

# Turnitin JTIK Febriyanti Syofiani

*by* Febriyanti Syofiani

---

**Submission date:** 05-Sep-2023 10:37AM (UTC+0700)

**Submission ID:** 2157972112

**File name:** Artikel\_Febriyanti\_Syofiani.pdf (737.21K)

**Word count:** 4050

**Character count:** 23582

## **Analisis Sentimen Penilaian Masyarakat Terhadap *Childfree* Berdasarkan Komentar di Youtube Menggunakan Algoritma Naïve Bayes**

**Febriyanti Syofiani<sup>1\*)</sup>, Syariful Alam<sup>2)</sup>, M. Imam Sulisty S.<sup>3)</sup>**

**2**  
<sup>1)2)3)</sup> Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana

<sup>\*)</sup>Correspondence Author: [febriyantisofiani31@wastukencana.ac.id](mailto:febriyantisofiani31@wastukencana.ac.id), Purwakarta, Indonesia

**DOI:**

### **Abstrak**

YouTube merupakan jejaring sosial tempat pengguna dapat menonton video yang telah diunggah dan memiliki panjang durasi tertentu dalam berbagai genre. Salah satu acara YouTube yang masih menjadi trending saat ini adalah *childfree*. *Childfree* adalah perjanjian sepasang kekasih, atau antara suami dan istri, untuk tidak memiliki anak selama pernikahan mereka. Dunia maya, yang begitu bebas dan sulit diatur, memberi kesan kepada setiap orang bahwa mereka memiliki pendapat atau pandangan mereka sendiri tentang masalah *childfree*, yang menghasilkan banyak opini publik, baik yang positif maupun negatif. Tujuan analisis sentimen adalah untuk menentukan respon *audience* terhadap masalah *Childfree*, mengidentifikasi dan mengklasifikasi polaritas suatu teks untuk menentukan apakah suatu dokumen memiliki nilai positif atau negatif menurut kategori tertentu. Teknik pembelajaran analisis sentimen adalah *Naïve Bayes Classifier* dan pengujian data menggunakan *confusion matrix* yang ada pada *tools orange*. Hasil dari reaksi orang terhadap *childfree* tergolong negatif dengan hasil persentase 97% pada akurasi (*accuracy*), 98% pada nilai *precision*, dan tingkat keberhasilan (*recall*) 96%. Berdasarkan nilai tersebut membuktikan sentimen penilaian masyarakat terhadap *childfree* berdasarkan komentar di YouTube tergolong negatif. Untuk penelitian selanjutnya, lebih bagus apabila peneliti mengambil komentar YouTube dari luar negeri seperti Jepang, Tiongkok, dan negara Arab Saudi. Dierkomendasikan juga untuk menggunakan algoritma lain untuk menganalisis sentimen, seperti algoritma SVM dan algoritma K-Nearest Neighbor.

**Kata kunci:** *Childfree*, *Naïve Bayes*, YouTube, *Orange*

### **Abstract**

YouTube is a social network where users can watch videos that have been uploaded and have a certain duration in various genres. One of the YouTube shows that is still trending at the moment is *childfree*. *Childfree* is an agreement between lovers, or between husband and wife, not to have children during their marriage. The virtual world, being so free and unruly, gives everyone the impression that they have their own opinion or views on the issue of *childfree*, which generates a lot of public opinion, both positive and negative. The purpose of sentiment analysis is to determine audience response to *Childfree* issues, identify and classify the polarity of a text to determine whether a document has positive or negative values according to certain categories. The sentiment analysis learning technique is the *Naïve Bayes Classifier* and data testing uses the confusion matrix in the orange tools. The results of people's reactions to *childfree* are classified as negative with a percentage of 97% for accuracy, 98% for precision, and a 96% recall rate. Based on these values, it proves that the public's assessment of *childfree* based on comments on YouTube is negative. For further research, it is better if researchers take YouTube comments from abroad such as Japan, China and Saudi Arabia. It is also recommended to use other algorithms to analyze sentiment, such as the SVM algorithm and the K-Nearest Neighbor algorithm.

**Keywords:** *Childfree*, *Naïve Bayes*, YouTube, *Orange*

## PENDAHULUAN

Modernisasi dan kemajuan teknologi menuntut segalanya untuk menjadi maju. Salah satunya adalah video, karena informasi audiovisual mudah dipahami. Media sosial YouTube berisi berbagai video (Buana et al., 2023). YouTube merupakan jejaring sosial tempat pengguna dapat menonton video yang telah diunggah dan memiliki panjang durasi tertentu dalam berbagai genre (Harpizon et al., 2022). Penonton dapat terhubung dengan situs YouTube dengan memberikan *like*, *dislike*, dan komentar pada video yang ditampilkan di media YouTube serta mengungkapkan opini mereka tentang video yang ditonton. Salah satu acara YouTube yang masih menjadi trending saat ini adalah *childfree*.

