

Analisis Sentimen Ulasan Pengguna Aplikasi Ajaib Menggunakan Metode Naïve Bayes

Hilmy Zhafran Muflih, Allif Rizki Abdillah, Firman Noor Hasan*

Fakultas Teknologi Industri dan Informatika, Teknik Informatika, Universitas Muhammadiyah Prof.Dr.Hamka, Jakarta, Indonesia

Email: ¹hilmyzhrnm@gmail.com, ²allifajah321@gmail.com, ³*firman.noorhasan@uhamka.ac.id

Email Penulis Korespondensi: firman.noorhasan@uhamka.ac.id

Abstrak—Perkembangan yang terjadi di bidang investasi membuat masyarakat tertarik untuk memulai berinvestasi. Pada saat ini investasi dapat dilakukan dengan mudah menggunakan aplikasi pada smartphone, seperti aplikasi Ajaib. Aplikasi Ajaib menawarkan kenyamanan dalam melakukan investasi dengan fitur-fitur, kemanan, serta kemudahan yang diberikan kepada pengguna, sehingga dari hal tersebut muncul berbagai ulasan dari pengguna dalam bentuk sentimen negatif atau positif setelah menggunakan aplikasi Ajaib. Dari berbagai ulasan yang diberikan masih ada beberapa orang yang masih ragu untuk menggunakan aplikasi Ajaib karena ulasan yang diberikan oleh orang yang telah menggunakan. Ulasan yang telah diberikan pengguna mendorong peneliti untuk melakukan penelitian mengenai pandangan pengguna setelah menggunakan aplikasi Ajaib dengan algoritma Naïve Bayes. Pelaksanaan penelitian dilakukan untuk mengetahui banyaknya sentimen negatif dan positif dari ulasan yang ada mengenai aplikasi Ajaib. Peneliti mengumpulkan 500 data melalui proses scraping pada google playstore di aplikasi Ajaib. Hasil yang diperoleh pada penelitian mengenai analisis sentimen dengan metode algoritma Naïve Bayes mendapatkan hasil yaitu 44,2% atau 221 sentimen positif dan 55,8% atau 279 sentimen negatif dari 500 data sentimen yang telah dianalisis. Proses evaluasi dilakukan dengan menggunakan confusion matrix untuk mengevaluasi metode algoritma Naïve Bayes. Hasil dari proses evaluasi memperoleh nilai akurasi yaitu 74,44%.

Kata Kunci: Analisis Sentimen; Ulasan; Pengguna; Naïve Bayes; Ajaib

Abstract—Developments occurring in the investment sector have made people interested in starting to invest. Currently, investments can be made easily using applications on smartphones, such as the Ajaib application. The Ajaib application offers comfort in making investments with the features, security and convenience provided to users, so that various reviews emerge from users in the form of negative or positive sentiments after using the Ajaib application. From the various reviews given, there are still some people who are still hesitant to use the Ajaib application because of the reviews given by people who have used it. The reviews given by users encouraged researchers to conduct research regarding user views after using the Ajaib application with the Naïve Bayes algorithm. The research was carried out to determine the number of negative and positive sentiments from existing reviews regarding the Ajaib application. Researchers collected 500 data through a scraping process on Google Playstore in the Ajaib application. The results obtained in research on sentiment analysis using the Naïve Bayes algorithm method were 44.2% or 221 positive sentiments and 55.8% or 279 negative sentiments from the 500 sentiment data that had been analyzed. The evaluation process was carried out using a confusion matrix to evaluate the Naïve Bayes algorithm method. The results of the evaluation process obtained an accuracy value of 74.44%.

Keywords: Sentiment Analysis; Review; User; Naïve Bayes; Ajaib

1. PENDAHULUAN

Investasi adalah tindakan mengalokasikan kekayaan yang dimiliki ke dalam berbagai pilihan investasi dengan harapan mendapatkan keuntungan di masa mendatang [1]. Investasi adalah strategi yang digunakan oleh entitas atau individu untuk melindungi nilai kekayaan mereka dan, pada saat yang sama, menciptakan sumber pendapatan pasif. Dengan kemajuan dan perkembangan di bidang investasi saat ini, semakin banyak masyarakat yang tertarik memulai investasi [2]. Sederhananya, investasi didefinisikan sebagai suatu tindakan menaruh dana pada satu aset atau lebih dengan niat untuk memperoleh laba [3]. Aplikasi adalah perangkat lunak yang digunakan pengguna untuk membantu melakukan tugas serta aktivitas pada sistem komputer. Aplikasi mobile, di sisi lain, dirancang untuk beroperasi tanpa terputus ketika pengguna berpindah ke tempat lain, menjaga kelangsungan komunikasi [4]. Aplikasi dirancang untuk menjalankan suatu tugas atau fungsi yang dapat digunakan oleh pengguna atau aplikasi lain sesuai dengan tujuannya [5]. Ajaib adalah aplikasi investasi yang memudahkan berinvestasi dalam reksa dana dan saham *online*. Aplikasi ini memberikan akses ke berbagai reksa dana terbaik dan platform perdagangan saham secara *real-time*, serta sumber belajar investasi untuk semua tingkat pengalaman. Ajaib merupakan produk dari Ajaib Group, yang juga mengelola Ajaib Reksa Dana dan Ajaib Sekuritas. Pada awalnya, aplikasi ini fokus pada investasi reksa dana daring.

Basis data adalah kelompok data yang terkait dan berhubungan dengan logis dan dibuat untuk melengkapi informasi yang diperlukan [6]. Data adalah informasi yang diperoleh secara empiris dan dikumpulkan oleh peneliti untuk digunakan dalam menyelesaikan permasalahan riset. Data riset dapat diperoleh dari berbagai sumber dan dikumpulkan melalui berbagai metode selama proses penelitian [7]. Pengguna dapat memberikan ulasan mengenai aplikasi Ajaib secara bebas setelah menggunakannya. Ulasan tersebut dikumpulkan untuk dijadikan basis data. Data yang telah dikumpulkan akan dijadikan sumber sentimen yang nantinya akan melalui proses analisis sentimen. Analisis sentimen adalah tahapan di mana data teks dianalisis secara otomatis dengan tujuan untuk memahami dan mengekstrak sentimen yang ada didalam sebuah ulasan [8]. Dengan menggunakan analisis sentimen, peneliti memiliki kemampuan untuk mengevaluasi polaritas dari pendapat yang dikumpulkan, yang dapat digunakan untuk meramalkan sentimen tersebut apakah bersifat negatif atau positif. Analisis sentimen merupakan sebuah penyelesaian yang memungkinkan peneliti untuk mengukur tingkat kepuasan pengguna. Melalui metode analisis sentimen, peneliti dapat menyimpulkan informasi dari data yang bersifat acak [9].

