

**PENERAPAN *BINARY PARTICLE SWARM OPTIMIZATION*
SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI
KOMENTAR *CYBERBULLYING* DI INSTAGRAM**

DEWI FORTUNA



**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2021**

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penerapan *Binary Particle Swarm Optimization Support Vector Machine* untuk Klasifikasi Komentar *Cyberbullying* di Instagram” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, September 2021

Dewi Fortuna
NIM G14170004

ABSTRAK

DEWI FORTUNA. Penerapan *Binary Particle Swarm Optimization Support Vector Machine* untuk Klasifikasi Komentar *Cyberbullying* di Instagram. Dibimbing oleh ITASIA DINA SULVIANTI dan GERRY ALFA DITO.

Kebebasan berpendapat pada media sosial adakalanya tidak sesuai dengan etika berkomunikasi sehingga mengarah kepada tindakan *cyberbullying*. Instagram merupakan media sosial yang paling umum digunakan dalam melakukan *cyberbullying*. Tindakan *cyberbullying* perlu diminimalisasi karena menimbulkan banyak dampak negatif. Salah satu cara yang dapat dilakukan yaitu dengan mengidentifikasi komentar *cyberbullying* agar komentar tersebut dapat dihapus secara otomatis. Metode yang digunakan pada penelitian ini yaitu klasifikasi teks menggunakan algoritme *Support Vector Machine* (SVM) dengan penerapan seleksi fitur *Binary Particle Swarm Optimization* (BPSO). Penelitian ini bertujuan untuk menyusun model klasifikasi komentar *cyberbullying* serta membandingkan performa kinerja model klasifikasi dengan dan tanpa penerapan seleksi fitur. Hasil penelitian menunjukkan bahwa pemodelan dengan SVM menghasilkan performa klasifikasi yang cukup akurat lebih dari 72% untuk semua performa klasifikasi pada setiap nilai C. Penerapan seleksi fitur BPSO dapat meningkatkan performa klasifikasi dengan meningkatnya nilai akurasi dan spesifisitas. Akan tetapi untuk kasus pada penelitian ini dipilih model tanpa seleksi fitur pada $C = 0,1$ karena memiliki nilai sensitivitas paling besar dengan akurasi dan spesifisitas yang cukup baik sehingga dapat mendeteksi komentar *cyberbullying* secara lebih akurat.

Kata kunci: *binary particle swarm optimization, cyberbullying, seleksi fitur, support vector machine, text mining.*

ABSTRACT

DEWI FORTUNA. Application of Binary Particle Swarm Optimization Support Vector Machine for Cyberbullying Comments Classification on Instagram. Supervised by ITASIA DINA SULVIANI and GERRY ALFA DITO.

Freedom of speech on social media is sometimes inappropriate with the ethics of communicating and has led to cyberbullying. According to the surveys from Ditch The Label, Instagram is the most commonly used social media in cyberbullying. Cyberbullying needs to be minimized because it has many negative effects. One way that can be done is by identifying cyberbullying comments so those comments can be deleted automatically. The method used in this study is text classification using Support Vector Machine (SVM) algorithm with the application of Binary Particle Swarm Optimization (BPSO) as features selection. The study aims to build a cyberbullying comments classification model and compare the classification model performance with and without the application of features selection. The experimental results showed that modeling with SVM produces a reasonably accurate classification performance over 72% for all classification performance on each C. The application of BPSO features selection can improve classification performance by increasing accuracy and specificity. However, the model without features selection on $C = 0,1$ is chosen in this study case because it has the highest sensitivity with good accuracy and specificity that can detect cyberbullying comments more accurately.

Keywords: binary particle swarm optimization, cyberbullying, feature selection, support vector machine, text mining.

© Hak Cipta milik IPB, tahun 2021¹
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

**PENERAPAN *BINARY PARTICLE SWARM OPTIMIZATION*
SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI
KOMENTAR *CYBERBULLYING* DI INSTAGRAM**

DEWI FORTUNA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2021**

Tim Penguji pada Ujian Skripsi:
1 Nama lengkap dan gelar

Judul Skripsi : Penerapan *Binary Particle Swarm Optimization Support Vector Machine* untuk Klasifikasi Komentar *Cyberbullying* di Instagram

Nama : Dewi Fortuna

NIM : G14170004

Disetujui oleh

Pembimbing 1:

Dra. Itasia Dina Sulvianti, M.Si.

Pembimbing 2:

Gerry Alfa Dito, S.Stat., M.Si.

Diketahui oleh

Ketua Program Studi:

Dr. Anang Kurnia, S.Si., M.Si

NIP. 197308241997021001

Tanggal Ujian:

2 September 2021

Tanggal Lulus:

(tanggal penandatanganan oleh Dekan
Fakultas/Sekolah ...)

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah *subhanaahu wa ta'ala* atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Februari 2021 sampai bulan Juli 2021 ini ialah *text mining*, dengan judul “Penerapan *Particle Swarm Optimization Support Vector Machine* untuk Klasifikasi Komentar *Cyberbullying* di Instagram”.

Terima kasih penulis ucapkan kepada pihak-pihak yang telah memberi dukungan dan arahan selama proses penulisan skripsi, diantaranya:

1. Bapak, kakak, adik, tante dan seluruh keluarga atas segala dukungan baik moril maupun materiil untuk penulis, serta ibu yang senantiasa mendoakan dari surga.
2. Ibu Dra. Itasia Dina Sulvianti, M.Si dan Bapak Gerry Alfa Dito, S.Stat., M.Si yang telah membimbing dan banyak memberi saran.
3. Ibu/Bapak.... selaku Dosen Penguji Skripsi Departemen Statistika FMIPA IPB.
4. Seluruh Dosen Departemen Statistika FMIPA IPB atas ilmu yang telah diberikan selama perkuliahan.
5. Seluruh Staf dan TU Departemen Statistika FMIPA IPB yang telah membantu dan melancarkan penulisan karya ilmiah ini.
6. Sahabat terdekat penulis yaitu Syifa, Prisda, Raihana, Yuniar, Riska, Hanniva, Asti, Fardilla, Nada yang telah banyak membantu selama perkuliahan.
7. Syahnita, Widia, Ais, Khusnul, Nada, Indah selaku teman satu bimbingan yang sudah banyak membantu memberikan saran dan doa.
8. Kim Nam Joon, Kim Seok Jin, Min Yoon Gi, Jung Ho Seok, Park Ji Min, Kim Tae Hyung, Jeon Jung Kook atas karya-karyanya yang sangat memotivasi penulis dan memberi semangat saat mengerjakan skripsi di masa pandemi.
9. Mahasiswa/i Departemen Statistika FMIPA IPB Angkatan 54.
10. Seluruh pihak terkait yang telah membantu penulis yang tidak dapat dituliskan satu persatu.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, September 2021

Dewi Fortuna

DAFTAR ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xi
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	2
II TINJAUAN PUSTAKA	3
2.1 <i>Text Mining</i>	3
2.2 <i>Cyberbullying</i>	3
2.3 <i>Support Vector Machine</i> (SVM)	4
2.4 <i>Synthetic Minority Oversampling Technique</i> (SMOTE)	6
2.5 <i>Term Frequency-Inverse Document Frequency</i> (TF-IDF)	7
2.6 <i>Binary Particle Swarm Optimization</i> (BPSO)	7
2.7 Evaluasi Model	9
III METODE	11
3.1 Data	11
3.2 Prosedur Analisis Data	11
IV HASIL DAN PEMBAHASAN	15
4.1 Eksplorasi Data	15
4.2 Model <i>Support Vector Machine</i>	17
4.3 Seleksi Fitur BPSO	18
V SIMPULAN DAN SARAN	22
5.1 Simpulan	22
5.2 Saran	22
DAFTAR PUSTAKA	23
LAMPIRAN	25
RIWAYAT HIDUP	31

DAFTAR TABEL

1	<i>Confusion matrix</i> dari dua kelas	9
2	Contoh data hasil <i>scraping</i>	11
3	Komentar tidak sesuai konteks	12
4	Ilustrasi komentar yang telah diberi label	12
5	Ilustrasi komentar sebelum dan setelah <i>pre-processing</i>	13
6	Perbandingan performa klasifikasi sebelum dan setelah seleksi fitur BPSO	19

DAFTAR GAMBAR

1	Proses <i>text mining</i>	3
2	<i>Hyperplane</i> terbaik yang memisahkan kedua kelas -1 dan +1	4
3	Prosedur algoritme PSO	8
4	Pemodelan dengan seleksi fitur BPSO menggunakan <i>10-fold cross-validation</i>	14
5	Grafik perbandingan jenis <i>bullying</i>	15
6	Proporsi label pada setiap akun	15
7	Distribusi banyak karakter tiap kelas	16
8	<i>Wordcloud</i> kelas (a) “0: Bukan <i>cyberbullying</i> ” (b) “1: <i>Cyberbullying</i> ”	16
9	Performa klasifikasi untuk setiap nilai C	17
10	Rata-rata banyaknya fitur terpilih untuk setiap nilai C	18
11	Performa pada setiap fold (a) C = 0.1 (b) C = 1 (c) C = 10	20
12	Perbandingan fitur (a) Sebelum seleksi fitur (b) Setelah seleksi fitur	21

DAFTAR LAMPIRAN

1	Diagram alir prosedur analisis data	26
2	Hasil percobaan parameter BPSO	27
3	Banyaknya fitur terpilih pada tiap <i>fold</i>	29
4	Daftar tautan <i>Text Pre-processing</i>	30

I PENDAHULUAN

1.1 Latar Belakang

Media sosial saat ini menjadi kebutuhan bagi setiap orang. Menurut laporan Hootsuite (*We are Social*), pengguna media sosial di Indonesia pada Januari 2020 mencapai 160 juta pengguna atau setara dengan 59% dari total penduduk Indonesia. Media sosial merupakan media *online* dimana para penggunanya dapat berbagi, berpartisipasi, dan menciptakan akun jejaring sosial menggunakan aplikasi berbasis internet yang didukung oleh teknologi. Salah satu manfaat media sosial yaitu sebagai media berdiskusi dan menyampaikan pendapat. Akan tetapi, adakalanya pendapat yang diberikan tidak diikuti dengan penerapan etika dalam berkomunikasi sehingga mengarah kepada tindakan *cyberbullying* (Rifauddin 2016).

Cyberbullying merupakan perilaku seseorang atau sekelompok orang dalam melakukan tindakan yang dapat menyakiti orang lain melalui komputer, telepon seluler, dan alat elektronik lainnya secara sengaja dan berulang kali (Rifauddin 2016). Menurut hasil survei yang dilakukan oleh APJII pada April 2019 terhadap pengguna internet Indonesia, sebanyak 42% dari 5900 responden pernah mengalami *cyberbullying* melalui media sosial di Indonesia.

