

## ANALISIS SENTIMEN MENGGUNAKAN METODE NAÏVE BAYES TERHADAP CHILDFREE

Amellia Veronica Agustin<sup>1</sup>, Fitria Nur Sa'adah<sup>2</sup>, Yuyun Umaidah<sup>3</sup>

<sup>1,2,3</sup>Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang

e-mail: <sup>1</sup>2010631250004@student.unsika.ac.id, <sup>2</sup>2010631250083@student.unsika.ac.id,

<sup>3</sup>yuyun.umaidah@staff.unsika.ac.id

### ABSTRAK

*Childfree* merujuk pada individu atau pasangan yang secara sukarela memutuskan untuk tidak memiliki keturunan. Fenomena *Childfree* semakin menarik perhatian masyarakat modern dan mempengaruhi keputusan hidup banyak individu. Data yang digunakan diperoleh dengan melakukan *crawling* data pada sosial media *twitter*. Dalam analisis sentimen ini, metode yang digunakan *naïve bayes*, data akan diflasifikasikan menjadi dua hasil, yaitu sentimen negatif dan positif. Langkah selanjutnya melibatkan penilaian data dengan mengaplikasikan pengujian menggunakan *confusion matrix* untuk mengukur tingkat akurasi, *recall*, dan *precision*. Hasil pengujian dengan *Rapidminer* menunjukkan tingkat akurasi *naïve bayes* yang tinggi sebesar 96%, untuk sentimen positif menghasilkan nilai *precision* 84,13%, dan nilai *recall* 100%. Sedangkan untuk sentimen negatif menunjukkan nilai *precision* 94,92%, dan *recall* 100%. Hasil klasifikasi tersebut menghasilkan 254 tweet negatif, dan 74 tweet positif. Analisis sentimen menggunakan metode *Naïve Bayes* dapat memberikan wawasan yang berharga tentang pandangan dan sentimen masyarakat terhadap *Childfree*. Hasil penelitian ini penting untuk memahami dan menghargai keragaman pandangan terkait topik ini. Hasil penelitian ini dapat memiliki dampak positif pada proses pengambilan keputusan. dalam konteks sosial, budaya, dan kebijakan yang berkaitan dengan keputusan *Childfree*.

**KATA KUNCI:** *Childfree, Naïve Bayes, Twitter, Rapidminer.*

### 1. PENDAHULUAN

Perkembangan teknologi informasi telah memberikan pengaruh signifikan dalam dunia komunikasi dan informasi. Saat ini, platform media sosial seperti *Twitter, Instagram, dan Facebook* menjadi pilihan bagi publik sebagai tempat untuk menyampaikan pendapat, termasuk mengenai topik yang sensitif seperti keputusan hidup *childfree*. Baru-baru ini *childfree* menjadi populer di kalangan masyarakat umum setelah Gita Savitri Devi (@gitasav) mengemukakan opininya mengenai dirinya dan suaminya memilih untuk tidak ingin mempunyai keturunan. Pernyataan itu menyebabkan munculnya berbagai komentar online dan membuat orang-orang yang hidup tanpa anak mulai berani berpendapat. Hal tersebut menyebabkan pro-kontra di media sosial mengenai kebebasan untuk memilih *childfree*. Ada beberapa alasan yang mendasari mengapa seseorang memutuskan untuk tidak memiliki anak dalam kehidupannya, diantaranya seperti khawatir tentang perkembangan anak, masalah lingkungan, masalah ekonomi, ataupun masalah pribadi [1]

Dikutip dari [2] berdasarkan penelitian *Childlessness in the United States* menyatakan untuk memilih *childfree* mengalami kenaikan sebesar 20% pada tahun 2000-an. Peningkatan tersebut terutama disebabkan oleh masalah keluarga dan pertimbangan terkait tanggung jawab mengasuh anak di masa depan. Faktor-faktor ini semakin diperkuat oleh adanya kesulitan ekonomi dan hambatan psikologis yang mempengaruhi kehidupan keluarga. Menurut para psikolog dan dokter di Amerika Serikat, keputusan pasangan suami istri untuk tidak memiliki anak dipengaruhi oleh alasan yang bersifat subyektif dari diri mereka sendiri.

Dalam penelitian ini, menitik beratkan pada penerapan metode *Naïve Bayes* dalam melakukan analisis sentimen terhadap pandangan dan opini yang terkait dengan gaya hidup *childfree*. Metode *Naïve Bayes* merupakan suatu pendekatan klasifikasi yang berdasarkan probabilitas, di mana sekelompok probabilitas dihitung dengan mengakumulasi frekuensi dan kombinasi nilai dari dataset yang diberikan [3]. *Naïve Bayes* merupakan sebuah algoritma klasifikasi yang banyak digunakan dalam analisis sentimen.

Analisis sentimen merupakan bagian dari klasifikasi teks yang tergolong dalam ranah komputasi linguistik, pengolahan bahasa alami, dan penambangan teks. Mempunyai tujuan untuk mengevaluasi

sentimen, ulasan, pendapat, sikap, penilaian, dan emosi yang dinyatakan oleh penulis atau pembaca terkait dengan suatu topik, produk, organisasi, layanan, individu, atau kegiatan tertentu [4].