Dalam penelitian, sentimen sering dimaksudkan sebagai big data karena volume data teksnya bertambah besar dan keragaman makna konteksnya semakin meningkat [10]. Proses analisis mencakup data dari tweet, ulasan teks, blog, forum, dan tahap preprocessing melibatkan langkah-langkah seperti menghapus kata-kata penghubung (stopword), mereduksi kata-kata menjadi bentuk dasarnya (stemming), menghilangkan elemen-elemen tertentu, mengidentifikasi sentimen, membagi teks menjadi token, dan mengklasifikasikan sentimen [11]. Ini bisa dijelaskan sebagai langkah-langkah dalam menganalisis teks untuk mengambil informasi yang memiliki nilai tambah sesuai dengan tujuan yang spesifik [12].

Beberapa macam metode analisis sentimen seperti *Random Forest*, *Naïve Bayes Classifier*, dan sejenisnya adalah pendekatan yang umumnya digunakan [13]. *Naïve Bayes Classifier* merupakan jenis dari algoritma *Machine Learning* yang tidak memiliki aturan, sehingga memiliki fleksibilitas. *Naïve Bayes* mengandalkan probabilitas untuk menentukan kemungkinan suatu kejadian dengan memperhitungkan frekuensi pada berbagai klasifikasi dalam data pelatihan [14]. Perkembangan penelitian dalam menerapkan *Naïve Bayes* untuk melakukan metode klasifikasi telah banyak menggambarkan berbagai variasi topik pembahasan [15]. Konsep mendasar dari *Naïve Bayes* melibatkan pemanfaatan Teorema Bayes, yang merupakan suatu prinsip dalam statistik untuk menghitung probabilitas. Dalam *Naïve Bayes*, probabilitas dihitung untuk mengaitkan suatu kelas dengan atribut yang ada, sehingga dapat menentukan kelas yang memungkinkan. Algoritma *Naïve Bayes* digunakan untuk menghitung serta untuk menghitung probabilitas yang tertinggi, yang kemudian digunakan untuk mengklasifikasikan data uji ke dalam bagian yang tepat [16]. Data tersebut dijalankan melalui alat RapidMiner untuk pengujian [17]. RapidMiner adalah alat yang mampu mengolah data dengan memanfaatkan algoritma dan prinsip data mining [18].

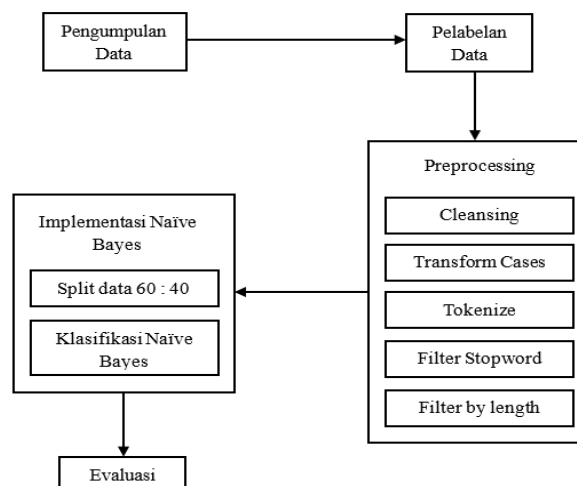
Penelitian mengenai analisis sentimen sudah banyak dilakukan sebelumnya, seperti penelitian mengenai analisis sentimen yang membahas mengenai aplikasi Canva dengan menggunakan metode algoritma KNN dan *Naïve Bayes*, accuracy yang didapatkan dari hasil penelitian sebesar 77.41% untuk algoritma *Naïve Bayes* dan 83,92% untuk algoritma KNN [19]. Penelitian terdahulu lainnya yaitu membahas tentang analisis sentimen yaitu penelitian analisis sentimen yang membahas tentang kualitas pelayanan dari aplikasi Gojek dengan algoritma *Naïve Bayes* untuk memprosesnya, hasil akurasi yang diperoleh yaitu sebesar 68% [20].

Berbagai respon masyarakat baik pro dan kontra dalam hal investasi secara daring melalui aplikasi Ajaib, sehingga peneliti melakukan penelitian mengenai tanggapan masyarakat setelah menggunakan aplikasi Ajaib. Tujuan dilakukannya penelitian ini untuk mengetahui respon para pengguna aplikasi Ajaib setelah menggunakannya dalam bentuk ulasan di google playstore. Ulasan yang diberikan oleh pengguna akan mencerminkan tingkat kepuasan mereka terhadap aplikasi Ajaib. Hal yang mempengaruhi kepuasan dari para pengguna terhadap aplikasi Ajaib yaitu dari fitur-fitur yang diberikan serta kemudahan dalam pemakaiannya. Ketika respon para pengguna aplikasi Ajaib yang didapat dari ulasan yang ada di google playstore banyak yang tidak puas diharapkan pihak dari aplikasi Ajaib bisa mengevaluasi serta membuat pembaharuan dari saran dan masukan yang diberikan di kolom ulasan. Perbedaan penelitian ini dari penelitian yang telah ada, penelitian ini mengumpulkan data dengan *web scraping* untuk dijadikan dataset, kemudian penelitian ini menggunakan tools RapidMiner untuk melakukan proses analisis terhadap dataset yang telah didapatkan dengan algoritma *Naïve Bayes*.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Metode yang digunakan peneliti yaitu menggunakan ulasan yang didapat dari *review/ulasan* aplikasi yang ada di *google playstore* dan menerapkan pada platform *RapidMiner* untuk memprediksi sentimen, sehingga dapat mempermudah untuk mengetahui sentimen bersifat positif atau negatif. Peneliti mengilustrasikan alur penelitian pada Gambar 1 berikut.



Gambar 1. Alur penelitian

2. 2 Pengumpulan Data

Proses *scraping* data dilakukan melalui berbagai metode dan teknik. Proses *scraping* data ini juga memiliki kemampuan untuk mengubah data yang tidak terstruktur menjadi terstruktur, sehingga dapat disimpan dalam basis data. Jadi, proses *scraping* data ini adalah menyimpan, mengumpulkan, dan memvalidasi data agar dapat diinterpretasikan sebagai informasi yang bermutu [21].