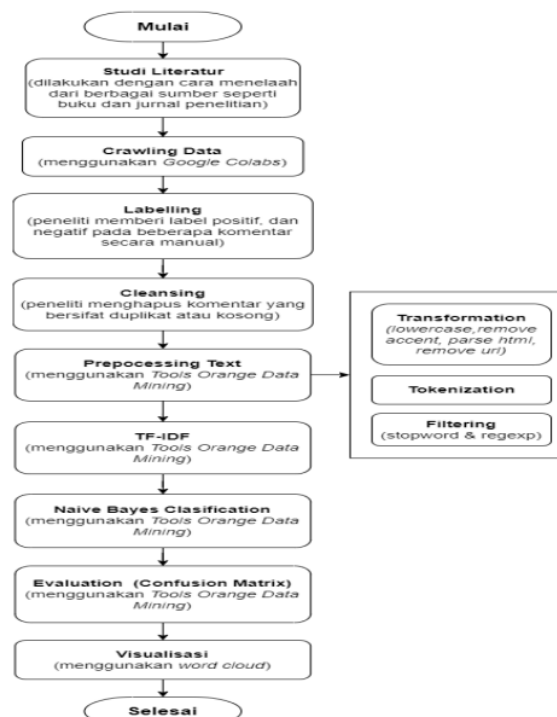
*Childfree* adalah perjanjian sepasang kekasih, atau antara suami dan istri, untuk tidak memiliki anak selama pernikahan mereka (Fadhilah, 2022). Ungkapan "childfree" pertama kali muncul dalam leksikon bahasa Inggris Merriam-Webster sebelum 1901 (Agama et al., 2022). *Childfree* menjadi perhatian publik di media sosial YouTube ketika seorang YouTuber ternama Indonesia menyatakan dirinya sebagai pengikut bebas anak di akun media sosialnya. Hal ini memunculkan berbagai diskusi dimana tubuh perempuan tidak memiliki kekuasaan dalam perbincangan sosial. Dunia maya, yang begitu bebas dan sulit diatur, memberi kesan kepada setiap orang bahwa mereka memiliki pendapat atau pandangan mereka sendiri tentang masalah *childfree*, yang menghasilkan banyak opini publik, baik yang positif maupun negatif. Selain itu *childfree* sangat berpengaruh terhadap angka kelahiran di Indonesia. Berdasarkan Badan Pusat Statistik tanggal 31 Maret 2023 mencatat, *total fertility rate* (TFR) Indonesia tahun 1971 berada pada angka 5,61%, pada tahun 1980 total kelahiran berada pada angka 4,68%, pada tahun 1990 total kelahiran berada pada angka 3,33%, pada tahun 2000 total kelahiran berada pada angka 2,34%, pada tahun 2010 total kelahiran berada pada angka 2,41%, dan pada tahun 2020 total kelahiran berada pada angka 2,18%. Dimana artinya pada tahun 2020 terdapat penurunan angka kelahiran di Indonesia.

Topik yang diangkat dalam penelitian ini adalah *trend childfree* di media sosial YouTube Indonesia dimana tubuh perempuan diperbincangkan dan menimbulkan banyak opini publik, tidak hanya positif tetapi juga negatif. Salah satu metode untuk mengamati reaksi terhadap komentar publik di YouTube yakni melalui analisis sentimen. Analisis sentimen merupakan penelitian tentang menganalisis pendapat, perasaan, dan emosi yang

terkandung dalam teks atau informasi. Tujuan analisis sentimen adalah untuk menentukan respon *audiens* terhadap masalah *Childfree*, mengidentifikasi dan mengklasifikasikan polaritas suatu teks untuk menentukan apakah suatu dokumen memiliki nilai positif atau negatif menurut kategori tertentu. Pendekatan klasifikasi Naive Bayes dapat digunakan untuk mengkategorikan data komentar YouTube. Metode ini merupakan klasifikasi yang paling sederhana dan paling umum digunakan serta menawarkan beberapa keunggulan seperti kesederhanaan, kecepatan dan akurasi yang tinggi. *Tools Orange data mining* dapat digunakan untuk aplikasi pemrosesan data karena mendukung berbagai algoritma dan metode lainnya.

