingga informasi mengenai kepuasan wisatawan dapat diketahui. Hasil analisis sentimen tersebut dapat dimanfaatkan oleh pemerintah daerah, pelaku wisata maupun stakeholder sebagai pendukung pengambilan keputusan dalam pengembangan dan pengelolaan pariwisata Lombok kedepannya (Putu, Ahmad Zuli Amrullah, and Ismarmiaty 2021).

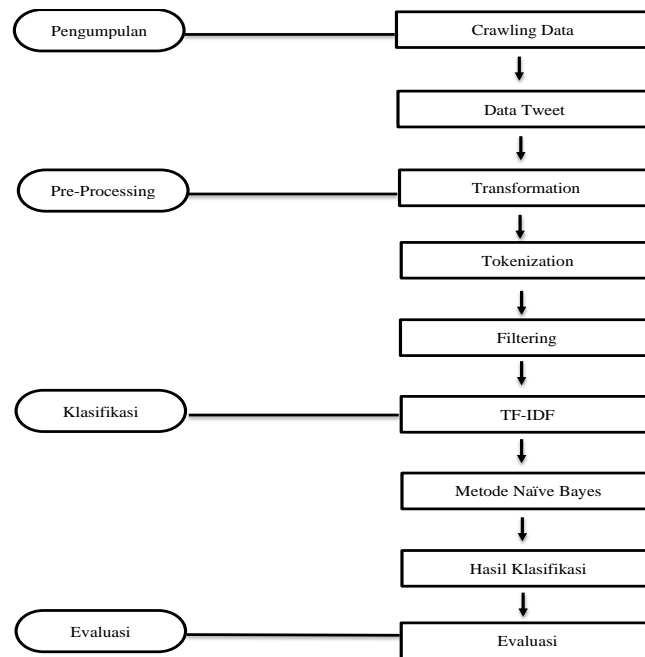
Berdasarkan penelitian yang telah dilakukan, banyak peneliti melakukan eksplorasi data teks melalui klasifikasi untuk menghasilkan informasi yang penting seperti penelitian yang dilakukan tentang analisis sentimen untuk penilaian tempat tujuan wisata kota Tegal menggunakan metode naive bayes dan decision tree (Somantri and Dairoh 2019). Penelitian ini bertujuan untuk mendapatkan model terbaik untuk diimplementasikan pada sistem. Hasil penelitian menunjukkan bahwa nilai akurasi naive bayes sebesar 77,50% lebih baik dibandingkan dengan decision tree dengan nilai akurasi sebesar 60,83%.

Penelitian lainnya ada juga menganalisis ulasan masyarakat tentang pariwisata kota Malang melalui analisis sentimen. Pada penelitian ini menggunakan metode naive bayes untuk klasifikasi teks dan metode query expansion ranking untuk mengurangi jumlah fitur pada proses klasifikasi. Hasil penelitian menghasilkan akurasi sebesar 86,6% (Fanissa, Fauzi, and Adinugroho 2018). Selanjutnya melakukan analisis sentimen twitter tentang pariwisata Lombok. Penelitian ini menggunakan metode naive bayes serta mutual information untuk seleksi fitur dengan nilai akurasi sebesar 97,9% (Ulfa, Irmawati, and Husodo 2018).

Hasil dari analisis sentimen ini bertujuan untuk menggunakan metode Naive Bayes sebagai metode klasifikasi analisis sentimen dan mengetahui nilai persentase dari analisis sentimen pendapat positif, negatif dan netral masyarakat Indonesia terhadap penggunaan aplikasi identitas kependudukan digital (IKD) untuk mengetahui besarnya dampak pengaruhnya terhadap masyarakat lain.

