

ALGORITMA *DECISION TREE* UNTUK ANALISIS SENTIMEN PUBLIC TERHADAP *MARKETPLACE* DI INDONESIA

Azril Tazidan Octa.N¹, Ma'mun Hasbullah², Miftahul Rizal³, Muhamad Fauzan Rajab⁴, Nova Agustina⁵
Program Studi Teknik Informatika^{1,2,3,4,5}
Universitas Adhirajasa Reswara Sanjaya^{1,2,3,4}, Sekolah Tinggi Teknologi Bandung⁵
nova@sttbandung.ac.id⁵

Abstrak

Pasar *online* di Indonesia semakin berkembang pesat dan menjadi salah satu destinasi utama bagi pengguna *internet* dan media sosial. Berbagai macam layanan *marketplace* telah tersedia dan diakses oleh mayoritas masyarakat Indonesia. Namun, seiring dengan perkembangan tersebut, kepuasan konsumen terhadap layanan *marketplace* juga beragam, mulai dari yang positif, negatif, hingga netral. Banyak di antara konsumen yang mengekspresikan reaksinya di media sosial, termasuk di Twitter. Dalam penelitian ini, dilakukan analisis opini terhadap *post* yang dilakukan pelanggan bisnis *online* di Indonesia di Twitter dari berbagai macam konsumen. Namun, karena jumlah komentar yang sangat banyak, sulit untuk menyimpulkan pendapat pelanggan tentang situs belanja *online* yang menawarkan layanan terbaik. Bahkan, trending topik di Twitter hanya menampilkan topik hangat yang banyak dibahas tanpa kesimpulan yang jelas. Untuk mengklasifikasikan data opini umum di Twitter dari situs *e-commerce*, langkah pertama adalah memproses data *tweet* menggunakan tools *Rapidminer* untuk mengenali data *tweet* tersebut. Kemudian, digunakan algoritma *decision tree* untuk mengelompokkan data opini. Hasil penelitian menunjukkan bahwa dengan menggunakan *cross validation*, algoritma *decision tree* mencapai akurasi sebesar 70,27 persen, sementara menggunakan *split validation* mencapai akurasi 66,95 persen. Dalam hal ini, akurasi yang lebih baik tercapai dengan menggunakan *cross validation*. Hasil penelitian ini dapat memberikan informasi yang bermanfaat bagi pelaku bisnis *online* di Indonesia untuk meningkatkan kualitas layanan mereka dan meningkatkan kepuasan pelanggan. Selain itu, penelitian ini juga memberikan gambaran tentang pentingnya memanfaatkan algoritma *decision tree* dalam mengelompokkan data opini di media sosial, terutama di Twitter, sebagai alat untuk menganalisis sentimen konsumen terhadap suatu layanan atau produk.

Kata kunci : *Cross validation, Decision Tree, Klasifikasi, Marketplace, Rapidminer, Split validation, Twiter.*

Abstract

Online marketplaces in Indonesia are growing rapidly and have become one of the main destinations for internet and social media users. Various marketplace services are available and accessed by the majority of Indonesians. However, with this development, consumer satisfaction with marketplace services varies, from positive, negative, to neutral. Many consumers express their reactions on social media, including Twitter. In this study, an analysis of opinions was conducted on posts by online business customers in Indonesia on Twitter from various consumers. However, due to the large number of comments, it is difficult to conclude the customer's opinion about online shopping sites that offer the best services. Even trending topics on Twitter only display hot topics that are widely discussed without clear conclusions. To classify general opinion data on Twitter from e-commerce sites, the first step is to process tweet data using Rapidminer tools to recognize the tweet data. Then, the decision tree algorithm is used to categorize opinion data. The results showed that using cross-validation, the decision tree algorithm achieved an accuracy of 70.27 percent, while using split validation, it achieved an accuracy of 66.95 percent. In this case, better accuracy was achieved using cross-validation. The results of this study can provide useful information for online businesses in Indonesia to improve the quality of their services and increase customer satisfaction. In addition, this study also provides an overview of the importance of utilizing the decision tree algorithm in categorizing opinion data on social media, especially on Twitter, as a tool for analyzing consumer sentiment towards a service or product.

Keyword : Cross validation, Decision Tree, Classification, Marketplace, Rapidminer, Split validation, Twiter.

I. PENDAHULUAN

Banyak pengguna internet dan media sosial di Indonesia yang telah mengalami kemajuan dalam mengakses pasar *online*. Pasar *online* ini merupakan destinasi populer bagi pengguna media sosial di Indonesia untuk membeli berbagai barang dan layanan. Selain itu, Twitter juga menjadi *platform* media sosial yang sangat populer di Indonesia. Banyak orang menggunakan Twitter untuk berbicara tentang pengalaman mereka saat melakukan transaksi belanja *online*. Dalam media sosial ini, pengguna dapat mengungkapkan pendapat mereka secara langsung dan *real-time* melalui pesan atau *tweet* yang diposting. Karena popularitasnya, Twitter telah menjadi *platform* yang penting dalam mengevaluasi dan menganalisis opini masyarakat tentang pasar *online* di Indonesia. Penelitian ini, dilakukan analisis terhadap opini pelanggan bisnis *online* di Indonesia melalui unggahan di Twitter. Oleh karena itu, Twitter menjadi

salah satu media sosial yang paling penting dalam mengukur kepuasan pelanggan dan kualitas layanan pasar *online* di Indonesia [1].

