KLASIFIKASI SENTIMEN NETIZEN DI MEDIA SOSIAL X TERHADAP ANCAMAN RESESI EKONOMI INDONESIA DENGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES

Endang Sri Palupi
Teknologi Informasi, Universitas Bina Sarana Informatika
Jalan Kramat Raya no 98, Kwitang, Jakarta Pusat
endang.epl@bsi.ac.id

ABSTRAK

Perekonomian dunia saat ini dibayangi oleh ketidakpastian dengan berbagai macam faktor. Indonesia sempat menghadapi resesi sebanyak tiga kali yakni pada 1963, 1998, dan 2020/2021. Ketiga krisis tersebut dipicu penyebab yang berbeda dan dengan dampak yang berbeda pula. Ancaman resesi yang dipicu faktor eksternal antara lain disebabkan ketidakstabilan geopolitik, kebijakan ekonomi Donald Trump di Amerika Serikat, serta perang dagang antara beberapa negara besar menjadi faktor utama yang dapat memengaruhi pertumbuhan ekonomi nasional. Sedangkan kondisi internal yang memperburuk situasi antara lain beban utang jatuh tempo pemerintah pada 2025 mencapai Rp 800,33 triliun. Ditandai dengan daya beli yang semakin menurun dan gelombang PHK yang semakin besar, komentar netizen di media sosial X dengan tagar Resesi Indonesia trending. Dalam penelitian ini, klasifikasi sentimen netizen X mengenai ancaman resesi ekonomi di Indonesia mengambil data komentar negatif dan positif dari netizen pada tahun 2022 sampai dengan 2023 dengan teknik crawling dan text mining. Hasil penelitian menunjukkan hasil akurasi sebesar 81.71% dengan *class recall true positif* 70,59% dan *true negatif* 89,58%. Hasil penelitian menyatakan metode naïve bayes classifier cukup akurat untuk bisa mengatasi klasifikasi sentimen masyarakat di media sosial X terhadap ancaman resesi ekonomi di Indonesia.

Kata kunci : *klasifikasi, naïve bayes, resesi*

1. PENDAHULUAN

Penurunan pertumbuhan ekonomi yang diperkirakan akan terjadi pada tahun 2023 baik pada negara berkembang atau negara maju, dan perekonomian global dapat berdampak negatif terhadap sektor riil. Hal tersebut dikarenakan menurunnya pertumbuhan ekonomi akan menyebabkan terjadinya penurunan omset penjualan yang akan menurunkan keuntungan perusahaan. Penurunan keuntungan dari perusahaan apabila terjadi terus menerus akan berdampak terhadap penutupan operasional perusahaan tersebut sehingga terjadi pemutusan hubungan kerja Hal tersebut akan berdampak terhadap penurunan konsumsi masyarakat.[1] Pengguna media sosial telah tersebar luas di berbagai lapisan masyarakat dan menyebar dengan sangat cepat. Tidak saja menjadi alat untuk bersosialisasi dan bercengkrama, namun juga untuk memberitahu keinginan dan memvisualisasikan peristiwa dan perasaan masyarakat, diantara jejaring sosial salah satu yang paling populer di masyarakat adalah Twitter (Saat ini berganti nama menjadi X). X adalah media sosial yang menghubungkan banyak orang di seluruh dunia dalam berbagai topik. Dengan 328 juta pengguna, X adalah media sosial paling populer di antara berbagai media sosial. Jumlah pengguna X dapat digunakan untuk mengetahui pendapat masyarakat tentang resiko resesi ekonomi. Perihal ini dapat dijadikan sebagai model guna mengetahui pendapat masyarakat tentang dampak masyarakat terhadap resesi ekonomi. [2]

Klasifikasi merupakan suatu proses pengkategorian yang digunakan untuk menentukan kelas dari data yang tidak di ketahui label class nya. Memprediksi label kelas adalah tujuan utama dari klasifikasi ini. Dua tahapan yang dimiliki pada poses klasifikasi, yaitu terlebih dahulu learning yang mana learning ini menganalisa sebuah data training dengan menggunakan algoritma klasifikasi, lalu proses klasifikasi yaitu memprediksi ketepatan dari klasifikasi dengan menggunakan data testing. Klasifikasi data dapat dimulai dengan membuat aturan klasifikasi tertentu menggunakan data latih dan data uji. Metode Naïve Bayes adalah metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yang berakar pada teorema bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. Naïve Bayes memiliki tingkat akurasi yang lebih baik dan keuntungan menggunakan metode ini kita hanya membutuhkan jumlah data training yang kecil untuk menentukan estimasi parameter yang dibutuhkan dalam proses klasifikasi. Algoritma Naïve Bayes sangat cocok untuk melakukan klasifikasi pada data set yang bertipe nominal. [3] Dalam penelitian ini penulis menggunakan *framework* khusus, karena semua fasilitas sudah disediakan. *Rapid Miner* dikhususkan untuk penggunaan *data mining*. *RapidMiner* memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. *RapidMiner* merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. *RapidMiner* ditulis dengan munggunakan bahasa java sehingga dapat bekerja di semua sistem operasi. *RapidMiner* sebelumnya