## METODE

Kerangka berpikir adalah logika penelitian yang diciptakan melalui pengamatan, fakta, dan tinjauan pustaka (Hadinata & Kurniawan, 2020). Dalam penelitian ini, dijelaskan dalam kerangka berpikir sebagai berikut :



Gambar 1. Kerangka Berpikir Penelitian

1. Tahap awal penelitian, hal pertama yang dilakukan yaitu *review* pada jurnal penelitian untuk memahami teori seputar topik yang akan diteliti. Studi literatur dilakukan dengan cara melakukan analisis dan menelaah berbagai sumber teori seperti buku-buku, jurnal penelitian dari internet yang berkaitan dengan analisis sentiment menggunakan algoritma Naïve Bayes untuk dijadikan referensi penelitian.
2. Pada tahap *Crawling data*, pengumpulan komentar (*crawling* komentar) diambil dengan cara *crawling data* pada sosial media YouTube dengan bantuan *API (Application Programming Interface)* YouTube. Pada proses *crawling data* ini dilakukan dengan penarikan komentar salah satu content creator YouTube yang bertopik mengenai *Childfree*. *Crawling data* dilakukan pada tanggal 11 Maret 2023 dengan menggunakan *Google Colabs*.
3. Pada tahap *labelling* peneliti secara manual memberi label sentimen positif dan negatif teks dokumen.
4. Pada tahap *cleansing* dilakukan penghapusan komentar yang bersifat duplikat atau kosong secara manual oleh peneliti.
5. Transformasi, tokenisasi, dan filtering dilakukan selama tahap *preprocessing* untuk menyiapkan dokumen atau dokumen perwakilan untuk diproses.
6. Selama tahap *Term Frequency - Inverse Document Frequency*, kata-kata dalam dokumen diberi bobot untuk menentukan bobot kata.
7. Tahap *naïve bayes* digunakan untuk mengklasifikasikan data pada langkah klasifikasi.
8. Tahap perhitungan akurasi, presisi, *recall* dengan *confusion matrix* dilakukan pada teks yang diklasifikasikan oleh algoritma NBC selama langkah evaluasi.
9. Pada langkah visualisasi, stopword atau kata-kata umum tanpa signifikansi, seperti contoh (yang, dalam, adalah), akan diterapkan ke data komentar yang diproses. Akibatnya, hanya kata-kata dengan makna, seperti kata kerja, kata benda, dan lainnya, yang akan hadir. Setelah kata dibersihkan dari stopwords, Word Cloud akan digunakan untuk memvisualisasikannya.

Analisis sentimen, juga dikenal sebagai pengumpulan opini, yaitu Proses memahami, menggali, dan mengolah informasi tekstual secara otomatis untuk memperoleh informasi opini yang dimasukkan dalam kalimat opini (Liu 2010). Tujuan utama analisis sentimen adalah untuk mengkategorikan fitur teks dalam kalimat dan opini, yang sering

diklasifikasikan sebagai positif atau negatif (Nurtikasari, Syariful Alam, and Teguh Iman Hermanto 2022).

*Naive Bayes* adalah klasifikasi berdasarkan teori probabilitas yang dibuat oleh Thomas Bayes seorang ilmuwan berasal dari Inggris, yang memberikan perkiraan untuk kemungkinan potensial berdasarkan kinerja masa lalu, yang dikenal sebagai Teorema Bayes. (Rachman & Handayani, 2021). *Naive Bayes Classifier (NBC)* merupakan algoritma berdasarkan perkiraan probabilistik sederhana yang menggunakan aturan Bayes dengan anggapan independensi yang kuat (naif). Dengan kata lain, "model fitur independen" dipakai di *Naive Bayes*. (Halim, 2018). Berikut persamaan NBC :

$$P(C|X) = \frac{P(X|C).P(c)}{P(x)}$$

Keterangan :

X : Jenis data yang tidak diketahui

C : Hipotesis yang termasuk dalam kategori tertentu

P(c|X) : Peluang hipotesis menurut kondisi

P(c) : Probabilitas hipotesis

P(X|c) : Probabilitas yang didasari oleh kondisi pada hipotesis

P(X) : Probabilitas c

Menurut (Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, 2018) keuntungan *Naive Bayes* adalah bahwa ketika tes dijalankan pada kategori data, setiap atribut berdiri sendiri. *Crawling* merupakan metode memperoleh informasi atau data dari sebuah *website* dengan menginput *Uniform Resource Locator* (URL) (Arsi & Waluyo, 2021). Proses *crawling data* pada penelitian ini dilakukan pada tanggal 11 Maret 2023 menggunakan *Google Colabs* dengan bantuan API YouTube dan menghasilkan total 4355 data.