Penerapan metode *Naïve Bayes* sudah diterapkan pada penelitian sebelumnya, seperti yang dilakukan oleh Arsi [5] melakukan klasifikasi sentimen dari twitter menggunakan metode *Naïve Bayes* mengenai pemindahan ibu kota. Hasil dari analisa dari penelitian ini menghasilkan nilai akurasi sebesar 94,33%, precision sebesar 0,87 dan recall 0,99. Lalu pada penelitian yang dilakukan oleh [6] terhadap tweet mention dari jasa ojek online pada platform twitter, yang menghasilkan nilai akurasi sebesar 80% dari 800 data. [4] melakuakn penelitian mengenai analisis sentiment data twitter tentang masalah obesitas di Indonesia yang meghasilkan sentiment positif lebih besar yaitu 22.246 (51,2%) tweet dan sentimen negative nya sebanyak 12.015 (27,7%) tweet, lalu sentimen netral berjumlah 9.174 (21,1%) dari 43.435 tweet. Dengan nilai akurasi sebesar 94%.

Berdasarkan informasi yang telah dijelaskan sebelumnya, tujuan dari penelitian ini mengelompokkan atau mengklasifikasikan komentar, atau pendapat yang terkait dengan *childfree* di bagi menjadi dua kategori utama yaitu sentiment positif dan sentiment negatif. Dengan menggunakan dataset yang dikumpulkan dari platform media sosial yaitu twitter yang mencakup berbagai macam opini tentang *childfree*. Harapannya, hasil penelitian ini dapat meningkatkan pemahaman terhadap pandangan masyarakat terhadap fenomena *childfree*. Selain itu, untuk menilai performa metode *Naïve Bayes* dalam klasifikasi sentimen pada dataset yang digunakan.

## 2. TINJAUAN PUSTAKA

### 2.1 *Childfree*

"*Childfree* adalah pilihan hidup yang diambil oleh pasangan untuk tidak mempunyai keturunan. Ini tidak hanya berkaitan dengan tidak memiliki anak biologis, tetapi juga mencakup ketidak inginan untuk memiliki anak angkat. Dalam arti yang lebih jelas, *childfree* menggambarkan keadaan di mana seseorang dan pasangannya secara aktif memilih untuk hidup tanpa anak, baik itu anak biologis maupun anak angkat, dan mereka merasa bebas dari tanggung jawab dan kewajiban yang terkait dengan membesarkan anak [1]

*Childfree* bukanlah keadaan yang sama dengan infertilitas atau ketidak mampuan untuk mempunyai anak. Orang-orang yang memilih untuk *childfree* dapat memiliki kemampuan biologis untuk memiliki anak, tetapi mereka secara sadar memilih untuk tidak melakukannya. Keputusan ini didasarkan pada preferensi dan nilai-nilai individu yang mengarahkan mereka untuk menjalani kehidupan tanpa tanggung jawab dan kewajiban orang tua. Pilihan *childfree* dapat berlaku untuk individu tunggal atau pasangan, dan hal ini tidak harus menjadi keputusan permanen. Beberapa orang mungkin memilih untuk menjadi *childfree* pada suatu periode dalam hidup mereka, sementara yang lain mungkin mempertimbangkan atau mengubah keputusan mereka di kemudian hari.

Menurut [7] Pasangan yang memilih jalur *childfree* merasakan dampak positifnya. Baru-baru ini, ada banyak orang tua yang menghadapi kesulitan dalam mendidik anak-anak mereka, yang berujung pada keterlibatan anak-anak dalam perilaku bebas seperti narkoba, putus sekolah, kehamilan di luar nikah, perilaku nakal, dan lain sebagainya. Karenanya, bagi pasangan yang belum siap secara mental dan finansial untuk memberikan pendidikan, perawatan, dan pengasuhan kepada anak, memilih gaya hidup *childfree* dianggap sebagai opsi yang cocok. Namun, perlu diingat bahwa keputusan *childfree* juga dapat membawa dampak negatif dengan memutuskan rantai keturunan, yang berarti tidak ada generasi berikutnya yang dapat menjaga keberlanjutan populasi manusia di masa mendatang.

### 2.2 *Analisis Sentimen*

Menurut [8] Analisis sentimen, yang sering juga disebut sebagai opinion mining, merupakan salah satu disiplin ilmu yang mempelajari dan menganalisis tentang pendapat, sentimen, evaluasi, penilaian, sikap, dan emosi masyarakat umum terhadap berbagai entitas seperti individu, layanan, produk, isu, peristiwa, atau organisasi. Tugas utama dalam analisis sentimen adalah mengkategorikan teks dalam suatu kalimat atau dokumen, dan kemudian menilai apakah pendapat yang terdapat dalam kalimat atau dokumen tersebut bersifat positif atau negatif. Analisis sentimen juga dapat mengidentifikasi perasaan emosional seperti kesedihan, kebahagiaan, atau kemarahan. Dengan menggunakan analisis sentimen, dapat mengetahui opini publik tentang produk, merek, atau orang tertentu dan menentukan apakah opini tersebut cenderung lebih condong ke positif atau negatif.