Menurut hasil survei dari Ditch The Label, Instagram merupakan media sosial yang paling umum digunakan dalam melakukan *cyberbullying*. Instagram merupakan media sosial yang digunakan untuk membagikan foto dan video yang di dalamnya terdapat fitur-fitur seperti *direct message*, komentar, suka, dan lain-lain. Tindakan *bullying* yang dilakukan yaitu memberi komentar negatif pada postingan tertentu, pesan pribadi tak bersahabat, atau menyebarkan postingan seseorang dengan mengolok-olok (Ditch The Label 2017). Menurut Anderson *et al.* (2014), *cyberbullying* dapat berdampak negatif terhadap beberapa aspek kesehatan mental seperti depresi, kecemasan sosial, bunuh diri, dan harga diri yang rendah.

Sebagai upaya meminimalisasi tindakan *cyberbullying* di Instagram, salah satu cara yang dapat dilakukan yaitu dengan mengidentifikasi suatu komentar apakah termasuk *cyberbullying* atau tidak. Hal ini dapat dilakukan dengan menerapkan analisis *text mining* dengan metode klasifikasi teks. Klasifikasi komentar dilakukan agar komentar *cyberbullying* dapat terdeteksi secara otomatis dengan bantuan *machine learning* sehingga komentar yang termasuk *cyberbullying* dapat terhapus secara otomatis. Metode klasifikasi yang akan digunakan pada penelitian ini yaitu metode *Support Vector Machine* (SVM).

Algoritme SVM merupakan salah satu metode yang umum digunakan dalam klasifikasi teks. Algoritme ini mempunyai akurasi yang tinggi dan lebih tidak rentan terhadap *overfitting* dibandingkan metode lainnya (Han *et al.* 2012). Penelitian yang dilakukan oleh Chrismanto dan Lukito (2017) menunjukkan SVM menghasilkan tingkat akurasi 78,49% yang lebih baik dari metode *Naïve Bayes* (77,25%) pada identifikasi komentar spam di Instagram. Penelitian lainnya juga dilakukan oleh Luqyana *et al.* (2018) mengenai analisis sentimen *cyberbullying* di Instagram menggunakan metode SVM yang menghasilkan akurasi sebesar 90%. Berdasarkan hal tersebut maka metode SVM dipilih pada penelitian ini.

Komentar teks pada Instagram terdiri dari banyak fitur/kata. Algoritme klasifikasi menggunakan semua fitur yang terdapat pada data untuk membangun sebuah model, padahal tidak semua fitur sesuai dengan hasil klasifikasi. Maka dari

itu, diperlukan seleksi fitur untuk memilih fitur penting yang relevan terhadap data (Maulida *et al.* 2016). Metode seleksi fitur yang akan diterapkan pada penelitian ini yaitu *Particle Swarm Optimization* (PSO) versi biner yang dinamakan *Binary Particle Swarm Optimization* (BPSO). Metode PSO dipilih karena mudah digunakan, serta cepat konvergen dengan sedikit parameter *tuning* (Kurniawan M 2017). Berdasarkan penelitian yang dilakukan oleh Pajri *et al.* (2020), penerapan seleksi fitur PSO dapat meningkatkan akurasi metode KNN dari 88,11% menjadi 97,9%. Dalam perbandingannya dengan metode lain, Arifin (2016) menyatakan bahwa metode PSO lebih baik digunakan pada seleksi fitur dibandingkan metode *Genetic Algorithm* (GA) pada analisis sentimen *review*. Berdasarkan hal tersebut, metode BPSO digunakan agar dapat meningkatkan performa klasifikasi pada penelitian ini. Penerapan BPSO untuk seleksi fitur dan SVM sebagai algoritme klasifikasi diharapkan dapat memberikan hasil maksimal dalam pengklasifikasian *cyberbullying* di Instagram.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disebutkan di atas, rumusan masalah pada penelitian ini yaitu:

1. Seberapa baik performa klasifikasi dari metode SVM yang diterapkan pada kasus *cyberbullying*?
2. Apakah penerapan seleksi fitur BPSO dapat memberikan hasil yang lebih baik dalam klasifikasi komentar *cyberbullying*?

1.3 Tujuan

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, tujuan penelitian ini yaitu menyusun model klasifikasi komentar *cyberbullying* serta membandingkan performa kinerja model klasifikasi sebelum dan setelah penerapan seleksi fitur.

1.4 Manfaat

Manfaat dari penelitian ini yaitu:

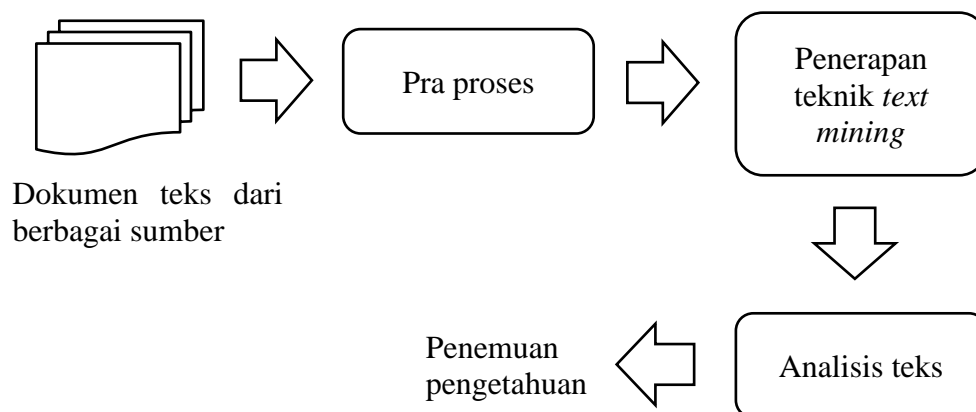
1. Model ini dapat menjadi pertimbangan bagi pengembang Instagram untuk membantu mendeteksi komentar Instagram secara otomatis menggunakan bantuan *machine learning*.
2. Mengetahui efektivitas dari penggunaan seleksi fitur BPSO pada metode SVM untuk klasifikasi komentar *cyberbullying* di Instagram.
3. Menambah informasi mengenai kata yang sering digunakan dalam *cyberbullying* sehingga pembaca dapat lebih bijak dalam berkomentar di media sosial.

II TINJAUAN PUSTAKA

2.1 Text Mining

Text Mining merupakan bidang interdisipliner yang mengacu pada pencarian informasi, penambangan data, pembelajaran mesin, statistika, dan linguistik komputasi. *Text mining* mirip seperti *data mining* hanya saja untuk teknik data mining didesain untuk bekerja pada data terstruktur dalam *database*, sedangkan *text mining* dapat bekerja pada data tidak terstruktur atau data semi-terstruktur. Sebagian besar informasi disimpan sebagai teks seperti artikel berita, perpustakaan digital, blog dan halaman web. Tujuan penting dari *text mining* yaitu untuk memperoleh informasi berkualitas tinggi dari sebuah teks. Jenis *text mining* yang umum digunakan meliputi kategorisasi teks, penggerombolan teks, konsep/entitas ekstraksi, produksi taksonomi granular, analisis sentimen, ringkasan dokumen, dan pemodelan hubungan entitas (Han *et al.* 2012).

Proses *text mining* diawali dengan pengumpulan dokumen dari berbagai sumber. Kemudian dilakukan *pre-processing* untuk mengecek format dan pengaturan karakter. Lalu dokumen akan melalui tahap analisis teks. Analisis teks adalah analisis semantik untuk memperoleh informasi berkualitas tinggi dari teks. Teknik *information retrieval* seperti pengindeksan teks telah dikembangkan untuk menangani dokumen yang tidak terstruktur (Gaikwad *et al.* 2014). Proses *text mining* ditunjukkan pada Gambar 1.



Gambar 1 Proses *text mining*

2.2 Cyberbullying

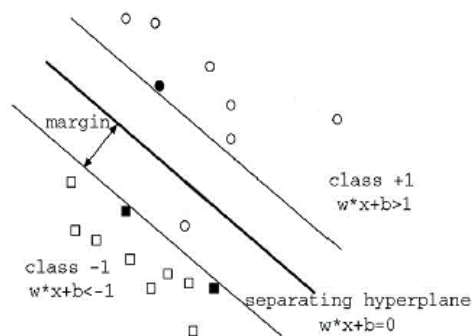
Cyberbullying merupakan perilaku seseorang atau sekelompok orang dalam melakukan tindakan yang dapat menyakiti orang lain melalui komputer, telepon seluler, dan alat elektronik lainnya secara sengaja dan berulang kali (Rifauddin 2016). Bentuk tindakan *cyberbullying* sangat beragam seperti mengunggah foto atau postingan yang memperlakukan korban, mengancam dan mengolok-olok korban, bahkan membuat situs web untuk menyebar fitnah. Motivasi pelaku dalam melakukan *cyberbullying* sangat beragam, diantaranya hanya sekedar bercanda, ingin mencari perhatian, marah, frustrasi, maupun balas dendam (Rifauddin 2016).

Price dan Dalgeish (2010) melakukan penelitian dengan mengadakan survei terhadap 548 remaja Australia yang pernah mengalami *cyberbullying*. Berdasarkan hasil penelitian tersebut, ditemukan beberapa bentuk *cyberbullying* yang sering dilakukan pada media internet antara lain:

- Called name* (Pemberian Nama Negatif): Memanggil atau memberi label kepada seseorang dengan nama negatif di media sosial seperti nama hewan (monyet, anjing, babi, tikus, buruk), nama makhluk halus (hantu, kuntilanak), dan panggilan fisik (wajah, keseluruhan fisik).
- Image of victim spread* (Penyebaran Foto): Menyebar foto aib orang lain yang dijadikan hinaan secara masif.
- Threatened physical harm* (Mengancam Keselamatan Fisik): Mengancam keselamatan orang lain di media sosial dengan memberi komentar yang berisi kata-kata seperti “mati” atau “bunuh”.
- Opinion slammed* (Pendapat yang Merendahkan): Merendahkan atau menghina keadaan atau penampilan orang lain di media sosial. Contohnya yaitu sebutan “alay”.

2.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan salah satu metode dalam *supervised learning* yang biasanya digunakan untuk klasifikasi. Ide dasar dari SVM adalah untuk menentukan pemisah (*hyperplane*) terbaik yang dapat memisahkan kelas yang berbeda pada ruang *input* (Taneja dan Sharma 2014).



Gambar 2 *Hyperplane* terbaik yang memisahkan kedua kelas -1 dan +1

Proses pencarian *hyperplane* terbaik yaitu dengan mencari *maximum marginal hyperplane* atau *hyperplane* dengan margin terbesar (Han *et al.* 2012). Margin adalah jarak antara *hyperplane* dengan titik terdekat dari masing-masing kelas. Titik terdekat tersebut disebut *support vector* yang pada Gambar 2 ditunjukkan dengan objek yang dihitamkan.