*Labelling* merupakan pemberian kelas pada data (Ahmad et al., 2023). Ketika label ditetapkan, identitas dan penjelasan data akan ditampilkan bagaimana kelas sentimen pada data. *Cleansing* merupakan proses memeriksa kualitas data dengan mengedit, memodifikasi, atau menghapus data yang berlebihan, tidak lengkap, atau salah, atau mempunyai tipe data atau file yang salah dalam database, untuk membuat data yang berkualitas (Darwis et al.,



2021). Proses ini dilakukan secara manual menggunakan *Microsoft Excel* dengan melihat setiap respon secara individual.

*Preprocessing text* adalah tahap dimana dokumen asli dikonversi sebagai input dan beberapa teknik dasar digunakan untuk mengganti atau menghapus bagian dokumen atau teks yang tidak lagi berfungsi dalam pemrosesan (Najjichah et al., 2019). Tahap-tahap *Preprocessing text* :

1. *Transformation*

Proses *transformasion* dapat mengubah input data, termasuk mengubah semua teks menjadi huruf kecil dan menghapus alamat web atau URL dari teks..

2. *Tokenizing*

*Tokenization* dikenal sebagai teknik menggabungkan simbol tanda baca dan memisahkan teks menjadi kata-kata.

3. *Filtering*

*Filtering* dilakukan untuk memfilter atau menghilangkan kata-kata yang tidak akan dibutuhkan pada proses berikutnya seperti “dan”, “yang”, “di”, dan “adalah”.

- a. *Stopword Removal* adalah teknik yang melibatkan penghapusan kata-kata yang berlebihan. Isi dalam pernyataan tidak akan berubah atau dihilangkan jika kata tersebut dihapus. Seperti tanda hubung “adalah”, “akan”, “dalam”, “pada”, “dan” (Kurniawan et al., 2020).
- b. *Regex* adalah menghapus istilah yang cocok dengan ekspresi reguler. Defaultnya adalah menghilangkan tanda baca simbol seperti @#/(.); (Wiguna & Rifai, 2021).

*Orange data mining* adalah perangkat lunak penambangan data atau teknik pembelajaran mesin *open source*. *Orange* dapat digunakan untuk eksplorasi dan visualisasi data. *Orange* digunakan dalam genetika, biomedis, bioinformatika, dan studi pendidikan. *Orange* selalu disukai dalam hal inovasi, kualitas, atau ketergantungan (Nurhafida & Sembiring, 2021).

## HASIL DAN PEMBAHASAN

Komentar YouTube yang telah melewati tahap *preprocessing* dapat diklasifikasikan berdasarkan kata-kata yang terkandung di dalamnya. Salah satu algoritma klasifikasi, *Naive Bayes* membuat prediksi kasus berdasarkan hasil klasifikasi. Menghitung Klasifikasi pada contoh kumpulan data yang akan digunakan meliputi 10 data komentar, termasuk 10 sampel yang dilabelkan dengan 5 nilai positif dan 5 nilai negatif, seperti yang ditunjukkan pada tabel 1 di bawah ini:

**Tabel 1.** Data Sampel

| Dokumen | Komentar YouTube                                                                                              | Labelling |
|---------|---------------------------------------------------------------------------------------------------------------|-----------|
| D1      | orang eropa pikir dukung <i>childfree</i>                                                                     | 2         |
| D2      | seru nonton baca komen dunia mutusin dukung <i>childfree</i>                                                  | 2         |
| D3      | <i>childfree</i> orang hebat                                                                                  | 2         |
| D4      | netizen gasiap ngapain anak mending ikut <i>childfree</i> sono                                                | 2         |
| D5      | <i>childfree</i> gih biar anak saing kerja jakarta mudah dukung penuh                                         | 2         |
| D6      | tolak mindset <i>childfree</i> coba ubah mindset anak takdir hidup                                            | 1         |
| D7      | tolak <i>childfree</i> percaya anak anugerah bentuk pertanggungjawaban tuhan                                  | 1         |
| D8      | potensial singgung nilai nilai orang telan sendiri jangan publikasi indonesia mayoritas anti <i>childfree</i> | 1         |
| D9      | anti <i>childfree</i> tujuan tuhan cipta ikut perintah tuhan                                                  | 1         |
| D10     | pantesan orang indonesia kerja jepang <i>childfree</i> indonesia punah                                        | 1         |

Untuk melakukan perhitungan menggunakan contoh dataset data tabel di atas menggunakan *Microsoft Excel*, langkah pertama adalah memastikan *term frequency* (TF) di setiap dokumen, yang ditunjukkan pada tabel 2 di bawah ini.