METODE PENELITIAN

Penelitian ini melakukan analisis sentimen terhadap Aplikasi Identitas Kependudukan Daerah (IKD). Pada Media Sosial melalui platform media sosial Twitter. Metode yang digunakan untuk analisis sentimen adalah algoritma Naive Bayes. Metode Naive Bayes adalah metode yang bisa memprediksi peluang terjadinya kejadian di masa depan berdasarkan data yang ada sebelumnya (Hakimi 2018). Kelebihan metode Naive Bayes ini adalah tingkat akurasinya yang tinggi dan juga waktu komputasinya yang lebih cepat. Metode ini dirasa cocok digunakan pada analisis sentimen dikarenakan metode ini bertujuan sebagai metode klasifikasi kedalam kategori positif, negatif maupun netral (Arvyantomo and Ratama 2023). Kajian ini melibatkan serangkaian langkah termasuk rencana untuk melakukan proses validasi dan evaluasi terhadap data yang dikumpulkan. Adapun tahapan penelitian yang dilakukan dalam analisis sentimen adalah sebagai berikut :



Gambaran Flowchart Penelitian

Pengumpulan Data

Pengumpulan data pada penelitian ini didapatkan dari teks komentar dalam bahasa Indonesia yang mengandung opini. Data tersebut diperoleh dari data komentar yang di-crawling dari kolom komentar pada video terkait aplikasi identitas kependudukan digital (IKD). Komentar diunduh dari kolom komentar video yang diunggah dari Januari hingga Desember 2023, sebanyak 1561 komentar. Crawling menggunakan algoritma yang dikembangkan peneliti dalam bahasa pemrograman Python beserta Application Programming Integration (API) Youtube untuk mengumpulkan data.

Pre-processing

Pre-processing adalah tahapan yang paling penting dalam proses text mining, tahapan ini adalah tahap pertama dalam pemrosesan text. Pada tahapan ini text akan dinormalisasikan sehingga nantinya dapat memudahkan pada tahap pembobotan. Pre-processing dilakukan untuk menghindari data yang kurang sempurna atau cacat, atau data-data yang tidak konsisten (Muzaki and Witanti 2021). Data dari Twitter yang telah dikumpulkan pada dasarnya adalah data mentah yang belum siap untuk dianalisis. Oleh karena itu dilakukan langkah pre-processing untuk memperoleh data yang siap dianalisis. Tahapan textpre-processingPada penelitian ini diantaranya.

- a) Transformasi adalah langkah kritis dalam analisis sentimen yang melibatkan manipulasi data untuk memastikan bahwa struktur dan distribusinya mendukung keperluan analisis tersebut. Dalam konteks ini, teknik seperti normalisasi atau

transformasi logaritmik dapat diterapkan untuk mengubah skala atau distribusi data. Tujuan utamanya adalah menyelaraskan data sehingga dapat memberikan hasil analisis sentimen yang lebih akurat dan relevan. Transformasi menjadi kunci untuk memastikan bahwa data yang digunakan dalam analisis sentimen dapat diinterpretasikan dan digunakan dengan efektif.

- b) Tokenizing adalah proses yang memisahkan data teks menjadi beberapa token. Dengan garis besar, tokenizing melibatkan pemecahan suatu teks menjadi kata-kata atau unit-unit lainnya. Penting untuk membedakan karakter-karakter tertentu yang dapat berfungsi sebagai pemisah kata atau bukan. Pemisahan ini menjadi langkah kunci dalam mempersiapkan data teks untuk analisis lebih lanjut.
- c) Filtering adalah tahapan dalam proses analisis teks yang bertujuan untuk menghilangkan kata-kata yang dianggap tidak penting dari hasil tokenizing sebelumnya. Pada tahap ini, kata-kata tersebut akan diidentifikasi dan dieliminasi berdasarkan jenis kata tertentu, seperti kata keterangan, kata seru, kata ganti, kata depan, dan kata hubung. Filtering menjadi langkah penting dalam menyederhanakan data teks dan fokus pada kata-kata yang memiliki kontribusi lebih besar dalam analisis sentimen. Selain itu juga dilakukan penghapusan tanda baca dengan menggunakan fungsi regex yang bisa menghilangkan tanda baca seperti berikut “\ . | , | : ; ! \ ? \ (\) \ | \ + | ' | " | ' | " | ' | \ ' | ... \ - | - | — \ \$ | & \ * | > | <” (Indah Nurhafida and Sembiring 2021).