Misalkan terdapat himpunan data berukuran n yaitu (x_1, x_2, \dots, x_n) dengan label kelas dinotasikan $y_i \in \{-1, +1\}$. Menurut Han *et al* (2017), *hyperplane* klasifikasi linear SVM dinotasikan:

$$w \cdot x + b = 0 \quad (1)$$

dimana \mathbf{w} merupakan vektor bobot yang tegak lurus terhadap *hyperplane*, \mathbf{x} adalah vektor data masukan, b adalah bilangan skalar yaitu bias yang merepresentasikan posisi bidang relatif terhadap pusat koordinat.

$$\begin{aligned} \mathbf{w} \cdot \mathbf{x} + b &\geq 1 && \text{untuk } y_i = +1 \\ \mathbf{w} \cdot \mathbf{x} + b &\leq -1 && \text{untuk } y_i = -1 \end{aligned}$$

Nilai margin antara dua kelas yaitu $m = \frac{2}{\|\mathbf{w}\|}$. Proses memaksimalkan margin dilakukan dengan menggunakan fungsi optimasi *Lagrange multiplier* sebagai berikut:

$$L = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i (y_i (\mathbf{x}_i \cdot \mathbf{w} + b) - 1) \quad (2)$$

dimana $\alpha = (\alpha_1, \dots, \alpha_n)$ merupakan koefisien pengali *Lagrange*. Dengan meminimumkan L terhadap \mathbf{w} dan b dapat diperoleh persamaan :

$$\frac{\partial L}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i = 0 \rightarrow \mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \quad (3)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^n \alpha_i y_i = 0 \quad (4)$$

dengan memasukkan persamaan (3) dan (4) ke dalam persamaan (2) dan dilakukan beberapa operasi aljabar, maka akan didapatkan persamaan :

$$L = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \quad (5)$$

Solusi dari persamaan (5) didapatkan dengan memaksimumkan α_i . Nilai α yang diperoleh digunakan untuk mencari \mathbf{w} . *Support vector* memiliki nilai $\alpha_i \geq 0$ sedangkan titik amatan lainnya memiliki nilai $\alpha_i = 0$. Maka dari itu, fungsi keputusan hanya dipengaruhi oleh *support vector*.

Setelah diperoleh nilai α_i , maka kelas dari data pengujian \mathbf{x} ditentukan dengan fungsi keputusan sebagai berikut:

$$f(\mathbf{x}_d) = \sum_{i=1}^{NS} \alpha_i y_i \mathbf{x}_i \cdot \mathbf{x}_d + b \quad (6)$$

dengan

\mathbf{x}_i = *support vector*

NS = banyaknya *support vector*

\mathbf{x}_d = data yang akan diklasifikasikan

Untuk menyelesaikan masalah *non linear*, maka dilakukan modifikasi dengan memasukkan fungsi kernel. Fungsi kernel dapat merepresentasikan pemetaan data ke dimensi yang lebih tinggi. Berikut beberapa jenis fungsi kernel yang umum digunakan.

1. Fungsi linear (tanpa kernel)

Berikut persamaan fungsi linear

$$K(\mathbf{x}, \mathbf{x}_i) = \mathbf{x}_i^T \cdot \mathbf{x}$$

2. Fungsi *Polynomial*

Berikut persamaan fungsi *polynomial*

$$K(\mathbf{x}, \mathbf{x}_i) = (\gamma \cdot \mathbf{x}_i^T \mathbf{x} + r)^d$$

3. Fungsi *Radial Basis Function* (RBF)

Berikut persamaan fungsi RBF

$$K(\mathbf{x}, \mathbf{x}_i) = \exp \left(-\gamma \|\mathbf{x}_i - \mathbf{x}\|^2 \right), \text{ dengan } \gamma = 1/2\sigma^2$$

Pada ketiga kernel di atas, \mathbf{x}_i adalah vektor dari setiap data, d yaitu jumlah derajat dari fungsi *polynomial*, dan γ menjelaskan ukuran rentangan.

2.4 *Synthetic Minority Oversampling Technique* (SMOTE)

Ketidakeimbangan data merupakan salah satu masalah utama pada data klasifikasi biner. Data tidak seimbang terjadi jika salah satu kelas memiliki objek yang jauh lebih banyak dari kelas lainnya. Kelas yang memiliki objek lebih sedikit disebut kelas minor sedangkan kelas yang memiliki objek lebih banyak disebut kelas mayor. Pengolahan data dengan data tidak seimbang dapat menyebabkan hasil klasifikasi didominasi oleh kelas mayor dan mengabaikan kelas minor. Pada kasus tertentu, kesalahan klasifikasi pada kelas minor dapat berakibat fatal dibandingkan kesalahan klasifikasi pada kelas mayor (Somasundaram dan Reddy 2016).

Salah satu cara untuk mengatasi data tidak seimbang yaitu menggunakan *Synthetic Minority Oversampling Technique* (SMOTE). Teknik ini merupakan metode *oversampling* kelas minoritas dengan membangkitkan data buatan (sintetik) pada kelas data minor sehingga proporsi kelas data minor dan mayor menjadi seimbang. Konsep *k*-tetangga terdekat dimanfaatkan pada proses *oversampling*. Penentuan *k*-tetangga terdekat dipilih secara acak tergantung banyaknya *oversampling* yang dibutuhkan. (Chawla *et al.* 2002). Pembuatan data buatan untuk data berskala numerik diukur jarak kedekatannya dengan jarak *Euclidean*, sedangkan untuk data kategorik menggunakan rumus *Value Difference Metric* (VDM). Data sintetik dibangkitkan dengan cara berikut: Hitung perbedaan antara vektor utama dengan *k*-tetangga terdekatnya. Kalikan perbedaan tersebut dengan bilangan acak antara 0 dan 1. Tambahkan perbedaan yang telah dikalikan pada vektor utama.

2.5 Term Frequency-Inverse Document Frequency (TF-IDF)

Pembobotan TF-IDF merupakan suatu proses untuk mengubah data teks menjadi bentuk numerik. Metode ini digunakan untuk mengetahui seberapa jauh keterhubungan kata (*term*) terhadap dokumen dengan memberikan bobot setiap kata (Deolika *et al.* 2019). Metode TF-IDF dikenal sebagai algoritme yang sederhana namun relevan dalam melakukan pencocokan kata pada sebuah dokumen. Metode ini merupakan gabungan antara frekuensi kemunculan sebuah kata pada dokumen dan *inverse* frekuensi dokumen yang mengandung kata tersebut. Perhitungan bobot TF-IDF sebagai berikut:

$$w_{ij} = tf_{ij} \times idf_i$$

$$idf_i = \log \left(\frac{N}{df_i} \right) \quad (7)$$

dimana w_{ij} adalah bobot kata- i pada dokumen- j , tf_{ij} adalah jumlah kemunculan kata- i pada dokumen- j , idf_i adalah *Inverse Document Frequency* untuk kata- i , N adalah banyaknya dokumen, dan df_i merupakan banyaknya dokumen yang mengandung kata- i .

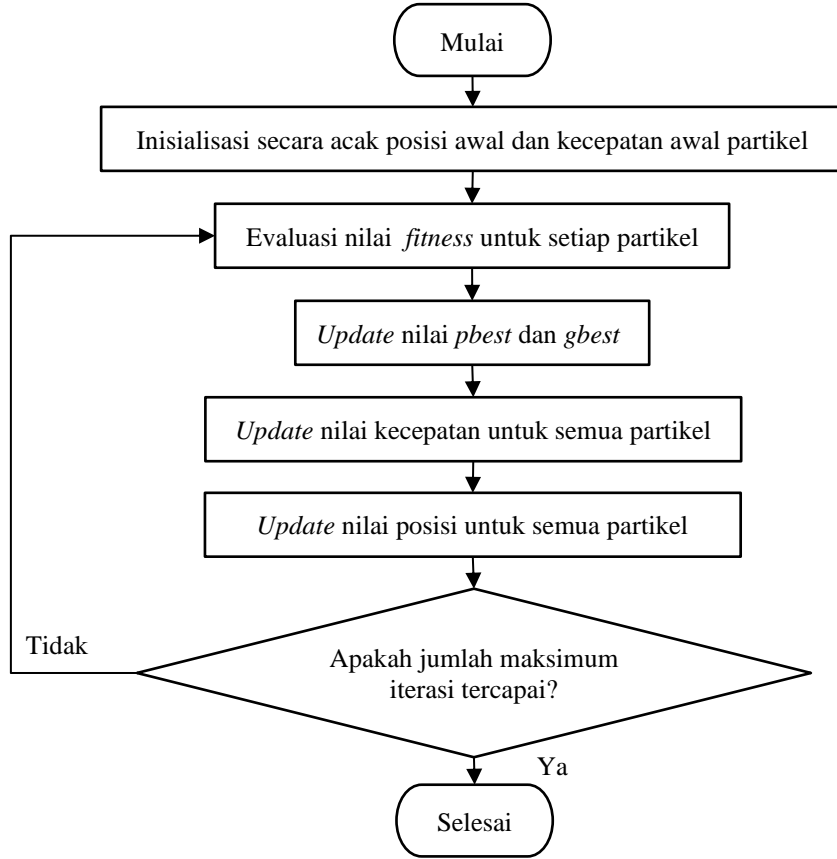
2.6 Binary Particle Swarm Optimization (BPSO)

Particle Swarm Optimization (PSO) merupakan algoritme *metaheuristic* berbasis populasi yang terinspirasi dari perilaku sosial kawanan burung yang mencoba mencapai tujuan yang tidak diketahui. Konsep dasar dari PSO yaitu bahwa pengetahuan dioptimalkan bukan hanya dari pengalaman individu tetapi dari interaksi sosial dalam populasi. Populasi disebut *swarm* dan individu disebut partikel. Setiap partikel (kandidat solusi) pada PSO adalah titik dalam ruang pencarian d -dimensi. Partikel memiliki memori tersendiri yang merekam pengalaman terbaiknya dalam ruang pencarian bersama dengan pengalaman terbaik yang diperoleh dari *swarm* secara keseluruhan untuk menemukan solusi terbaik. Setiap solusi individu pada PSO bergerak di ruang pencarian dengan kecepatan yang disesuaikan secara dinamis yang dipengaruhi oleh pengalamannya sendiri dan pengalaman global partikel lain (Mafarja *et al.* 2018).

Populasi awal partikel pada *swarm* terdistribusi secara acak pada ruang pencarian. Posisi setiap partikel dilambangkan sebagai vektor dimana D adalah dimensi ruang pencarian. Saat setiap partikel bergerak $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ di ruang pencarian untuk menemukan solusi optimal, kecepatan saat pencarian $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$. Selama pencarian, partikel memperbarui posisi dan kecepatannya sesuai dengan pengalaman dari partikel itu sendiri dan tetangga terdekatnya. Setiap partikel mempunyai memori untuk merekam posisi dimana partikel tersebut menemukan pengalaman terbaiknya dan dilambangkan sebagai *pbest*. Global terbaik merupakan posisi terbaik di antara semua partikel yang dilambangkan sebagai *gbest*.