**Tabel 2. Term Frequency**

| Dokumen   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| orang     | 1   | 0   | 1   | 0   | 0   | 0   | 0   | 1   | 0   | 1   |
| eropa     | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| pikir     | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| dukung    | 1   | 1   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   |
| childfree | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| seru      | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| nonton    | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| baca      | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| komen     | 0   | 1   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   |
| dunia     | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| hebat     | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| netizen   | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| gasiap    | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| ngapain   | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| anak      | 0   | 0   | 0   | 1   | 1   | 1   | 1   | 0   | 0   | 0   |
| mending   | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| ikut      | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 1   | 0   |
| .....     | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| .....     | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| .....     | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| perintah  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   |
| pantesan  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   |
| jepang    | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   |
| punah     | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   |

Jumlah kemunculan kata unik dan jumlah di setiap kategori kemudian dihitung.

Tabel 3 menggambarkan urutan penghitungan :

**Tabel 3. Jumlah Kata Unik**

| Kategori         | Jumlah dokumen di setiap kategori | Jumlah dokumen seluruhnya | P(Kategori) | Jumlah Kata Setiap Kategori |
|------------------|-----------------------------------|---------------------------|-------------|-----------------------------|
| Positif          | 5                                 | 10                        | 0.50        | 33                          |
| Negatif          | 5                                 | 10                        | 0.50        | 42                          |
| Jumlah Kata Unik |                                   | 52                        |             |                             |

Istilah yang sering muncul di setiap dokumen dikenal sebagai kata-kata unik. Setelah mengidentifikasi jumlah kata unik dan jumlah kata tiap kategori, probabilitas di setiap kategori dihitung.

$$P(\text{Kata}) = \frac{\text{jumlah dari satu kata di satu kategori} + 1}{\text{jumlah kata tiap kategori} + \text{jumlah keseluruhan kata unik}}$$

Dari tabel 3 di atas dihitung jumlah kata, dalam kategori positif 33 kata, negatif 42 kata dengan jumlah kata unik sebanyak 52 kata. Cara menghitung probabilitas dicontohkan untuk kata *childfree* dimana 1 tambah 5 lalu dibagi 33 yang didapat dari jumlah setiap kategori positif lalu di tambah 52 yang didapat dari jumlah kata unik. Setelah didapatkan jumlah kata unik dan jumlah kata pada tiap kategori, selanjutnya menghitung probabilitas pada setiap kategori. Untuk menghitung probabilitas pada setiap kategori adalah sebagai berikut :

$$(\text{childfree} | \text{positif}) \frac{5 + 1}{33 + 52} = \frac{6}{85} = 0,07059$$

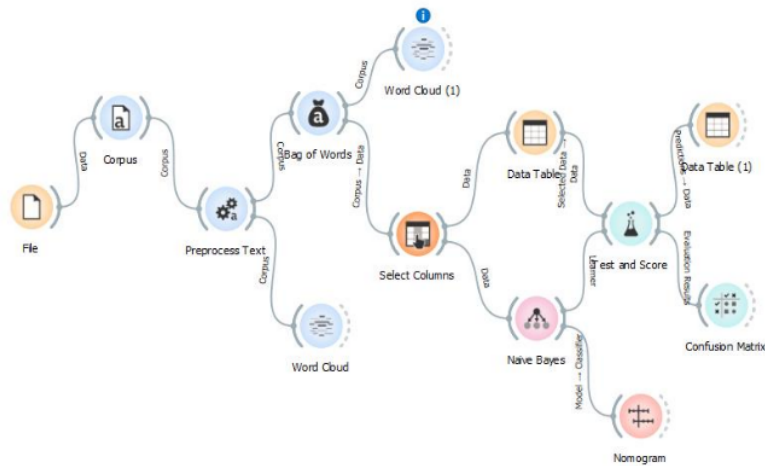
$$(\text{childfree} | \text{negatif}) \frac{5 + 1}{42 + 52} = \frac{6}{94} = 0,0638$$

Pada tahapan ini, dilakukan perhitungan pada dataset, seperti pada tabel 4 di bawah ini:

