



## *Sentiment Analysis of Childfree Campaign on X Social Media Using NBC and SVM Algorithms*

### **Analisis Sentimen Masyarakat Mengenai Gerakan Childfree di Media Sosial X Menggunakan Algoritma NBC dan SVM**

Moh. Azlan Shah Putra<sup>1</sup>, Inggih Permana<sup>2</sup>, M. Afdal<sup>4</sup>

<sup>1,2,3,4</sup>Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,  
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

e-mail: <sup>1</sup>11753101087@students.uin-suska.ac.id,  
<sup>2</sup>inggihpermana@uin-suska.ac.id, <sup>4</sup>mafdal@uin-suska.ac.id

Received Apr 25th 2024; Revised Jun 7th 2024; Accepted Jun 27th 2024  
Corresponding Author: Moh. Azlan Shah Putra

#### **Abstract**

Children are one of the common entities in forming a family, but in recent years there has been discussion about childfree. With so many pro-con debates about childfree, it is necessary to analyze sentiment related to this issue. This research aims to analyze public sentiment regarding the childfree movement on social media X using the Naïve Bayes Classifier (NBC) and Support Vector Machine (SVM) algorithms. Sentiment is divided into 3 classes: positive, negative, and neutral. This research collects data by crawling data on social media X with the keyword childfree. The data obtained is raw text data so that a pre-process stage is needed. The pre-process stage carried out is tokenizing, case folding, stopword filter, stemming, TF-IDF, and data balancing. Based on simulation, the performance of the NBC algorithm is: accuracy = 56.36%, precision = 56.41%, and recall = 56.35%, while the performance of the SVM algorithm is: 76.12% accuracy, 76.36% precision, and 76.13% recall. So it can be concluded that SVM has better performance than NBC in sentiment analysis in this study. In this study, it was found that the majority of people on social media X are still against the childfree movement, but many people are more concerned about who is doing and how to communicate their opinions about childfree.

**Keywords:** Childfree, NBC, Sentiment Analysis, Social Media X, SVM

#### **Abstrak**

Anak merupakan salah satu entitas yang umum dalam membentuk sebuah keluarga, namun dalam beberapa tahun kebelakang muncul pembahasan mengenai childfree. Dengan banyaknya perdebatan pro-kontra mengenai childfree, perlu dilakukannya sentimen analisis terkait isu ini. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat mengenai gerakan childfree di media sosial X menggunakan algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM). Sentimen dibagi menjadi 3 kelas yaitu positif, negatif, dan netral. Penelitian ini mengumpulkan data dengan crawling data pada media sosial X dengan keyword childfree. Data yang diperoleh merupakan data teks mentah sehingga dibutuhkan tahap pra proses. Tahap pra proses yang dilakukan adalah tokenizing, case folding, filter stopword, stemming, TF-IDF, dan data balancing. Berdasarkan simulasi, performa algoritma NBC adalah: akurasi = 56,36%, presisi = 56,41%, dan recall = 56,35%, sedangkan performa algoritma SVM adalah: akurasi 76,12%, presisi 76,36%, dan recall 76,13%. Sehingga dapat disimpulkan bahwa SVM memiliki performa yang lebih baik dari pada NBC pada analisis sentimen di penelitian ini. Pada penelitian ini, didapatkan bahwa mayoritas masyarakat pada media sosial X masih bersifat kontra mengenai gerakan childfree, namun banyak khalayak yang lebih mempermasalahkan siapa yang melakukan dan cara mengkomunikasikan opininya mengenai childfree.

**Kata kunci:** Analisis Sentimen, Childfree, Media Sosial X, NBC, SVM

## **1. PENDAHULUAN**

Pernikahan adalah suatu institusi yang berdiri atas dasar cinta kasih, kesetiaan, dan saling percaya antara seorang pria dan seorang wanita sesuai dengan agama dan hukum yang berlaku, serta bertujuan memiliki keturunan yang baik dan membentuk keluarga yang sehat, sakinah, mawaddah, dan rahmah [1]. Pengertian perkawinan menurut UU No.1 Tahun 1974 adalah ikatan lahir batin seorang pria dan seorang wanita sebagai

suami istri dengan tujuan membentuk keluarga yang bahagia dan kekal berdasarkan Ketuhanan Yang Maha Esa [2]. Salah satu entitas yang umum dalam membentuk sebuah keluarga ialah anak-anak.

Memiliki anak dalam pernikahan dianggap sebagai suatu bentuk pemenuhan tanggung jawab sosial dan spiritual bagi pasangan suami istri [3]. Anak-anak merupakan anugerah yang dititipkan oleh Allah sebagai pemerkaya kehidupan keluarga serta menjadi amal jariyah yang terus berlanjut setelah kepergian orang tua. Melalui kehadiran anak, pasangan suami-istri dapat memperkuat ikatan keluarga, mengembangkan nilai-nilai agama, serta membentuk generasi penerus yang akan membangun masyarakat yang lebih baik [4]. Namun pada pembahasan sosial media belakangan, muncul sebuah gerakan baru yang ramai diperdebatkan dalam ranah pernikahan yaitu *Childfree*.

*Childfree* adalah istilah yang mulai populer di Indonesia dalam beberapa tahun kebelakang, yang mana *childfree* merupakan sebuah keputusan atau pilihan hidup yang dibuat oleh suatu pasangan untuk tidak memiliki anak. Terdapat beberapa alasan dan kondisi yang menjadi dasar gerakan ini antara lain seperti kebebasan hak, gaya hidup, kesadaran lingkungan untuk masa depan, kendala kesehatan dan kesetaraan *gender* [5]. Opini tentang *childfree* masih menjadi pembahasan yang kontroversial dan terkadang menimbulkan reaksi yang ekstrem dari masyarakat Indonesia. Selain itu, kepala Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) Tahun 2019-2024 ikut berpendapat tentang isu *childfree* ini, beliau menghimbau bagi calon suami-istri untuk mengikuti kelas pra-nikah di Kantor Urusan Agama (KUA) agar mereka mengetahui konsep ideal pernikahan.

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, maka dilakukanlah penelitian untuk menganalisis sentimen masyarakat tentang gerakan *childfree* dan diharapkan dapat berperan sebagai referensi dari fenomena resesi anak dan juga analisis sentimen pada media sosial X. Sentimen tersebut akan dibagi menjadi komentar positif, negatif, dan netral. Melalui penelitian ini, diharapkan dengan informasi sentimen masyarakat tentang *childfree* ini dapat membantu pihak KUA dalam melakukan bimbingan perkawinan di kelas pra-nikah sesuai dengan permasalahan sosial yang terjadi saat ini. Sentimen masyarakat mengenai gerakan *childfree* ini dianalisis melalui media sosial X.