Kennedy dan Eberhart mengembangkan versi biner dari algoritme PSO yaitu *Binary Particle Swarm Optimization* (BPSO). Pada BPSO, nilai peubah x_{id} , *pbest*, dan *gbest* hanya dapat memiliki nilai 0 atau 1. Kecepatan akan menunjukkan probabilitas sebuah partikel dalam vektor posisi untuk mengambil

nilai 1. Posisi partikel ‘1’ merepresentasikan fitur terpilih sedangkan ‘0’ untuk fitur yang tidak dipilih.



Gambar 3 Prosedur algoritme PSO

Posisi dan kecepatan diperbarui untuk setiap partikel. Baik $pbest$ dan $gbest$ membantu dalam memperbarui posisi dan kecepatan setiap partikel sesuai dengan persamaan berikut :

$$v_{id}(t+1) = w * v_{id}(t) + c_1 * r_1 * (p_{id} - x_{id}(t)) + c_2 * r_2 * (p_{gd} - x_{id}(t)) \quad (8)$$

dimana t merupakan iterasi ke- t dengan $d \in D$ adalah dimensi ke- d dalam ruang pencarian. Peubah w adalah bobot inersia yang mengontrol dampak dari kecepatan sebelumnya pada kecepatan saat ini. Peubah c_1 dan c_2 adalah konstanta percepatan, $r_1, r_2 \in [0,1]$ adalah bilangan acak berdistribusi seragam. Peubah p_{id} adalah *local best* ($pbest$) dan peubah p_{gd} adalah *global best* ($gbest$) pada dimensi ke- d . Posisi partikel saat ini diperbarui pada persamaan (9) berdasarkan nilai probabilitas $T(Vt)$ yang diperoleh dari persamaan (10).

$$x(t+1) = \begin{cases} 1 & \text{Jika } rand < S(v(t+1)) \\ 0 & \text{lainnya} \end{cases} \quad (9)$$

dimana $rand$ adalah bilangan acak $\in [0,1]$ dan $S(v(t))$ merupakan fungsi Sigmoid sebagai berikut:

$$S(v(t)) = \frac{1}{1 + e^{-v(t)}} \quad (10)$$

Tujuan optimasi pada penelitian ini yaitu untuk meminimumkan fungsi *fitness*. Fungsi *fitness* yang digunakan pada penelitian ini yaitu sebagai berikut:

$$Fitness = 1 - P \quad (11)$$

dimana P merupakan nilai *balanced accuracy* dengan formula sebagai berikut:

$$Balanced\ accuracy = \frac{Sensitivitas + Spesifisitas}{2} \quad (12)$$

2.7 Evaluasi Model

Model klasifikasi yang dihasilkan diharapkan memiliki kinerja yang baik, yang berarti model dapat memprediksi data dengan benar. Terdapat beberapa parameter yang menjadi dasar untuk mengukur kinerja klasifikasi yaitu akurasi, sensitivitas (*recall*), spesifisitas, *precision*, F_1 , dan F_β (Han *et al.* 2012). Evaluasi model yang digunakan pada penelitian ini yaitu akurasi, sensitivitas, dan spesifisitas. Untuk mengukur ketiga parameter tersebut digunakan *confusion matrix*. *Confusion matrix* merupakan alat ukur berbentuk matriks yang digunakan untuk mengetahui jumlah ketepatan klasifikasi terhadap kelas yang berbeda (Han *et al.* 2012). Berikut tampilan dari *confusion matrix*.

Tabel 1 *Confusion matrix* dari dua kelas

<i>Confusion Matrix</i>		Kelas hasil prediksi	
		Positif	Negatif
Kelas sebenarnya	Positif	<i>True Positif</i> (TP)	<i>False Negatif</i> (FN)
	Negatif	<i>False Positif</i> (FP)	<i>True Negatif</i> (TN)

Keterangan:

TP : jumlah data dengan kelas positif dan hasil prediksi benar

TN : jumlah data dengan kelas negatif dan hasil prediksi benar

FP : jumlah data dengan kelas positif dan hasil prediksi salah

FN : jumlah data dengan kelas negatif dan hasil prediksi salah

Akurasi merupakan persentase banyaknya prediksi benar pada semua data yang diklasifikasikan. Sensitivitas merupakan proporsi dari kelas positif yang diidentifikasi dengan benar. Spesifisitas merupakan proporsi dari kelas negatif yang

diidentifikasi dengan benar (Han *et al.* 2012). Kelas positif pada kasus penelitian ini yaitu “komentar *cyberbullying*”, sedangkan kelas negatif yang dimaksud yaitu “komentar bukan *cyberbullying*”. Berikut perhitungan dari akurasi, sensitivitas dan spesifisitas.

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \quad (13)$$

$$Sensitivitas = \frac{TP}{TP + FN} \quad (14)$$

$$Spesifisitas = \frac{TN}{TN + FP} \quad (15)$$

III METODE

3.1 Data

Data yang digunakan dalam penelitian ini merupakan data hasil *scraping* komentar-komentar pada postingan beberapa akun instagram. Akun yang dipilih merupakan akun dari beberapa pengguna instagram yang pernah menjadi korban *cyberbullying* yaitu @viavallen, @riaricis1795, @aurelie.hermansyah, @rahmawatikekeyiputricantikka23, dan @cimoyluv.

Data yang diambil dari setiap akun merupakan komentar dari beberapa postingan saja yaitu postingan pada saat akun tersebut mengalami *cyberbullying* sehingga periode postingan yang diambil dari setiap akun berbeda-beda. Banyaknya postingan yang diambil dari setiap akun masing-masing dua postingan, kecuali pada akun @viavallen yang hanya diambil satu postingan. Hal ini karena telah dihapusnya beberapa postingan pada akun tersebut pada masa pemberitaan *cyberbullying*.

Proses pengambilan data komentar Instagram dilakukan melalui proses *scraping* dengan bantuan tools *Phantombuster.com*. Berdasarkan hasil *scraping*, dihasilkan data sebanyak 7866 komentar dari 9 postingan.

Tabel 2 Contoh data hasil *scraping*

No.	Komentar
1	liat Kk Cantik Nyaaa
2	Cantikâ• ¢• â• ¢• â• ¢• â• ¢•
3	Kak Icis, Please FollbcaK Aku @riaricis1795 ðŸ•

3.2 Prosedur Analisis Data

Proses analisis data dilakukan dengan menggunakan *Software* R 4.0.2 dan Python 3.7.10. Diagram alir dari prosedur analisis data ditampilkan pada Lampiran 1. Berikut merupakan penjelasan dari prosedur analisis data yang dilakukan.

1. Pemilahan data

Pemilahan data diawali dengan menghapus data duplikat yang disebabkan oleh kesalahan pada proses *scraping*. Setelah itu, dilakukan pemilahan data sesuai konteks penelitian. Komentar yang tidak sesuai dengan konteks penelitian yaitu komentar yang berisi promosi/jualan serta komentar yang tidak mempunyai arti yang jelas. Ilustrasi data komentar yang tidak sesuai konteks ditampilkan pada tabel berikut.

Tabel 3 Komentar tidak sesuai konteks

No	Komentar
1	BASO ACI TERCIDUK 14rb aja ðŸ”¥ðŸ”¥ðŸ”¥ðŸ”¥
2	Dddhd

Komentar pada data terdiri dari dua jenis yaitu komentar secara langsung terhadap isi postingan dan komentar yang berupa jawaban terhadap komentar orang lain. Komentar yang digunakan pada penelitian ini hanya komentar terhadap postingan akun, sehingga komentar yang hanya menjawab komentar orang lain tidak diikutsertakan dalam penelitian.

Karena pada saat proses *scraping* emotikon tidak terbaca dengan jelas, maka komentar yang hanya berisi emotikon saja dan tidak mengandung huruf a-z tidak akan diikutsertakan dalam penelitian. Proses pada tahap pemilahan data menyisakan sebanyak 4383 data.

2. Pelabelan data

Proses pelabelan data dilakukan secara manual oleh seorang mahasiswa Departemen Sains Komunikasi dan Pengembangan Masyarakat (SKPM) IPB University. Setiap komentar Instagram akan diberi satu label atau kelas. Komentar akan diberi label “1” jika komentar tersebut mengandung kalimat yang mengarah kepada *cyberbullying* dan diberi label “0” apabila komentar tidak mengandung unsur *cyberbullying*. Untuk keperluan eksplorasi data, komentar yang termasuk *cyberbullying* diberi label jenis *bullying* apakah termasuk *Called Name* (cn), *Threatened Physical Harm* (tp), atau *Opinion Slammed* (os).

Tabel 4 Ilustrasi komentar yang telah diberi label

No.	Komentar	Label	Jenis <i>bullying</i>
1	Bantet bunggg	1	cn
2	Klo ada yg ngebuli kk jangan di masukin ke hati masukin aja ke idung	1	os
3	Gemes pengen bunuh	1	tp
4	Cantik nya	0	-

Setelah diberi label, dihasilkan 539 data yang termasuk *cyberbullying* berlabel “1” dan 3844 data yang tidak termasuk *cyberbullying* berlabel “0”.

3. Text pre-processing

Proses ini terdiri dari beberapa tahapan yaitu sebagai berikut:

a. Case folding

Case folding yaitu mengkonversi keseluruhan teks dalam dokumen menjadi huruf kecil.

- b. *Filtering*
Digunakan untuk menghapus *hashtag*, url, *mention*, emotikon, tanda baca, angka, dan karakter yang tidak relevan terhadap proses klasifikasi.
- c. Normalisasi
Digunakan untuk mengubah kata tidak baku seperti *slang*, kata yang disingkat, penulisan kata dengan angka, dan kesalahan penulisan kata menjadi kata baku sesuai dengan Kamus Besar Bahasa Indonesia (KBBI). Normalisasi dilakukan menggunakan kamus dari *colloquial-indonesian-lexicon* yang berisi pasangan kata formal dan informal dalam Bahasa Indonesia.
- d. *Stemming*
Merupakan proses mengubah bentuk kata menjadi bentuk kata dasar dengan menghilangkan imbuhan sehingga proses penambangan teks menjadi lebih optimal (Munkova *et al.* 2013). Proses *stemming* dilakukan menggunakan *Library Python Sastrawi*.
- e. Penghapusan *Stopwords*
Stopwords merupakan kata yang memiliki informasi rendah pada sebuah teks. Penghapusan *stopwords* Bahasa Indonesia dilakukan menggunakan *Library Python Sastrawi* dengan beberapa modifikasi sedangkan untuk *stopwords* dalam Bahasa Inggris berasal dari *Natural Language Toolkit* (NLTK).