2. 3 Pelabelan data

Pada tahap pelabelan data, keputusan diambil untuk mengelompokkan data menjadi dua kategori: positif dan negatif. Kategori positif digunakan untuk menggambarkan emosi positif seperti santai, senang, rileks, ceria, dan emosi yang baik. Sementara itu, kategori negatif diterapkan pada emosi negatif seperti tertekan, sedih, takut, marah, dan emosi yang mengarah ke rasa penderitaan. Proses pengklasifikasian ini dilakukan secara manual dan bersifat subjektif [22].

2. 4 Preprocessing Data

Tahap *preprocessing* data dilakukan pada tahap ini ialah untuk mengolah data yang telah dikumpulkan dalam bentuk yang tidak teratur. *Preprocessing* data dilakukan untuk mengganti beberapa kata asing ke dalam Bahasa Indonesia serta menghilangkan karakter yang tidak diperlukan [23]. Pada processing data terdapat 5 tahapan, yaitu *cleansing, transform cases, tokenize, filter stopwords, filter by length*.

2. 5 Naïve Bayes

Naïve Bayes Classifier adalah sebuah metode algoritma untuk melakukan klasifikasi yang didasarkan pada konsep probabilitas. Metode ini merujuk pada Teorema Bayes sebagai dasar konsepnya. Naïve Bayes merupakan salah satu instrumen klasifikasi yang sangat efektif [24]. Keunggulan Naïve Bayes Classifier terletak pada kemampuannya yang sederhana namun menghasilkan kinerja yang baik dalam klasifikasi teks [25].

2. 6 Evaluasi

Pada tahap evaluasi dilakukan dengan menggunakan *confusion matrix* untuk mengetahui performa dari algoritma *naïve bayes* yang telah dibuat untuk melakukan proses klasifikasi. Hasil yang terdapat pada *confusion matrix* berupa nilai *accuracy, recall, dan precision*.

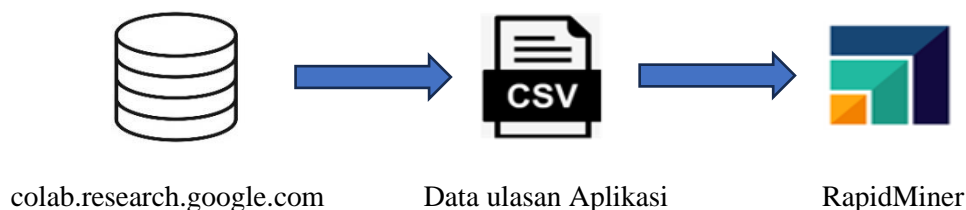
2. 7 RapidMiner

RapidMiner pada penelitian ini digunakan untuk mengolah data serta melakukan proses klasifikasi dengan algoritma *naïve bayes*. RapidMiner merupakan perangkat lunak sumber terbuka yang dapat diakses oleh semua individu. RapidMiner berfungsi sebagai alat analisis data yang menggunakan berbagai metode, termasuk metode deskriptif dan prediktif [26].

3. HASIL DAN PEMBAHASAN.

3. 1 Pengumpulan Data

Pada penelitian ini, peneliti mengumpulkan data mengenai ulasan aplikasi Ajaib dengan menggunakan *web scraping*. Data yang dikumpulkan yaitu mulai dari Desember 2020 hingga Agustus 2023 dengan total jumlah 500 ulasan untuk dijadikan *dataset*. Data tersebut akan disimpan dalam format *csv*. Data yang telah dikumpulkan pada penelitian ini didapat dari aplikasi Ajaib yang ada di *website google playstore*. Berikut Gambar 2 adalah ilustrasi pada proses pengumpulan data.



Gambar 2. Ilustrasi pada proses pengumpulan data

3. 2 Pelabelan Dataset

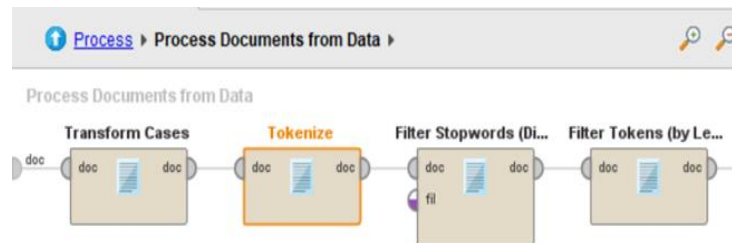
Setelah proses pengumpulan data dengan *web scraping*, maka akan dilanjutkan dengan proses pelabelan terhadap data ulasan aplikasi Ajaib yang dilakukan oleh peneliti. Peneliti akan memberikan label sentimen positif dan negatif pada data ulasan dan proses pelabelan akan dilakukan pada *Miscrosoft Excel*. Pelabelan yang dilakukan oleh peneliti dengan algoritma *Naïve Bayes* ini akan dijadikan untuk data latih. Data yang diberi label sentimen oleh peneliti yaitu sebanyak 300 data dari total data yaitu 500 data. Berikut Gambar 3 merupakan hasil pelabelan data ulasan aplikasi Ajaib.

‡ aplikasi cocok buat pemula/amateur. tidak ramah terlalu banyak iklan sangat mengganggu ! bintang 3	negatif
‡ Min bisa tambah menunya dong, tambah menu Pesanan IPO kita, gak tampak kalau kita sudah pesan atau belum. Harus di klik dulu saham IPO baru nampak.	negatif
‡ Di versi sebelumnya ada menu untuk melihat berapa total profit dan loss nya, versi terakhir kok malah dihilangi	negatif
‡ tolong dong min diperbaiki mau lihat chart di trading view ko lama banget loading nya. terus kalo bisa full screen layar HP dong	negatif
‡ Saat d jalankan sering Lemot, terkadang melakukan order jadi telat. Untuk fiturnya lengkap & bisa dipahami bagi pemula.	positif
‡ Hadeh kecewa banget sama ni aplikasi udh daftar eh malah gagal membuka akun alhasil hanya saja investasi reksadana doang giliran minta bantuan cs dan email pun juga gabis	negatif
‡ Aplikasi ajaib cocok banget untuk pemula untuk belajar investasi, mudah digunakan, dan fitur lengkap.	positif
‡ Sebenarnya ini aplikasi sangat user friendly untuk pemula dan mudahhntuk dipahami namun ada 1 kelemahan yang mungkin bisa di benahi dari tim ajaib menambahkan fitur ahli	positif
‡ Cocok untuk pemula seperti saya yg baru belajar crypto, fitur fitur nya mudah di pahami, dan transaksi nya pun realtime, saya suka !!!	positif
‡ Sekedar saran untuk pengembang aplikasi, untuk tampilan mode gelap, warna grafik mungkin bisa dibuat lebih gelap supaya tidak terlalu kontras dengan warna latar belakang, dan	positif
‡ Menunya keren2tp Terlalu berat utk buka chart. Dan tdk semua hp bs buka, terutama android versi lama. Tolong dibuat ke versi yg bs support atau bs buka di android lama krn tdk	positif