Banyak orang menggunakan situs media sosial untuk mengekspresikan pendapat mereka dari suatu tempat. Ada banyak situs media sosial populer seperti Facebook, Twitter, Instagram, dan masih banyak. Salah satunya media sosial yang paling sering digunakan untuk menyampaikan pendapat adalah Twitter. Berbagai macam konsumen tidak dapat menyimpulkan pendapat pelanggan tentang situs belanja *online* di Twitter. Situs E-niaga mana yang menawarkan layanan terbaik, karena jumlahnya yang banyak. Komentar di Twitter bahkan trending topik di Twitter hanya menampilkan topik hangat banyak dibahas tanpa kesimpulan.

Analisis sentimen juga dikenal sebagai sebuah proses. Anda akan menemukan pendapat pengguna tentang berbagai topik atau teks yang dikirimkan oleh pengguna atau proses menentukan apakah tulisan itu positif, negatif atau netral. Studi ini berkaitan dengan pembuatan analisis opini dari data opini Pelanggan pasar *online* di Twitter digunakan untuk menentukan apakah ada data opini. Berikan umpan balik positif atau negatif sehingga nantinya dapat digunakan untuk menentukan halaman pembelian dinilai baik secara *online*[2].

Penelitian ini menggunakan algoritma *decision tree* karena salah satu algoritma yang paling populer dan banyak digunakan oleh peneliti. Algoritma pohon keputusan merupakan algoritma yang dapat digunakan untuk membentuk pohon keputusan. Sebuah pohon keputusan adalah sebuah metode cukup mudah bagi orang untuk menafsirkan. Pohon Keputusan adalah model prediksi yang menggunakan struktur-struktur pohon atau struktur hirarki. Konsep pohon keputusan mengubah data menjadi pohon keputusan dan aturan keputusan. Keuntungan utamanya adalah kerusakan mempermudah proses pengambilan keputusan yang kompleks lebih mudah bagi pengambil keputusan. Lebih mudah menemukan solusi untuk masalah yang ada[3].

Penelitian ini bertujuan untuk mengimplementasikan Algoritma *decision tree* sebagai model klasifikasi sentimen pengguna media sosial Twitter dalam konteks acara belanja di *marketplace*, dengan kategori opini positif, negatif, dan netral. Model yang dikembangkan dalam penelitian ini dapat memberikan gambaran tentang tren opini publik di setiap pasar, sehingga dapat membantu pelaku bisnis *online* dalam memahami persepsi pelanggan terhadap layanan mereka. Penelitian ini, data opini publik diambil dari Twitter dan diproses menggunakan alat Rapidminer untuk mengidentifikasi pola opini yang muncul. Algoritma *decision tree* digunakan untuk mengklasifikasikan opini tersebut ke dalam kategori positif, negatif, atau netral. Dengan demikian, model yang dikembangkan dapat memberikan informasi yang berguna bagi pelaku bisnis *online* dalam memahami sentimen pelanggan dan mengambil tindakan yang tepat untuk meningkatkan kualitas layanan mereka. Selain itu, penelitian ini juga diharapkan dapat memberikan wawasan tentang penggunaan algoritma *decision tree* sebagai alat untuk menganalisis data opini di media sosial, khususnya di Twitter, dalam rangka memahami tren opini publik terkait suatu produk atau layanan [4].

II. TINJAUAN PUSTAKA

1. Analisis Sentimen

Analisis sentiment atau disebut juga opinion mining dalam bahasa Indonesia adalah teknik atau cara digunakan untuk mengidentifikasi bagaimana pendapat diungkapkan melalui teks dan bagaimana perasaan ini dapat diklasifikasikan sebagai perasaan positif, negatif dan netral[5]. Dalam bidang penelitian *data mining* ditujukan untuk menganalisis, memahami, mengolah, dan mengekstraksi data tekstual yang dapat berupa opini, sentimen, evaluasi, sikap, produk, jasa, organisasi tertentu, individu, masalah [6].

Fungsi analisis sentimen menggolongkan setiap kutub teks di berbagai sumber Internet dan media sosial dalam bentuk dokumen atau kalimat dan menentukan apakah kata-kata termasuk dalam kategori positif, netral, atau negatif. Analisis sentimen juga dapat mengungkapkan emosi seperti kesedihan, kegembiraan, dan kemarahan. Hal pertama yang harus dilakukan selama proses dokumen adalah membagi satu karakter menjadi kata-kata, yang biasa dikenal sebagai "*encoding*" Dalam tokenisasi. Sebuah kalimat dibagi menjadi beberapa bagian dengan kata-kata. Bagian kata ini disebut *token* [7].