Berikut merupakan ilustrasi komentar sebelum dan setelah *pre-processing*.

Tabel 5 Ilustrasi komentar sebelum dan setelah *pre-processing*

No	Komentar sebelum <i>pre-processing</i>	Komentar setelah <i>pre-processing</i>
1	Cantikâ• ¢Œ• â• ¢Œ• â• ¢Œ• â• ¢Œ• • â• ¢Œ•	cantik
2	Haii kaka ricissðŸ’–	kakak ricis
3	Itu Di pim ya KAK @riaricis1795	pim kakak

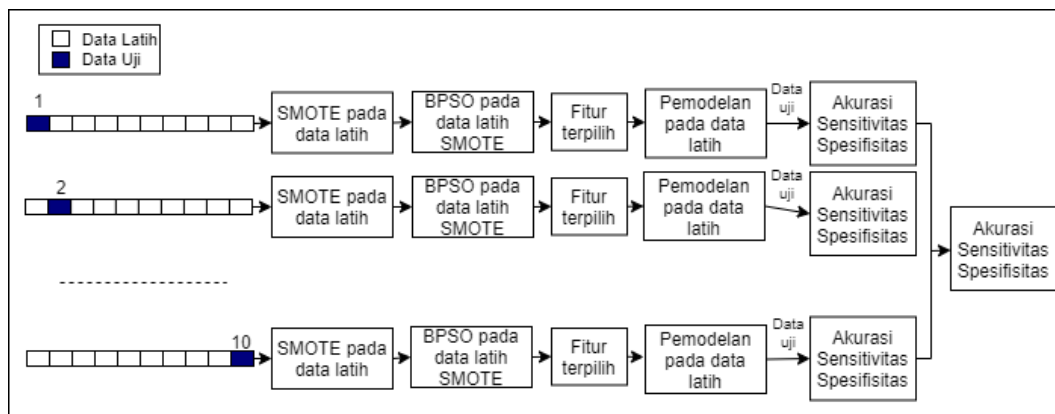
4. Eksplorasi Data
Setiap data komentar dikelompokkan berdasarkan nama akun untuk melihat distribusi label pada setiap akun. Selain itu, dilakukan juga eksplorasi jenis *bullying* dan eksplorasi pada panjang teks komentar. Data komentar yang telah melalui *text pre-processing* masuk ke tahap pembentukan awan kata untuk mengetahui kata-kata yang berkaitan dengan komentar *cyberbullying* dan bukan *cyberbullying*.
5. Pembobotan TF-IDF
Data yang sudah melalui *text pre-processing* akan dibuat menjadi vektor menggunakan perhitungan TF-IDF. Pembobotan dilakukan pada masing-masing data latih dan data uji pada tahap *cross validation*. Banyaknya fitur yang diperoleh pada tahap ini berbeda pada setiap *fold*. Pada tahap ini rata-rata fitur yang diperoleh sebanyak 2333 fitur atau kata. Fitur-fitur ini yang akan digunakan dalam pemodelan.

6. Seleksi Fitur Menggunakan BPSO

Fitur atau kata yang telah diberi bobot TF-IDF selanjutnya masuk pada tahap seleksi fitur dengan *Binary Particle Swarm Optimization* (BPSO). Fungsi *fitness* yang digunakan pada penelitian ini terdapat pada persamaan (11). Langkah-langkah pemilihan fitur dan pemodelan ditampilkan pada Gambar 4. Nilai parameter BPSO yang digunakan ditentukan dengan melihat penelitian sebelumnya dengan beberapa penyesuaian. Yan *et al.* (2016) menyatakan bahwa bobot inersia (ω) yang umum digunakan berkisar dari 0,4 sampai 0,9. Mafarja *et al.* (2018) pada penelitiannya menggunakan 10 partikel dan 100 iterasi menghasilkan hasil yang baik.

7. Klasifikasi menggunakan SVM

Pemodelan klasifikasi dilakukan dengan dua kondisi yaitu tanpa seleksi fitur dan dengan seleksi fitur BPSO. Pemodelan dilakukan menggunakan metode *K-fold cross validation* dengan $k=10$. Metode ini membagi data menjadi 10 bagian/*fold* dimana setiap *fold* berkesempatan menjadi data latih dan data uji. Proses penentuan data latih dan data uji pada setiap *fold* dilakukan secara *stratified*. *Stratified* merupakan teknik dalam *cross-validation* untuk memastikan distribusi label pada setiap *fold* menyerupai data awal. Menurut Joachims (1997), sebagian besar masalah klasifikasi teks adalah *linearly separable*. Maka dari itu, pada penelitian ini digunakan kernel linear dengan parameter C. Beberapa nilai parameter C yang akan diujikan berdasarkan penelitian dari Laia dan Setiawan (2020) yaitu 0,1; 1, dan 10. Karena data pada penelitian ini tidak seimbang, maka dilakukan penanganan ketidakseimbangan data menggunakan metode *Synthetic Minority Over-sampling Technique* (SMOTE). Prosedur pemodelan tanpa seleksi fitur BPSO sama seperti Gambar 4 hanya saja tanpa melibatkan seleksi fitur BPSO pada data latih SMOTE.



Gambar 4 Pemodelan dengan seleksi fitur BPSO menggunakan 10-fold cross-validation

8. Evaluasi model

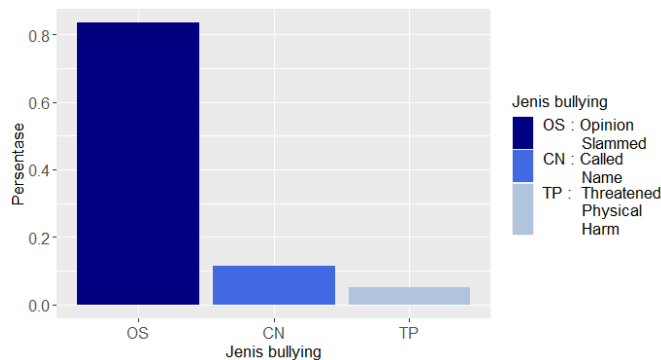
Menghitung ketepatan klasifikasi berdasarkan tingkat akurasi, spesifisitas, dan sensitivitas untuk model tanpa seleksi fitur dan dengan seleksi fitur.

9. Interpretasi dan penarikan kesimpulan

IV HASIL DAN PEMBAHASAN

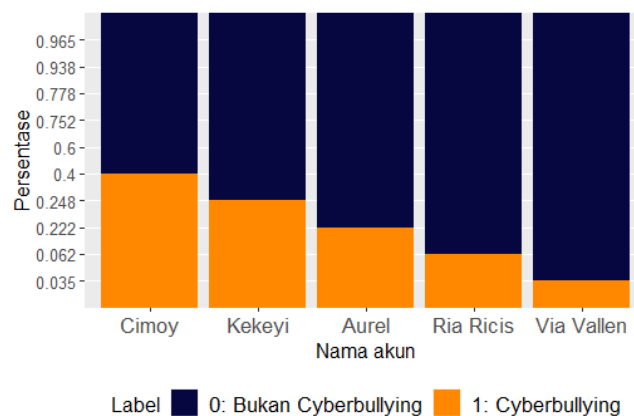
4.1 Eksplorasi Data

Komentar yang termasuk ke dalam *cyberbullying* dikelompokkan menurut jenis *bullying* yang dilakukan apakah termasuk *Called Name* (cn), *Threatened Physical Harm* (tp), atau *Opinion Slammed* (os).



Gambar 5 Grafik perbandingan jenis *bullying*

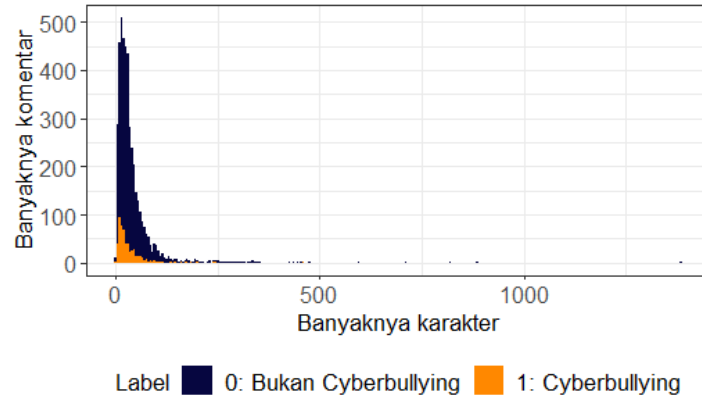
Berdasarkan Gambar 5 dapat dilihat bahwa jenis *bullying* yang paling banyak dilakukan yaitu berbentuk *opinion slammed* (pendapat yang merendahkan) dengan persentase sebanyak 83% dari banyaknya komentar *bullying* yang ada. Kata-kata yang berkaitan dengan hinaan seperti sebutan “tante” mendominasi kategori ini. Terdapat juga kata-kata merendahkan lainnya seperti “menor”, “sensasi”, ”plastik”, “gendut”, “jelek”, dan sebagainya. Jenis *bullying* yang paling jarang dilakukan yaitu *threatened physical harm* (mengancam keselamatan fisik) dengan persentase 0.05% dari banyaknya komentar *bullying* yang ada. Kata-kata seperti “bunuh” dan “pergi” mendominasi kategori ini.



Gambar 6 Proporsi label pada setiap akun

Distribusi jenis label pada setiap akun beraneka ragam. Berdasarkan Gambar 6 dapat dilihat bahwa proporsi komentar yang termasuk *cyberbullying* tidak mendominasi pada setiap akun. Akun yang paling banyak mengalami *cyberbullying* yaitu @cimoyluv dengan persentase komentar *cyberbullying* mencapai 40% dari

banyaknya komentar pada akun tersebut. Kata-kata yang berkaitan dengan hinaan fisik wajah dan sebutan negatif berupa nama hewan mendominasi pada akun ini. Akun yang paling sedikit mengalami *cyberbullying* yaitu @viavallen dengan persentase komentar *cyberbullying* hanya sebesar 3%.



Gambar 7 Distribusi banyak karakter tiap kelas

Gambar 7 menggambarkan seberapa banyak karakter yang dituliskan oleh setiap akun pada saat berkomentar. Berdasarkan gambar tersebut dapat dilihat bahwa komentar *cyberbullying* jauh lebih sedikit daripada komentar yang bukan *cyberbullying*. Dengan kata lain, proporsi label pada data ini tidak seimbang. Tidak terdapat perbedaan banyak karakter yang terlalu signifikan antara label “0” dan “1”. Sebagian besar data berisi komentar dengan panjang karakter yang tidak terlalu banyak. Komentar bukan *cyberbullying* banyak memiliki komentar dengan panjang karakter yang cukup besar. Terdapat pencilan komentar yang terdiri dari 1387 karakter pada kelas ini. Hal ini menunjukkan orang akan cenderung menulis lebih panjang untuk komentar yang tidak termasuk *cyberbullying*.