Analisis sentimen dapat diterapkan dalam berbagai konteks, termasuk analisis opini produk, pengawasan media sosial, pemantauan reputasi merek, pemahaman umpan balik pelanggan, dan analisis sentimen politik. Dengan menggunakan teknik analisis sentimen, organisasi dan individu dapat memperoleh wawasan berharga tentang bagaimana publik merespons dan merasakan suatu entitas, sehingga dapat dimanfaatkan dalam pengambilan keputusan terkait perbaikan produk atau layanan, atau manajemen reputasi.

### 2.3 Twitter

Twitter diperkenalkan sebagai platform mikroblogging pada bulan Maret 2006, yang memungkinkan pengguna untuk mengirimkan pembaruan status dengan batasan 140 karakter yang dikenal sebagai 'tweet'. Sejak peluncurannya, Twitter berhasil membangun komunitas pengguna yang besar, dan hingga Juni 2011, jumlah pengguna telah mencapai lebih dari 300 juta. Skala besar pengguna ini memiliki dampak yang signifikan pada masyarakat, terutama dalam hal menyampaikan penilaian melalui pendapat atau komentar terkait peristiwa yang telah terjadi atau sedang berlangsung. [9]

Twitter menjadi populer sebagai platform berita dan jejaring sosial, di mana pengguna dapat mengikuti akun resmi media, tokoh terkenal, perusahaan, organisasi, dan individu lainnya untuk mendapatkan informasi terkini dan berpartisipasi dalam percakapan publik. Twitter juga telah menjadi tempat bagi kampanye pemasaran, aktivisme, dan diskusi umum mengenai berbagai topik, termasuk politik, olahraga, hiburan, dan isu-isu sosial.

Twitter juga menawarkan API (Application Programming Interface) yang memungkinkan pengembang dan peneliti untuk mengakses data Twitter dan melakukan analisis lebih lanjut, termasuk analisis sentimen, tren topik, dan jaringan sosial.

### 2.4 Naïve Bayes

Naive Bayes merupakan suatu teknik data mining yang simpel digunakan dan memiliki proses pemrosesan yang efisien. Metode ini dapat diterapkan dengan struktur yang sederhana namun efektif. Pendekatan klasifikasi Naive Bayes adalah suatu metode statistik untuk mengelompokkan data, memungkinkan prediksi probabilitas keanggotaan dalam suatu kelas tertentu. Klasifikasi ini bergantung pada teorema Bayes dan memiliki kemampuan klasifikasi sebanding dengan decision tree dan neural network. Keunggulan metode ini terlihat dari tingkat akurasi dan kecepatannya yang tinggi, terutama ketika diaplikasikan pada database dengan dataset yang besar. [10] Berikut ini persamaan yang umum dari Naïve Bayes.

$$P(A|X) = \frac{p(X|A).P(A)}{P(X)} \quad (1)$$

Keterangan:

X = Sampel data yang mempunyai class atau label yang tidak diketahui.

A = Hipotesis kalau X merupakan data class atau label.

P(A) = Probability dari hipotesis A.

P(X) = Peluang yang asalnya dari data sampel yang diamati (probability A).

P(X|A) = Probability X berdasar kondisi hipotesis A.

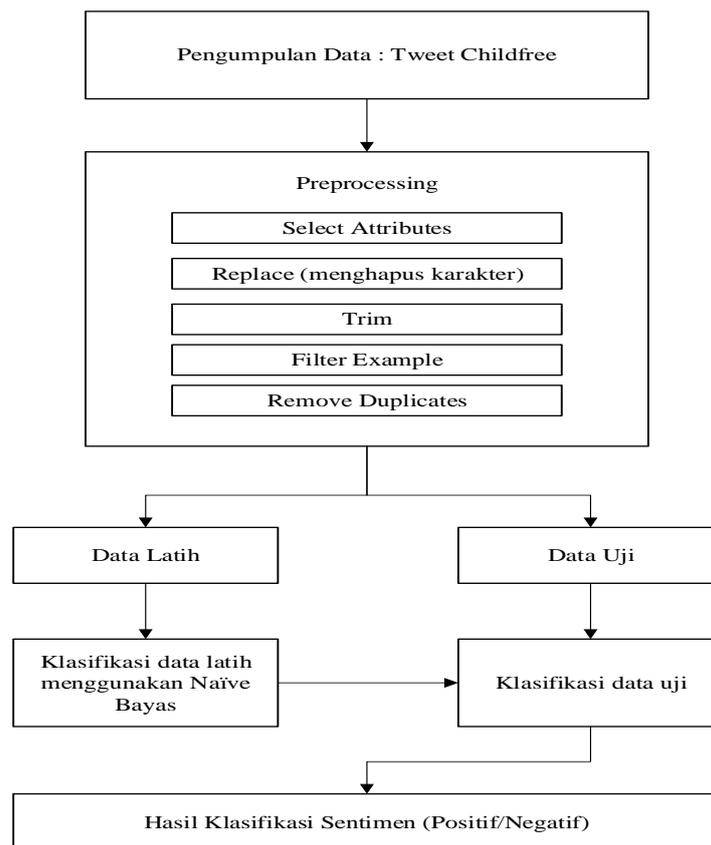
P(A|X) = Probability hipotesis A berdasarkan pada kondisi X.