Analisis sentimen untuk proses pemahaman, penggalian, dan pengolahan data teks secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam opini. Penelitian ini, analisis sentimen dilakukan untuk mengetahui apakah opini seseorang merupakan ujaran kebencian. Ujaran kebencian dapat dimasukkan dalam kategori opini negatif. Besarnya dampak dan manfaat analisis sentimen membuat penelitian dan aplikasi berbasis analisis sentimen berkembang pesat. Ada sekitar 20 hingga 30 perusahaan di Amerika Serikat yang berfokus pada analisis sentimen [8].

2. Klasifikasi

Klasifikasi adalah proses pencarian pola berdasarkan kelas yang digunakan untuk membedakan antara satu kelas dengan kelas lainnya. Pohon keputusan merupakan aplikasi metode klasifikasi paling populer yang memungkinkan objek dikelompokkan dan dimodelkan dalam pohon keputusan sehingga mudah dipahami[9]. Klasifikasi adalah bentuk analisis data yang mengekstraksi model untuk mendeskripsikan kelas data yang penting. Model seperti itu disebut klasifikasi. Tujuannya adalah untuk memprediksi kelas label kategori (diskrit, tidak terurut)[6]. Klasifikasi adalah teknik penambangan data yang memetakan data ke dalam kelompok atau kelas yang telah ditentukan sebelumnya.

Klasifikasi adalah metode pembelajaran terawasi yang memerlukan data pelatihan berlabel untuk menghasilkan aturan yang mengklasifikasikan data uji ke dalam grup atau kelas yang telah ditentukan sebelumnya. Metode klasifikasi mengacu pada penerapan algoritma yang diketahui ke gudang data yang disurvei untuk membentuk grup data [10]. Metode ini berguna untuk proses bisnis yang memerlukan informasi kategori seperti pemasaran atau penjualan. Berbagai algoritma seperti tetangga terdekat, pembelajaran pohon keputusan, dapat digunakan. Pohon keputusan juga berguna untuk mengeksplorasi data dan menemukan hubungan tersembunyi antara sekumpulan kandidat variabel input dan variabel target. Pohon keputusan menggabungkan eksplorasi data dan pemodelan dan merupakan langkah awal yang bagus dalam proses pemodelan, bahkan ketika digunakan sebagai model akhir untuk teknik lain [10].

3. *Decision Tree*

Algoritma *decision tree* merupakan salah satu algoritma yang sering digunakan dalam pengambilan keputusan. Pohon keputusan mencari solusi untuk masalah dengan mengubah kriteria menjadi node yang terhubung untuk membentuk struktur seperti pohon. Sebuah pohon keputusan juga merepresentasikan model prediktif untuk sebuah keputusan dengan struktur hirarki atau pohon. Setiap pohon memiliki cabang, cabang tersebut mewakili atribut yang harus dipenuhi untuk mencapai cabang berikutnya hingga berakhir pada daun (tidak ada cabang lagi)[11].

Decision tree teknik pemodelan prediktif yang dapat digunakan untuk klasifikasi dan prediksi tugas. Pohon keputusan menggunakan teknik "bagi dan taklukkan" untuk membagi ruang pencarian masalah menjadi kumpulan masalah. Proses pohon keputusan adalah memformat ulang data tabular menjadi model pohon. Pohon model menghasilkan aturan dan disederhanakan[10].

Decision tree adalah struktur data yang terdiri dari *node* dan *edge*. Simpul pohon dibagi menjadi tiga bagian, yaitu simpul akar (akar/simpul), simpul cabang/internal (cabang/internal simpul) dan simpul daun. Pohon keputusan adalah presentasi sederhana metode klasifikasi untuk banyak kelas, di mana simpul internal dan simpul akar ditandai dengan nama atribut, rusuk-rusuk diberi diberi label nilai atribut yang mungkin dan simpul daun ditandai dengan kelas- kelas yang berbeda [12].

4. *Twitter*

Twitter adalah salah satu media sosial yang paling banyak digunakan untuk komunikasi dan informasi. Melalui media sosial ini, pengguna dapat mengungkapkan berbagai jenis pendapat dan komentar tentang suatu topik. Opini dan komentar pengguna di *tweet*. *Twitter* media dan layanan *social microblogging* yang memberdayakan penggunaannya untuk mengirim berita *real-time*. *Tweet* adalah Pesan singkat dengan panjang karakter terbatas maksimal 140 karakter. Pembatasan karakter yang akan ditulis, *tweet* sering mengandung singkatan, Bahasa-bahasa gaul dan kesalahan ketik[13].

Media sosial adalah tempat orang berbagi tentang diri mereka sendiri dan perspektif mereka tentang kehidupan. Menyadari bahwa banyak informasi dapat digali dari media sosial, beberapa penelitian telah dilakukan dengan menggunakan media sosial sebagai objek. Ini juga mencakup penelitian tentang penyalahgunaan media sosial. Dengan kemajuan informasi yang pesat, orang biasa dipermudah oleh teknologi yang dapat memperburuk keadaan. Pesan lelucon dan ujaran kebencian dapat menyebar dengan sangat mudah karena kehadiran media sosial seperti Facebook, Twitter, atau aplikasi ponsel seperti *WhatsApp*, *Line*[14].