bernama YALE (*Yet Another Learning Environment*), dimana versi awalnya mulai dikembangkan pada tahun 2001 oleh RalfKlinkenberg, Ingo Mierswa, dan Simon Fischer di *Artificial Intelligence* Unit dari University of Dortmund. RapidMiner didistribusikan di bawah lisensi AGPL (*GNU Affero General Public License*) versi 3. Hingga saat ini telah ribuan aplikasi yang dikembangkan mengunakan *RapidMiner* di lebih dari 40 negara. *RapidMiner* sebagai *software open sourc*e untuk *data mining* tidak perlu diragukan lagi karena *software* ini sudah terkemuka di dunia. [4]

Algoritma *Naïve bayes* memiliki banyak manfaat sehingga banyak digunakan di berbagai aspek kehidupan. Antara lain sebagai berikut:

- *Real time prediction*
- *Multiclass prediction*, disini *Naïve Bayes* dapat memprediksi probabilitas beberapa kelas variabel target
- *Text classification*, selain itu *Naïve Bayes* juga sering digunakan utnuk *membuat text classification* karena memiliki tingkat keberhasilan yang lebih tinggi dibandingkan dengan algoritma lain
- *Recommendation system*, *Naïve Bayes* juga langganan digunakan untuk proyek *data mining* untuk menyaring informasi yang tidak terlihat dan memprediksi apakah pengguna menginginkan sumber daya yang diberikan atau tidak.[5]

Penelitian ini bertujuan untuk mengklasifikasi sentiment masyarakat Indonesia khususnya netizen media sosial X mengenai ancaman resesi ekonomi di Indonesia menggunakan algoritma *Naïve Bayes*. Dalam konteks ini, metode *Naïve Bayes* dapat diterapkan untuk mengklasifikasikan reaksi netizen terhadap resesi ekonomi di Indonesia menjadi kategori positif atau negatif. Analisis ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai bagaimana pandangan dari masyarakat secara umum, serta dapat memberikan masukan bagi pembuat kebijakan dalam merancang program terkait perekonomian di masa yang akan datang.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Pada penelitian yang berjudul Penerapan Algoritma Naïve Bayes Untuk Analisis Sentimen Pada Media Sosial X Terhadap Performa Tim Nasional Sepakbola Indonesia Di Era Kepemimpinan Shin Tae-Yong menunjukkan algoritma Naïve Bayes mencapai akurasi yang sangat tinggi yaitu 98,73%, dari 393 tweet yang dianalisis, sentiment positif lebih dominan dengan 50.95 atau 200 tweet, sedangkan sentiment negatif sebanyak 49,1% atau sebanyak 193 tweet yang menunjukkan bahwa opini terhadap topik yang diteliti cukup seimbang walaupun condong kearah positif. Data diproses melalui tahapan seperti *cleaning, tokenize, transform cases, stopword removal* dan *filtering* untuk meningkatkan akurasi. Dalam penelitian ini, penulis juga melakukan tahap tersebut dalam *pre-processing data* sehingga data siap untuk diolah menggunakan *framework RapidMiner*.[6]

Penelitian menganalisis respon netizen twitter terhadap program makan siang gratis menggunakan Nlp metode naïve bayes ditulis pada tahun 2024 pengumpulan data dilakukan menggunakan proses crawling menggunakan bahasa pemrograman *python*, sedangkan penelitian ini pengumpulan datanya menggunakan teknik *scrapping* dan *text mining* untuk pengolahan datanya menggunakan *framework Rapidminer*. Hasilnya dari hasil evaluasi pelatihan dan pengujian, dapat menyimpulkan bahwa model cenderung memiliki kinerja yang lebih baik pada data pelatihan daripada pada data uji. Ada indikasi adanya *overfitting* pada data pelatihan, terutama dalam klasifikasi sentimen negatif dan netral. [7]

Dalam penelitian yang ditulis oleh Yuda dan rekan pada tahun 2023 yang berjudul Klasifikasi Sentimen Masyarakat di Twitter terhadap Ganjar Pranowo dengan Metode Modified K-Nearest Neighbor, setelah melewati tahap dari pengambilan, pelabelan data, preprocessing, feature weighting, feature selection, MK-NN dan evaluasi akurasi mendapatkan nilai akurasi tertinggi di 83,8% dengan perbandingan 90:10 dengan nilai k=3. Penelitian tersebut menggunakan bahasa pemrograman *phyton*, sedangkan dalam penelitian ini menggunakan *RapidMiner* dengan klasifikasi algoritma *Naïve Bayes*. [8]