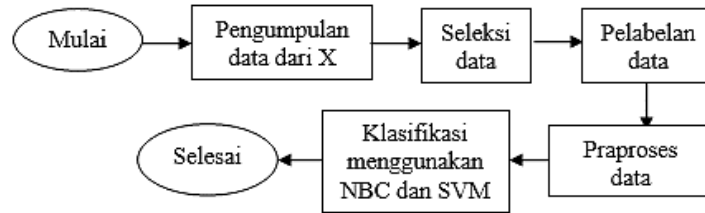
X (sebelumnya bernama Twitter) telah menjadi media sosial yang sangat populer untuk mengungkapkan opini terhadap suatu topik yang sedang hangat diperbincangkan, baik itu berupa opini positif maupun negatif melalui komentar dan kutipan. X memungkinkan pengguna untuk berinteraksi dengan orang lain secara mudah dan cepat. Berbeda dengan media sosial lain, pengguna X tetap dapat saling terhubung dan berinteraksi secara bebas tanpa harus berteman [6]. Dengan banyaknya orang yang berbagi opini di X, media sosial ini dapat dimanfaatkan sebagai sebuah analisis sentimen terhadap isu-isu yang hangat sehingga menjadi sebuah informasi sentimen masyarakat tentang isu tertentu [7].

Analisis sentimen yang dilakukan pada penelitian ini menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM). NBC adalah salah satu algoritma klasifikasi yang umum digunakan dalam *machine learning* untuk menentukan probabilitas kelas dari suatu data tertentu [8]. NBC merupakan salah satu algoritma yang sederhana dan mudah untuk diimplementasikan, dan NBC terbukti efektif dalam mengklasifikasikan data dalam berbagai aplikasi, termasuk analisis sentimen dan pengenalan teks [9]. Sedangkan SVM bekerja dengan membangun model yang dapat memisahkan dua kelas data dengan mencari *hyperplane* terbaik yang memiliki jarak terdekat antara *hyperlane* dengan data dari masing-masing kelas [10]. SVM memiliki kelebihan dalam menciptakan batas keputusan yang optimal sehingga dapat mengklasifikasikan data yang tidak diketahui dengan akurasi tinggi [11]. Disini penulis tertarik untuk membandingkan Algoritma NBC yang sederhana untuk melihat performanya terhadap sentimen-sentimen pada media sosial X dengan kalimatnya yang bersayap, dibandingkan dengan SVM yang dapat membangun model untuk memisahkan kelas-kelas yang berbeda.

Terdapat beberapa penelitian sebelumnya yang pernah melakukan analisis sentimen menggunakan Algoritma NBC dan SVM, seperti penelitian sentimen terkait vaksinasi [12], pada penelitian tersebut NBC memiliki akurasi yang lebih tinggi. Kemudian analisis sentimen terkait pembelajaran daring [13], pada penelitian ini menghasilkan SVM sebagai algoritma terbaik. Selanjutnya penelitian Enda Esyudha Pratama [14] membandingkan Algoritma NBC, SVM, dan decision tree terkait pembelajaran daring mahasiswa, pada penelitian ini mendapatkan NBC sebagai akurasi terbaik. Pada penelitian Deni Gunawan dkk [15] menganalisa sentimen calon gubernur 2018-2023 menggunakan NBC dan SVM, pada penelitian ini SVM memiliki akurasi terbaik. Dikarenakan selalu adanya perbedaan hasil algoritma terbaik antara NBC dan SVM dalam analisis sentimen, maka penelitian ini membandingkan performa antara Algoritma NBC dan SVM untuk analisis sentimen pada topik gerakan *childfree* dengan menggunakan parameter *grid search* pada setiap kernel Algoritma SVM, sehingga didapatkan titik *grid search* yang optimal.

**2. METODOLOGI PENELITIAN**

Alur tahap-tahap penelitian ini dapat dilihat pada Gambar 1.



**Gambar 1.** Alur tahap penelitian

**2.1. Pengumpulan Data**

Dataset pada penelitian ini, dikumpulkan dari komentar-komentar pengguna mengenai *childfree* di media sosial X. *Keyword* yang digunakan adalah “*childfree*”. Teknik pengumpulan data yang digunakan adalah *crawling data*. Teknik ini menggunakan bahasa pemrograman *Python* dengan *library pandas*. Rentang waktu komentar yang dipilih adalah Februari 2023 hingga Agustus 2023.

**2.2. Seleksi Data**

Pada tahapan ini, dilakukan pembersihan komentar-komentar yang tidak diperlukan pada penelitian yang dilakukan. Bentuk komentar yang tidak diperlukan adalah:

1. Komentar yang hanya menggunakan keyword *childfree* tetapi tidak membahas *childfree*.
2. Komentar yang bukan Bahasa Indonesia.
3. Komentar yang duplikat.
4. Komentar yang bermajas (contoh: emang ada baiknya kalian *childfree* biar dunia di masa depan damai dan tentram)

Selain itu, hasil dari *crawling data* menghasilkan beberapa atribut seperti *created\_at*, *id\_str*, *full\_text*, *quote\_count*, *reply\_count*, *retweet\_count*, *favorite\_count*, *language*, *user\_id\_str*, *conversation\_id\_str*, *username*, dan *tweet\_url*. Pada penelitian ini hanya diambil atribut *full\_text* saja.

**2.3. Pelabelan Data**

Dataset yang telah dikumpulkan melalui media sosial X merupakan dataset yang belum berlabel, sehingga perlu dilakukan pelabelan pada dataset agar dapat dipelajari. Pelabelan data dilakukan oleh seorang *annotator*. *Annotator* merupakan seseorang yang bertanggungjawab untuk pemberian label (positif, negatif, dan netral) pada masing-masing komentar [16]. *Annotator* pada penelitian ini adalah seseorang yang memiliki latar belakang pada bidang Pendidikan Bahasa Indonesia.