Media sosial khususnya Twitter, telah menjadi alat komunikasi yang sangat populer di kalangan pengguna internet karena memberikan kebebasan untuk mengirimkan pesan dalam jumlah karakter yang terbatas. Pada konferensi pengembang resmi Twitter Chirp 2010, perusahaan merilis statistik tentang situs Twitter dan penggunaannya yang menunjukkan pertumbuhan yang pesat. Menurut statistik ini, pada April 2010 Twitter memiliki 106 juta akun dengan lebih dari 180 juta pengguna unik per bulan. Pengguna Twitter memiliki kecenderungan untuk membagikan informasi dengan cepat dan mudah hal ini menyebabkan Twitter menjadi sumber informasi yang sangat populer. Dengan perkiraan pertumbuhan sebanyak 300.000 pengguna setiap hari, Twitter akan terus menjadi *platform* komunikasi yang penting dalam masyarakat digital [6].

5. *RapidMiner*

RapidMiner adalah aplikasi atau perangkat lunak yang sedang berjalan sebagai sarana pembelajaran dalam ilmu *data mining* yang sangat populer dan diakui secara internasional. Aplikasi ini dikembangkan oleh perusahaan yang berdedikasi untuk menyediakan solusi dalam langkah-langkah pengolahan data besar dalam operasi bisnis, penelitian, pendidikan, pelatihan, dan pembelajaran. Selain itu, *RapidMiner* memiliki lebih dari 100 solusi pembelajaran, termasuk pengelompokan data (*clustering*), klasifikasi, dan analisis regresi, yang menjadikannya salah satu perangkat lunak terbaik dalam industri *data mining* [15].

Rapidminer adalah *platform* penambangan data Informasi yang dikembangkan oleh perusahaandengan nama yang sama yang menyediakan lingkungan dikombinasikan dengan pembelajaran mesin (*machine learning*), pembelajaran mendalam, penambangan teks (*Text Mining*) dan analisis prediktif (Prediktif analitik). Aplikasi ini digunakan untuk aplikasi komersial dan ekonomi serta penelitian, pendidikan, pelatihan, pembuatan prototipe

cepat untuk pengembangan aplikasi dan mendukung semua fase teknik pembelajaran mesin, termasuk persiapan data, visualisasi, validasi dan optimalisasi hasil. *Rapidminer* dikembangkan menggunakan model *open-core* [16]. *Rapidminer* adalah perangkat lunak pengolah data. Menggunakan prinsip data dan algoritma *rapidminer* mengekstrak pola dari *big data* dengan menggabungkan metode statistik, kecerdasan buatan dan *database*. *Rapidminer* memudahkan pengguna menghitung data dalam jumlah besar menggunakan operator. Operator ini digunakan untuk mengubah data. Data digabungkan operator simpul lalu tambahkan ke simpul hasil untuk melihat hasilnya [17].

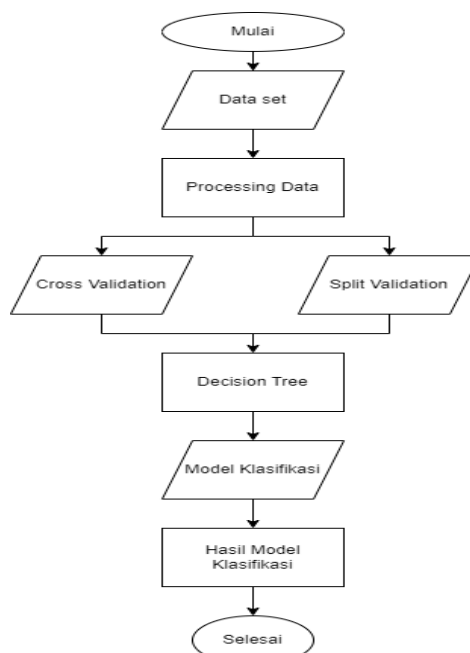
6. Marketplace

Marketplace adalah model bisnis baru yang berkembang dengan pesatnya perkembangan infrastruktur teknologi. Pasar ini dirancang untuk meminimalkan proses bisnis yang kompleks menjadi efisiensi. Dengan adanya *marketplace* ini, siapapun dapat melakukan jual beli dengan mudah, cepat dan murah karena tidak ada batasan ruang atau jarak dan waktu. Secara tradisional, pasar memiliki beberapa fungsi, termasuk memfasilitasi transaksi dan menyediakan infrastruktur [18]. *Marketplace* adalah pasar lingkungan *online* (berbasis web), tempat untuk melakukan bisnis dan transaksi di antara mereka pembeli dan penjual. Pembeli dapat mencari pemasok semaksimal mungkin sesuai dengan kriteria yang diinginkan sesuai dengan harga pasar. Menurut Brun Jensen dan Skovgaard, *marketplace* adalah forum bisnis interaktif Elektronik yang menawarkan pasar dimana perusahaan dapat berpartisipasi dalam belanja *online* antara B2B atau toko elektronik lainnya. Dari beberapa definisi disebutkan bahwa pasar merupakan sarana pemasaran produk elektronik yang menyatukan banyak pemasok Pembeli melakukan bisnis satu sama lain [19].