Gambar 3. Pelabelan ulasan

3.3 Preprocessing Data

Proses selanjutnya peneliti melakukan preprocessing pada dataset ulasan aplikasi Ajaib dengan menggunakan aplikasi *RapidMiner*. Pada proses preprocessing, data yang berisi teks ulasan aplikasi Ajaib yang tidak beraturan akan diubah menjadi teks yang beraturan sehingga tidak terjadi kesalahan. Pada tahap proses preprocessing ini ada tahap beberapa tahap yaitu *cleaning*, *tokenize*, *transform cases*, *stopword*, *filter tokens (by length)*. Selain itu, pada penelitian ini peneliti menggunakan operator di aplikasi RapidMiner yaitu operator *process document* untuk melakukan penyesuaian terhadap operator yang akan digunakan pada tahapan *tokenize*, *transform cases*, *stopword*, *filter tokens (by length)* seperti pada Gambar 4 berikut.



Gambar 4. Tahap Preprocessing

Cleansing merupakan proses yang dilakukan untuk melakukan proses pembersihan dengan pada teks ulasan dengan menghapus berbagai simbol atau karakter seperti : ,!/ dan seterusnya [27]. Berikut contoh ulasan sebelum dan sesudah melewati proses *cleansing* pada Tabel 1.

Tabel 1. Proses *cleansing*

Sebelum <i>cleansing</i>	Sesudah <i>cleansing</i>
Cocok untuk pemula seperti saya yang baru belajar crypto, fitur fitur nya mudah dipahami, dan transaksi nya pun realtime, saya suka !!!	Cocok untuk pemula seperti saya yang baru belajar crypto fitur fitur nya mudah dipahami dan transaksi nya pun realtime saya suka
Saya baru update...tapi kenapa order saya tidak muncul ??????? di riwayat pun tidak???????? tadi saya salah pasang harga mau Amand tapi tidak muncul...itu bagaimana?	Saya baru update tapi kenapa order saya tidak muncul di riwayat pun tidak tadi saya salah pasang harga mau Amand tapi tidak muncul itu bagaimana
Klo bsa WD nya jangan H+3 karena klo kita butuh bgt! Masa harus nunggu 3 hari ? Idealnya instant WD, tp H+1 masih ok, lah . Klo h+1 pasti yg Invest pada nambah nominal deposit ratusan juta! Gak masalah! Toh WD nya cpt! Klo h+3 mntok 10-20jutaan karena klo lagi butuh banget ada deadline, yah lama jg 3 hari mah !	Klo bsa WD nya jangan H3 karena klo kita butuh bgt Masa harus nunggu 3 hari Idealnya instant WD tp H1 masih ok lah Klo h1 pasti yg Invest pada nambah nominal deposit ratusan juta Gak masalah Toh WD nya cpt Klo h3 mntok 10 20jutaan karena klo lagi butuh banget ada deadline yah lama jg 3 hari mah

Transform cases merupakan proses untuk merubah setiap huruf besar dataset untuk diubah menjadi huruf kecil [28]. Hal tersebut dilakukan untuk mencegah kesalahan pada saat *tokenize*. Berikut merupakan contoh dari beberapa data sebelum dan sesudah melalui proses *transform cases* pada Tabel 2.

Tabel 2. Proses *transform cases*

Sebelum <i>transform cases</i>	Sesudah <i>transform cases</i>
Cocok untuk pemula seperti saya yang baru belajar crypto fitur fitur nya mudah dipahami dan transaksi nya pun realtime saya suka	cocok untuk pemula seperti saya yang baru belajar crypto fitur fitur nya mudah dipahami dan transaksi nya pun realtime saya suka

<p>Saya baru update tapi kenapa order saya tidak muncul di riwayat pun tidak tadi saya salah pasang harga mau Amand tapi tidak muncul itu bagaimana</p> <p>Klo bsa WD nya jangan H3 karena klo kita butuh bgt Masa harus nunggu 3 hari Idealnya instant WD tp H1 masih ok lah Klo h1 pasti yg Invest pada nambah nominal deposit ratusan juta Gak masalah Toh WD nya cpt Klo h3 mntok 10 20jutaan karena klo lagi butuh banget ada deadline yah lama jg 3 hari mah</p>	<p>saya baru update tapi kenapa order saya tidak muncul di riwayat pun tidak tadi saya salah pasang harga mau amand tapi tidak muncul itu bagaimana</p> <p>klo bsa wd nya jangan h3 karena klo kita butuh bgt masa harus nunggu 3 hari idealnya instant wd tp h1 masih ok lah klo h1 pasti yg invest pada nambah nominal deposit ratusan juta Gak masalah toh wd nya cpt Klo h3 mntok 10 20jutaan karena klo lagi butuh banget ada deadline yah lama jg 3 hari mah</p>
--	--

Tokenizing merupakan tahapan di mana teks dipisah menjadi bagian-bagian yang dinamakan sebagai token untuk dilakukannya analisis nantinya [29]. Berikut merupakan contoh ulasan sebelum dan sesudah melewati proses tokenize pada Tabel 3.

Tabel 3. Proses *tokenizing*

Sebelum <i>tokenizing</i>	Sesudah <i>tokenizing</i>
cocok untuk pemula seperti saya yang baru belajar crypto fitur fitur nya mudah dipahami dan transaksi nya pun realtime saya suka	['cocok', 'untuk', 'pemula', 'seperti', 'saya', 'yang', 'baru', 'belajar', 'crypto', 'fitur', 'fitur', 'mudah', 'dipahami', 'dan', 'transaksi', 'realtime', 'saya', 'suka']
saya baru update tapi kenapa order saya tidak muncul di riwayat pun tidak tadi saya salah pasang harga mau amand tapi tidak muncul itu bagaimana	['saya', 'baru', 'tapi', 'kenapa', 'saya', 'tidak', 'muncul', 'di', 'riwayat', 'pun', 'tidak', 'tadi', 'saya', 'salah', 'pasang', 'harga', 'mau', 'amand', 'tapi', 'tidak', 'muncul', 'itu', 'bagaimana']
klo bsa wd nya jangan h3 karena klo kita butuh bgt masa harus nunggu 3 hari idealnya instant wd tp h1 masih ok lah klo h1 pasti yg invest pada nambah nominal deposit ratusan juta Gak masalah toh wd nya cpt Klo h3 mntok 10 20jutaan karena klo lagi butuh banget ada deadline yah lama jg 3 hari mah	['klo', 'bsa', 'wd', 'nya', 'jangan', 'h3', 'karena', 'klo', 'kita', 'butuh', 'bgt', 'masa', 'harus', 'nunggu', '3hari', 'idealnya', 'instan', 'wd', 'tp', 'h1', 'masih', 'ok', 'lah', 'klo', 'h1', 'pasti', 'yg', 'inves', 'pada', 'nambah', 'nominal', 'deposit', 'ratusan', 'juta', 'gak', 'masalah', 'toh', 'wd', 'nya', 'cpt', 'klo', 'h3', 'mntok', '1020jutaan', 'jutaan', 'karena', 'klo', 'lagi', 'butuh', 'banget', 'ada', 'deadline', 'yah', 'lama', 'jg', '3hari', 'mah']