Pembobotan Kata(TF-IDF)

Data hasil preprocessing yang berupa kata akan diubah ke dalam bentuk angka dengan dilakukan proses pembobotan kata yang bertujuan untuk menghitung bobot pada masing-masing kata yang akan digunakan sebagai fitur, semakin banyak dokumen yang akan diproses maka semakin banyak fitur. Pada tahapan ini terdapat dua bagian proses yaitu TF (Term Frequency) dan IDF (Inverse Document Frequency), TF adalah jumlah kemunculan tiap kata pada sebuah dokumen semakin banyak kata muncul pada tiap dokumen maka semakin besar nilai TF. IDF adalah jumlah nilai dokumen pada tiap kata yang berbanding terbalik yaitu apabila suatu kata jarang muncul pada sebuah dokumen maka nilai IDF lebih besar daripada kata yang sering muncul (Septian, Fachrudin, and Nugroho 2019).

Algoritma Naïve Bayes Classifier

Naïve bayes adalah salah satu algoritma yang populer digunakan untuk keperluan data mining karena kemudahan penggunaannya serta pemrosesan yang cepat, mudah diimplementasikan dengan strukturnya yang cukup sederhana dan tingkat efektifitas yang tinggi (Khotimah and Utami 2022). Algoritma Naive Bayes relatif sederhana dan cepat. Ini membuatnya mudah diimplementasikan dan efisien dalam hal waktu komputasi, terutama pada dataset besar.

Evaluasi

Pada tahap ini dilakukan untuk mengevaluasi mengenai hasil dan nilai performa model dalam penggunaan metode klasifikasi yang diterapkan, adapun evaluasi model yang digunakan yaitu dengan melakukan proses perhitungan pada nilai akurasi, presisi, f1-score, dan recall pada hasil model yang digunakan (Atmadja 2022).

Analisa Sentimen

Sentimen analisis merupakan sebuah teknik menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi mengenai layanan, produk, individu atau kegiatan tertentu. Tujuan dari analisis sentimen untuk menentukan opini dari seseorang dengan memperhatikan topik tertentu. Langkah yang dilakukan dalam analisis sentimen adalah mendefinisikan domain dataset, preprocessing, featured selection, pelabelan, klasifikasi dan evaluasi (Suryani, Linawati, and Saputra 2019).

HASIL DAN PEMBAHASAN

Skenario Penelitian

Berikut merupakan alur pengujian Algoritma *Naïve Bayes* menggunakan tools *Orange Data Mining* yang dapat dilihat pada gambar berikut :



Gambar 1. Skenario Penelitian menggunakan tools *Orange Data Mining*

Penelitian dilakukan dengan memanfaatkan *Tools Orange Data Mining* dalam melakukan klasifikasi *Naïve Bayes* terhadap data yang digunakan.

Proses Crawling Data

Penarikan data Youtube *coment* dilakukan melalui *Google Colab* menggunakan bahasa pemrograman *Python* dengan memanfaatkan *library “discovery.build”*, yang kemudian membutuhkan *access* terhadap Youtube *API* dalam melakukan pengambilan data.