Marketplace memiliki konsep yang kurang lebih sesuai dengan pasar tradisional. *Marketplace* tidak bertanggung jawab atas barang yang dijual, karena *marketplace* hanya menyediakan tempat bagi penjual yang ingin berjualan dan membantu penjual untuk bertemu pelanggan dan memfasilitasi transaksi. Bisnis diatur oleh pasar, setelah itu penjual mengirimkan produk ke pembeli setelah menerima pembayaran. Salah satu alasan mengapa *marketplace* terkenal adalah kemudahan penggunaan dan kenyamanannya. Banyak yang menggambarkan toko *online* sebagai department store [20].

III. METODE DAN LANGKAH PENELITIAN

Hasil tes analisis sentimen yang terbaik, terdapat beberapa tahapan yang harus dilakukan. Pertama, pengumpulan data opini atau pendapat publik yang terkait dengan topik yang ingin dianalisis. Kedua, pemrosesan data untuk membersihkan data dari *noise* atau informasi yang tidak relevan dengan topik. Selanjutnya, dilakukan klasifikasi sentimen untuk menentukan apakah pendapat publik bersifat positif, negatif, atau netral. Tahapan kelima, dilakukan evaluasi performa algoritma yang digunakan dengan menggunakan metrik evaluasi seperti akurasi, *presisi* dan *recall*. Akhirnya, dilakukan interpretasi hasil analisis sentimen untuk mengambil kesimpulan dan informasi yang berguna dalam pengambilan keputusan atau pengembangan produk atau layanan. Semua tahapan ini sangat penting dilakukan secara sistematis dan teliti untuk menghasilkan hasil analisis sentimen yang akurat dan dapat diandalkan.



Gambar 1. Metode klasifikasi Sentimen *Public* Terhadap *Marketplace* Di Indonesia

1. Pengumpulan Data

Langkah pertama adalah melakukan proses analisis sentiment mengumpulkan informasi. Di dalam penelitian ini, data berasal dari Twitter dan mendapatkan informasi dari Twitter.

2. Pelabelan Data

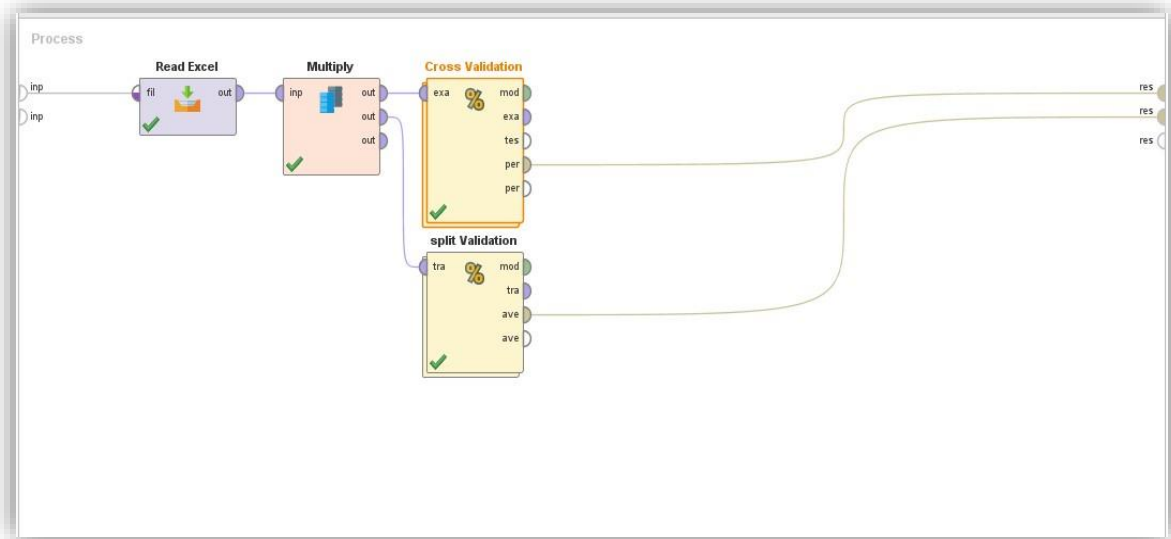
Setelah melakukan proses analisis sentiment mengumpulkan informasi, langkah berikutnya adalah melakukan pengumpulan data dari berbagai sumber yang relevan dengan tujuan penelitian. Dalam penelitian ini, data dikumpulkan dari Twitter menggunakan teknik pengambilan sampel acak dari sejumlah tweet yang berkaitan dengan tema acara belanja di *marketplace*. Proses pengumpulan data ini dilakukan dengan memanfaatkan API Twitter dan menggunakan kata kunci tertentu untuk mengidentifikasi *tweet* yang relevan dengan topik penelitian. Data yang diperoleh kemudian diolah dan disimpan dalam format yang dapat diakses oleh algoritma *decision tree* untuk melakukan analisis sentiment. Kategori sentiment yang banyak digunakan yaitu ada tiga kelas yaitu positif, negatif dan netral. *Dataset* yang dikumpulkan dapat dilihat pada Tabel I.