Filter stopwords adalah tahap menghilangkan kata yang tidak memiliki arti [30]. Peneliti menggunakan *stopwords* Bahasa Indonesia yang didapat dari www.kaggle.com. Berikut merupakan contoh ulasan yang telah melalui proses filter *stopwords* pada Tabel 4.

Tabel 4. Proses *stopword*

Sebelum <i>stopword</i>	Sesudah <i>stopword</i>
cocok untuk pemula seperti saya yang baru belajar crypto fitur fitur nya mudah dipahami dan transaksi nya pun realtime saya suka	cocok pemula baru belajar crypto fitur mudah dipahami transaksi realtime suka
saya baru update tapi kenapa order saya tidak muncul di riwayat pun tidak tadi saya salah pasang harga mau amand tapi tidak muncul itu bagaimana	baru update tidak order muncul riwayat salah pasang harga mau amand tidak muncul
klo bsa wd nya jangan h3 karena klo kita butuh bgt masa harus nunggu 3 hari idealnya instant wd tp h1 masih ok lah klo h1 pasti yg invest pada nambah nominal deposit ratusan juta Gak masalah toh wd nya cpt Klo h3 mntok 10 20jutaan karena klo lagi butuh banget ada deadline yah lama jg 3 hari mah	jangan kerana butuh masa nunggu ideal instan masih invest nambah deposit masalah karena lagi butuh banget deadline lama

Filter tokens by length merupakan tahapan untuk menghapus kata yang telah ditetapkan jumlah hurufnya oleh peneliti. Berikut merupakan contoh ulasan yang telah melalui proses *filter tokens by length* pada Tabel 5.

Tabel 5. Proses *filter by length*

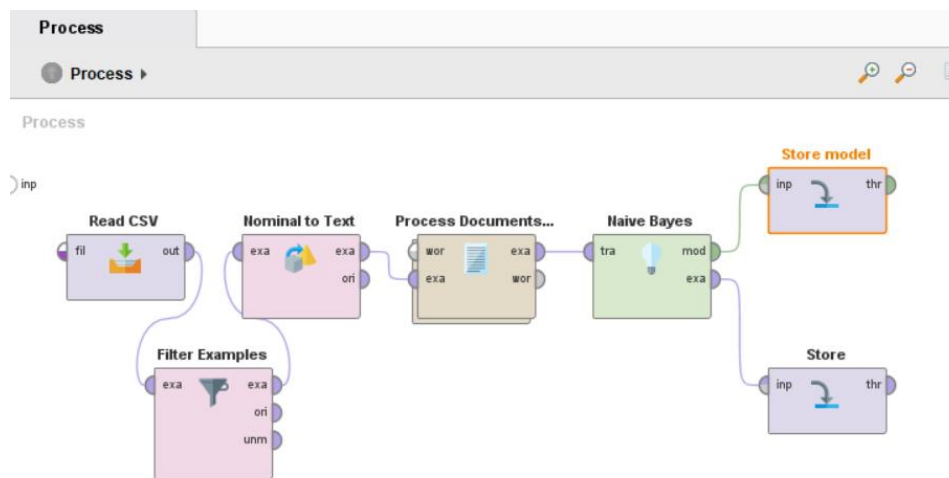
Sebelum <i>filter tokens by length</i>	Sesudah <i>filter tokens by length</i>
cocok pemula baru belajar crypto fitur mudah dipahami transaksi realtime suka	cocok pemula baru belajar crypto fitur mudah dipahami transaksi realtime suka
baru update tidak order muncul riwayat salah pasang harga mau amand tidak muncul	baru update tidak order muncul riwayat salah pasang harga mau amand tidak muncul

klo bsa wd nya jangan h3 karena klo kita butuh bgt masa harus nunggu 3 hari idealnya instant wd tp h1 masih ok lah klo h1 pasti yg invest pada nambah nominal deposit ratusan juta Gak masalah toh wd nya cpt Klo h3 mntok 10 20jutaan karena klo lagi butuh banget ada deadline yah lama jg 3 hari mah

jangan karena kita butuh masa harus nunggu hari idealnya instan masih pasti invest pada nambah nominal deposit ratusan juta masalah mntok 20jutaan karena lagi butuh banget ada deadline lama hari