title	publishedAt	authorDisplayName	textDisplay	ReCount
1	2023-01-29T09:...	@ImamWaryNahwanda	Untuk keamanan data masyarakat, scan QR Code untuk aktivasi IKD dilakukan secara tatap muka dengan operator di dukcapil dan dilakukan hanya sekali saj...	55
2	2023-02-04T13:...	@kicadangan138	minta code QR dukcapil bang, malas ke kantor dukcapil	5
3	2023-02-05T01:...	@rohimaheem3726	Bisa lihat akte kelahiran gak ya	2
4	2023-02-06T01:...	@arecom46527	Masih harus ke kantor capil, Hm cape deh.	1
5	2023-02-06T03:...	@iamsyriaa256	Saya harus ke dukcapil bdg, katanya harus ke dukcapil domisili dulu	2
6	2023-02-07T02:...	@rickianar2757	@@iamseptianak256 waduh tetap harus ke dukcapil domisili ya om	0
7	2023-12-19T13:...	@sianfanda533	Digitalisasi 🇲🇵	0
8	2023-12-12T02:...	@hukanglauskabumi	Apakah untuk KTP atau kk itu di aplikasinya bisa di cetak sendiri bang	0
9	2023-12-05T05:...	@nofiad6123	Ibu sama aja kita ke kantor dukcapil juga, katanya ngk perlu kesana bisa di rumah dll, ah ay	0
10	2023-12-02T13:...	@user-g9fo4c1v	Langsung berikan blanko ktp digitalnya	0
11	2023-12-01T06:...	@fefahendra7525	Big gimna kalo mau liat ktp kita sendiri kan udh terdptar. Dna ny. Tpi ktp kita yg udh jadi foto ny rusak gk klihatan lagi. Maka ny mau liat yg dna ny. Kih tau c...	0
12	2023-11-24T05:...	@pengrajinalmunam657	Ko ktp hilang untuk mengetahui nik gmn bang.	0
13	2023-11-24T07:...	@ImamWaryNahwanda	Datang ke Dukcapil untuk dicek biometriknya	0
14	2023-11-19T14:...	@muhammadarifuddin3190	Bang mau tanya, apakah bisa buat akta kelahiran online di aplikasi ini?	1
15	2023-11-19T15:...	@ImamWaryNahwanda	Bisa	0
16	2023-11-17T01:...	@HikmahMaulana2710	ribet kali harus datang ke dukcapil.. wasting time.	0
17	2023-11-17T02:...	@ImamWaryNahwanda	ketika saatnya wajib juga pasti akan diurus juga..	0
18	2023-11-15T07:...	@kamikisan4678	Apakah bisa verifikasi qr code di kelurahan setempat	1
19	2023-11-15T15:...	@ImamWaryNahwanda	Bisa jika kelurahan di daerah anda diberikan akses oleh dukcapil	1
20	2023-11-15T02:...	@user-vm1ro9bhbo	1 aplikasi bisa di pakai berapa aktifitas kak.kalau saya daftar terus istri saya gak itu gimana solusinya	1
21	2023-11-15T02:...	@ImamWaryNahwanda	Setiap orang memiliki akun masing-masing, bisa 1 HP dua orang tapi logout dulu	0
22	2023-11-15T02:...	@user-vm1ro9bhbo	Tidak sama seperti JKN ya kak.kalau JKN satu akun sudah ada semua.katu sama nama masing masing	0
23	2023-11-15T02:...	@user-vm1ro9bhbo	Jadi enak kita tinggal tunjukkan saja.	0
24	2023-11-15T02:...	@ImamWaryNahwanda	@user-vm1ro9bhbo beda	0
25	2023-11-15T03:...	@ImamWaryNahwanda	JKN juga setiap orang juga bisa beda akun, kecuali anak anak.	0
26	2023-11-14T08:...	@deasabrina4718	Ka maaf ltn tanya kalo masalahnya ktp/kk kita rusak bisa di unduh gak sih sofflennya biar cetak sendiri ??	1
27	2023-11-14T11:...	@ImamWaryNahwanda	Bisa, sudah sy jelaskan pada video lain, cek playlist DURCAPIL pada YouTube saya disini 	0
28	2023-11-14T14:...	@deasabrina4718	@@ImamWaryNahwanda makasih ka jawabannya 🙏.	0

Gambar 2. Hasil Crawling Data

Pada gambar 2 diatas merupakan sebuah kumpulan data yang diperoleh pada proses *crawling* data tentang “Identitas Kependudukan Digital”, dimana data tersebut diambil dari seluruh *comment* dari salah satu video Youtube yang berjumlah 1561. Data yang sudah diperoleh merupakan data yang kotor dan perlu dilakukan *preprocessing text* dalam melakukan pembersihan data.

Preprocessing Text

Seluruh data yang sudah didapatkan melalui *crawling* data akan disimpan menjadi sebuah file csv, yang kemudian akan dilakukan sebuah process pembersihan data sebelum lanjut ke tahap selanjutnya. Tahapan *preprocessing text* dilakukan dengan menggunakan *tools Orange Data Mining*, berikut adalah hasil dari tampilan data setelah dilakukan *preprocessing text*.



Gambar 3. Tampilan *wordcloud* hasil *preprocessing* data

Gambar diatas merupakan kumpulan data hasil preprocessing, dengan melakukan pemebersihan terhadap karatker-karakter yang tidak digunakan seperti URL, @, #, dan symbol yang tidak memiliki kegunaan dalam proses selanjutnya.