### 2.5 RapidMiner

*RapidMiner* merupakan perangkat lunak sumber terbuka yang menyediakan solusi untuk melakukan analisis pada data mining, text mining, dan analisis prediksi. Dengan menggunakan berbagai teknik deskriptif dan prediktif, *RapidMiner* bertujuan memberikan pemahaman kepada pengguna agar mereka dapat membuat keputusan optimal. Perangkat lunak ini dapat diunduh melalui situs web resminya, *RapidMiner*, yang sebelumnya dikenal sebagai *YALE (Yet Another Learning Environment)*. Awalnya dikembangkan pada tahun 2001. [11].

## 3. METODE PENELITIAN

Berikut adalah langkah-langkah dalam melakukan penelitian analisis sentimen terhadap *childfree* dengan menggunakan metode klasifikasi *Naïve Bayes*.



Gambar 1. Tahapan Penelitian Menggunakan Klasifikasi *Naïve Bayes*

### 3.1 Pengumpulan Data

Pada penelitian ini menggunakan data dari twitter dengan menggunakan kata kunci *childfree* yang berada di Indonesia. Data tersebut dikumpulkan dari [www.twitter.com](http://www.twitter.com) dan dibantu dengan framework rapidminer.

### 3.2 Preprosessing

Pada tahapan *preprosessing* yaitu proses menyiapkan data yang akan diolah lebih lanjut, tujuan yaitu untuk memastikan kebersihan data dan mengurangi gangguan, dilakukan proses data preprocessing agar data menjadi lebih bersih [12]. Proses data preprocessing memiliki peran penting dalam analisis sentimen terhadap data yang diperoleh dari media sosial. Beberapa tahapan dalam data preprocessing yang dilakukan dalam penelitian ini meliputi:

#### a. *Select Atribut*

Operator *Select Attributes* dalam *RapidMiner* berfungsi untuk memilih atribut tertentu dari suatu dataset dan menghapus atribut yang tidak relevan. Operator ini menawarkan berbagai jenis filter yang memudahkan dalam pemilihan atribut, seperti pemilihan langsung, pemilihan dengan menggunakan regular expression, atau pemilihan hanya atribut yang tidak memiliki nilai yang hilang (missing value).

#### b. *Replace (Menghapus karakter pada tweet)*

Operator *Replace* pada *RapidMiner* digunakan untuk menghilangkan simbol dan karakter yang biasanya ditemukan dalam teks di platform Twitter, seperti retweet, @username, #hashtag, link serta karakter khusus lainnya yang tidak relevan dan tidak diperlukan dalam analisis.

#### c. *Trim*

Operator *Trim* dalam *RapidMiner* digunakan untuk menghilangkan spasi kosong yang terdapat di awal dan akhir nilai pada atribut nominal tertentu. Dengan menggunakan operator ini, dapat dibuat atribut baru yang telah di-trim. Operator ini memiliki manfaat dalam membersihkan data dari spasi kosong yang tidak diinginkan pada nilai atribut nominal.

#### d. *Filter Example*

Operator *Filter Example* dalam *RapidMiner* digunakan untuk untuk melakukan penghapusan atau pemfilteran baris atau contoh tertentu yang tidak diinginkan dari dataset. Tujuan dari penggunaan fitur ini adalah untuk mengurangi jumlah contoh atau baris dalam dataset, menghilangkan data yang

tidak relevan atau duplikat, serta meningkatkan kualitas dataset sebelum melanjutkan proses selanjutnya.

#### e. *Remove Duplicates*

Operator *Remove Duplicates* pada *RapidMiner* digunakan untuk menghapus baris duplikat dari sebuah dataset. Operator ini berguna untuk membersihkan data dari baris duplikat yang tidak digunakan dan memastikan bahwa setiap baris dalam dataset unik.

### 3.3 Pelabelan Data

Proses pelabelan data memiliki tujuan untuk mengategorikan data ke dalam beberapa kategori sentimen. Dalam penelitian ini, kategori sentimen dibagi menjadi dua kelompok, yaitu sentimen *positif* dan *negatif* [3]. Tahap selanjutnya melibatkan pemisahan data menjadi data latih dan data uji sesuai dengan rasio yang telah ditetapkan sebelumnya. Data latih digunakan untuk melatih model *Naive Bayes*, sedangkan data uji digunakan untuk menguji seberapa akurat model yang telah dilatih.

### 3.4 Proses Klasifikasi

Setelah melakukan pelabelan sentimen pada data latih, dilakukan ekstraksi fitur menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode TF-IDF digunakan untuk menghitung bobot atau nilai penting dari setiap kata yang terdapat dalam tweet pada data latih. Bobot tersebut kemudian digunakan sebagai *input* dalam model *Naive Bayes*.

Selanjutnya, model *Naive Bayes* yang telah dilatih dengan data latih akan diuji menggunakan data uji dengan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Evaluasi tersebut bertujuan untuk mengevaluasi performa model dan menentukan apakah model tersebut cukup baik untuk digunakan dalam analisis sentimen terhadap data tweet tentang *childfree*.