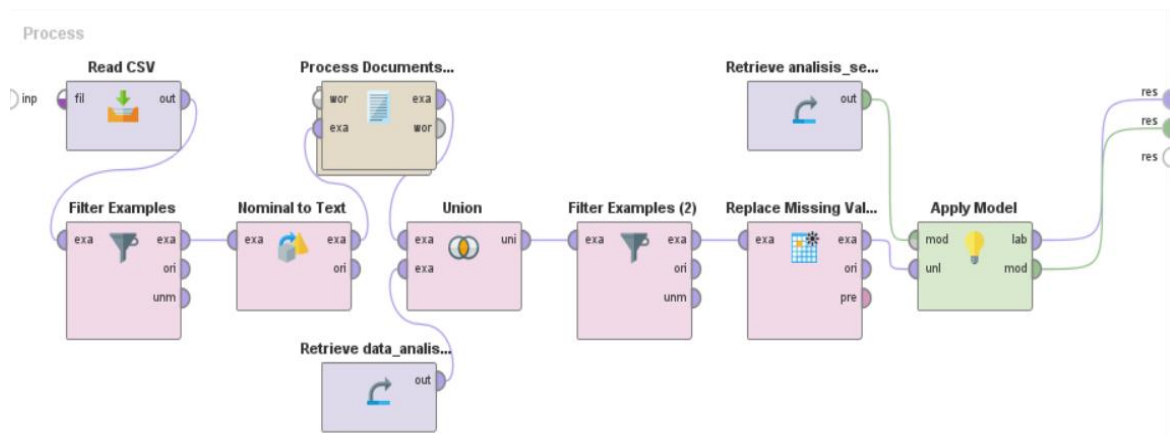
3.4 Implementasi Naïve Bayes

Setelah tahap *preprocessing* maka akan dilanjut ke tahap pengimplementasian algoritma *Naïve Bayes* terlebih dahulu. Proses implementasi *Naïve Bayes* ini dilakukan dengan menggunakan aplikasi *RapidMiner*. Hal pertama yang dilakukan oleh peneliti yaitu membuat model algoritma *Naïve Bayes* dan juga data latih di aplikasi *RapidMiner* yaitu dengan menghubungkan operator *process document* dari tahap *preprocessing* ke operator algoritma *Naïve Bayes*, kemudian difilter terlebih dahulu data yang berisi ulasan aplikasi investasi *Ajaib* yang sudah diberi label sentimen oleh peneliti. Perbandingan yang peneliti gunakan untuk data latih dan data uji yang dipakai peneliti yaitu 60 : 40. Berikut Gambar 5 merupakan alur pemrosesannya.



Gambar 5. Proses pembuatan model *Naïve Bayes* dan data latih.

Tahap selanjutnya setelah memiliki model *Naïve Bayes* dan juga data latih yaitu dilanjut dengan tahap klasifikasi algoritma *Naïve Bayes*. Data yang berisi Ulasan Aplikasi *Naïve Bayes* yang akan dijadikan data uji terlebih dahulu difilter untuk diambil data yang sentimennya masih kosong, lalu dilanjut dengan proses *preprocessing* yang ada di dalam operator *process document*. Hasil dari operator *process document* akan digabungkan oleh data latih yang telah kita dapatkan sebelumnya dengan menggunakan operator *union*, hal ini dilakukan karena adanya perbedaan kata-kata yang terdapat pada data uji dan data latih. Operator *union* digunakan oleh peneliti untuk menyesuaikan atribut yang ada diantara data latih dan data uji. Setelah digabungkan akan difilter lagi terlebih dahulu untuk diambil data yang sentimennya masih kosong. Setelah data uji sudah siap untuk dilakukan proses klasifikasi dengan menggunakan model algoritma *Naïve Bayes*, peneliti menghubungkan model dari algoritma *Naïve Bayes* dengan *apply* untuk melakukan proses klasifikasi pada data uji, setelah selesai melakukan proses klasifikasi *Naïve Bayes* akan didapatkan hasil berupa data yang sudah dilabeli sentimennya secara otomatis dengan algoritma *Naïve Bayes*. Berikut adalah gambar 6 proses pemberian label secara otomatis dengan menggunakan metode algoritma *Naïve Bayes*.



Gambar 6. Proses pelabelan secara otomatis dengan *Naïve Bayes* di aplikasi *RapidMiner*

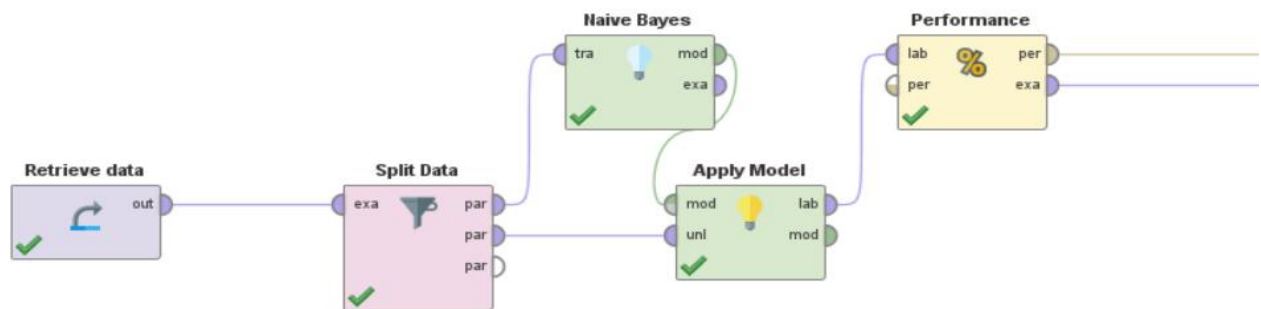
Setelah melalui proses pemberian label secara otomatis dengan menggunakan metode algoritma *Naïve Bayes*, maka akan muncul hasil pemberian label dari *Naïve Bayes* dalam bentuk *prediction* sentimen. Pada proses pemberian label secara otomatis oleh algoritma *Naïve Bayes*. Algoritma tersebut telah di latih menggunakan data latih yang telah disiapkan oleh peneliti, berupa data yang sebelumnya telah diberi label secara manual. Setelah dilatih, algoritma *Naïve Bayes* dapat memahami pola yang telah dipelajari dari data latih, sehingga mampu memprediksi data yang akan diprediksi ke dalam dua kelas yaitu positif dan negatif. Berikut gambar 7 hasil dari pemberian label secara otomatis dengan *Naïve Bayes*.

text	prediction(S...
ajaib cocok banget pemula pengen investasi aplikasinya gampang fiturnya lengkap keren pokoknya	positif
bagus histori total gain dapatkan repot kalo buka email ngurangi semoga fitur ajaib stockbit	positif
perbaharui harapan kdepan nyaman lelet omong jaringan bagus login udah masuk fortopollo data masuk aplikasi perbaharui berjalan lancar tolong kakbang udah terlanjur nyaman pakai ajaib	negatif
parah harga turun update update lihat harga turunupdateanya bagus sayang pengguna aplikasi ajaib nggunakan handphone versi keatasকাশian versi kebawah grafik harga terima kasih	negatif
aplikasinya ribet rekomended pemula investasi suka banget pake aplikasi ajaib	positif
fitur disajikan bagus cocok belajar investasi penggunaan mudah tampilan menarik	positif
pemula paham aplikasi cocok tampilan simple ngga ribet ajaib namanya	positif

Gambar 7. Hasil dari proses pemberian label secara otomatis dengan *Naïve Bayes*

3. 5 Pengujian Evaluasi

Tahapan selanjutnya setelah klasifikasi dengan algoritma *Naïve Bayes* yaitu melakukan proses pengujian/evaluasi dari algoritma *Naïve Bayes*. Pada pengujian evaluasi ini *confusion matrix* digunakan untuk melakukan pengujian evaluasi. Berikut Gambar 8 merupakan alur dilakukannya pengujian algoritma *Naïve Bayes*.