**2.4. Praproses Data**

Tahapan pra proses dilakukan menggunakan *tools* Rapidminer versi 9.10. Pada tahap ini dilakukan 6 tahapan, yaitu:

1. *Tokenizing*, yaitu proses memisahkan teks komentar menjadi perkata berdasarkan penggunaan spasi.
2. *Case Folding*, yaitu proses mengubah seluruh komentar menjadi huruf kecil.
3. *Filter Stopwords*, yaitu proses untuk membuang kata-kata dianggap tidak memiliki makna (contoh: ini, itu, mau, yang). Proses ini menggunakan kamus *stopword* Bahasa Indonesia [17].
4. *Stemming*, yaitu proses mengubah kata berimbuhan yang ada menjadi kata dasar. Proses ini menggunakan kamus *stemming* yang telah dibuat oleh peneliti dan terdiri dari 5.928 kata. Kamus tersebut dapat dilihat pada link <https://bit.ly/4a65Qbq>.
5. *Term Frequency dan Inverse Document Frequency (TF-IDF)*, yaitu proses untuk menentukan keterikatan kata (*term*) terhadap dokumen dengan memberikan nilai pada setiap kata [18]. Hasil dari proses ini adalah tabel TF-IDF.
6. *Data balancing*, yaitu proses untuk menyeimbangkan jumlah class pada data. Penelitian ini menggunakan metode *Random Under Sampling (RUS)* untuk *balancing data*. Dengan melakukan RUS, maka dataset dari kelas mayoritas akan berkurang jumlahnya sehingga menjadi seimbang dengan kelas minoritas [19].

Contoh *tokenizing*, *case folding*, *filter stopword* dan *stemming* dapat dilihat pada Tabel 1.

**Tabel 1.** Hasil praproses data sentimen

Tahapan	Hasil Praproses Data
Data Awal	Childfree tidak ingin punya anak adalah sebuah kekonyolan. Bukankah anak adalah PENERUS

Tahapan	Hasil Praproses Data
Tokenize	PERJUANGAN kita dan sebab kita mendapatkan pahala jariyah walaupun kita telah mati [Childfree] [tidak] [ingin] [punya] [anak] [adalah] [sebuah] [kekonyolan] [Bukankah] [anak] [adalah] [PENERUS] [PERJUANGAN] [kita] [dan] [sebab] [kita] [mendapatkan] [pahala] [jariyah] [walaupun] [kita] [telah] [mati]
Case Folding	[childfree] [tidak] [ingin] [punya] [anak] [adalah] [sebuah] [kekonyolan] [bukankah] [anak] [adalah] [penerus] [perjuangan] [kita] [dan] [sebab] [kita] [mendapatkan] [pahala] [jariyah] [walaupun] [kita] [telah] [mati]
Filter Stopword	[childfree] [anak] [sebuah] [kekonyolan] [sebuah] [anak] [penerus] [perjuangan] [sebab] [pahala] [jariyah] [mati]
Stemming	[childfree] [anak] [sebuah] [konyol] [sebuah] [anak] [terus] [juang] [sebab] [pahala] [jariyah] [mati]
Data Akhir	childfree anak sebuah konyol sebuah anak terus juang sebab pahala jariyah mati

## 2.5. Klasifikasi Menggunakan NBC dan SVM

Tahap ini dilakukan dengan menggunakan *tools* Rapidminer versi 9.10. Pada algoritma SVM, parameter-parameter yang dapat dilihat pada Tabel 2. *Tuning* parameter menggunakan *grid search* yang berdasarkan referensi dari penelitian Nugraha [20] dan kapabilitas dari *tools* yang digunakan oleh penulis. Metode validasi yang digunakan adalah *K-Fold Cross Validation* dengan nilai  $k = 10$ . Performa model klasifikasi dilihat dari *confusion matrix*, dan metrik-metrik pengukuran performa seperti akurasi, precision, dan recall (untuk lebih jelasnya lihat Tabel 3).

**Table 2.** Parameter Algoritma SVM

No	Parameter	Nilai
1	Kernel	Linear
	Cost (C)	$10^{-2}, 10^{-1}, 10^0, 10^1$
2	Kernel	Polynomial
	Cost (C)	$10^{-2}, 10^{-1}, 10^0, 10^1$
	Gamma (y)	$2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3$
	Degree (d)	1, 2, 3
3	Kernel	RBF
	Cost (C)	$2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3$
	Gamma (y)	$2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3$
4	Kernel	Sigmoid
	Cost (C)	$10^{-2}, 10^{-1}, 10^0, 10^1$
	Gamma (y)	$2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3$

**Tabel 3.** Metrik Pengukuran Performa

Metrik	Rumus
Akurasi	$\frac{TP+TN+TNg}{TP+TN+TNg+FP+FN+FNg} \times 100\%$ (1)
Presisi	$\frac{TP}{TP+TN+TNg} \times 100\%$ (2)
Recall	$\frac{TP}{TP+FN+FNg} \times 100\%$ (3)

## 3. HASIL DAN PEMBAHASAN

### 3.1. Hasil Pengumpulan Data dan Seleksi Data

Berdasarkan hasil pengumpulan data dari media sosial X pada Februari 2023 hingga Agustus 2023, didapatkan dataset sebanyak 3.202 komentar tentang *childfree*. Setelah dilakukan seleksi data tersisa dataset sebanyak 1.217 komentar. Pada tabel TF-IDF terlihat didapat 2.507 kata (Lihat Tabel 4). Jumlah *class* positif, netral, dan negatif, sebelum dan setelah RUS dapat dilihat pada Tabel 5.

**Tabel 4.** Hasil TF-IDF

No	1	2	3	4	5	6	...	2.507
	Adab	Anak	Bebas	Bocah	Mandul	Nikah	...	Zaman
1	0,129	0,025	0	0,505	0,107	0,167	...	0
2	0	0,036	0,439	0,237	0,138	0,105	...	0,229
...	...	...	...	...	...	...	...	...
1.217	0,602	0,369	0,299	0,669	0,429	0,141	...	0

**Tabel 5. Balancing Data**

Label	Sebelum RUS	Setelah RUS
Negatif	415	398
Netral	398	398
Positif	404	398

**3.2. Klasifikasi Menggunakan Algoritma NBC dan SVM**

Hasil percobaan-percobaan yang dilakukan sesuai dengan parameter pada Tabel 2 dapat dilihat Tabel 6 dan Tabel 7.