Labelling Data

Sebelum dilakukan pelabelan dan klasifikasi data, nilai sentimen terhadap data harus diketahui terlebih dahulu. Dalam menentukan sentimen tersebut dapat menggunakan *tools* yang ada pada *Orange Data Mining*. Berikut adalah gambar data hasil dari pemberian nilai sentimen dan labelling data.

compound	textDisplay	positive	negative	neutral
1	netral	0.0	0.0	1.0
2	netral	0.0	0.0	1.0
3	netral	0.0	0.0	1.0
4	netral	0.0	0.0	1.0
5	netral	0.0	0.0	1.0
6	netral	0.0	0.0	1.0
7	netral	0.0	0.0	1.0
8	positif	0.172	0.0	0.828
9	netral	0.0	0.0	1.0
10	netral	0.0	0.0	1.0
11	netral	0.0	0.0	1.0
12	netral	0.0	0.0	1.0
13	netral	0.0	0.0	1.0
14	netral	0.0	0.0	1.0
15	netral	0.0	0.0	1.0
16	negatif	0.0	0.278	0.722
17	netral	0.0	0.0	1.0
18	netral	0.0	0.0	1.0
19	netral	0.0	0.0	1.0
20	netral	0.0	0.0	1.0
21	netral	0.0	0.0	1.0
22	netral	0.0	0.0	1.0
23	netral	0.0	0.0	1.0
24	netral	0.0	0.0	1.0
25	netral	0.0	0.0	1.0
26	netral	0.0	0.0	1.0
27	netral	0.0	0.0	1.0
28	netral	0.0	0.0	1.0
29	netral	0.0	0.0	1.0

Gambar 4. Tampilan data hasil pelabelan nilai sentimen

Hasil yang didapatkan pada gambar diatas diperoleh dari pemberian nilai sentimen menggunakan metode vader yang terdapat pada *tools Orange Data Mining* untuk menghasilkan *feature/class* sentimen berupa *positive*, *negative*, *neutral*, dan *compound* sebagai target hasil pengumpulan nilai yang terdapat pada tiap fitur.

Pembobotan TF-IDF

Pada tahapan ini dilakukan pemberian terhadap nilai dari masing-masing teks dalam menentukan total dari nilai akhir sentimen, tahapan ini dilakukan menggunakan *tools Orange Data Mining* yang memiliki kesamaan fungsi dalam memberikan nilai bobot, berikut adalah hasil dari pemberian bobot.