Jurnal yang berjudul Analisis Sentimen Terhadap Film “Dirty Vote” Pada Media Sosial Xdan Youtube dengan Algoritma Naive Bayes dan SVM, menggunakan metode yang sama dengan penelitian ini, yaitu tahapan preprocessing yang dilakukan mencakup Cleansing, Transform Cases, Tokenizing, Stopwords Removal, dan Stemming. Data yang diperoleh kemudian diklasifikasikan ke dalam kategori sentimen positif dan negatif. Evaluasi model dilakukan menggunakan Confusion Matrix yang meliputi accuracy, precision, dan recall. Perbedaannya penulis hanya menggunakan satu algoritma Naïve bayes tidak ada algoritma lain sebagai pembanding hasil akurasi.[9]

Penelitian sebelumnya yang berjudul Klasifikasi Sentimen Masyarakat di Twitter terhadap Kenaikan Harga BBM dengan Metode K-NN, penerapan metode *K-Nearest Neighbor* (K-NN), Feature Weighting (TF-IDF), dan Feature Selection (Threshold) akan dilakukan implementasi dengan menggunakan *tools* yaitu *Google Collab*. Berdasarkan hasil pengujian metode K-NN menggunakan *confusion matrix* pada 10 nilai K yang berbeda (3,5,7,9,11,13,15,17,19,21) dengan mekanisme perbandingan yang digunakan 70:30, 80:20, dan 90:10 diperoleh akurasi paling tinggi sebesar 83,3% pada K=13 dan K=15 untuk perbandingan data training dan testing 90:10. Perhitungannya akurasinya dilakukan secara manual, sedangkan dalam jurnal ini penulis menggunakan *RapidMiner*.[10]

2.2. Algoritma Naïve Bayes

Dalam konteks ini, penggunaan algoritma *Naïve Bayes Classifier* dapat menjadi solusi yang potensial. Meskipun asumsi bahwa atribut dalam data adalah independen, kinerja klasifikasi *Naïve Bayes* tetap cukup tinggi, *Naïve Bayes Classifier* adalah metode klasifikasi yang populer dalam analisis data. Algoritma ini bekerja berdasarkan teorema bayes, yang menggunakan probabilitas kondisional untuk memprediksi kategori atau kelas dari suatu data.[11] *Naïve Bayes* merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki kemampuan dan akurasi tinggi. Naïve bayes classifer merupakan Metode klasifikasi yang memiliki beberapa fase penyelesaian yaitu dimulai dari *Training Data, Learning Algorithm, Model*, Test Data dan diakhiri dengan proses Testing sehingga dihasilkan sebuah keputusan yang akurat. Model *Naïve Bayes* mudah dibuat dan sangat berguna untuk kumpulan data yang sangat besar. *Naïve Bayes* juga merupakan perhitungan yang dapat mengelompokkan variabel tertentu menggunakan kemungkinan dan strategi faktual. *Naïve Bayes* menggunakan bagian ilmu matematika yang dikenal sebagai hipotesis kemungkinan untuk melacak kemungkinan terbaik dari pesanan potensial, dengan memeriksa pengulangan setiap karakterisasi dalam data training. Keuntungan dari penggunaan metode *Naïve Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil saja untuk menentukan estimasi paremeter yang diperlukan dalam proses pengklasifikasian dan dapat bekerja jauh lebih baik dalam situasi dunia nyata yang kompleks.[12]

Berikut adalah bentuk persamaan Teorema Naïve Bayes :

$$P\,(H|X) = \frac{P(X|H).P(H)}{P\,(X)} \quad (1)$$

Keterangan :

P (H|X) = Probabilitas bersyarat H yang diberikan oleh X

P (X|H) = Probabilitas bersyarat X yang diberikan oleh H

P(H) = Probabilitas kejadian H

P(X) = Probabilitas kejadian X

Confusion Matrix : menunjukkan jumlah dokumen yang diklasifikasikan dengan benar dan salah untuk setiap kelas.

Tabel 1. Tabel Confusion Matrix

	Positif (Aktual)	**Negatif (Aktual)**
Positif (Prediksi)	TP	FP
Negatif (Prediksi)	FN	TN

Keterangan :

TP (True Positive), yaitu data positif yang diprediksi benar sebagai data positif.

FP (False Positive), yaitu data negatif yang diprediksi salah sebagai data posifif.

FN (False Negative), yaitu data positif yang diprediksi salah sebagai data negatif.

TN (True Negative), yaitu data negatif yang diprediksi benar sebagai data negatif.