**Tabel 6. Hasil Percobaan Kernel Polynomial**

Iterasi	Degree	Cost	Gamma	Akurasi	Presisi	Recall
1	1.0	10 <sup>-2</sup>	2 <sup>-3</sup>	0.401	0.231	0.406
2	1.0	10 <sup>-2</sup>	2 <sup>-2</sup>	0.399	0.227	0.404
3	1.0	10 <sup>-2</sup>	2 <sup>-1</sup>	0.391	0.219	0.396
4	1.0	10 <sup>-2</sup>	2 <sup>0</sup>	0.401	0.227	0.406
5	1.0	10 <sup>-2</sup>	2 <sup>1</sup>	0.395	0.228	0.400
6	1.0	10 <sup>-2</sup>	2 <sup>2</sup>	0.413	0.240	0.418
7	1.0	10 <sup>-2</sup>	2 <sup>3</sup>	0.459	0.735	0.462
8	1.0	10 <sup>-1</sup>	2 <sup>-3</sup>	0.399	0.227	0.404
9	1.0	10 <sup>-1</sup>	2 <sup>-2</sup>	0.400	0.230	0.405
10	1.0	10 <sup>-1</sup>	2 <sup>-1</sup>	0.403	0.230	0.408
11	1.0	10 <sup>-1</sup>	2 <sup>0</sup>	0.627	0.717	0.629
12	1.0	10 <sup>-1</sup>	2 <sup>1</sup>	0.721	0.732	0.721
13	1.0	10 <sup>-1</sup>	2 <sup>2</sup>	0.738	0.744	0.738
14	1.0	10 <sup>-1</sup>	2 <sup>3</sup>	0.754	0.760	0.754
15	1.0	10 <sup>0</sup>	2 <sup>-3</sup>	0.704	0.715	0.705
16	1.0	10 <sup>0</sup>	2 <sup>-2</sup>	0.729	0.738	0.729
17	1.0	10 <sup>0</sup>	2 <sup>-1</sup>	0.739	0.748	0.739
18	1.0	10 <sup>0</sup>	2 <sup>0</sup>	0.737	0.743	0.737
19	1.0	10 <sup>0</sup>	2 <sup>1</sup>	0.729	0.733	0.729
20	1.0	10 <sup>0</sup>	2 <sup>2</sup>	0.708	0.711	0.709
21	1.0	10 <sup>0</sup>	2 <sup>3</sup>	0.698	0.703	0.698
22	1.0	10 <sup>1</sup>	2 <sup>-3</sup>	0.737	0.743	0.737
23	1.0	10 <sup>1</sup>	2 <sup>-2</sup>	0.723	0.727	0.722
24	1.0	10 <sup>1</sup>	2 <sup>-1</sup>	0.716	0.721	0.716
25	1.0	10 <sup>1</sup>	2 <sup>0</sup>	0.703	0.708	0.703
26	1.0	10 <sup>1</sup>	2 <sup>1</sup>	0.687	0.698	0.687
27	1.0	10 <sup>1</sup>	2 <sup>2</sup>	0.709	0.716	0.709
28	1.0	10 <sup>1</sup>	2 <sup>3</sup>	0.704	0.711	0.705
29	2.0	10 <sup>-2</sup>	2 <sup>-3</sup>	0.377	0.222	0.382
30	2.0	10 <sup>-2</sup>	2 <sup>-2</sup>	0.382	0.240	0.387
31	2.0	10 <sup>-2</sup>	2 <sup>-1</sup>	0.371	0.213	0.376
32	2.0	10 <sup>-2</sup>	2 <sup>0</sup>	0.367	0.223	0.372
33	2.0	10 <sup>-2</sup>	2 <sup>1</sup>	0.380	0.226	0.385
34	2.0	10 <sup>-2</sup>	2 <sup>2</sup>	0.366	0.194	0.371
35	2.0	10 <sup>-2</sup>	2 <sup>3</sup>	0.714	0.728	0.714
36	2.0	10 <sup>-1</sup>	2 <sup>-3</sup>	0.375	0.232	0.380
37	2.0	10 <sup>-1</sup>	2 <sup>-2</sup>	0.360	0.239	0.365
38	2.0	10 <sup>-1</sup>	2 <sup>-1</sup>	0.373	0.216	0.378
39	2.0	10 <sup>-1</sup>	2 <sup>0</sup>	0.372	0.232	0.377
40	2.0	10 <sup>-1</sup>	2 <sup>1</sup>	0.592	0.714	0.594
41	2.0	10 <sup>-1</sup>	2 <sup>2</sup>	0.718	0.723	0.718
42	2.0	10 <sup>-1</sup>	2 <sup>3</sup>	0.723	0.729	0.722
43	2.0	10 <sup>0</sup>	2 <sup>-3</sup>	0.386	0.232	0.391
44	2.0	10 <sup>0</sup>	2 <sup>-2</sup>	0.372	0.197	0.377
45	2.0	10 <sup>0</sup>	2 <sup>-1</sup>	0.387	0.487	0.392
46	2.0	10 <sup>0</sup>	2 <sup>0</sup>	0.721	0.728	0.721
47	2.0	10 <sup>0</sup>	2 <sup>1</sup>	0.721	0.727	0.721
48	2.0	10 <sup>0</sup>	2 <sup>2</sup>	0.729	0.734	0.729
49	2.0	10 <sup>0</sup>	2 <sup>3</sup>	0.718	0.724	0.718
50	2.0	10 <sup>1</sup>	2 <sup>-3</sup>	0.383	0.233	0.388
51	2.0	10 <sup>1</sup>	2 <sup>-2</sup>	0.722	0.730	0.722
52	2.0	10 <sup>1</sup>	2 <sup>-1</sup>	0.724	0.733	0.724
53	2.0	10 <sup>1</sup>	2 <sup>0</sup>	0.719	0.726	0.720
54	2.0	10 <sup>1</sup>	2 <sup>1</sup>	0.724	0.737	0.725