Gambar 8. alur pengujian algoritma *Naïve Bayes* dengan *confusion matrix*

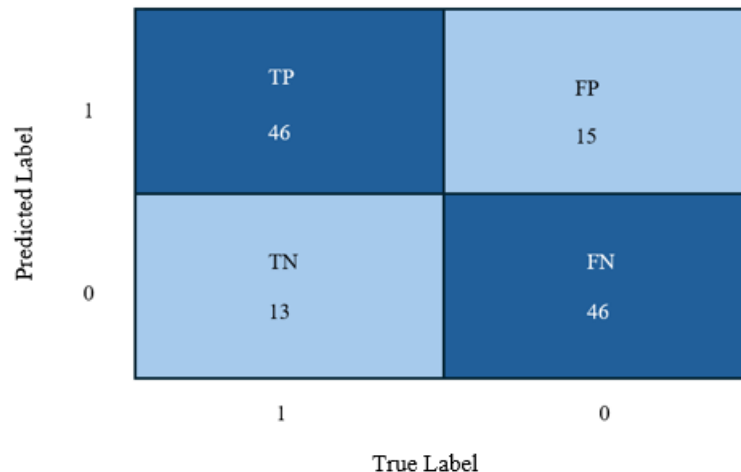
Pada operator split data peneliti akan menggunakan split data dengan perbandingan yaitu sebesar 60:40. Hasil dari proses pengujian algoritma *Naïve bayes* ini berupa nilai *Precision*, *recall*, dan *accuracy*. Berikut Gambar 9 merupakan *confusion matrix* hasil dari proses pengujian algoritma *Naïve Bayes*.

accuracy: 76.67%

	true positif	true negatif	class precision
pred. positif	46	15	75.41%
pred. negatif	13	46	77.97%
class recall	77.97%	75.41%	

Gambar 9. *Confusion matrix* dari hasil pengujian algoritma *Naïve Bayes*

Berdasarkan hasil *confusion matrix* dari algoritma *Naïve Bayes* mendapatkan nilai *accuracy* sebesar 76,67%. Kemudian sentimen positif pada proses pengujian juga memperoleh recall 77,97% dan *precision* sebesar 75,41%. Sedangkan sentimen negatif memperoleh *recall* 75,41% dan *precision* sebesar 77,97% . Berikut Gambar 10 merupakan visualisasi dari *confusion matrix*.



Gambar 10. Visualisasi confusion matrix

4. KESIMPULAN

Berdasarkan hasil dari pengolahan data aplikasi *Ajaib* dengan algoritma *naïve bayes*. Terdapat 221 data sentimen bernilai positif dan 279 data sentimen yang bernilai negatif dari 500 data yang didapatkan dari Desember 2020 hingga Agustus 2023 dan telah dianalisis. Dengan jumlah sentimen yang bernilai negatif lebih tinggi dari sentimen yang bernilai positif dapat disimpulkan para pengguna aplikasi *Ajaib* merasa kurang puas karena fitur yang diberikan serta penggunaan dari aplikasi *Ajaib* terlihat dari banyaknya ulasan negatif yang diberikan. Akan tetapi sebagian pengguna juga merasa puas atas fitur yang diberikan. Dilihat dari jumlah sentimen positif dan negatif yang perbedaannya jumlah banyaknya tidak terlalu jauh. Hasil akurasi pengujian algoritma *naïve bayes* yang didapatkan dari proses evaluasi dengan *confusion matrix* menunjukkan bahwa akurasi algoritma *naïve bayes* adalah sebesar 76,67%. Pada proses evaluasi sentimen positif didapatkan nilai precision mencapai 75,41% dengan recall sebesar 77,97%, sementara untuk sentimen negatif mendapatkan nilai precision 77,97% dengan recall sebesar 75,41%. Kekurangan dari penelitian ini adalah kurangnya jumlah sampel data yang digunakan. Peneliti hanya melabeli 300 data untuk dijadikan data latih oleh karena itu, masih terdapat beberapa data yang kurang akurat dalam proses klasifikasi *naïve bayes*. Saran untuk penelitian selanjutnya yang diberikan oleh peneliti, pada saat proses pelabelan pada dataset diharapkan dapat dilakukan lebih banyak lagi sehingga pada saat proses klasifikasi *naïve bayes* hasil yang didapatkan bisa lebih akurat.

REFERENCES

- [1] F. Ainayah, I. Yuliana, and M. I. Prajawati, "Sikap Investasi dan Norma Subjektif Terhadap Minat Investasi dengan Literasi Keuangan Sebagai Variabel Moderasi," *Ekon. dan Bisnis*, vol. 8, no. 8(1), pp. 39–45, 2021, doi: 10.35590/jeb.v10i1.6284.
- [2] M. A. Khakim, "Pengaruh Investasi Dalam Perekonomian," *J. AKSES*, no. 2085, pp. 1–10, 2022, doi: 10.47329/jurnalakses.v14i2.892.
- [3] M. Rinestu, I. P. Made Indra, B. Marsanto, and S. Trisakti, "Classification Of Investment Decisions During Covid-19 Pandemic Using Naive Bayes," *Manag. Stud. Entrep. J.*, vol. 3, no. 4, pp. 1784–1796, 2022, [Online]. Available: <http://journal.yrpicu.com/index.php/msej>.
- [4] A. Voutama and E. Novalia, "Perancangan Aplikasi M-Magazine Berbasis Android Sebagai Sarana Mading Sekolah Menengah Atas," *J. Tekno Kompak*, vol. 15, no. 1, p. 104, Feb. 2021, doi: 10.33365/jtk.v15i1.920.
- [5] Aditya Fajar Ramadhan, Ade Dwi Putra, and Ade Surahman, "Aplikasi Pengenalan Perangkat Keras Komputer Berbasis Android Menggunakan augmented Reality (Ar)," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, pp. 24–31, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSl>.
- [6] EllyahHelmud, "Optimasi Basis Data Oracle Menggunakan Complex View Studi Kasus : PT. Berkat Optimis Sejahtera (PT.BOS) Pangkalpinang," *J. Informanika*, vol. 7, no. 1, pp. 80–86, 2021, doi: 10.52233/informanika.v7i01.204.
- [7] A. Wahab, A. Syahid, and J. Junaedi, "Penyajian Data Dalam Tabel Distribusi Frekuensi Dan Aplikasinya Pada Ilmu Pendidikan," *Educ. Learn. J.*, vol. 2, no. 1, p. 40, 2021, doi: 10.33096/eljour.v2i1.91.
- [8] F. Sidik, I. Suhada, A. H. Anwar, and F. N. Hasan, "Analisis Sentimen Terhadap Pembelajaran Daring Dengan Algoritma Naive Bayes Classifier," *J. Linguist. Komputasional*, vol. 5, no. 1, p. 34, Apr. 2022, doi: 10.26418/jlk.v5i1.79.
- [9] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia," *J. Manaj. Inform.*, vol. 12, no. 2, pp. 89–99, 2022, doi: 10.34010/jamika.v12i2.7764.
- [10] M. A. Amrustian, W. Widayat, and A. M. Wirawan, "Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 535, 2022, doi: 10.30865/mib.v6i1.3527.
- [11] W. A. Prabowo and C. Wiguna, "Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 1, p. 149, Jan. 2021, doi: 10.30865/mib.v5i1.2604.
- [12] N. Ramdhani and R. H. Al-Fadillah, "Analisis Sentimen Pengguna Twitter Terhadap Belajar Daring Selama Pandemi Covid-19 Dengan Deep Learning," *J. Siliwangi*, vol. 7, no. 2, p. 2021, 2021.
- [13] P. Arsi, R. Wahyudi, and R. Waluyo, "Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia,"