Accuracy : mengukur sejauh mana model berhasil untuk klasifikasi dokumen – dokumen dengan benar pada seluruh dataset.

$$Accuracy = \frac{TP + TN}{TP+TN+FP+FN}\ x\ 100\% \quad (2)$$

Precision : mengukur sejauh mana dokumen dokumen yang diprediksi sebagai suatu kelas benar – benar termasuk ke dalam tersebut.

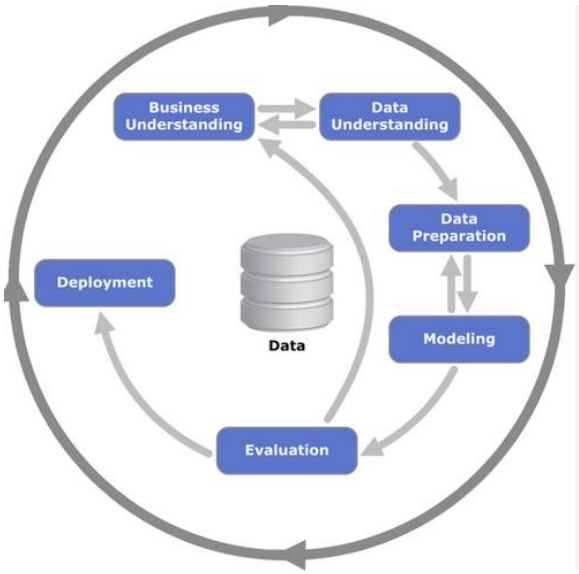
$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Recall : mengukur sejauh mana dokumen – dokumen dari suatu kelas berhasil diidentifikasi oleh model.

$$Recall\ = \frac{TP}{TP+FN} \quad (4)$$

2.3. Data Mining

Data mining, atau penemuan pengetahuan dalam database (KDD), adalah proses untuk menemukan pola dan hubungan dalam kumpulan data besar. Data historis digunakan untuk memprediksi masa depan dan membantu pengambilan keputusan. [13]



Gambar 1. Tahapan *Data Mining*

Tahapan *data mining* yaitu :

a. *Business Understanding*

Ini adalah tahap pertama dalam CRISP-DM dan termasuk bagian yang cukup vital. Pada tahap ini membutuhkan pengetahuan dari objek bisnis, bagaimana membangun atau mendapatkan data, dan bagaimana untuk mencocokan tujuan pemodelan untuk tujuan bisnis sehingga model terbaik dapat dibangun. Kegiatan yang dilakukan antara lain: menentukan tujuan dan persyaratan dengan jelas secara keseluruhan, menerjemahkan tujuan tersebut serta menentukan pembatasan

dalam perumusan masalah *data mining*, dan selanjutnya mempersiapkan strategi awal untuk mencapai tujuan tersebut

b. Data Understanding
Secara garis besar untuk memeriksa data, sehingga dapat mengidentifikasi masalah dalam data. Tahap ini memberikan fondasi analitik untuk sebuah penelitian dengan membuat ringkasaan (*summary*) dan mengidentifikasi potensi masalah dalam data. Tahap ini juga harus dilakukan secara cermat dan tidak terburu-buru, seperti pada visualisasi data, yang terkadang *insight*-nya sangat sulit didapat dika dihubungkan dengan *summary data* nya. Jika ada masalah pada tahap ini yang belum terjawab, maka akan menggangu pada tahap *modeling*.

c. Data Preparation
Secara garis besar untuk memperbaiki masalah dalam data, kemudian membuat *variabel derived*. Tahap ini jelas membutuhkan pemikiran yang cukup matang dan usaha yang cukup tinggi untuk memastikan data tepat untuk algoritma yang digunakan.

d. Modeling
Secara garis besar untuk membuat model prediktif atau deskriptif. Pada tahap ini dilakukan metode statistika dan *Machine Learning* untuk penentuan terhadap teknik *data mining*, alat bantu *data mining*, dan algoritma *data mining* yang akan diterapkan. Lalu selanjutnya adalah melakukan penerapan teknik dan algoritma data mining tersebut kepada data dengan bantuan alat bantu. Jika diperlukan penyesuaian data terhadap teknik data mining tertentu, dapat kembali ke tahap *data preparation*.