Iterasi	Degree	Cost	Gamma	Akurasi	Presisi	Recall
55	2.0	10 <sup>1</sup>	2 <sup>2</sup>	0.725	0.729	0.725
56	2.0	10 <sup>1</sup>	2 <sup>3</sup>	0.724	0.732	0.724
57	3.0	10 <sup>-2</sup>	2 <sup>-3</sup>	0.337	0.210	0.342
58	3.0	10 <sup>-2</sup>	2 <sup>-2</sup>	0.332	0.210	0.3375
59	3.0	10 <sup>-2</sup>	2 <sup>-1</sup>	0.334	0.210	0.339
60	3.0	10 <sup>-2</sup>	2 <sup>0</sup>	0.333	0.177	0.338
61	3.0	10 <sup>-2</sup>	2 <sup>1</sup>	0.337	0.244	0.343
62	3.0	10 <sup>-2</sup>	2 <sup>2</sup>	0.387	0.695	0.389
63	3.0	10 <sup>-2</sup>	2 <sup>3</sup>	0.682	0.688	0.682
64	3.0	10 <sup>-1</sup>	2 <sup>-3</sup>	0.335	0.177	0.340
65	3.0	10 <sup>-1</sup>	2 <sup>-2</sup>	0.331	0.193	0.336
66	3.0	10 <sup>-1</sup>	2 <sup>-1</sup>	0.335	0.210	0.340
67	3.0	10 <sup>-1</sup>	2 <sup>0</sup>	0.332	0.210	0.337
68	3.0	10 <sup>-1</sup>	2 <sup>1</sup>	0.443	0.705	0.443
69	3.0	10 <sup>-1</sup>	2 <sup>2</sup>	0.689	0.691	0.682
70	3.0	10 <sup>-1</sup>	2 <sup>3</sup>	0.680	0.685	0.680
71	3.0	10 <sup>0</sup>	2 <sup>-3</sup>	0.331	0.210	0.336
72	3.0	10 <sup>0</sup>	2 <sup>-2</sup>	0.332	0.210	0.338
73	3.0	10 <sup>0</sup>	2 <sup>-1</sup>	0.337	0.244	0.343
74	3.0	10 <sup>0</sup>	2 <sup>0</sup>	0.673	0.678	0.673
75	3.0	10 <sup>0</sup>	2 <sup>1</sup>	0.679	0.687	0.679
76	3.0	10 <sup>0</sup>	2 <sup>2</sup>	0.681	0.686	0.681
77	3.0	10 <sup>0</sup>	2 <sup>3</sup>	0.679	0.690	0.679
78	3.0	10 <sup>1</sup>	2 <sup>-3</sup>	0.335	0.210	0.340
79	3.0	10 <sup>1</sup>	2 <sup>-2</sup>	0.332	0.210	0.338
80	3.0	10 <sup>1</sup>	2 <sup>-1</sup>	0.676	0.681	0.676
81	3.0	10 <sup>1</sup>	2 <sup>0</sup>	0.687	0.692	0.687
82	3.0	10 <sup>1</sup>	2 <sup>1</sup>	0.688	0.694	0.689
83	3.0	10 <sup>1</sup>	2 <sup>2</sup>	0.683	0.689	0.683
84	3.0	10 <sup>1</sup>	2 <sup>3</sup>	0.672	0.676	0.672

Tabel 7. Hasil Percobaan Kernel Linear, RBF dan Sigmoid

Iterasi	Kernel	Cost	Gamma	Akurasi	Presisi	Recall
1	Linear	10 <sup>-2</sup>	-	0.394	0.255	0.399
2	RBF	10 <sup>-2</sup>	2 <sup>-3</sup>	0.749	0.754	0.749
3	Sigmoid	10 <sup>-2</sup>	2 <sup>-3</sup>	0.746	0.751	0.7463
4	RBF	10 <sup>-2</sup>	2 <sup>-2</sup>	0.740	0.746	0.740
5	Sigmoid	10 <sup>-2</sup>	2 <sup>-2</sup>	0.741	0.749	0.741
6	RBF	10 <sup>-2</sup>	2 <sup>-1</sup>	0.739	0.745	0.740
7	Sigmoid	10 <sup>-2</sup>	2 <sup>-1</sup>	0.748	0.751	0.748
8	RBF	10 <sup>-2</sup>	2 <sup>0</sup>	0.761	0.765	0.761
9	Sigmoid	10 <sup>-2</sup>	2 <sup>0</sup>	0.738	0.744	0.738
10	RBF	10 <sup>-2</sup>	2 <sup>1</sup>	0.734	0.740	0.733
11	Sigmoid	10 <sup>-2</sup>	2 <sup>1</sup>	0.734	0.741	0.734
12	RBF	10 <sup>-2</sup>	2 <sup>2</sup>	0.681	0.706	0.682
13	Sigmoid	10 <sup>-2</sup>	2 <sup>2</sup>	0.719	0.726	0.719
14	RBF	10 <sup>-2</sup>	2 <sup>3</sup>	0.367	0.599	0.367
15	Sigmoid	10 <sup>-2</sup>	2 <sup>3</sup>	0.691	0.696	0.691
16	Linear	10 <sup>-1</sup>	-	0.656	0.721	0.658
17	RBF	10 <sup>-1</sup>	2 <sup>-3</sup>	0.732	0.738	0.732
18	Sigmoid	10 <sup>-1</sup>	2 <sup>-3</sup>	0.744	0.748	0.744
19	RBF	10 <sup>-1</sup>	2 <sup>-2</sup>	0.738	0.744	0.738
20	Sigmoid	10 <sup>-1</sup>	2 <sup>-2</sup>	0.747	0.754	0.747
21	RBF	10 <sup>-1</sup>	2 <sup>-1</sup>	0.742	0.745	0.742
22	Sigmoid	10 <sup>-1</sup>	2 <sup>-1</sup>	0.734	0.740	0.734
23	RBF	10 <sup>-1</sup>	2 <sup>0</sup>	0.738	0.745	0.738
24	Sigmoid	10 <sup>-1</sup>	2 <sup>0</sup>	0.734	0.740	0.734
25	RBF	10 <sup>-1</sup>	2 <sup>1</sup>	0.733	0.736	0.733
26	Sigmoid	10 <sup>-1</sup>	2 <sup>1</sup>	0.726	0.736	0.726
27	RBF	10 <sup>-1</sup>	2 <sup>2</sup>	0.702	0.724	0.702
28	Sigmoid	10 <sup>-1</sup>	2 <sup>2</sup>	0.712	0.719	0.712
29	RBF	10 <sup>-1</sup>	2 <sup>3</sup>	0.358	0.647	0.358
30	Sigmoid	10 <sup>-1</sup>	2 <sup>3</sup>	0.678	0.685	0.679
31	Linear	10 <sup>0</sup>	-	0.745	0.750	0.745
32	RBF	10 <sup>0</sup>	2 <sup>-3</sup>	0.742	0.746	0.742
33	Sigmoid	10 <sup>0</sup>	2 <sup>-3</sup>	0.739	0.743	0.739