### 3.5 Hasil Klasifikasi Sentimen

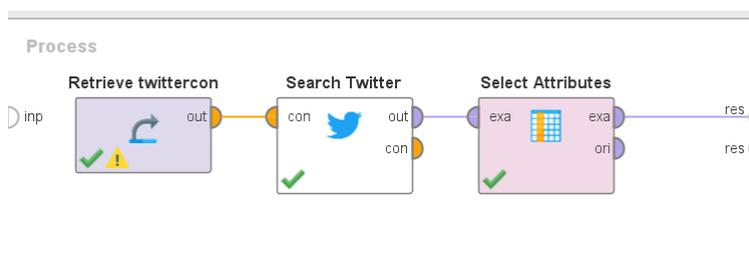
Setelah proses klasifikasi sentimen dilakukan, hasilnya adalah prediksi sentimen untuk setiap data yang diuji. Hasil klasifikasi sentimen menunjukkan klasifikasi atau label sentimen yang diberikan oleh model terhadap setiap data uji. Misalnya, hasil klasifikasi sentimen dapat mencakup klasifikasi sebagai sentimen *positif* atau *negative*.

## 4. HASIL DAN PEMBAHASAN

Hasil dan pembahasan berfungsi untuk memberikan gambaran lengkap tentang apa yang telah ditemukan dalam penelitian dan bagaimana hal itu berkaitan dengan pertanyaan penelitian yang diajukan. Bagian ini penting karena membantu pembaca memahami kontribusi penelitian terhadap bidang ilmu yang relevan dan menyajikan kesimpulan yang didukung oleh data. Dalam penelitian ini, kami akan menyajikan hasil dan pembahasan secara rinci mengenai penggunaan metode klasifikasi teks menggunakan Algoritma Naive Bayes [13].

### 4.1 Pengumpulan Data

Pengumpulan data ini menggunakan kata kunci '*Childfree*' pada tweet menggunakan bahasa Indonesia. Proses pengambilan data dilakukan secara manual dengan menggunakan akses token dari Twitter API, dan alat yang digunakan adalah *Rapidminer* untuk mendapatkan informasi dan data yang diinginkan. Jumlah data yang digunakan untuk penelitian ini sebanyak 325 data, kemudian data tersebut dikelompokkan menjadi 2 yaitu, bila tweet berisi penolakan atau tidak mendukung adanya *childfree* maka data tersebut masuk kedalam sentiment *positif* dan bila tweet berisi dukungan atau setuju dengan adanya *childfree* maka masuk kedalam sentimen *negatif*.



Gambar 2 Proses *Crawling* Data Menggunakan tool *Rapidminer*

Pada gambar 2 memperlihatkan langkah-langkah dalam mengambil data menggunakan *Rapid Miner*. Beberapa elemen operator yang difungsikan meliputi "*search Twitter*" untuk terhubung dengan

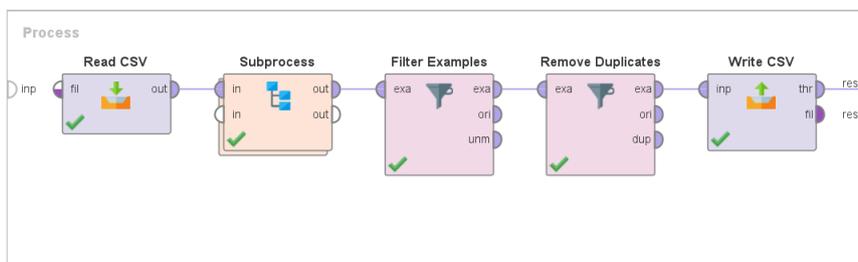
Twitter dan menggunakan kode akses token yang diperoleh dari Twitter API. Operator selanjutnya adalah "select attributes" seperti yang ditampilkan pada gambar 2, yang dimanfaatkan untuk menentukan atribut yang relevan, seperti teks. Kemudian operator tambahan yang digunakan adalah "write excel" untuk menyimpan file data dalam bentuk format Excel. Hasil pengambilan data menggunakan Rapidminer ditampilkan dalam Tabel 1 dengan atribut-atribut seperti nama pengguna dan teks.

No	Teks
1	@najjaevm hadehh, kayaknya aku mau childfree deh nanti kalau ada niatan nikah wkwkwkw
2	semakin mantab mengikuti jejak mba gitasav untuk childfree <a href="https://t.co/VtK02OUcxP">https://t.co/VtK02OUcxP</a>
3	gw pernah kepikiran childfree sih tapi abis itu ngga lagi,,,
4	knp sih 2023 msh pada mau pny anak lu ga liat ini depok udah sesempit margonda childfree kek elah
5	untungnya ga childfree @najamila

Tabel 1 Data Crawling

#### 4.2 Preprocessing

Kemudian langkah selanjutnya yaitu tahap *preprocessing*, Preprocessing dilaksanakan untuk mengatasi ketidaksempurnaan data, gangguan data, dan ketidak-konsistenan data. Hal ini bertujuan untuk memperbaiki kualitas data sebelum dilakukan analisis lebih lanjut [14].