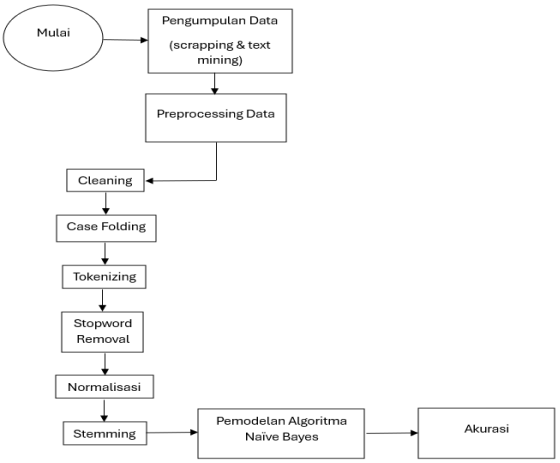
e. Evaluation
Melakukan interpretasi terhadap hasil dari data mining yang dihasilkan dalam proses pemodelan pada tahap sebelumnya. Evaluasi dilakukan terhadap model yang diterapkan pada tahap sebelumnya dengan tujuan agar model yang ditentukan dapat sesuai dengan tujuan yang ingin dicapai dalam tahap pertama.

f. Deployment
Tahap *deployment* atau rencana penggunaan model adalah tahap yang paling dihargai dari proses CRISP-DM. Perencanaan untuk *Deployment* dimulai selama *Business Understanding* dan harus menggabungkan tidak hanya bagaimana untuk menghasilkan nilai model, tetapi juga bagaimana mengkonversi skor keputusan, dan bagaimana untuk menggabungkan keputusan dalam sistem operasional.[14]

3. METODE PENELITIAN

3.1. Jenis Penelitian

Penelitian ini menggunakan metode *kuantitatif deskriptif*, Penelitian *deskriptif kuantitatif* adalah mendeskripsikan, meneliti, dan menjelaskan sesuatu yang dipelajari apa adanya, dan menarik kesimpulan dari fenomena yang dapat diamati dengan menggunakan angka-angka. Penelitian *deskriptif kuantitatif* adalah penelitian yang hanya menggambarkan isi suatu variabel dalam penelitian, tidak dimaksudkan untuk menguji hipotesis tertentu. Dengan demikian dapat diketahui bahwa penelitian *deskriptif kuantitatif* adalah penelitian yang menggambarkan, mengkaji dan menjelaskan suatu fenomena dengan data (angka) apa adanya tanpa bermaksud menguji suatu hipotesis tertentu.[15]



Gambar 2. Tahapan Penelitian

Tahapan penelitian yang dilakukan pada gambar 2, berikut penjelasannya :

a. Pengumpulan data
Mengambil data dengan *teknik crapping* dan *text mining* pada media sosial X dengan kata kunci : "Resesi Indonesia" atau tagar #ResesiIndonesia. Data akan melalui seleksi data, atribut yang digunakan untuk membersihkan data dan untuk klasifikasi selanjutnya akan dipilih pada tahapan ini. Dalam rangka analisis ini, digunakan dataset yang terdiri dari 6850 dataset dan eksekusi dilakukan melalui Google Colab. Data diambil mulai tanggal 1 Juli 2022 hingga 30 Desember 2023. Proses pelabelan data dilaksanakan secara manual dengan membaca tweet satu persatu lalu diberikan label positif atau negatif yang digunakan untuk melatih dan mengklasifikasikan data. [16] Data mentah terdiri dari 10 atribut yaitu : *retweetCount*, *replyCount*, *quoteCount*, *retweetedTweet*, *quotedTweet*, *mentionedUsers*, *ConcersationId*, *source*, media, *user*.

b. Preprocessing data
Beberapa tahapan yang dilakukan pada *preprocessing* yaitu perubahan bentuk kata dasar, menghapus kata yang tidak penting, menghapus imbuhan, dan konjungsi dari dokumen X. Selanjutnya data yang telah melewati tahap *preprocessing* siap untuk dilakukan pembuatan model analisis sentimen yang berguna dalam pengambilan keputusan terhadap permasalahan tersebut. Melakukan *Preprocessing Text Mining* untuk memecahkan berbagai jenis masalah penelitian dalam terkait penambangan data seperti penambangan Text, penambangan web,

penambangan gambar, penambangan pola sekuensial, penambangan spasial, penambangan medis, penambangan multimedia, penambangan struktur dan penambahan grafik. Pada teknik *Preprocessing* dalam penelitian ini menggunakan beberapa tahap diantaranya : *cleaning, case folding, tokenizing, stopword removal, normalisasi, stemming*.[17]

c. Pemodelan Algoritma *Naïve Bayes*
Pemodelan menggunakan *framework RapidMiner* di implementasikan dengan *cross validation* agar kinerja model yang lebih akurat dan mencegah *overfitting* dan *confusion matrix* sebagai alat evaluasi kinerja dalam klasifikasi *machine learning*.

d. Akurasi
Setelah proses pemodelan diimplementasikan dengan algoritma *Naïve Bayes* maka akan mendapatkan nilai akurasi, hasil nilai akurasi adalah sebesar sebesar 81.71% dengan *class recall true positif* 70,59% dan *true negatif* 89,58%.