Gambar 8 Wordcloud kelas (a) “0: Bukan *cyberbullying*” (b) “1: *Cyberbullying*”

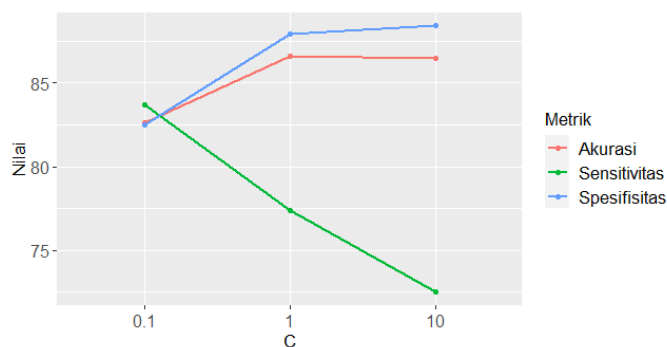
Kata-kata yang merepresentasikan tiap kelas dapat dilihat pada Gambar 8. Kata yang ditampilkan merupakan 100 kata dengan frekuensi kemunculan terbanyak pada tiap label. Berdasarkan gambar tersebut, dapat dilihat bahwa kata-

kata yang sering muncul pada kelas bukan *cyberbullying* yaitu “kakak”, “cantik”, “follback”. Sebagian besar komentar pada kelas ini berisi sapaan, pujian, dan permintaan untuk diberi *follow back* oleh akun terkait. Terdapat juga kata-kata positif lainnya seperti “semangat”, “sukses”, “keren”, dan sebagainya.

Pada kelas *cyberbullying*, kata-kata yang sering muncul yaitu kata yang berkaitan dengan komentar buruk tentang penampilan fisik seperti “wajah”, “tante”, “umur”. Berdasarkan kata-kata tersebut, dapat diketahui bahwa komentar tersebut mengarah kepada ejekan atas bentuk fisik seseorang yang tidak sesuai dengan umur mereka. Berdasarkan eksplorasi pada Gambar 5, diketahui bahwa jenis *bullying* yang paling banyak dilakukan yaitu *opinion slammed* atau pendapat yang merendahkan. Dapat dilihat pada *wordcloud cyberbullying* bahwa terdapat banyak kata yang merendahkan seperti sebutan “tante”, “jelek”, “tua”, “alay”, “menor”, “aneh” dan sebagainya. Terdapat juga pemberian nama negatif (*called name*) seperti kata “anjing”, “monyet”, dan “lonte”. Pada kelas ini juga terdapat kata “cantik”. Akan tetapi setelah dilakukan penelusuran terhadap data, kata cantik yang dimaksud bukan pujian melainkan bentuk sindiran seperti “Sok cantik lu”, “Cantik operasi”, dan sebagainya.

4.2 Model Support Vector Machine

Proporsi label pada data penelitian ini tidak seimbang yaitu terdiri dari 3830 data (88%) label “0” yang bukan *cyberbullying* dan 539 data (12%) label “1” yang termasuk *cyberbullying*. Ketidakseimbangan data ini akan berdampak pada performa hasil klasifikasi yaitu menghasilkan nilai sensitivitas yang kecil. Sensitivitas menggambarkan seberapa baik model dapat mendeteksi komentar yang termasuk *cyberbullying* dengan benar. Apabila nilai sensitivitas kecil, maka akan banyak komentar yang termasuk *cyberbullying* diklasifikasikan secara salah menjadi kelas bukan *cyberbullying* (Somasundaram dan Reddy 2016). Hal ini tentunya tidak diharapkan karena tujuan pemodelan ini kedepannya agar dapat digunakan untuk mendeteksi komentar *cyberbullying* secara otomatis. Apabila banyak komentar *cyberbullying* yang tidak terdeteksi, maka akan menjadi kurang bermanfaat. Untuk itu, perlu dilakukan penanganan ketidakseimbangan data menggunakan SMOTE agar dapat memperbaiki performa klasifikasi dan meningkatkan nilai sensitivitas. Berikut merupakan ringkasan performa klasifikasi pada data uji untuk berbagai nilai C.

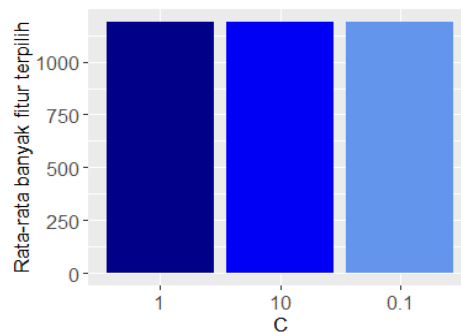


Gambar 9 Performa klasifikasi untuk setiap nilai C

Setelah penerapan SMOTE, nilai sensitivitas pada sebagian besar nilai C terbilang cukup akurat yaitu di atas 72%. Nilai sensitivitas menurun seiring bertambahnya nilai C, sedangkan nilai akurasi dan spesifisitas cenderung meningkat. Sensitivitas tertinggi terdapat pada saat $C = 0,1$. Meskipun nilai spesifisitas dan akurasi pada C ini paling rendah diantara C lainnya, akan tetapi nilai tersebut masih akurat yaitu di atas 80%. Performa klasifikasi pada semua nilai C terbilang cukup akurat, dibuktikan dengan nilai akurasi dan spesifisitas yang dihasilkan di atas 80% dengan nilai sensitivitas di atas 72%. Karena klasifikasi ini berguna untuk mendeteksi *cyberbullying*, maka dibutuhkan nilai sensitivitas yang cukup tinggi tanpa mengabaikan performa lainnya. Maka dari itu, performa terbaik terdapat pada saat $C = 0,1$ dengan nilai akurasi sebesar 82,6%, sensitivitas 83,7%, dan spesifisitas 82,5%.

4.3 Seleksi Fitur BPSO

Setiap kombinasi fitur yang terdapat pada TF-IDF akan dievaluasi menggunakan fungsi *fitness* pada BPSO untuk menentukan banyaknya fitur terbaik. Penerapan BPSO dilakukan dengan bantuan *Library Python Pyswarms*. Nilai parameter yang digunakan pada penelitian ini yaitu $c_1 = 1,774$, $c_2 = 1,902$, $w = 0,544$, $k = 7$ dengan 10 partikel dan 100 iterasi. Parameter tersebut dipilih setelah dilakukan beberapa percobaan parameter *tuning*.



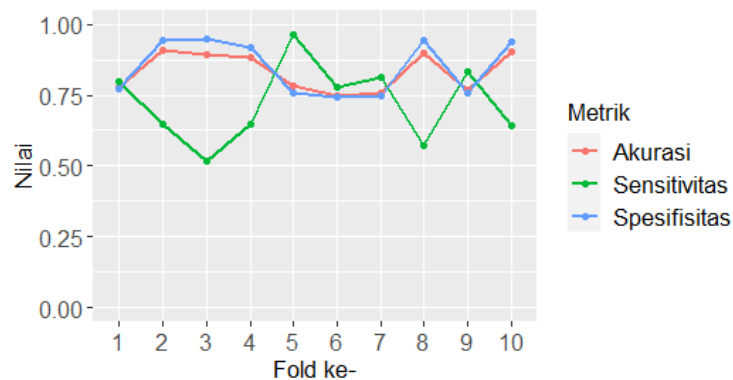
Gambar 10 Rata-rata banyaknya fitur terpilih untuk setiap nilai C

Dikarenakan terdapat 10 kali pengulangan pemodelan pada saat *cross validation*, maka kombinasi fitur terbaik yang diperoleh pada setiap nilai C sebanyak 10 kombinasi fitur. Banyaknya fitur terpilih pada setiap *fold* dapat dilihat pada Lampiran 3. Gambar 10 menunjukkan rata-rata banyaknya fitur terpilih pada setiap nilai C. Rata-rata banyaknya fitur terpilih pada setiap nilai C tidak terlalu berbeda secara signifikan yaitu berada pada kisaran 1187 – 1191 fitur. Jumlah ini berada pada kisaran 50% dari banyak fitur sebelumnya. Selain dapat mengurangi banyak fitur, penerapan seleksi fitur BPSO juga dapat mengurangi waktu komputasi pada saat pemodelan untuk setiap nilai C.

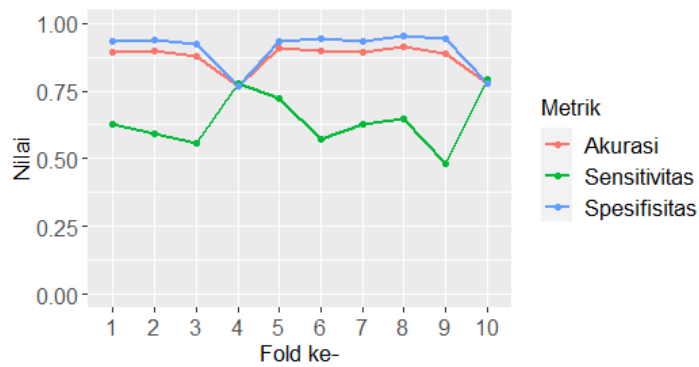
Tabel 6 Perbandingan performa klasifikasi sebelum dan setelah seleksi fitur BPSO

Nilai C	Sebelum seleksi fitur			Setelah seleksi fitur		
	Akurasi (%)	Sensitivitas (%)	Spesifisitas (%)	Akurasi (%)	Sensitivitas (%)	Spesifisitas (%)
0,1	82,6	83,7	82,5	83,1	72,2	84,7
1	86,6	77,4	87,9	87,2	64	90,5
10	86,5	72,5	88,4	86,6	68,3	89,1

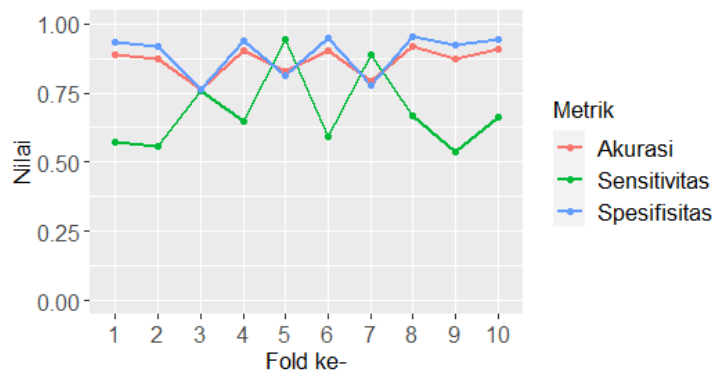
Tabel 6 menampilkan nilai rata-rata performa klasifikasi dari *10-fold cross-validation*. Setelah penerapan seleksi fitur, nilai akurasi dan spesifisitas mengalami kenaikan pada setiap nilai C, sedangkan nilai sensitivitas mengalami penurunan. Nilai sensitivitas mengalami penurunan paling besar pada saat C = 1 yaitu sebesar 13,4%. Nilai spesifisitas mengalami peningkatan lebih dari 2% pada saat C = 1 dan C = 0,1. Peningkatan nilai akurasi pada kedua nilai C ini juga tidak terlalu berbeda jauh. C = 10 mengalami peningkatan nilai akurasi dan spesifisitas yang paling rendah diantara nilai C lainnya. Penerapan seleksi fitur BPSO menghasilkan performa yang baik pada setiap nilai C dengan nilai akurasi dan spesifisitas yang cukup besar. Nilai ini lebih baik daripada model tanpa seleksi fitur. Nilai rata-rata sensitivitas yang diperoleh pada seleksi fitur BPSO masih cukup akurat di atas 60% hanya saja mengalami penurunan daripada model tanpa seleksi fitur.