TABEL I
DATASET DAN LABEL DATA

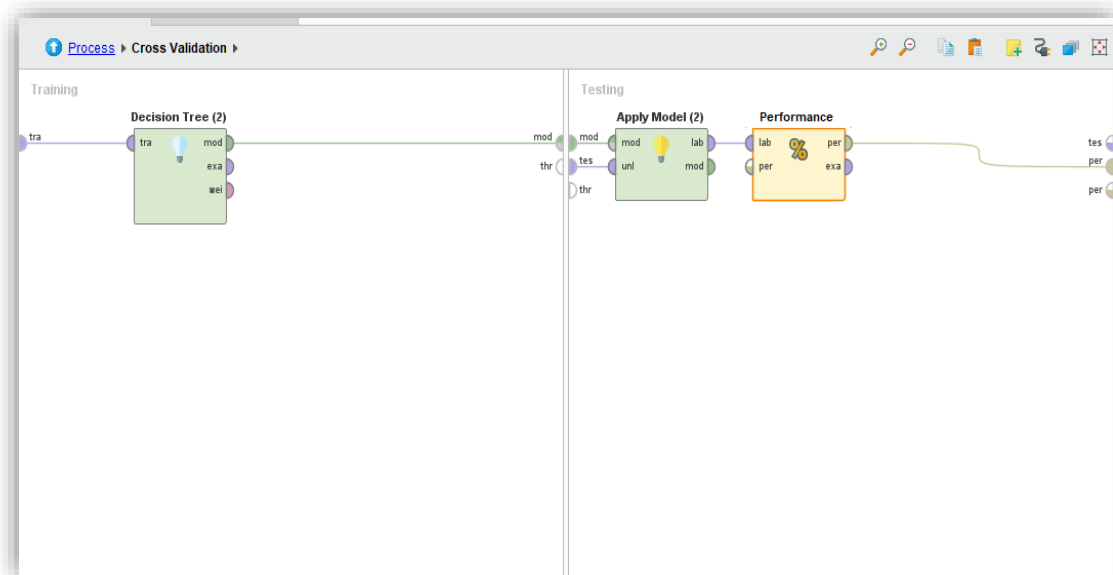
ID	ANALISIS SENTIMEN	POST TWEET
1	POSITIVE	Saran aku beli di <i>official</i> storenya jgn beli di marketplace karena ga bisa jamin keaslian dan <i>expdate</i> nya, mereka cuma jual official di http://hdistore.com atau di center HDI seluruh Indonesia. Googling aja sesuai dom kalian
2	POSITIVE	Sebagai perusahaan teknologi lokal buatan Indonesia, Tokopedia berusaha untuk mengakselerasi transformasi digital UMKM di tanah air, termasuk di daerah.
3	POSITIVE	Mari ramaikan dengan memberi dukungan kepada UMKM Indonesia lewat UMKM EXPO (RT) Brilianpreneur 2022 dari tanggal 14-18 Desember 2022 di JCC atau juga bisa beli produknya lewat <i>marketplace</i> loh seperti Blibli.
4	POSITIVE	Kartu kredit terbaru ini hadir untuk melengkapi kebutuhan alat pembayaran bagi penggunamarketplace di seluruh Indonesia
5	POSITIVE	Mewakili Bapak Presiden RI '@Jokowi' saya hadir pada event UMKM EXPO (RT) Brilianpreneur 2022. Alhamdulillah, saat ini sudah lebih dari 21 juta total UMKM yang onboarding ke marketplace.
6	POSITIVE	Sobat, selain lebih praktis, perkembangan marketplace yang begitu pesat, dan beragam jenis pembayaran yang terjamin membuat aktivitas belanja <i>online</i> di Indonesia terus meningkat.
7	POSITIVE	Pos Indonesia punya kurir perempuan, namanya O Ranger Mawar. Gesit, cepat, aman & tepat waktu. Komitmen @PosAjaOfficial utk membuat konsumen nyaman. Ada baiknya @tiktok @ShopeeID @LazadaID menjadikan Pos Aja sebagai mitra,
8	NEGATIVE	Sudah cukup kelihatan kalau partai ini tengah menjaring wajah-wajah "centang biru" baru. Kalacuma berusaha menggaet suara baru, mungkin bisa dicoba cara-cara <i>marketplace online</i> di Indonesia: coba ajak member K-Pop atau aktor/aktris K-Drama.
9	POSITIVE	yuk mulai dari sekarang, kita dukung produk UMKM Indonesia dengan beli produk UMKM Brilianpreneur 2022 di <i>marketplace</i> kesukaan kalian!
10	POSITIVE	Lalu untuk produk <i>business to consumer</i> (B2C), yaitu aplikasi Goers dan <i>website</i> http://Goersapp.com , yang merupakan <i>marketplace</i> tiket destinasi wisata serta event sebagai <i>directory activity</i> , berbagai promo & tour yang dikurasi berdasarkan preferensi pada aplikasi
...
1165	POSITIVE	tokopedia agar memudahkanku mengelola jualan <i>online</i> di marketplace dan socmed, selain itu speknya Huawei juga no kaleng-kaleng udah gak perlu diraguin lagi ??