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Tabel 2. Data cuitan X

Date	Time	Username	Hashtags	Text
06/03/2023	22:05:00	@BonekaCantik	#ResesiIndonesia	Daya beli terus menurun disini, bgmn kalau
06/03/2023	21:55:08	@Presidenbone	#IndonesiaResesiEkonomi	Badai layoff sudah byk didpn masyarakat yg
06/03/2023	21:54:01	@an3kasaf4r!	#ResesiIndonesia	Berasa bgt skrg duit udah makin sulit sekali
06/03/2023	21:53:00	@BonekSBY	#ResesiIndonesia	Ya Allah dimana2 byk bgt yg kena lay off dan
06/03/2023	21:52:07	@HariMingguKe	#ResesiIndonesia	Jadi mau ikutan tren kaburdulu udah terasa
06/03/2023	21:50:00	@Mayangsariba	#ResesiIndonesia	Percuma demo juga karena ekomomi sdh
06/03/2023	21:49:01	@Aurelia	#ResesiIndonesia	Amerika juga sdg mengalami krisis apalagi
06/03/2023	21:48:10	@kodokkorek	#IndonesiaResesi	Rate dollar semakin naik ekonomi semakin a
06/03/2023	21:47:05	@Binsarrajulius	#IndonesiaResesi	Negara kita diambang kehancuran dan krisis
06/03/2023	21:46:03	@Joybintang	#IndonesiaResesiEkonomi	Indonesia udah mau kolaps anj, hade pusing

Tabel 2 adalah hasil pengumpulan data dari media sosial X dengan *teknik crapping* dan *text mining* pada media sosial X dengan kata kunci : “Resesi Indonesia” atau tagar #ResesiIndonesia # IndonesiaResesiEkonomi.

4.2. *Preprocessing Data*

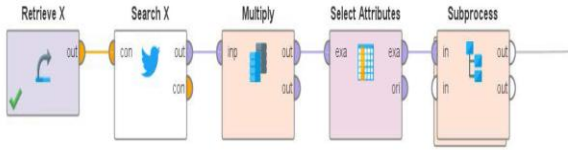
Tabel 3. Proses *Text Processing*

Text Processing		Sebelum	Sesudah
Cleaning	Penghapus kata yang tidak perlu contohnya : simbol, tanda baca, emotion dll	Aktivitas ekonomi melambat, pengangguran naik, bisnis rugi.Bagaimana pak @Jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di Indonesia.	Aktivitas ekonomi melambat, pengangguran naik, bisnis rugi.Bagaimana pak Jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di Indonesia.
Case Folding	Seluruh kata dirubah menggunakan huruf kecil	Aktivitas ekonomi melambat, pengangguran naik, bisnis rugi.Bagaimana pak Jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di Indonesia.	aktivitas ekonomi melambat, pengangguran naik, bisnis rugi.bagaimana pak jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di indonesia.
Tokenizing	Proses untuk membagi teks, menjadi token-token/bagian-bagian tertentu	aktivitas ekonomi melambat, pengangguran naik, bisnis rugi.bagaimana pak jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di indonesia.	"aktivitas" "ekonomi" "melambat" "pengangguran" "naik" "bisnis" "rugi" "bagaimana" " pak" "Jokowi" "kebijakan" "pemerintah" "banyak" "menjadi" "pemicu" "resesi" "di" "Indonesia"
Stopword Removal	proses dalam pemrosesan bahasa alami (NLP) yang menghilangkan kata-kata yang sangat umum dan tidak informatif (seperti "dan", "yang", "di") dari teks	aktivitas ekonomi melambat, pengangguran naik, bisnis rugi.bagaimana pak jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di indonesia.	aktivitas ekonomi melambat, pengangguran naik, bisnis rugi. Presiden Jokowi, kebijakan pemerintah menjadi pemicu resesi di Indonesia.
Normalisasi	Memperbaiki ejaan yang salah pada kata yang terdapat dalam kalimat	Aktivitas ekonomi melambat, pengangguran naik, bisnis rugi.Bagaimana pak Jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di Indonesia.	Aktivitas ekonomi melambat, pengangguran naik. Ini fase sulit. Bagaimana pak Jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di Indonesia.

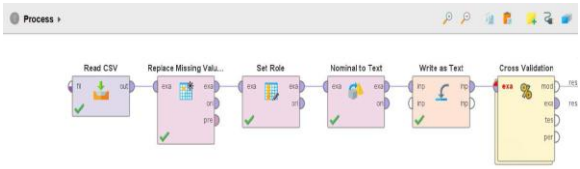
Text Processing		Sebelum	Sesudah
Stemming	proses dalam pemrosesan bahasa alami (NLP) yang bertujuan untuk mereduksi kata-kata berimbuhan menjadi bentuk dasarnya, yang disebut "stem" atau "akar kata"	Aktivitas ekonomi melambat, pengangguran naik, bisnis rugi.Bagaimana pak Jokowi, kebijakan pemerintah banyak menjadi pemicu resesi di Indonesia.	Aktivitas ekonomi lambat, ganggur naik. Bagaimana pak presiden Jokowi, bijak pemerintah banyak jadi pemicu resesi di Indonesia.