Iterasi	Kernel	Cost	Gamma	Akurasi	Presisi	Recall
34	RBF	10 <sup>0</sup>	2 <sup>-2</sup>	0.734	0.740	0.734
35	Sigmoid	10 <sup>0</sup>	2 <sup>-2</sup>	0.744	0.750	0.744
36	RBF	10 <sup>0</sup>	2 <sup>-1</sup>	0.746	0.759	0.746
37	Sigmoid	10 <sup>0</sup>	2 <sup>-1</sup>	0.738	0.743	0.738
38	RBF	10 <sup>0</sup>	2 <sup>0</sup>	0.753	0.755	0.753
39	Sigmoid	10 <sup>0</sup>	2 <sup>0</sup>	0.739	0.743	0.740
40	RBF	10 <sup>0</sup>	2 <sup>1</sup>	0.743	0.748	0.743
41	Sigmoid	10 <sup>0</sup>	2 <sup>1</sup>	0.713	0.721	0.713
42	RBF	10 <sup>0</sup>	2 <sup>2</sup>	0.685	0.713	0.685
43	Sigmoid	10 <sup>0</sup>	2 <sup>2</sup>	0.711	0.717	0.711
44	RBF	10 <sup>0</sup>	2 <sup>3</sup>	0.354	0.588	0.354
45	Sigmoid	10 <sup>0</sup>	2 <sup>3</sup>	0.693	0.699	0.693
46	Linear	10 <sup>1</sup>	-	0.699	0.706	0.699
47	RBF	10 <sup>1</sup>	2 <sup>-3</sup>	0.744	0.749	0.744
48	Sigmoid	10 <sup>1</sup>	2 <sup>-3</sup>	0.732	0.738	0.732
49	RBF	10 <sup>1</sup>	2 <sup>-2</sup>	0.738	0.744	0.738
50	Sigmoid	10 <sup>1</sup>	2 <sup>-2</sup>	0.745	0.751	0.746
51	RBF	10 <sup>1</sup>	2 <sup>-1</sup>	0.748	0.751	0.748
52	Sigmoid	10 <sup>1</sup>	2 <sup>-1</sup>	0.733	0.736	0.733
53	RBF	10 <sup>1</sup>	2 <sup>0</sup>	0.750	0.755	0.750
54	Sigmoid	10 <sup>1</sup>	2 <sup>0</sup>	0.746	0.752	0.746
55	RBF	10 <sup>1</sup>	2 <sup>1</sup>	0.742	0.746	0.742
56	Sigmoid	10 <sup>1</sup>	2 <sup>1</sup>	0.714	0.723	0.714
57	RBF	10 <sup>1</sup>	2 <sup>2</sup>	0.685	0.706	0.686
58	Sigmoid	10 <sup>1</sup>	2 <sup>2</sup>	0.721	0.728	0.721
59	RBF	10 <sup>1</sup>	2 <sup>3</sup>	0.358	0.494	0.358
60	Sigmoid	10 <sup>1</sup>	2 <sup>3</sup>	0.681	0.691	0.681

Pada Tabel 7 di atas, didapatkan hasil percobaan Kernel RBF terbaik pada C: 10-2, dan y: 20, dengan akurasi = 76,1%, presisi = 76,5%, recall = 76,1%. Pada Kernel Sigmoid, hasil percobaan terbaik didapatkan pada C: 10-2, dan y: 2-1, dengan akurasi = 74,8%, presisi = 75,1%, recall = 74,8%. Dan pada Kernel Linear, hasil percobaan terbaik didapatkan pada C: 100 dengan akurasi = 74,5%, presisi = 75,0%, recall = 74,5%. Pada Kernel RBF dan Kernel Sigmoid, nilai gamma yang besar cenderung mendapatkan hasil yang lebih rendah daripada nilai gamma yang kecil dan hal tersebut terlepas dari nilai cost yang kecil ataupun besar. Sedangkan pada Kernel Linear, nilai cost yang kecil cenderung mendapatkan hasil yang lebih rendah.

Tabel 8 merupakan *confusion matrix* untuk hasil model klasifikasi dari Algoritma NBC. Pada tabel tersebut terlihat banyak terjadi salah klasifikasi dalam klasifikasi sentimen. Sehingga menyebabkan performa model klasifikasi yang dihasilkan oleh NBC rendah.

**Tabel 8.** *Confusion Matrix* NBC

		Aktual		
		Netral	Negatif	Positif
Prediksi	Netral	165	63	70
	Negatif	98	253	73
	Positif	135	82	255

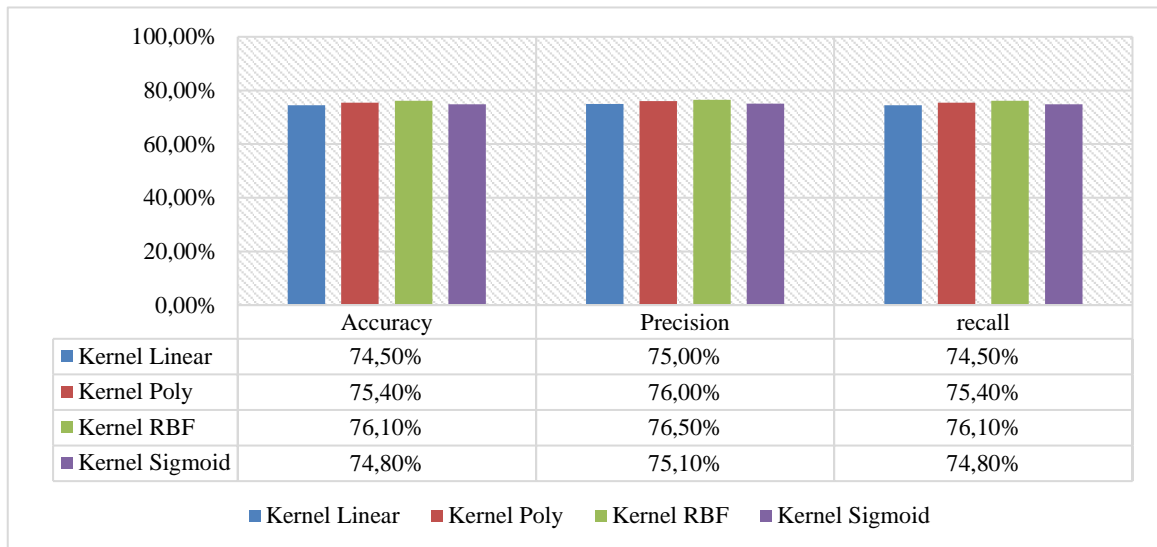
**Tabel 9.** *Confusion Matrix* SVM

		Aktual		
		Netral	Negatif	Positif
Prediksi	Netral	314	58	65
	Negatif	50	298	36
	Positif	34	42	297