3. Analisis dan Pengujian Data

Setelah data berhasil diproses, langkah selanjutnya dalam penelitian ini adalah melakukan eksperimen dan pengujian dengan menggunakan metode *decision tree*. Pengujian ini, digunakan dua jenis metode yaitu *cross validation* dan *split validation* untuk menguji akurasi dari model klasifikasi yang digunakan. Aplikasi yang digunakan dalam pengujian ini adalah Rapidminer, yang merupakan *platform* analisis data yang populer. Dengan menggunakan metode *decision tree* dan aplikasi *Rapidminer*. Diharapkan dapat ditemukan model klasifikasi yang paling akurat dalam mengklasifikasikan sentimen pengguna media sosial Twitter terhadap tema acara belanja di *marketplace*, dengan kategori opini positif, negatif, dan netral.



Gambar 2. Pengujian dengan *Cross Validation* dan *Split Validation*



Gambar 3. Pemodelan dengan Metode *Decision Tree*.

IV. HASIL PENELITIAN

Pada gambar 4, terlihat proses evaluasi dan validasi metode klasifikasi *Decision Tree* yang terdiri dari dua bagian dan tiga tampilan proses. Bagian pertama adalah *Training*, dimana proses ini menggunakan metode *Decision Tree* untuk mengklasifikasikan data opini ke dalam kategori positif, negatif, dan netral. Bagian kedua adalah *Testing*, dimana model yang sudah dilatih diuji dan performanya dievaluasi untuk mendapatkan nilai akurasi yang optimal. Proses pengujian dilakukan dengan menggunakan metode *cross validation* dan *split validation* untuk memastikan kehandalan model klasifikasi yang telah dibuat.

PerformanceVector (cross Performance)

Table View Plot View

accuracy: 70.27% +/- 3.46% (micro average: 70.27%)

	true POSITIVE	true NEGATIVE	true NETRAL	true POSITIVE	true NEGATIVE	true NETRAL	true NEGATIVE'	class precision
pred. POSITIVE	711	151	21	0	0	0	0	80.52%
pred. NEGATIVE	64	84	54	1	0	0	1	41.18%
pred. NETRAL	9	9	5	0	0	0	0	21.74%
pred. POSITIVE	0	0	0	2	4	5	0	18.18%
pred. NEGATIVE	1	1	0	8	14	10	0	41.18%
pred. NETRAL	0	0	1	2	4	2	0	22.22%
pred. NEGATIVE'	0	0	0	0	0	0	0	0.00%
class recall	90.57%	34.29%	6.17%	15.38%	63.64%	11.76%	0.00%	

Gambar 4. Hasil Akurasi *Cross Validation*

PerformanceVector (split performance)

Table View Plot View

accuracy: 66.95%

	true POSITIVE	true NEGATIVE	true NETRAL	true POSITIVE	true NEGATIVE	true NETRAL	true NEGATIVE'	class precision
pred. POSITIVE	196	31	1	0	0	0	0	85.96%
pred. NEGATIVE	31	30	16	4	7	5	0	32.26%
pred. NETRAL	8	12	7	0	0	0	0	25.93%
pred. POSITIVE	0	0	0	0	0	0	0	0.00%
pred. NEGATIVE	0	0	0	0	0	0	0	0.00%
pred. NETRAL	0	0	0	0	0	0	0	0.00%
pred. NEGATIVE'	0	0	0	0	0	0	0	0.00%
class recall	83.40%	41.10%	29.17%	0.00%	0.00%	0.00%	0.00%	

Gambar 5. Hasil Akurasi *Split Validation*

Pada gambar 4 dan 5 yaitu menunjukkan hasil akurasi setelah dilakukan pengujian dengan metode *Decision Tree* yang menghasilkan nilai akurasi pada *Cross Validation* sebesar 70,27% dan nilai akurasi pada *Split Validation* sebesar 66,95%. Dengan begitu nilai akurasi pada *Cross Validation* lebih unggul dibandingkan dengan *Split validation*.

V. KESIMPULAN

Bahwa *marketplace* merupakan pasar yang di mana para penjual dan para pembeli menukar jasa dan barang- barang untuk uang atau untuk jasa dan barang-barang lainnya. Twitter termasuk salah satu media sosial yang paling banyak digunakan untuk menyampaikan pendapat tentang transaksi pemasaran, media sosial ini juga yang memungkinkan pengguna untuk memposting pesan secara *real time*. Twitter menjadi salah satu media sosial yang masuk ke sebuah proses analisis sentimen sehingga bisa menemukan pendapat pengguna tentang berbagai topik atau teks yang dikirimkan oleh pengguna atau proses menentukan apakah tulisan itu positif, negatif atau netra. Data opini menjadi langkah utama untuk mengklasifikasikan nya sehingga algoritma yang di pakainya adalah algoritma *decision tree*. Hasil dari algoritma tersebut menggunakan *Cross validation* mencapai akurasi sebesar 70,27% dan menggunakan *Split validation* mencapai akurasi sebesar 66,95%. Nilai akurasi pada *Cross validation* lebih unggul dibandingkan dengan *Split validation*.