Tabel 2 adalah proses *text processing* yang dilakukan sebelum proses pemodelan menggunakan *framework RapidMiner*.

4.3. Pemodelan Algoritma *Naïve Bayes*

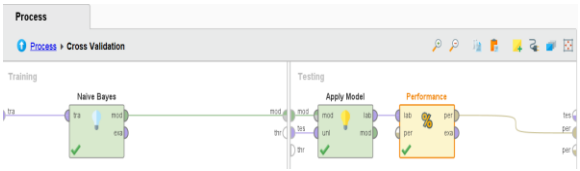


Gambar 3. Proses data X



Gambar 4. Proses pemodelan Naïve Bayes

Gambar 3 dan 4 adalah proses pemodelan data dari cuitan media sosial X menggunakan *framework RapidMiner*.



Gambar 5. Algoritma Naïve Bayes

Gambar 5 adalah pemodelan menggunakan *Algoritma Naïve Bayes* menggunakan *framework RapidMiner.*

Tabel 4. Hasil Akurasi Algoritma Naïve Bayes

Accuracy 81,71%			
	true POSITIF	true NEGATIF	class precision
pred. POSITIF	24	5	82,76%
pred. NEGATIF	10	43	81,13%
class recall	70,59%	89,58%	

$$Accuracy = \frac{TP+TN}{TP + TN + FP + FN} \; x100\%$$

$$Accuracy = \frac{24+43}{25 + 43 + 10 + 5} \; x100\%$$

$$Accuracy = \frac{67}{82} \; x100\% = 0{,}81707 = 81{,}71\%$$

Pada Tabel 3 adalah hasil akurasi Algoritma Naïve Bayes dengan hasil akurasi 81,71%, dengan jumlah *class precision* 82.76% untuk *pred positif* dan 81,13% untuk pred negatif. Sedangkan untuk *class recall* 70,59% untuk *true positif* dan 89,58% untuk *true negatif.* Hasil akurasi pengujian dengan perbandingan data dengan *confusion matrix* 90:10, 80:20 dan 70:30 dengan menggunakan nilai k=3 dan menggunakan *cross validation* pada proses pemodelan Naïve Bayes sehingga hasil penelitian termasuk dalam kategori *good classification*.

5. KESIMPULAN DAN SARAN

Penelitian ini menghasilkan klasifikasi sentimen netizen di media sosial X terhadap ancaman resesi ekonomi di Indonesia. Dalam pengambilan data di media sosial X menggunakan *teknik crapping* dan *text mining* pada media sosial X dengan kata kunci : "Resesi Indonesia" atau tagar #ResesiIndonesia. Data dibersihkan melalui proses *text processing* dan *processing data*, sehingga data siap diolah. Klasifikasi *Algoritma Naïve Bayes* mendapatkan hasil akurasi sebesar 81.71% dengan *class recall true positif* 70,59% dan *true negatif* 89,58%. Hasil penelitian menyatakan metode naïve bayes classifier cukup akurat untuk bisa mengatasi klasifikasi sentimen masyarakat di media sosial X terhadap ancaman resesi ekonomi di Indonesia. Namun jika data dikurangi dan klasifikasi datanya tidak balance maka akurasinya menjadi lebih rendah. Saran untuk penelitian selanjutnya, dapat menggunakan metode klasifikasi lain untuk melihat hasil dari kinerja metode tersebut menggunakan permasalahan yang sama yaitu klasifikasi sentimen netizen terhadap ancaman resesi ekonomi di Indonesia dan data bisa diambil dari platform media sosial lainnya seperti facebook atau instagram.

DAFTAR PUSTAKA

[1] E. F. Zakiyah and A. B. Prayoga Kasmo, "Peran Dan Fungsi Usaha Mikro Kecil Dan Menengah (UMKM) Dalam Memitigasi Resesi Ekonomi Global 2023," J. Cakrawala Ilm., vol. 2, no. 4, pp. 349–365, 2022, [Online]. Available: http://bajangjournal.com/index.php/JCI

[2] D. R. Sari, "Klasifikasi Sentimen Masyarakat di Twitter terhadap Ancaman Resesi Ekonomi 2023 Dengan Metode Naive Bayes Classifier," J. Sist. Komput. dan Inform., vol. 4, no. e-ISSN 2685-998X, pp. 577–585, 2023, doi: 10.30865/json.v4i4.6276.