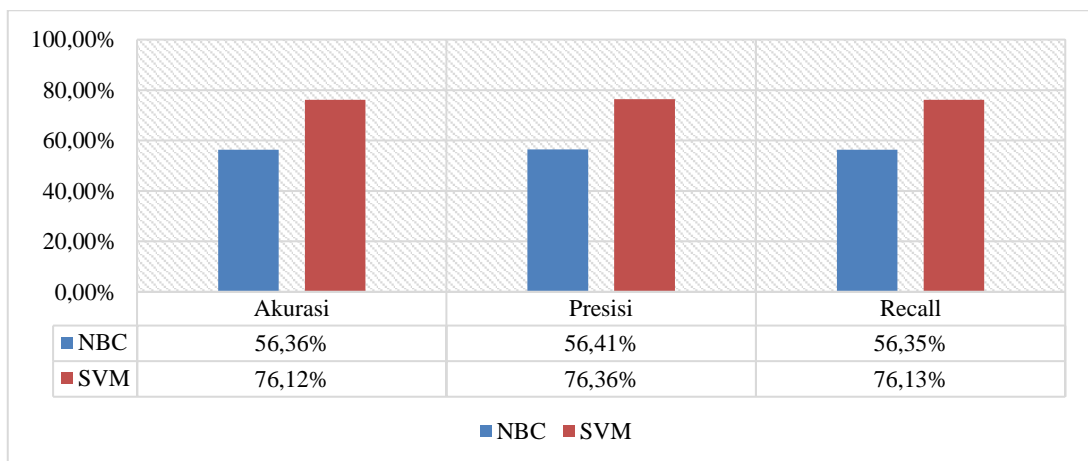
Tabel 9 merupakan *confusion matrix* untuk model klasifikasi dari Algoritma SVM (Kernel: RBF, C: 10<sup>-2</sup>, y: 2<sup>0</sup>). Pada *confusion matrix* tersebut terlihat bahwa model klasifikasi yang dihasilkan SVM lebih baik dari model klasifikasi yang dihasilkan oleh NBC. Meskipun pada SVM masih terdapat beberapa kesalahan klasifikasi.

Gambar 2 merupakan perbandingan performa model klasifikasi algoritma SVM dengan berbagai penggunaan kernel. Terlihat SVM dengan Kernel RBF memiliki performa terbaik, yaitu dengan akurasi = 76,1%, presisi = 76,5%, recall = 76,1%. Keempat kernel memiliki perbedaan performa yang tidak terlalu signifikan, sehingga disarankan untuk menggunakan kernel yang paling sederhana yaitu Kernel Linear. Kernel

Linear melakukan transformasi data ke dimensi yang lebih tinggi dengan cara yang paling sederhana dan langsung, yaitu dengan tanpa transformasi apapun [21].



**Gambar 2.** Perbandingan Performa Kernel SVM



**Gambar 3.** Perbandingan Performa Algoritma

Gambar 3 merupakan perbandingan performa model klasifikasi antar Algoritma NBC dan SVM. Pada gambar tersebut terlihat SVM memiliki performa yang jauh lebih baik dari pada NBC, baik dari segi akurasi, presisi, ataupun recall. Sehingga disarankan menggunakan SVM dalam klasifikasi sentimen dalam topik *childfree* pada sosial media X.



**Gambar 4.** Wordcloud Childfree



**Gambar 5.** Wordcloud Sentimen Negatif





Gambar 6. Wordcloud Sentimen Netral



Gambar 7. Wordcloud Sentimen Positif

#### 4. KESIMPULAN

Penelitian ini telah berhasil melakukan sentimen analisis pada media sosial X terkait childfree menggunakan Algoritma Support Vector Machine (SVM) dan Naïve Bayes Classifier (NBC). Berdasarkan pelabelan sentimen yang ada, masih banyak pro dan kontra yang terjadi dalam topik childfree ini, dan masih banyak masyarakat yang memberikan sentimen negatif terkait gerakan childfree ini seperti yang terlihat pada tabel 5. Namun secara keseluruhan, banyak masyarakat di media sosial X masih bersifat kontra mengenai gerakan childfree ini. Namun banyak juga diantaranya memaklumi, mempertanyakan, dan berdiskusi mengenai tindakan childfree yang dilakukan dalam berbagai situasi dan kondisi. Dan mayoritas masyarakat lebih memperlakukan siapa dan bagaimana cara menegkomunikasikan pandangannya mengenai childfree itu sendiri.

Pada penelitian ini dilakukan penanganan pada data yang tidak seimbang menggunakan metode Random Under Sampling (RUS). Pengklasifikasian pada Algoritma NBC menghasilkan akurasi 56,36%, presisi 56,41%, dan recall 56,35%. Sedangkan pengklasifikasian pada algoritma SVM, pada Kernel RBF menghasilkan akurasi = 76,1%, presisi = 76,5%, recall = 76,1%; pada Kernel Polynomial dengan akurasi = 75,4%, presisi = 76%, recall = 75,4%; pada Kernel Sigmoid dengan akurasi = 74,8%, presisi = 74,8%, recall = 75,1%; dan pada Kernel Linear dengan akurasi = 74,5%, presisi = 75%, recall 74,5%. Dalam penelitian ini, tidak ada perbedaan yang signifikan pada hasil akurasi kernel-kernel SVM. Oleh karenanya, peneliti menyarankan untuk menggunakan Kernel Linear sebagai kernel yang paling sederhana diantara kernel lainnya. Dan secara keseluruhan, Algoritma SVM memiliki kinerja yang lebih baik dalam melakukan klasifikasi sentimen pada topik childfree di media sosial X. Sedangkan Algoritma NBC tidak disarankan untuk analisis sentimen pada media sosial X dikarenakan banyaknya penggunaan kalimat yang bermajas oleh user media sosial X.