row	compound	textDisplay	(...)
1	netral	Untuk keamanan data masyarakat, scan QR Code untuk aktivasi IKD dilakukan.	neutral=1.0, aktivasi=5.70654, akun=2.6526, code=3.44106, data=2.50105, dukcapil=1.53895, hilang=3.59188, ...
2	netral	minta code QR dukcapil bang, malas ke kantor dukcapil	neutral=1.0, code=3.44106, dukcapil=3.0779, kantor=3.61541, qr=2.19979
3	netral	Bisa lihat akte kelahiran gak ya	neutral=1.0, kelahiran=4.50049
4	netral	Masih harus ke kantor capil, Hm cape deh.	neutral=1.0, capil=4.09499, kantor=3.61541
5	netral	Saya barusan ke dukcapil bdg. Katanya harus ke dukcapil domisili dulu	neutral=1.0, dukcapil=3.0779
6	netral	@hansseptiana0756 waduh tetep harus ke dukcapil domisili ya om	neutral=1.0, dukcapil=1.53895
7	netral	Digitalisasi	neutral=1.0
8	netral	Apakah untuk KTP atau kk itu di aplikasinya bisa di cetak sendiri bang	posible=0.172, neutral=-0.628, aplikasinya=3.95188, cetak=4.21750, ktp=1.93698
9	netral	Ru sama aja kita ke kantor dukcapil juga, katanya ngk perlu kesana bisa d rum.	neutral=1.0, d=4.50049, dukcapil=1.53895, kantor=3.61541
10	netral	Langsung berikan blanko ktp digitalnya	neutral=1.0, digitalnya=3.48188, ktp=1.93698, langsung=2.85327
11	netral	Rg gimna kalo mau list ktp kita sendiri kan udh trdptar. Dns ny. Tpi ktp kita y...	neutral=1.0, kalo=3.0904, ktp=3.87386, mau=4.51987
12	netral	Klo ktp hing untuk mengetahui nik gmn bang.	neutral=1.0, gmn=4.58049, klo=3.63951, ktp=1.93698, nik=3.79773
13	netral	Datang ke Dukcapil untuk cetak biometriknya	neutral=1.0, dukcapil=1.53895
14	netral	Bang mau tanya, apakah bisa buat akte kelahiran online di aplikasi ini?	neutral=1.0, akta=4.51987, aplikasi=2.47788, kelahiran=4.58049, online=2.71835
15	netral	Bisa	neutral=1.0
16	negatif	ribet kali harus datang ke dukcapil, wasting time.	negative=-0.278, neutral=0.722, dukcapil=1.53895, kali=4.51987, ribet=3.68952
17	netral	ketika saatnya wajib juga pasti akan diurus juga.	neutral=1.0
18	netral	Apakah bisa verifikasi qr code di kelurahan setempat	neutral=1.0, code=3.44106, qr=2.19979, verifikasi=3.38279
19	netral	Bisa jika kelurahan di daerah anda diberikan akses oleh dukcapil	neutral=1.0, akses=4.51987, daerah=4.09499, dukcapil=1.53895
20	netral	I aplikasi bisa di pakai berapa aplikasi kkk klo saya daftar terus ltrn saya gak	neutral=1.0, i=2.91909, aplikasi=2.85327, aplikasi=2.47788, daftar=2.99637, gimana=3.34575, pakai=3.88735
21	netral	Setiap orang memiliki akun masing-masing, bisa 1 Hp dua orang tapi login.	neutral=1.0, i=2.91909, akun=2.6526, hp=2.91043, memiliki=4.64503, orang=5.79747
22	netral	Tidak sama seperti JKN ya kakikalo JKN satu akun sudah ada semua.katu sam.	neutral=1.0, akun=2.6526
23	netral	Jadi enak kita tinggal tunjukkin saja.	neutral=1.0, tinggal=3.95188
24	netral	@user:vmT00bHdo beda	neutral=1.0, beda=3.88735
25	netral	JKN juga setiap orang juga bisa beda akun, kecuali anak anak	neutral=1.0, akun=2.6526, beda=3.88735, orang=2.89873
26	netral	Ka maaf izin tanya kalo misalnya ktp/kr kita rusak bisa bisa di unduh gak sih s...	neutral=1.0, cetak=4.21750, kalo=3.0904, ktp=1.93698, maaf=4.26204
27	netral	Bisa mudah di lakukan untuk cetak ktp, kalo rusak bisa di unduh di dukcapil	neutral=1.0, i=2.91909, bisa=4.26204, dukcapil=1.53895, cetak=4.26204

Gambar 5. Hasil pemberian nilai bobot teks

Dari gambar diatas terlihat hasil dari pemberian nilai bobot terhadap kata yang sering digunakan, semakin banyak kata yang digunakan maka akan semakin tinggi nilai bobotnya.

Evaluasi Model

Tahapan berikutnya adalah melakukan evaluasi terhadap model *naïve bayes* dengan menggunakan *test and score* seperti yang dapat dilihat pada tabel berikut.

Tabel.1 Hasil pengujian akurasi naïve bayes

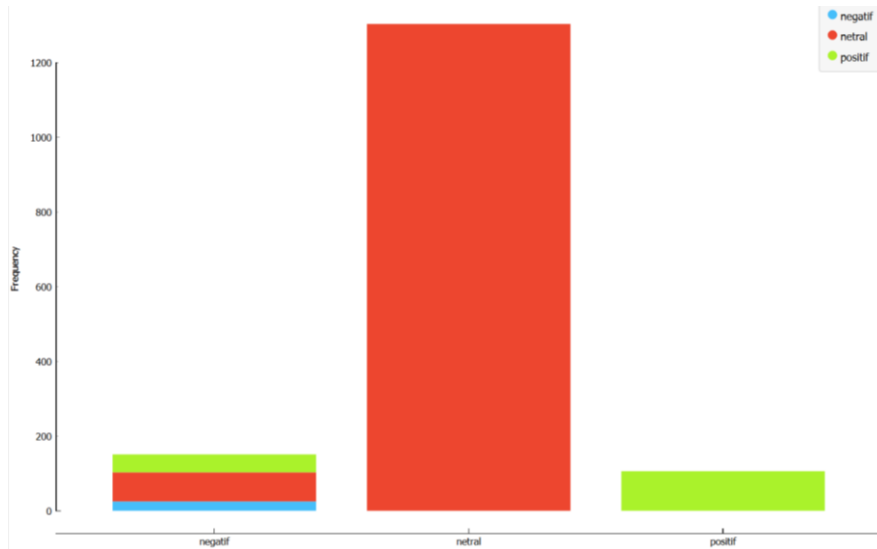
Model	CA	F1	Prec	Recall
Naïve Bayes	0.919	0.945	0.987	0.919