Gambar 3 Tahap Preprocessing

Pada Gambar 3 diatas ini merupakan tahanan *preprocessing* ini diantaranya yaitu *Select Attributes*, menghapus URL, @username, #hashtag, link serta karakter khusus lainnya yang tidak relevan dan tidak diperlukan dalam analisis, membersihkan data dari spasi kosong yang tidak diinginkan pada nilai atribut nominal, *Filter Examples* yaitu untuk menghilangkan data yang tidak relevan atau duplikat serta *Remove Duplicates* menghapus baris duplikat dari sebuah dataset.

#### 4.3 Pelabelan Data

Dalam proses ini pelabelan data diberi label secara manual. Setelah melalui tahap *preprocessing* data dan penempatan sentimen secara manual, hasil yang di dapat ditampilkan pada Tabel 2. Selanjutnya yaitu melibatkan Pengelompokan data menjadi data pelatihan dan data pengujian sesuai dengan rasio yang sudah ditetapkan, data yang sudah dilabeli secara manual termasuk kedalam data latih.

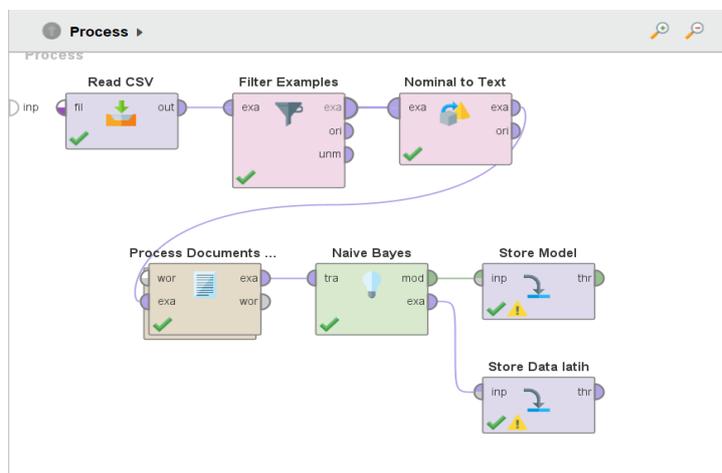
No	Teks	Sentimen
1	Kalo aku jawabnya, Enggak sayang, soalnya aku penganut childfree	Negatif
2	Dulu kepikiran childfree tapi setelah liat lucunya sutan langsung kepikiran mau bgt punya anak cewe udh gitu gue kayak ada magnet sama anak kecil	Positif
3	Semoga ketemu laki <sup>2</sup> yg menjunjung childfree	Negatif
4	Makanya lebih baik childfree aja daripada anak yg dijadikan objek pelampiasan	Negatif
5	Ihk takut banget ada childfree	Positif

Tabel 2 Pelabelan Data

Proses pelabelan data yaitu dilakukan untuk menentukan *tweet* tersebut termasuk ke dalam sentimen *positif* yang tidak mendukung atau menolak adanya *childfree* di Indonesia seperti ingin menikah dan memiliki keturunan. Sedangkan sentiment *negative* berisi dukungan terhadap *childfree* dan memutuskan tidak ingin memiliki anak.

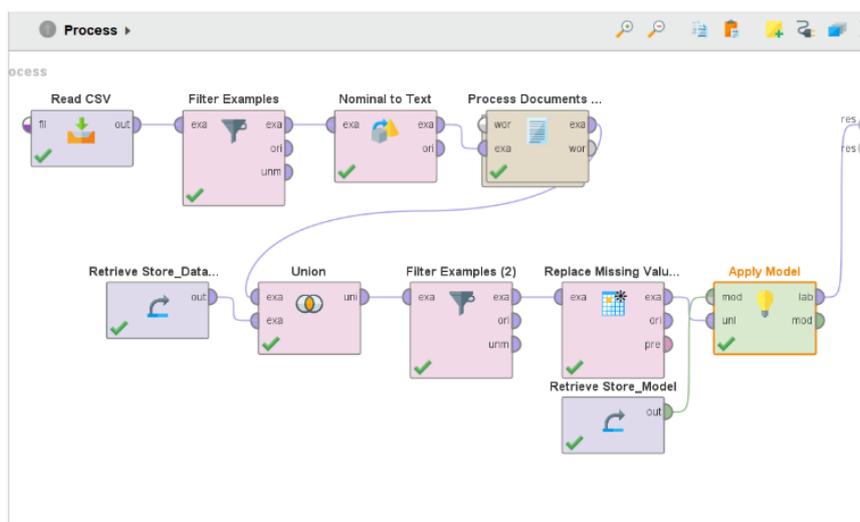
#### 4.4 Proses Klasifikasi

Pada tahapan ini, data yang sudah diproses akan dimanfaatkan untuk membangun konstruksi model klasifikasi yang dapat memprediksi sentimen dari teks yang sudah dilabelkan, sebagaimana ditunjukkan pada Gambar 4. Pada bagian ini tahapan penelitian menggunakan metode Naïve Bayes untuk membangun model klasifikasi dalam analisis sentimen ini. Kemudian, data yang telah diproses dan sudah dibagi menjadi dua bagian, yaitu data latih dan data uji. Data latih sejumlah 250 data, sementara data uji sebanyak 75 data. Data ini digunakan dalam pengujian keakuratan model klasifikasi yang telah dikonstruksi.