**Tabel 3.** Hasil Perhitungan *Naive Bayes*

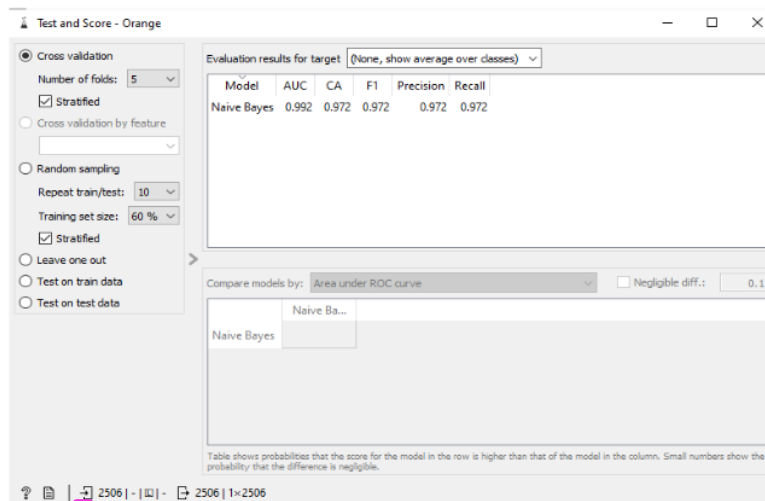
| Kata             | Positif | Negatif |
|------------------|---------|---------|
| orang            | 0.03529 | 0.03191 |
| eropa            | 0.02353 | -       |
| pikir            | 0.02353 | -       |
| dukung           | 0.04706 | -       |
| <i>childfree</i> | 0.07059 | 0.06383 |
| .....            | .....   | .....   |
| .....            | .....   | .....   |
| .....            | .....   | .....   |
| perintah         | -       | 0.02128 |
| pantesan         | -       | 0.02128 |
| jepang           | -       | 0.02128 |
| punah            | -       | 0.02128 |

Berikut adalah tahapan pengujian *Naive bayes* menggunakan *tools orange* yang ditunjukkan pada gambar 2 di bawah ini :



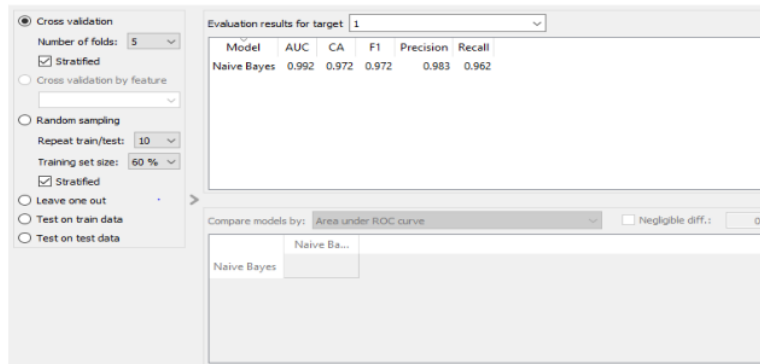
**Gambar 2.** Alur Pengujian *Orange*

Hasil Klasifikasi Algoritma *Naïve Bayes* adalah sebagai berikut, perhitungan untuk *accuracy*, *precision*, dan *recall* yang menggabungkan sistemisasi digunakan untuk menunjukkan seberapa baik algoritma *Naïve Bayes* dilakukan saat menggunakan *tools orange*. Hasil klasifikasi ditampilkan pada gambar 3 sampai 5 di bawah ini:



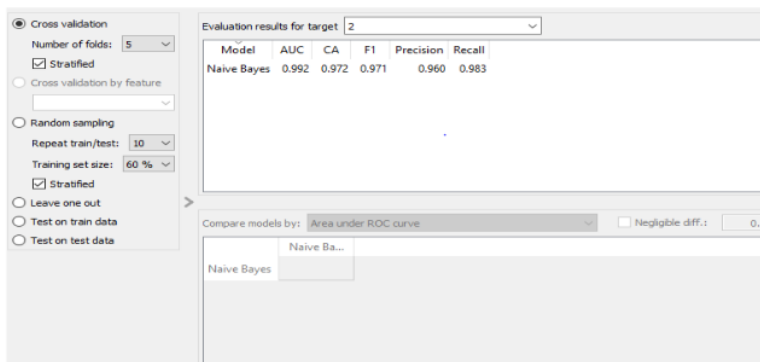
**Gambar 3.** Klasifikasi Rata - Rata Dari Tiap Kelas Sentimen

Pada gambar 3 di atas merupakan hasil perhitungan *accuracy*, *precision*, dan *recall*, pada rata-rata tiap kelas sentimennya.