Dari hasil akurasi yang didapatkan dari hasil pengujian data menggunakan model *naïve bayes* mendapatkan *accuracy* sebesar 91.9%, dengan *F1 Score* 94.5%, *Precision* 98.7%, dan *Recall* 91.9%. Kemudian untuk hasil *confussion matrix* bisa dilihat pada gambar 6 berikut.

		Predicted			Σ
		negatif	netral	positif	
Actual	negatif	25	0	0	25
	netral	78	1304	0	1382
	positif	48	0	106	154
Σ		151	1304	106	1561

Gambar 6. Hasil *confussion matrix*

Pada gambar 6 dijelaskan bahwa ada terdapat beberapa data yang *misclassified* atau salah klasifikasi. Berdasarkan hasil evaluasi, total data yang mengalami *misclassified* berjumlah 126 data. Dimana letak data *misclassified* terdapat pada label negatif. Terakhir terdapat grafik bar dari hasil prediksi sentimen dengan metode *naïve bayes* seperti berikut ini.



Gambar 7. Grafik hasil prediksi *naïve bayes*

Jika dilihat pada gambar grafik diatas ada terdapat 3 nilai sentimen, diantaranya adalah positif, netral, dan negatif. Tampilan hasil dari grafik bar sentimen tersebut didapatkan dari hasil pengujian *test and score* terhadap data yang digunakan dengan metode algoritma *naïve bayes*.

KESIMPULAN

Kesimpulan dari penelitian ini adalah metode klasifikasi *Naive Bayes* terbukti dapat dianalisis secara otomatis. Performa ini dicapai dengan menguji data *real-time* menggunakan aplikasi *Orange Data Mining*. Setiap kata dikategorikan berdasarkan persepsi positif, negatif, dan netral. Performa yang dihasilkan pada pengujian yang dilakukan oleh *Orange Data Mining* pada 1.561 data pengujian menghasilkan nilai *accuracy* 91.9%, dengan nilai *F1 Score* 94.5%, *precision* 98.7% dan *Recall* 91.9%.