Gambar 4 Model Klasifikasi

Pada gambar diatas langkah yang dilakukan pertama adalah *input read* data yang sudah dilabelkan dan dibersihkan sebelumnya, langkah kedua yaitu *filter examples* digunakan untuk memilih data latih yang sudah dilabeli dengan sentiment yang nantinya akan diproses, langkah ketiga dilakukan TF-IDF yaitu digunakan untuk pengukuran bobot kata yang digunakan dalam pemrosesan bahasa alami dan informasi pada teks. Langkah selanjutnya yaitu menyambungkan operator Naïve Bayes untuk pembuatan model, yang nantinya akan dimasukkan kedalam dua operator store yang digunakan untuk menyimpan model dan data latih.



Gambar 5 Evaluasi Perfoma Model

Di lihat pada Gambar 5 diatas ini merupakan proses evaluasi dilakukan untuk menentukan sejauh mana kemampuan model klasifikasi atau regresi dalam melakukan prediksi secara akurat dan reliabel. Pada gambar diatas, terlihat bahwa terdapat filter yang digunakan untuk memilih data uji yang belum memiliki sentiment yang terisi. Setelah itu, dilakukan proses yang serupa seperti sebelumnya. Setelah itu, model yang telah dibuat diterapkan pada data uji.

Sentimen	prediction(S...	text
?	Negatif	kesehatan dirpd childfree manfaatin duit cowo lain ngasih meki pengen horang ttep hdup susah
?	Positif	bangke juga childfree bisa party tiap hari malah
?	Negatif	mungkin saja pilihan untuk menjadi childfree adalah langkah putus solusi yang paling mungkin karena tidak pernah
?	Negatif	lagi cari username trus buka gugel words starts with nomor satu childfree kwkwkwk okeh banh
?	Negatif	yang heteroseksual childfree yang ngurusin anak orang
?	Negatif	terima kasih
?	Positif	pernah dimarahin gara gara bilang childfree padahal punya uang buat nyekolahin sampe internasional school
?	Negatif	plih punya anak atau childfree nder
?	Negatif	keren juga ngeluarin analisa isuisu terkini mulai dari childfree sampai citayam fashion week
?	Negatif	selamat malam tementemen saya sedang menyusun skripsi mencari informan penelitian dengan kriteria berikut
?	Negatif	jodoh akan selalu menemani kalau jodoh namanya jaman skrg bnyak skali pasangan childfree mereka tenang jika ...
?	Negatif	cari lakilaki childfree nder yang emang tidak berniat punya anak melihat perempuan bukan dari badannya
?	Negatif	banyak sihbuktinya yang pillg childfree kankalau sender childfree bukan karena ingin tapi karena kondisi
?	Negatif	mungkin childfree yang penting mungkin enaenanya

Gambar 6 Hasil Prediksi Pelabelan

Setelah melakukan evaluasi kinerja model, ditemukan hasil yang direpresentasikan dalam Gambar 6. Pada tabel tersebut, setiap data sentimen yang sebelumnya tidak terisi, secara otomatis diisi dengan menggunakan pembelajaran dari data latih yang digunakan untuk membangun model.

#### 4.5 Hasil Klasifikasi Sentimen

Setelah melakukan pemrosesan evaluasi performa model dan mendapatkan data dari hasil prediksi pelabelan tersebut, maka diperoleh hasil yang tercatat dalam Tabel 3.

	True Negatif	True Positif
Pred. Negatif	187	0
Pre. Positif	10	53

Tabel 3 Confusion Matrix

Setelah didapatkan hasil dari prediksi pelabelan maka perhitungan yang dilakukan, diperoleh temuan bahwa dari data yang dikumpulkan secara manual maupun dari data yang dihasilkan oleh sistem, ditemukan bahwa 77,44% pengguna Twitter mendukung adanya *childfree*, sedangkan 22,55% pengguna Twitter tidak mendukungnya adanya *childfree*. Dalam penelitian ini, dilakukan pengukuran hasil menggunakan confusion matrix dari 328 data tweet pada data uji. Sistem berhasil mengklasifikasikan 254 tweet sebagai sentimen negatif dan 74 tweet sebagai sentimen positif. Dalam rincian data tersebut, terdapat 187 tweet dengan sentimen negatif yang diprediksi dengan benar oleh sistem, sedangkan 10 tweet diprediksi dengan salah. Untuk sentimen positif, sebanyak 53 tweet diprediksi dengan benar oleh sistem, dan tidak ditemukan adanya prediksi yang salah. Dengan menggunakan metode Naive Bayes dalam teknik pemodelan klasifikasi, penelitian ini memiliki dua jenis klasifikasi, yaitu sentimen positif dan negatif. Dari hasil pengujian, dapat disimpulkan bahwa tingkat akurasi algoritma Naive Bayes mencapai 96%. Selain itu, ditemukan hasil dari sentimen positif dengan tingkat presisi sebesar 84,13% dan recall 100%. Sedangkan untuk sentimen negatif, diketahui tingkat presisinya mencapai 94,92% dan recall 100%.