**Gambar 4.** Klasifikasi Kelas Sentimen Negatif

Pada gambar 4 di atas merupakan hasil perhitungan *accuracy*, *precision*, dan *recall*, pada kelas sentimen negatif.



**Gambar 5.** Klasifikasi Kelas Sentimen Positif

Pada gambar 5 diatas merupakan hasil perhitungan *accuracy*, *precision*, dan *recall*, pada kelas sentimen positif.

Penilaian dilakukan untuk mengevaluasi keberhasilan dan ketepatan model klasifikasi. Metode *cross validation* digunakan dalam penelitian ini untuk menetapkan akurasi, presisi, dan recall, yang kemudian akan diukur dengan *Confusion Matrix*. Tabel 5 menunjukkan hasil *confusion matrix* yang diperoleh setelah menguji seluruh data komentar YouTube menggunakan *tools orange*.

**Tabel 4.** *Confusion Matrix*

|         | Negatif | Positif |
|---------|---------|---------|
| Negatif | 1249    | 50      |
| Positif | 21      | 1186    |

Perhitungan Akurasi

$$A = \frac{TP + TN}{\text{Jumlah Data}} = \frac{1249 + 1186}{2506} = \frac{2435}{2506} = 0,97$$

Perhitungan Presisi

**Tabel 5.** Perhitungan Presisi

|           | Negatif              | Positif              |
|-----------|----------------------|----------------------|
| TP        | 1249                 | 1186                 |
| FP        | 50                   | 21                   |
| Precision | $1249 / (1249 + 50)$ | $1186 / (1186 + 21)$ |
|           | 0,96                 | 0,98                 |

$$\text{All Precision} = \frac{\text{Precision A} + B + C}{\text{Jumlah Kelas}} = \frac{0,96 + 0,98}{2} = \frac{1,94}{2} = 0,97$$

Perhitungan Recall

**Tabel 6.** Perhitungan Recall

|           | Negatif              | Positif              |
|-----------|----------------------|----------------------|
| TP        | 1249                 | 1186                 |
| FN        | 21                   | 50                   |
| Precision | $1249 / (1249 + 21)$ | $1186 / (1186 + 50)$ |
|           | 0,98                 | 0,96                 |

$$\text{Recall} = \frac{\text{Recall A} + B + C}{\text{Jumlah Kelas}} = \frac{0,98 + 0,96}{2} = \frac{1,94}{2} = 0,97$$

Visualisasi adalah pembuatan gambar, grafik, dan animasi untuk tujuan menampilkan informasi. (Simanjuntak et al., 2022). *Word cloud* menggambarkan sentimen







**Gambar 8.** *Word Cloud* Kategori Negatif

## KESIMPULAN DAN REKOMENDASI

Kesimpulan yang didapat dari penelitian ini yaitu :

1. Dari hasil penelitian mengenai analisis sentimen penilaian masyarakat terhadap *childfree* berdasarkan komentar di YouTube diperoleh 2506 data, yang menjalani penyaringan tahap *preprocess text* terdiri tokenisasi, transformasi, dan klasifikasi menggunakan algoritma *naïve bayes*, serta evaluasi data dengan *confusion matrix* menggunakan *tools orange*.
2. Dapat dikatakan bahwa hasil dari reaksi orang terhadap *childfree* tergolong negatif dengan hasil persentase 97% pada akurasi (*accuracy*), 98% pada nilai *precision*, dan tingkat keberhasilan (*recall*) 96%. Nilai ini menunjukkan bahwa klasifikasi algoritma *naïve bayes* dinilai cukup baik dalam pemrosesan dokumen komentar YouTube, dikarenakan hasil persentase akurasi sebesar 97%.
3. Berdasarkan nilai tersebut membuktikan sentimen penilaian masyarakat terhadap *childfree* berdasarkan komentar di YouTube tergolong negatif.

Rekomendasi untuk penelitian analisis sentimen ini adalah :

1. Alangkah lebih bagus apabila peneliti mengambil komentar YouTube dari luar negeri seperti Jepang, Tiongkok, dan negara Arab Saudi.
2. Menggunakan algoritma lainnya dalam menganalisis sentimen seperti algoritma SVM dan algoritma K-Nearest Neighbour.