DAFTAR REFERENSI

- Ardianto, Rian, Tri Rivanie, Yuris Alkhalifi, Fitra Septia Nugraha, and Windu Gata. 2020. "Sentiment Analysis on E-Sports for Education Curriculum Using Naive Bayes and Support Vector Machine." *Jurnal Ilmu Komputer Dan Informasi* 13, no. 2: 109–22. <https://doi.org/10.21609/jiki.v13i2.885>.
- Arvyantomo, M, and N Ratama. 2023. "Analisis Sentimen Masyarakat Indonesia Terhadap Invasi Russia Di Ukraina Menggunakan Metode Naive Bayes Pada Media Sosial Facebook." *LOGIC: Jurnal Ilmu Komputer ...* 1, no. 4: 705–17. <https://www.journal.mediapublikasi.id/index.php/logic/article/view/2142>.
- Atmadja, Boby Rizki. 2022. "Analisis Sentimen Bahasa Indonesia Pada Tempat Wisata Di Kabupaten Sukabumi Dengan Naive Bayes Classifier." *Elkom : Jurnal Elektronika Dan Komputer* 15, no. 2: 371–82. <https://doi.org/10.51903/elkom.v15i2.872>.
- Cendana, Maya, and Silvester Dian Handy Permana. 2019. "Pra-Pemrosesan Teks Pada Grup Whatsapp Untuk Pemodelan Topik." *Jurnal Mantik Penusa* 3, no. 3: 107–16.
- Fanissa, Shima, M Ali Fauzi, and Sigit Adinugroho. 2018. "Analisis Sentimen Pariwisata Di Kota Malang Menggunakan Metode Naive Bayes Dan Seleksi Fitur Query Expansion Ranking." *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer* 2, no. 8: 2766–70. <http://j-ptiik.ub.ac.id>.
- Hakimi, Fajar Darwis Dzikril. 2018. "Sistem Analisis Sentimen Publik Tentang Opini Pemilihan Kepala Daerah Jawa Timur 2018 Pada Dokumen Twitter Menggunakan Naive Bayes Classifier." *Skripsi Universitas Islam Negeri Sunan Ampel Surabaya*.
- Indah Nurhafida, Selva, and Falentino Sembiring. 2021. "Analisis Text Clustering Masyarakat Di Twitter Mengenai Mcdonald'Sxbs Menggunakan Orange Data Mining." *SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika)*, 28–35.
- Khotimah, Aristin Chusnul, and Ema Utami. 2022. "Perbandingan Algoritma Naive Bayes Classifier, K-Nearest Neighbor Dan Support Vector Machine Dalam Klasifikasi Karakter Individu Pada Akun Twitter." *Jurnal Teknik Informatika (Jutif)* 3, no. 3: 673–80. <http://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/254>.
- Muzaki, Akhmad, and Arita Witanti. 2021. "Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier." *Jurnal Teknik Informatika (Jutif)* 2, no. 2: 101–7. <https://doi.org/10.20884/1.jutif.2021.2.2.51>.
- Prasetyo, Tri, Hadi Zakaria, and Pandu Wiliantoro. 2022. "Analisis Layanan Pelanggan PT PLN Berdasarkan Media Sosial Twitter Dengan Menggunakan Metode Naive Bayes Classifier." *OKTAL : Jurnal Ilmu Komputer Dan Sains* 1, no. 6: 573–82. <https://journal.mediapublikasi.id/index.php/oktal>.
- Putu, Ni Luh Putu Merawati, Ahmad Zuli Amrullah, and Ismarmiaty. 2021. "Analisis Sentimen Dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes Dan Latent Dirichlet Allocation." *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)* 5, no. 1: 123–31. <https://doi.org/10.29207/resti.v5i1.2587>.

- Septian, Jeremy Andre, Tresna Maulana Fachrudin, and Aryo Nugroho. 2019. "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF Dan K-Nearest Neighbor." *Journal of Intelligent System and Computation* 1, no. 1: 43–49. <https://doi.org/10.52985/insyst.v1i1.36>.
- Somantri, Oman, and Dairoh Dairoh. 2019. "Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining." *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)* 5, no. 2: 191. <https://doi.org/10.26418/jp.v5i2.32661>.
- Suryani, Putu Sri Merta, Linawati Linawati, and Komang Oka Saputra. 2019. "Penggunaan Metode Naïve Bayes Classifier Pada Analisis Sentimen Facebook Berbahasa Indonesia." *Majalah Ilmiah Teknologi Elektro* 18, no. 1: 145. <https://doi.org/10.24843/mite.2019.v18i01.p22>.
- Ulfa, Maria Arista, Budi Irmawati, and Ario Yudo Husodo. 2018. "Twitter Sentiment Analysis Using Naïve Bayes Classifier with Mutual Information Feature Selection." *Journal of Computer Science and Informatics Engineering (J-Cosine)* 2, no. 2: 106–11. <https://doi.org/10.29303/jcosine.v2i2.120>.
- Wahyuningsih, N. 2023. "Perbandingan Metode Klasifikasi Dalam Analisis Sentimen Masyarakat Terhadap Identitas Kependudukan Digital (IKD)" 8, no. 4: 1218–27. [https://repository.uksw.edu/handle/123456789/30436%0Ahttps://repository.uksw.edu/bitstream/123456789/30436/8/T1_672019296_Daftar Pustaka.pdf](https://repository.uksw.edu/handle/123456789/30436%0Ahttps://repository.uksw.edu/bitstream/123456789/30436/8/T1_672019296_Daftar%20Pustaka.pdf).