[3] Admin, "Algoritma Klasifikasi pada Data Mining," bktaruna.uma.ac.id, 2022.

https://bktaruna.uma.ac.id/algoritma-klasifikasi-pada-data-mining/

[4] A. Septiansyah, A. Yuliawati, and I. Santoso, "Analisis Sentimen Terhadap Layanan Internet Di Indonesia Menggunakan Metode K-Nearest Neighbor," J. IKRAITH-INFORMATIKA, vol. 2, no. 1, pp. 98–104, 2023, [Online]. Available: https://journals.upi-yai.ac.id/index.php/ikraith-informatika/issue/archive%0AAnalisis

[5] Rian Tineges, "Mengenal Naive Bayes Sebagai Salah Satu Algoritma Data Science," https://dqlab.id/mengenal-naive-bayes-sebagai-salah-satu-algoritma-data-science, 2022. https://dqlab.id/mengenal-naive-bayes-sebagai-salah-satu-algoritma-data-science

[6] W. Alfiyani, D. A. Fatah, and F. Irhamni, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pada Media Sosial X Terhadap Performa Tim Nasional Sepak Bola Indonesia Di Era Kepemimpinan Shin Tae-Yong," J. Mhs. Tek. Inform., vol. 9, no. 3, pp. 3969–3977, 2025.

[7] F. N. Zaman, M. A. Fadhilah, M. A. Ulinuha, and K. Umam, "Menganalisis Respons Netizen Twitter Terhadap Program Makan Siang Gratis Menerapkan Nlp Metode Naïve Bayes," J. Sist. Informasi, Teknol. Inf. dan Komput., vol. 14, no. 3, pp. 150–233, 2024, [Online]. Available: https://jurnal.umj.ac.id/index.php/just-it/index

[8] Y. Z. fadhlan, Yusra, and M. Fikry, "Klasifikasi Sentimen Masyarakat di Twitter terhadap Ganjar Pranowo dengan Metode Modified K-Nearest Neighbor," J. Inform. Univ. Pamulang, vol. 8, no. 2, pp. 191–198, 2023, doi: 10.32493/informatika.v7i2.30686.

[9] K. H. Sasongko and A. M. Hilda, "Analisis Sentimen Terhadap Film ' Dirty Vote ' Pada Media Sosial X dan Youtube dengan Algoritma Naive Bayes dan SVM," vol. 6, no. 1, pp. 73–85, 2024, doi: 10.47065/josyc.v6i1.6150.

[10] T. D. Arista, M. Fikry, and O. Lola, "Klasifikasi Sentimen Masyarakat di Twitter terhadap Kenaikan Harga BBM Dengan Metode K-NN," JUKI J. Komput. dan Inform., vol. 5, no. 1, pp. 140–150, 2023.

[11] D. Nuryansah and M. Ary, "Implementasi Algoritma Naïve Bayes Classifier Untuk Memprediksi Tingkat Produktivitas Kinerja Karyawan," JIKA (Jurnal Informatics) Univ. Muhammdiyah Tangerang, vol. 8, no. 3, pp. 297–303, 2024, doi: 10.31000/jika.v8i3.11125.

[12] E. Martantoh and N. Yanih, "Implementasi Metode Naïve Bayes Untuk Klasifikasi Karakteristik Kepribadiaan Siswa Di Sekolah MTS Darussa'adah Menggunakan Php Mysql," J. Teknol. Sist. Inf., vol. 3, no. 2, pp. 166–175, 2022, doi: 10.35957/jtsi.v3i2.2896.

[13] K. Ananda Mustari, P. Assiroj, B. Hartati, and F. Samuel, "Implementasi Data Mining Pada Instansi Pemerintahan," JATI (Jurnal Mhs. Tek. Inform., vol. 8, no. 3, pp. 3137–3142, 2024, doi: 10.36040/jati.v8i3.9618.

[14] T. Mauritsius and F. Binsar, "Cross-Industry Standard Process for Data Mining (CRISP-DM)," binus ac.id, 2020. https://mmsi.binus.ac.id/2020/09/18/cross-industry-standard-process-for-data-mining-crisp-dm/

[15] W. Sulistyawati, Wahyudi, and T. Sabekti, "Analisa (Deskriptif Kuantitatif) Motivasi Belajar Siswa Dengan Model Blended Learning Di Masa Pandemi Covid19," Kadikma, vol. 13, no. 1, pp. 2–7, 2022.

[16] R. Hidayat, M. Fikry, Y. Yusra, F. Yanto, and E. P. Cynthia, "Penerapan Naïve Bayes Classifier dalam Klasifikasi Sentimen Publik di Twitter terhadap Puan Maharani," JUKI J. Komput. dan Inform., vol. 6, no. 1, pp. 100–108, 2024, doi: 10.53842/juki.v6i1.479.

[17] D. Rifaldi, A. Fadlil, and Herman, "Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet 'Mental Health,'" J. Pendidik. Teknol. Inf., vol. 3, no. 1, pp. 161–169, 2023, doi: http://dx.doi.org/10.51454/decode.v3i2.131.