#### REFERENSI

- [1] Q. Shihab, *Kisah-Kisah Inspiratif Seputar Pernikahan dan Keluarga*. Jakarta: Lentera Hati, 2013.
- [2] “Undang-Undang Republik Indonesia Nomor 1 Tahun 1974 Tentang Perkawinan.”
- [3] Q. Shihab, *Pandangan Hidup Muslim: Dari Masalah Akhirat sampai Ziarah Kubur*. Jakarta: Lentera Hati, 2003.
- [4] A. Azra, *Memahami Keluarga Islami: Tantangan & Harapan*. Jakarta: Gema Insani Press, 2011.
- [5] Y. D. Damayanti, A. A. Refiana, and M. F. A. Nuary, “Fenomena Childfree di Twitter pada Generasi Millennial,” in *Prosiding Konferensi Ilmiah Pendidikan*, Kota Pekalongan: FKIP Universitas Pekalongan, 2022, pp. 879–882. Accessed: Jul. 01, 2024. [Online]. Available: <https://proceeding.unikal.ac.id/index.php/kip>
- [6] A. M. Pravina, I. Cholissodin, and P. P. Adikara, “Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM),” 2019. Accessed: Jul. 01, 2024. [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/4793>
- [7] F. Syadid, “Analisis Sentimen Komentar Netizen Terhadap Calon Presiden Indonesia 2019 dari Twitter Menggunakan Algoritman Term Frequency-Invers Document Frequency dan Metode Multi Layer Perceptron Neural Network,” *UIN Syarif Hidayatullah*, UIN Syarif Hidayatullah Jakarta, Jakarta, pp. 1–74, 2019. Accessed: Jul. 01, 2024. [Online]. Available: <http://repository.uinjkt.ac.id/dspace/handle/123456789/48184>
- [8] A. Pratama, D. M. Midyanti, and S. Bahri, “Penerapan Naïve Bayes Classifier Dengan Algoritma Stemming Nazief Dan Adriani Untuk Aplikasi Deteksi Ujaran Kebencian Berbasis Web,” *Coding : Jurnal Komputer dan Aplikasi*, vol. 8, no. 1, pp. 227–236, 2020, doi: <https://dx.doi.org/10.26418/coding.v8i1.39457>.
- [9] A. McCallum, K. Nigam, and L. H. Ungar, “Efficient Clustering of High-Dimensional Data Sets with Application to Reference Matching,” in *Proceeding of the Sixth ACM SIGKDD International*

- Conference on Knowledge Discovery and Data Mining*, Mar. 2000, pp. 169–178. doi: <http://dx.doi.org/10.1145/347090.347123>.
- [10] C. J. C. Burges, “A Tutorial on Support Vector Machines for Pattern Recognition,” *Data Min Knowl Discov*, vol. 2, no. 2, pp. 121–167, Jun. 1998, doi: <https://doi.org/10.1023/A:1009715923555>.
- [11] S. Bumbungan, “Penerapan Particle Swarm Optimization (PSO) dalam Pemilihan Parameter Secara Otomatis pada Support Vector Machine (SVM) untuk Prediksi Kelulusan Mahasiswa Politeknik Amamapare Timika,” *Jurnal Teknik AMATA*, vol. 4, no. 1, pp. 81–93, 2023, doi: <https://doi.org/10.55334/jtam.v4i1.77>.
- [12] B. Laurensz, A. Sentimen, and E. Sedyono, “Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19 (Analysis of Public Sentiment on Vaccination in Efforts to Overcome the Covid-19 Pandemic),” 2021. doi: <https://doi.org/10.22146/jnteti.v10i2.1421>.
- [13] Ariansyah and M. Kusmira, “Komparasi Algoritma Naive Bayes dan Support Vector Machine untuk Memprediksi Pengaruh Pembelajaran Daring Terhadap Motivasi Belajar,” *Faktor Exacta*, vol. 14, no. 3, pp. 100–107, Oct. 2021, doi: [10.30998/faktorexacta.v14i3.10325](https://doi.org/10.30998/faktorexacta.v14i3.10325).
- [14] E. E. Pratama, “Perbandingan Algoritma untuk Klasifikasi Komentar Mahasiswa pada Pembelajaran Daring,” *CYBERNETICS*, vol. 6, no. 01, pp. 1–9, 2022, doi: <http://dx.doi.org/10.29406/cbn.v6i01.3988>.
- [15] D. Gunawan, D. Riana, D. Ardiansyah, F. Akbar, and S. Alfarizi, “Komparasi Algoritma Support Vector Machine dan Naive Bayes dengan Algoritma Genetika pada Analisis Sentimen Calon Gubernur Jabar 2018-2023,” *Jurnal Teknik Komputer AMIK BSI*, vol. 6, no. 1, pp. 121–129, 2020, doi: [10.31294/jtk.v4i2](https://doi.org/10.31294/jtk.v4i2).
- [16] W. D. Asfi, I. Permana, R. Novita, and A. Marsal, “Analisis Sentimen Pengguna Transportasi Online Maxim Pada Instagram Menggunakan Naive Bayes Classifier dan K-Nearest Neighbor,” *Jurnal Media Informatika Budidarma*, vol. 7, no. 3, pp. 1134–1143, 2023, doi: [10.30865/mib.v7i3.6336](https://doi.org/10.30865/mib.v7i3.6336).
- [17] L. Agusta, “Perbandingan Algoritma Stemming Porter dengan Algoritma Nazief & Adriani untuk Stemming Dokumen Teks Bahasa Indonesia,” in *Konferensi Nasional Sistem dan Informatika*, Bali: STMIK STIKOM BALI, Nov. 2009, pp. 196–201. Accessed: Jul. 01, 2024. [Online]. Available: <https://www.academia.edu/download/54053917/perbandingan-algoritma-stemming-porter-dengan-algoritma-nazief-adriani-untuk-stemming-dokumen-teks-bahasa-indonesia.pdf>
- [18] P. Fremmuzar and A. Baita, “Uji Kernel SVM dalam Analisis Sentimen Terhadap Layanan Telkomsel di Media Sosial Twitter,” *Komputika : Jurnal Sistem Komputer*, vol. 12, no. 2, pp. 57–66, Sep. 2023, doi: [10.34010/komputika.v12i2.9460](https://doi.org/10.34010/komputika.v12i2.9460).
- [19] W. I. Sabilla and C. B. Vista, “Implementasi SMOTE dan Under Sampling pada Imbalanced Dataset untuk Prediksi Kebangkrutan Perusahaan,” *Jurnal Politeknik Caltex Riau*, vol. 7, no. 2, pp. 329–339, 2021, doi: <https://dx.doi.org/10.35143/jkt.v7i2.5027>.
- [20] Y. Rizki and A. Nugraha, “Implementasi Algoritma Support Vector Machine untuk Klasifikasi Penyesuaian Uang Kuliah Tunggal Terdampak Pandemi Covid-19 (Studi Kasus: Universitas Siliwangi),” *Jurnal Teknik Informatika Kaputama (JTIK)*, vol. 5, no. 2, pp. 210–218, 2021, doi: <https://doi.org/10.59697/jtik.v5i2.545>.
- [21] C. J. S.-T. Nello, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press, 2000. doi: <https://doi.org/10.1017/CBO9780511801389>.