- J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 10, pp. 231–237, 2021, doi: <https://doi.org/10.29207/resti.v5i2.2698>.
- [14] E. Fitriani, “Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan,” *Sistemasi*, vol. 9, no. 1, p. 103, 2020, doi: 10.32520/stmsi.v9i1.596.
- [15] T. A. Sari, E. Sinduningrum, and F. N. Hasan, “Analisis Sentimen Ulasan Pelanggan Pada Aplikasi Fore Coffee Menggunakan Metode Naïve Bayes,” *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 6, pp. 773–779, 2023, doi: 10.30865/klik.v3i6.884.
- [16] A. Tangkelayuk, “The Klasifikasi Kualitas Air Menggunakan Metode KNN, Naïve Bayes, dan Decision Tree,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 2, pp. 1109–1119, Jun. 2022, doi: 10.35957/jatisi.v9i2.2048.
- [17] C. Nas, “Data Mining Prediksi Minat Calon Mahasiswa Memilih Perguruan Tinggi Menggunakan Algoritma C4.5,” *J. Manaj. Inform.*, vol. 11, no. 2, pp. 131–145, 2021, doi: 10.34010/jamika.v11i2.5506.
- [18] K. F. Irnanda, A. P. Windarto, and I. S. Damanik, “Optimasi Particle Swarm Optimization Pada Peningkatan Prediksi dengan Metode Backpropagation Menggunakan Software RapidMiner,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 1, p. 122, 2022, doi: 10.30865/jurikom.v9i1.3836.
- [19] D. Pratmanto, F. Fandi, and D. Imaniawan, “Computer Science (CO-SCIENCE) Analisis Sentimen Terhadap Aplikasi Canva Menggunakan Algoritma Naive Bayes Dan K-Nearest Neighbors,” vol. 3, no. 2, pp. 110–117, 2023, doi: <https://doi.org/10.31294/coscience.v3i2.1917>.
- [20] K. D. Indarwati and H. Februariyanti, “Analisis Sentimen Terhadap Kualitas Pelayanan Aplikasi Gojek Menggunakan Metode Naive Bayes Classifier,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 10, no. 1, 2023, doi: 10.35957/jatisi.v10i1.2643.
- [21] J. Florensius Sianipar, Y. R. Ramadhan, and I. Jaelani, “Analisis Sentimen Pembangunan Kereta Cepat Jakarta-Bandung di Media Sosial Twitter Menggunakan Metode Naive Bayes,” *Media Online*, vol. 4, no. 1, pp. 360–367, 2023, doi: 10.30865/klik.v4i1.1033.
- [22] D. Oktavia, Y. R. Ramadahan, and M. Minarto, “Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM),” *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 407–417, 2023, doi: 10.30865/klik.v4i1.1040.
- [23] A. F. I. Muhammad Imam Ghozali, Wibowo Harry Sugiharto, “Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes,” *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 33, no. 1, pp. 1–12, 2022, doi: <https://doi.org/10.30865/klik.v3i6.936>.
- [24] E. W. Sholeha, S. Yunita, R. Hammad, V. C. Hardita, and K. Kaharuddin, “Analisis Sentimen Pada Agen Perjalanan Online Menggunakan Naïve Bayes dan K-Nearest Neighbor,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 3, no. 4, pp. 203–208, 2022, doi: 10.35746/jtim.v3i4.178.
- [25] M. Hudha, E. Supriyati, and T. Listyorini, “Analisis Sentimen Pengguna Youtube Terhadap Tayangan #Matanajwamenantiterawan Dengan Metode Naïve Bayes Classifier,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 1, pp. 1–6, 2022, doi: 10.33387/jiko.v5i1.3376.
- [26] L. L. Kaamilah, “Analisis Kelompok Lansia Berdasarkan Kategori Usia Dengan Metode K-Means Clustering,” *J. Ris. Ilmu Akunt.*, vol. 2, no. 2, pp. 1–14, 2023.
- [27] M. W. A. Putra, Susanti, Erlin, and Herwin, “Analisis Sentimen Dompot Elektronik Pada Twitter Menggunakan Metode Naïve Bayes Classifier,” *IT J. Res. Dev.*, vol. 5, no. 1, pp. 72–86, 2020, doi: 10.25299/itjrd.2020.vol5(1).5159.
- [28] A. Noviriandini, H. Hermanto, and D. Ayu Ambarsari, “Analisis Tingkat Kepuasan Pengguna Aplikasi JMO (Jamsostek Mobile) Menggunakan Algoritma Naive Bayes Classifier,” *Reputasi J. Rekayasa Perangkat Lunak*, vol. 4, no. 1, pp. 33–37, 2023, doi: 10.31294/reputasi.v4i1.1986.
- [29] A. Halim and Andri Safuwani, “Analisis Sentimen Opini Warganet Twitter Terhadap Tes Screening Genose Pendeteksi Virus Covid-19 Menggunakan Metode Naïve Bayes Berbasis Particle Swarm Optimization,” *J. Inform. Teknol. dan Sains*, vol. 5, no. 1, pp. 170–178, 2023, doi: 10.51401/jinteks.v5i1.2229.
- [30] M. F. El Firdaus, N. Nurfaizah, and S. Sarmini, “Analisis Sentimen Tokopedia Pada Ulasan di Google Playstore Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 5, p. 1329, 2022, doi: 10.30865/jurikom.v9i5.4774.