REFERENSI

- [1] I. Saputra *dkk.*, “Analisis Sentimen Pengguna Marketplace Bukalapak dan Tokopedia di Twitter Menggunakan Machine Learning,” *Faktor Exacta*, vol. 13, no. 4, hlm. 200, Feb 2021, doi: 10.30998/faktorexacta.v13i4.7074.
- [2] A. Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes Muljono, D. Putri Artanti, A. Syukur, A. Prihandono, dan D. I. Rosal Moses Setiadi, “Konferensi Nasional Sistem Informasi 2018 STMIK Atma Luhur Pangkalpinang,” 2018. [Daring]. Available: <http://twitter.com>
- [3] A. H. Nasrullah, “IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI PRODUK LARIS,” vol. 7, no. 2, 2021, [Daring]. Available: <http://ejournal.fikom-unasman.ac.id>
- [4] J. Homepage, A. Harun, dan D. P. Ananda, “MALCOM: Indonesian Journal of Machine Learning and Computer Science Analysis of Public Opinion Sentiment About Covid-19 Vaccination in Indonesia Using Naïve Bayes and Decision Tree Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naïve Bayes dan Decision Tree,” vol. 1, hlm. 58–63, 2021.
- [5] T. Fadiyah Basar, D. E. Ratnawati, dan I. Arwani, “Analisis Sentimen Pengguna Twitter terhadap Pembayaran Cashless menggunakan ShopeePAY dengan Algoritma Random Forest,” 2022. [Daring]. Available: <http://j-ptiik.ub.ac.id>
- [6] F. N. Zuhri, A. Alamsyah, dan S. Si, “ANALISIS SENTIMEN MASYARAKAT TERHADAP BRAND SMARTFREN MENGGUNAKAN NAIVE BAYES CLASSIFIER DI FORUM KASKUS PUBLIC SENTIMENT ANALYSIS OF SMARTFREN BRAND USING NAIVE BAYES CLASSIFIER ON KASKUS FORUM.”
- [7] J. Minfo Polgan *dkk.*, “Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naïve Bayes Classifier”, [Daring]. Available: www.tokopedia.com
- [8] G. A. Buntoro, “ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE,” Sep 2016.
- [9] D. Sartika dan D. I. Sensuse, “Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian,” 2017.
- [10] S. Bahri dan A. Lubis, “METODE KLASIFIKASI DECISION TREE UNTUK MEMREDIKSI JUARA ENGLISH PREMIER LEAGUE,” vol. 2, no. 1, 2020.
- [11] J. Fadlil, W. Firdaus, dan J. Matematika, “Pembuatan Sistem Rekomendasi Menggunakan Decision Tree dan Clustering.” [Daring]. Available: www.moviens.org.
- [12] A. H. Nasrullah, “IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI PRODUK LARIS,” vol. 7, no. 2, 2021, [Daring]. Available: <http://ejournal.fikom-unasman.ac.id>
- [13] S. Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen di Twitter dan U. Gajah Mada, “Adaptive online learning system based on gaze behavior View project Nurrun Muchammad Shiddieqy Hadna.” [Daring]. Available: <https://www.researchgate.net/publication/292831965>
- [14] B. I. Nugroho, Z. Ma'arif, dan Z. Arif, “Tinjauan Pustaka Sistematis: Penerapan Data Mining Metode Klasifikasi Untuk Menganalisa Penyalahgunaan Sosial Media,” 2022. [Daring]. Available: www.journal.peradaban.ac.id
- [15] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, dan C. Lauw, “Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Linear Regression,” *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 7, no. 1, hlm. 8–17, Mei 2021, doi: 10.25077/teknosi.v7i1.2021.8-17.
- [16] E. Fadilah, “Implementasi Metode Profile Matching Terhadap Sistem Pendukung Keputusan Penerimaan Dana Zakat pada Badan Amil Zakat Pertamina (BAZMA),” *MATICS*, vol. 10, no. 2, hlm. 39, Mar 2019, doi: 10.18860/mat.v10i2.5745.
- [17] Ainurrohmah, “Akurasi Algoritma Klasifikasi pada Software Rapidminer dan Weka,” 2021.
- [18] R. Yustiani, R. Yunanto, P. Studi Manajemen, dan P. Studi Komputerisasi Akuntansi, “PERAN MARKETPLACE SEBAGAI ALTERNATIF BISNIS DI ERA TEKNOLOGI INFORMASI,” *Ilmiah Komputer dan*, vol. 6, no. 2, 2017.
- [19] D. Apriadi, A. Y. Saputra, S. Informasi, S. Bina, dan N. Jaya, “E-Commerce Berbasis Marketplace Dalam Upaya Mempersingkat Distribusi Penjualan Hasil Pertanian,” vol. 1, no. 2, hlm. 131–136, 2017, [Daring]. Available: <http://jurnal.iaii.or.id>
- [20] E. Iskandar, R. Cokro Buwono, S. Ogi Nindiya Putri, dan P. Studi Informatika STMIK Akakom Jln Raya Janti, “IMPLEMENTASI PROGRESSIVE WEB APPS PADA MARKETPLACE.”