(a)



(b)



(c)

Gambar 11 Performa pada setiap fold (a) $C = 0,1$ (b) $C = 1$ (c) $C = 10$

Gambar 11 menunjukkan nilai performa klasifikasi pada setiap *fold*. Berdasarkan gambar tersebut dapat dilihat bahwa nilai performa klasifikasi pada setiap *fold* beraneka ragam. Terdapat *trade-off* antara nilai sensitivitas dan spesifisitas pada sebagian besar *fold*, yaitu jika nilai spesifisitas besar maka nilai sensitivitas cenderung kecil begitupun sebaliknya. Pada saat $C = 0,1$, semua nilai performa klasifikasi cenderung mengalami fluktuasi. Terdapat dua *fold* dengan nilai sensitivitas yang cukup besar di atas 83% yaitu pada *fold* ke-5 dan ke-9. Pada saat $C = 1$, nilai akurasi dan spesifisitas cenderung stabil pada nilai yang cukup tinggi hanya pada dua *fold* terjadi penurunan, sedangkan nilai sensitivitas cenderung fluktuatif pada nilai yang cukup rendah yaitu 48% - 79%. Nilai sensitivitas terbesar diperoleh pada *fold* ke-10 yaitu 79% dengan nilai akurasi dan spesifisitas sebesar 78%. Pada saat $C = 10$, semua nilai performa klasifikasi cenderung mengalami fluktuasi pada setiap *fold*. Nilai sensitivitas mengalami fluktuasi yang cukup beragam pada rentang 53% – 94% dengan nilai sensitivitas terbesar terdapat pada *fold* ke-5. Hasil performa klasifikasi BPSO menghasilkan nilai akurasi dan spesifisitas yang cukup akurat di atas 70% pada setiap fold untuk semua nilai C . Nilai sensitivitas setelah seleksi fitur BPSO menghasilkan nilai yang kurang akurat di bawah 60% pada beberapa *fold* sehingga menghasilkan rata-rata sensitivitas yang cukup kecil.

Penurunan nilai sensitivitas dapat disebabkan karena BPSO memiliki kecenderungan untuk mengalami konvergensi prematur yang menyebabkan proses pencarian solusi terjebak dalam optimum lokal sehingga tidak dapat menemukan

solusi optimal (Vieira *et al.* 2013). Untuk mengatasi hal tersebut dapat digunakan model yang lebih kompleks dengan melakukan modifikasi terhadap metode dan parameter BPSO. Terdapat beberapa pendekatan yang dapat dilakukan diantaranya menggunakan operator mutasi, melakukan reset *swarm* terbaik jika nilai *fitness* mengalami stagnasi, atau menggunakan mekanisme gangguan (Vieira *et al.* 2013). Selain itu juga dapat dilakukan modifikasi pada beberapa parameter BPSO dengan merubah nilai bobot inersia di tiap iterasi dengan metode *Linearly Decreasing Inertia Weight Particle Swarm Optimization* (LDIW-PSO) atau merubah nilai koefisien akselerasi c_1 dan c_2 di tiap iterasi dengan metode *Time-Varying Acceleration Coefficients Particle Swarm Optimization* (TVAC-PSO).

Penyebab lainnya yang mungkin terjadi yaitu dapat dilihat pada Gambar 7 pada tahap eksplorasi data, komentar *cyberbullying* memiliki banyak karakter yang lebih sedikit daripada komentar bukan *cyberbullying*. Pemangkasan fitur dapat menghilangkan beberapa informasi pada komentar *cyberbullying* sehingga penggunaan seluruh fitur menghasilkan nilai sensitivitas yang lebih baik.



Gambar 12 Perbandingan fitur (a) Sebelum seleksi fitur (b) Setelah seleksi fitur

Gambar 12 menunjukkan kata-kata yang sering muncul berdasarkan fitur terpilih saat $C = 0,1$. Fitur-fitur terpilih ini didasarkan pada fitur yang sering terpilih pada setiap *fold*. Berdasarkan gambar tersebut, dapat dilihat bahwa fitur yang sering muncul sebelum seleksi fitur seperti “kakak”, “cantik”, dan “follback” banyak terpilih sebagai fitur pada setiap *fold*. Berdasarkan Gambar 8, kata “kakak” dan “cantik” banyak digunakan baik pada komentar *cyberbullying* maupun bukan *cyberbullying*, sedangkan kata “follback” banyak digunakan pada komentar bukan *cyberbullying*. Fitur yang sering terpilih pada setiap *fold* untuk komentar *cyberbullying* diantaranya yaitu “jelek”, “tante”, “wajah”, “umur”, “pamit”, “alay”, “norak”, dan “menor”. Proporsi fitur pada Gambar 12 cenderung seimbang untuk komentar *cyberbullying* maupun bukan *cyberbullying*.

V SIMPULAN DAN SARAN

5.1 Simpulan

Pemodelan klasifikasi *cyberbullying* dengan SVM menghasilkan performa klasifikasi yang cukup akurat lebih dari 72% untuk semua performa klasifikasi pada setiap nilai C. Penerapan seleksi fitur BPSO dapat meningkatkan performa klasifikasi dengan meningkatnya nilai akurasi dan spesifisitas. Akan tetapi untuk kasus pada penelitian ini dipilih model tanpa seleksi fitur pada $C = 0,1$ karena memiliki nilai sensitivitas paling besar dengan akurasi dan spesifisitas yang cukup baik sehingga dapat mendeteksi komentar *cyberbullying* secara lebih akurat.

5.2 Saran

Pada penelitian ini tidak dilakukan konversi emotikon karena data tidak terbaca saat proses *scraping*. Penggunaan metode *scraping* yang lebih baik dapat dilakukan pada penelitian selanjutnya agar data emotikon dapat terbaca dan dilibatkan dalam analisis. Salah satu metode *scraping* yang dapat dicoba yaitu menggunakan *library selenium* python. Untuk mencegah konvergensi prematur dan meningkatkan performa klasifikasi pada BPSO, dapat diterapkan modifikasi metode pada BPSO seperti *Linear Decreasing Inertia Weight PSO* (LDIW-PSO), *Modified Binary PSO* (MBPSO), *Time-Varying Acceleration Coefficients PSO* (TVAC-PSO), dan lain sebagainya.