## REFERENSI

- Agama, I., Negeri, I., & Ponorogo, I. (2022). *Childfree Dalam Skripsi*.
- Ahmad, M. R. S. A., Syukur, M., Nisa, K., Putra Satria Mu, M., & Fadhilah Kadir, N. (2023). Dampak Pelabelan Terhadap Motivasi Belajar Siswa Impact of labeling on student motivation. *Comaserva: Jurnal Penelitian Dan Pengabdian Masyarakat*, 3(1), 313–318. <https://doi.org/10.59141/comserva.v3i1.756>
- Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, F. A. D. (2018). Metode-metode Klasifikasi. *Prosiding Seminar Ilmu Komputer Dan Teknologi Informasi*, 3(1), 134.
- Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(1), 147. <https://doi.org/10.25126/jtiik.0813944>
- Buana, U. M., Khusus, D., & Jakarta, I. (2023). Analisis Topik dan Perbandingan Klasifikasi pada Kolom Komentar Video Youtube Edukasi Indonesia Menggunakan Pendekatan Latent Dirichlet Allocation. 05(03), 7418–7429.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131. <https://doi.org/10.33365/jtk.v15i1.744>
- Fadhilah, E. (2022). Childfree Dalam Pandangan Islam. *Al-Mawarid Jurnal Syariah Dan Hukum (JSYH)*, 3(2), 71–80. <https://doi.org/10.20885/mawarid.vol3.iss2.art1>
- Hadinata, N., & Kurniawan, K. (2020). Analisis Pola Pembelian Produk Makanan Ringan Menggunakan Algoritma Apriori. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 9(1), 1–7. <https://doi.org/10.32736/sisfokom.v9i1.623>
- Halim, M. (2018). Pembuatan Aplikasi Pendukung Keputusan Beasiswa Tidak Mampu Dengan Metode Naive Bayes. *JSAI (Journal Scientific and Applied Informatics)*, 1(1), 24–31. <https://doi.org/10.36085/jsai.v1i1.6>
- Harpizon, H. A. R., Kurniawan, R., Iwan Iskandar, Salambue, R., Budianita, E., & Syafria, F. (2022). Analisis Sentimen Komentar Di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes. ... *Di YouTube ...*, 5(1), 131–140. <http://repository.uin-suska.ac.id/59746/>

- Kurniawan, M. A., Jalan, M. T., Transportasi, P., Bali, D., Samsam, D., Keselamatan, M., Jalan, T., Keselamatan, P., & Jalan, T. (2020). *Tingkat ketertarikan masyarakat terhadap transportasi online, angkutan pribadi dan angkutan umum berdasarkan persepsi*. 1(2), 51–58.
- Liu, B. (2010). *Analisis Sentimen dan Subjektivitas*. 1–38.
- Najjichah, H., Syukur, A., & Subagyo, H. (2019). Pengaruh Text Preprocessing Dan Kombinasinya Pada Peringkat Dokumen Otomatis Teks Berbahasa Indonesia. *Jurnal Teknologi Informasi*, XV(1), 1–11.
- Nurhafida, S. I., & Sembiring, F. (2021). *Analisis Text Clustering Masyarakat Di Twiter Mengenai Mcdonald's xpbs Menggunakan Orange Data Mining*. 28–35.
- Nurtikasari, Y., Syariful Alam, & Teguh Iman Hermanto. (2022). Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes. *INSOLOGI: Jurnal Sains Dan Teknologi*, 1(4), 411–423. <https://doi.org/10.55123/insologi.v1i4.770>
- Rachman, R., & Handayani, R. N. (2021). Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM. *Jurnal Informatika*, 8(2), 111–122. <https://doi.org/10.31294/ji.v8i2.10494>
- Simanjuntak, A. Y., Simatupang, I. S. E. S., & Anita, A. (2022). Implementasi Data Mining Menggunakan Metode Naï• Ve Bayes Classifier Untuk Data Kenaikan Pangkat Dinas Ketenagakerjaan Kota Medan. *Journal of Science and Social Research*, 5(1), 85. <https://doi.org/10.54314/jssr.v5i1.804>
- Wiguna, R. A. raffaidy, & Rifai, A. I. (2021). Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining. *Journal of Information Systems and Informatics*, 3(1), 1–12. <https://doi.org/10.33557/journalisi.v3i1.78>

ORIGINALITY REPORT

15%

SIMILARITY INDEX

12%

INTERNET SOURCES

7%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Perguruan Tinggi Pelita Bangsa

Student Paper

8%

2

[www.researchgate.net](http://www.researchgate.net)

Internet Source

4%

3

[journal.thamrin.ac.id](http://journal.thamrin.ac.id)

Internet Source

3%

Exclude quotes Off

Exclude bibliography On

Exclude matches < 3%