Selanjutnya dilakukan pengelompokkan data untuk mencari frekuensi kata yang paling sering muncul. Frekuensi kata ini akan direpresentasikan dalam bentuk *wordcloud positif* dan *negatif* seperti yang ditampilkan pada Gambar 7.



Dengan demikian, hasil analisis sentimen terhadap topik Childfree menggunakan metode Naïve Bayes menunjukkan hasil yang memuaskan dan dapat memberikan wawasan yang berharga tentang pandangan pengguna Twitter terhadap konsep tersebut.

## 6. SARAN

Agar mencapai hasil yang optimal, disarankan untuk meningkatkan analisis dengan menggunakan data yang lebih baik atau lebih terstruktur. Hal ini dapat dilakukan dengan memperhatikan tahapan preprocessing yang lebih terperinci. Selain itu, memperluas jumlah data yang digunakan atau dengan penggabungan metode lain juga dapat meningkatkan akurasi hasil analisis.

## DAFTAR PUSTAKA

- [1] S. Ajeng Wijayanti and N. Neneng, “Analisis Fenomena Childfree di Indonesia,” *Bandung Conference Series: Islamic Family Law*, vol. 2, no. 2, pp. 64–70, 2022, doi: 10.29313/bcsifl.v2i2.2684.
- [2] H. Tiara, “Konstruksi Masyarakat Tentang Hidup Tanpa Anak Setelah Menikah,” *Jurnal Analisa Sosiologi*, vol. 11, no. 1, pp. 126–136, 2022, doi: 10.20961/jas.v11i1.56920.
- [3] W. Bimananda, R. Insan, D. Karina, N. Rani, S. Theresa, and D. Yosherina, “Analisis Text Mining dari Cuitan Twitter Mengenai Infrastruktur di Indonesia dengan Metode Klasifikasi Naïve Bayes,” *Eigen Mathematics Journal*, vol. 2, no. 2, pp. 92–101, 2019, doi: 10.29303/emj.v1i2.36.
- [4] F. Fridom Mailo *et al.*, “Analisis Sentimen Data Twitter Menggunakan Metode Text Mining Tentang Masalah Obesitas di Indonesia,” 2021.
- [5] P. Arsi, B. A. Kusuma, and A. Nurhakim, “Analisis Sentimen Pindah Ibu Kota Berbasis Naive Bayes Classifier,” *JURNAL INFORMATIKA UPGRIS*, 2021.
- [6] D. Garbian Nugroho, Y. Herry Chrisnanto, A. Wahana Jurusan Informatika, and F. Matematika dan Ilmu Pengetahuan Alam Universitas Jenderal Achmad Yani Jalan Terusan Jenderal Sudirman, *ANALISIS SENTIMEN PADA JASA OJEK ONLINE MENGGUNAKAN METODE NAÏVE BAYES*.
- [7] M. Rindu Fajar Islamy, K. Siti Komariah, D. Mayadiana Suwarma, and A. Hafidzani Nur Fitria, “Fenomena Childfree di Era Modern: Studi Fenomenologis Generasi Z serta Pandangan Islam terhadap Childfree di Indonesia,” *Sosial Budaya*, vol. 19, no. 2, pp. 81–89, 2022, [Online]. Available: <http://dx.doi.org/10.24014/sb.v19i2.16602>
- [8] A. G. DARUSSALAM, “Analisis Sentimen Menggunakan Text Mining Dengan Metode Naïve Bayes Dan Regresi Logistik,” vol. 4, no. 1, pp. 1–23, 2021.
- [9] M. M. Mala Olhang, S. Achmadi, and F. X. A. Wibisono, “Analisis Sentimen Pengguna Twitter Terhadap Covid-19 Di Indonesia Menggunakan Metode Naive Bayes Classifier (Nbc),” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 4, no. 2, pp. 214–221, 2020, doi: 10.36040/jati.v4i2.2695.
- [10] E. Fitri, “Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine,” *Jurnal Transformatika*, vol. 18, no. 1, p. 71, 2020, doi: 10.26623/transformatika.v18i1.2317.
- [11] W. Khofifah, D. N. Rahayu, and A. M. Yusuf, “Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps,” *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 16, no. 4, pp. 28–38, 2022, doi: 10.35969/interkom.v16i4.192.
- [12] A. Sentimen, “Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor”.
- [13] N. Ruhyana, “ANALISIS SENTIMEN TERHADAP PENERAPAN SISTEM PLAT NOMOR GANJIL/GENAP PADA TWITTER DENGAN METODE KLASIFIKASI NAIVE BAYES.” [Online]. Available: [www.situs.com](http://www.situs.com)
- [14] I.Hemalatha Dr. G. P Saradhi Varma Dr. A.Govardhan, “Preprocessing the Informal Text for efficient Sentiment Analysis,” 2012, [Online]. Available: [www.Ecstasy.com](http://www.Ecstasy.com)