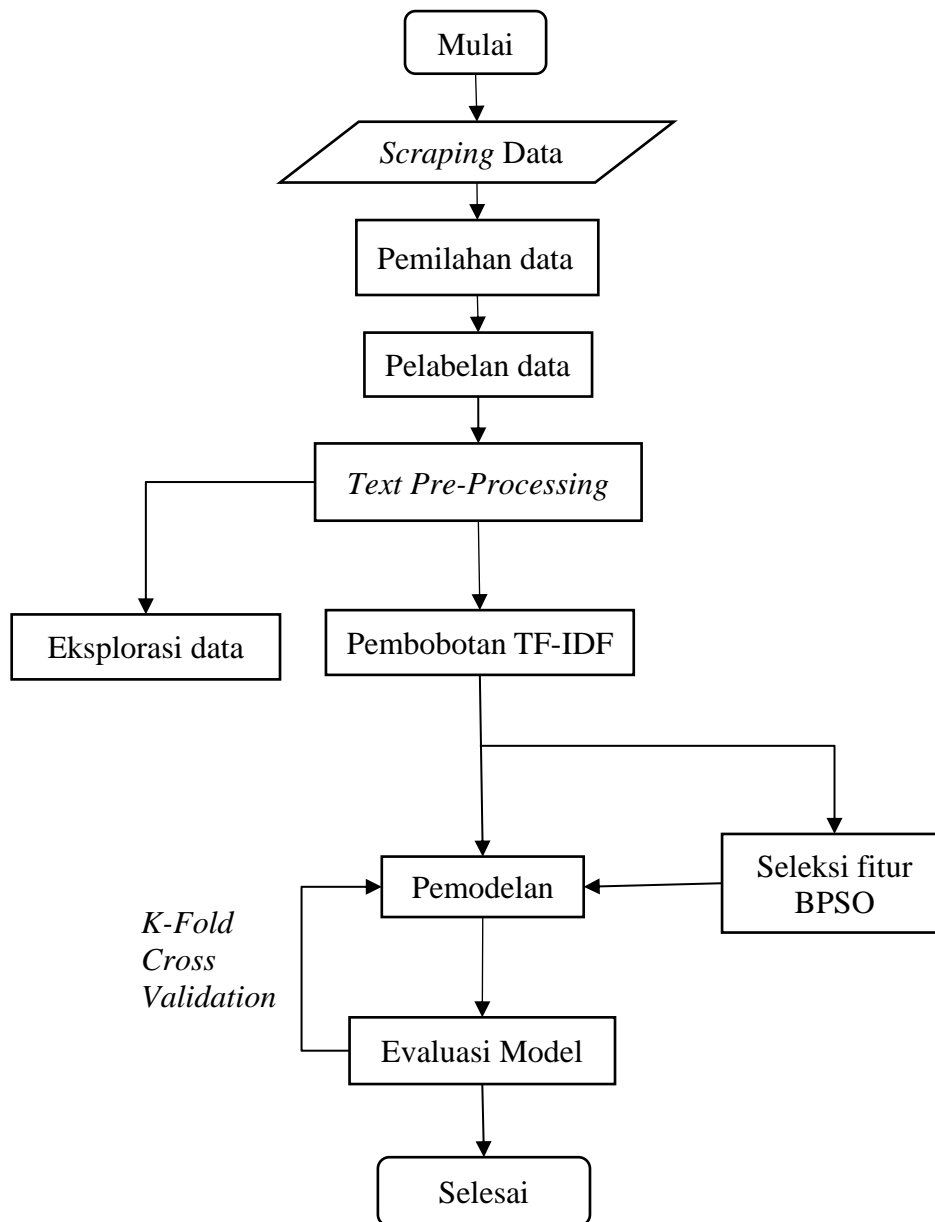
DAFTAR PUSTAKA

- Anderson J, Bresnahan M, Musatics C. 2014. Combating weight-based cyberbullying on facebook with the dissenter effect. *Cyberpsychol Behav Soc*. 17: 281–286. doi:10.1089/cyber.2013.0370.
- [APJII] Asosiasi Penyelenggara Jasa Internet Indonesia (APJII). 2019. Penetrasi & profil pelaku pengguna internet Indonesia tahun 2018. [diakses 2021 Jan 25]. www.apjii.or.id.
- Arifin YT. 2016. Komparasi fitur seleksi pada algoritme *support vector machine* untuk analisis sentimen review. *Informatika*. 3: 191-199.
- Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. 2002. SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*. (16):321-357.
- Chrismanto AR, Lukito Y. 2017. Identifikasi komentar spam pada Instagram. *Lontar Komputer*. 8(3): 219-231. Doi: 10.24843/LKJITI.2017.v08.i03.p08.
- Ditch The Label. 2018. 7 facts about *cyberbullying*. [diakses 2020 Des 28]. <https://www.ditchthelabel.org/7-facts-cyberbullying/>.
- Deolika A, Kusriani, Luthfi ET. 2019. Analisis pembobotan kata pada klasifikasi *text mining*. *Jurnal Teknologi Informasi*. 3(2): 179-184.
- Gaikwad SV, Chaugule A, Patil P. 2014. Text mining methods and techniques. *International Journal of Computer Applications*. 17(85): 42-45.
- Han J, Kamber M, Pei J. 2012. *Data Mining Concepts and Techniques*. Third Edition. USA: Elsevier.
- Hootsuite (*We are Social*). 2020. Digital in 2020. [diakses 2021 Jan 26]. <https://wearesocial.com/digital-2020>.
- Joachims T. 1997. *Text Categorization with Support Vector Machines: Learning with Many Relevant Features*. Germany: Springer.
- Kurniawan M. 2017. Optimasi parameter premise pada adaptive-network-based fuzzy inference system dengan modifikasi kombinasi particle swarm optimization dan genetic algorithm [tesis]. Surabaya: Institut Teknologi Sepuluh November.
- Laia ML, Setyawan Y. 2020. Perbandingan hasil klasifikasi curah hujan menggunakan metode svm dan nbc. *Jurnal Statistika Industri dan Komputasi*. 5(2): 51-61.
- Luqyana WA, Cholissodin I, Perdana RS. 2018. Analisis sentimen *cyberbullying* pada komentar Instagram dengan metode klasifikasi *support vector machine*. *J-PTIHK*. 2(11): 4704-4713.
- Mafarja M, Jarrar R, Ahmad S, Abusnaina AA. 2018. Feature selection using binary particle swarm optimization with time varying inertia weight strategies. *Proceedings of the 2nd International Conference on Future Networks and Distributed Systems*: 1-9.
- Maulida I, Suyatno A, Hatta HR. 2016. Seleksi fitur pada dokumen abstrak teks Bahasa Indonesia menggunakan metode *information gain*. *JSM STMIK Mikroskil*. 17(2): 249-258.
- Munkova D, Munk M, Vozar M. 2013. Data pre-processing evaluation for text mining: transaction/sequence model. *Procedia Computer Science*. 18:1198-1207.doi:10.1016/j.procs.2013.05.28.

- Pajri D, Umaidah Y, Padilah TN. 2020. K-nearest neighbor berbasis particle swarm optimization untuk analisis sentiment terhadap Tokopedia. *Jurnal Teknik Informatika dan Sistem Informasi*. 6(2): 242-253.
- Price M, Dalgleish M. 2010. Cyberbullying: experiences, impacts and coping strategies as described by Australian young people. *Youth Studies Australia*. 29(2): 51-59.
- Rifauddin M. 2016. Fenomena *cyberbullying* pada remaja: studi analisis media sosial facebook. *Jurnal Ilmu Perpustakaan, Informasi, dan Kearsipan Khazanah Al-Hikmah*. 4(1):35-44.
- Somasundaram A, Reddy US. 2016. Data imbalance: Effects and solutions for classification of large and highly imbalanced data. *ICRECT*. 28-34.
- Taneja T, Sharma B. 2014. Text classification using PSO & other technique. *International Journal of Recent Development in Engineering and Technology (IJRDT)*. 3(1): 15-20.
- Vieira SM, Mendonca LF, Farinha GJ, Sousa JM. 2013. Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients. *Applied Soft Computing*. 13(8): 3494-3504.
- Yan HE, Wei JMA, Ji PZ. 2016. The parameters selection of PSO algorithm influencing on performance of fault diagnosis. *MMME*. 63.

LAMPIRAN

Lampiran 1 Diagram alir prosedur analisis data



Lampiran 2 Hasil percobaan parameter BPSO

C = 0.1

c_1	c_2	w	k	Akurasi (%)	Sensitivitas (%)	Spesifisitas (%)
1,988	2,300	0,827	8	84	69,6	86
1,861	2,167	0,408	6	83,3	70,7	85
2	2	0,4	6	82,2	68,9	84,1
1,946	2,269	0,708	3	83,5	69	85,5
1,562	2,465	0,753	9	83,4	71,6	85
2	2	0,8	10	81,8	75,3	82,7
1,774	1,902	0,544	7	83,1	72,2	84,7
2	2	0,9	3	81,8	71,8	83,2
2	2	0,7	3	76,7	80,7	76,2
2	2	0,9	6	79,8	74	80,6
2	2	0,7	8	77,5	79	77,3

C = 1

c_1	c_2	w	k	Akurasi (%)	Sensitivitas (%)	Spesifisitas (%)
1,988	2,300	0,827	8	87,9	61,8	91,6
1,861	2,167	0,408	6	84,3	64,2	87,1
2	2	0,4	6	82,1	70,3	83,7
1,931	2,261	0,777	8	88	64,4	91,3
1,986	2,393	0,907	3	88	62	91,7
2	2	0,8	10	85,4	64,4	88,4
1,774	1,902	0,544	7	87,2	64	90,5
2	2	0,9	3	84,5	70	86,5
2	2	0,7	3	89,3	60,5	93,3
2	2	0,9	6	86,3	68,1	88,9
2	2	0,7	8	85,8	64,2	88,9

C=10

c_1	c_2	w	k	Akurasi (%)	Sensitivitas (%)	Spesifisitas (%)
1,988	2,300	0,827	8	83,5	67,6	85,7
1,861	2,167	0,408	6	81,7	68,8	83,5
1,809	2,318	0,952	6	81,2	71,2	82,6
1,067	1,613	2,069	6	84,4	70,9	86,3
2	2	0,8	10	80,2	73,1	81,2
1,774	1,902	0,544	7	86,6	68,3	89,1
2	2	0,9	3	83	66,6	85,4
2	2	0,7	3	79,9	71,4	81,1
2	2	0,9	6	80,4	75,7	81,1
2	2	0,7	8	86	67	88,6

Lampiran 3 Banyaknya fitur terpilih pada tiap *fold*

C	Fold ke-	Banyak fitur terpilih
0,1	1	1179
0,1	2	1129
0,1	3	1159
0,1	4	1188
0,1	5	1193
0,1	6	1211
0,1	7	1198
0,1	8	1196
0,1	9	1209
0,1	10	1214
1	1	1190
1	2	1196
1	3	1193
1	4	1186
1	5	1181
1	6	1221
1	7	1186
1	8	1172
1	9	1158
1	10	1220
10	1	1160
10	2	1167
10	3	1175
10	4	1212
10	5	1223
10	6	1214
10	7	1165
10	8	1173
10	9	1211
10	10	1191

Lampiran 4 Daftar tautan *Text Pre-processing*

No.	Jenis	Link
1.	<i>Stopwords</i> Bahasa Indonesia	https://github.com/audhiaprilliant/Indonesia-Public-Election-Twitter-Sentiment-Analysis/blob/master/Helpers/Data%20Helper/stopwords-id.txt
2.	<i>Stopwords</i> Bahasa Inggris	https://gist.github.com/sebleier/554280
3.	Normalisasi (<i>colloquial-indonesian-lexicon</i>)	https://github.com/audhiaprilliant/Indonesia-Public-Election-Twitter-Sentiment-Analysis/blob/master/Helpers/Data%20Helper/colloquial-indonesian-lexicon.txt
4.	Tambahan Normalisasi	https://docs.google.com/spreadsheets/d/1BkGsoaYQC01pnZi3pItUvQtY-bEgLjNj6J0necMhIs/edit?usp=sharing

RIWAYAT HIDUP

Penulis dilahirkan di kota Sumedang pada 27 November 1998 sebagai anak kedua dari pasangan bapak Cece Tohidi dan Almh. Nena Maemunah. Penulis menempuh pendidikan sekolah menengah atas (SMA) di SMA Negeri Situraja dan lulus pada tahun 2017. Setelah lulus SMA, tepatnya tahun 2017, penulis diterima sebagai mahasiswa program sarjana (S-1) di Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam IPB.

Selama mengikuti program S-1, penulis aktif menjadi asisten responsi mata kuliah Metode Statistika (Agustus 2020 - Desember 2020 dan Agustus 2019 – Desember 2019). Selain bidang akademik, penulis juga aktif dalam beberapa kegiatan organisasi dan kepanitiaan baik di dalam maupun luar kampus. Tahun 2020 penulis menjadi pengurus Badan Eksekutif Mahasiswa (BEM) Fakultas MIPA sebagai Kepala Biro Eksternal dan Riset. Penulis juga menjadi pengurus dari Ikatan Himpunan Mahasiswa Statistika Indonesia (IHMSI) sebagai Staff Badan Usaha pada tahun 2018-2020. Pada tahun 2019, penulis aktif sebagai Kepala Divisi Internal di Organisasi Mahasiswa Daerah Wapemala. Beberapa kepanitiaan yang pernah diikuti oleh penulis yaitu sebagai Sekretaris Divisi Humas dan Publikasi pada acara Pengena;an Fakultas (G-Family), manajer cabang atletik pada acara FMIPA Goes to OMI, Staff Divisi Kompetisi pada Pesta Sains Nasional, Staff Divisi *Liaison Officer* pada acara Statistika Ria, dan lain sebagainya. Tahun 2020, penulis berkesempatan melakukan praktik lapang di *Iris Worldwide Indonesia* pada Divisi *Customer Relationship Management* (CRM) Subdivisi Data dan Analisis.

Penulis pernah mengikuti beberapa perlombaan yang berkaitan dengan statistika dan sains data. Beberapa pencapaian yang diperoleh yaitu menjadi semifinalis pada Kompetisi Statistika Nasional Satria Data 2020, semifinalis *Data Analysis Competition* PRS ITS 2020, dan semifinalis *Statistics Team Competition* (STC) LOGIKA UI 2020.