Journal of Information Technology and Computer Science (INTECOMS)
Volume 8 Nomor 2, Tahun 2025
e-ISSN : 2614-1574
p-ISSN : 2621-3249

COMPARATIVE ANALYSIS OF SENTIMENT CLASSIFICATION ON CRIME INDICATED OPINION CYBERBULLYING ON TWITTER USING SVM AND NAÏVE BAYES METHODS

ANALISIS KOMPARASI KLASIFIKASI SENTIMEN PADA *CRIME INDICATED* OPINION *CYBERBULLYING* DI TWITTER MENGGUNAKAN METODE SVM DAN NAÏVE BAYES

Atika Soraya[1], Abdiansah[2], Ermatita[3]
Jurusan Magister Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya[1,2,3]
atikasoraya87@gmail.com[1]

ABSTRACT

Cyberbullying is one of the actions that violate the ITE Law where the crime is committed on social media applications such as Twitter. This action is difficult to detect if no one is reporting the tweet. Cyberbullying tweet identification aims to classify tweets that contain bullying content. The classification process uses Support Vector Machine and Naïve Bayes methods where the method aims to find a comparison of the accuracy values of each method. The system process starts from text preprocessing with the stages of case folding, tokenization, stopword removal, stemming and scoring. The next process is to classify based on bullying and non-bullying data labeling aimed at facilitating the process of finding the accuracy value of dataset classification using the Support Vector Machine and Naïve Bayes methods. The results obtained using the Support Vector Machine method are 82.29% better than the Naïve Bayes method with a yield of 80.84%.

Keywords*: Naïve Bayes, Support Vector Machine, Cyberbullying, Bullying, Classification*

ABSTRAK

Cyberbullying merupakan salah satu tindakan yang melanggar UU ITE dimana kejahatan ini dilakukan di media sosial salah satunya aplikasi Twitter. Tindakan ini sulit terdeteksi jika tidak ada yang mereport tweet tersebut. Identifikasi tweet *Cyberbullying* bertujuan untuk mengklasifikasikan tweet yang mengandung konten *bullying*. Klasifikasi dilakukan dengan menggunakan metode *Support Vector Machine* dan *Naïve Bayes* dimana metode tersebut bertujuan untuk mencari perbandingan nilai akurasi dari setiap metode. Proses sistem dimulai dari text *preprocessing* dengan tahapan *case folding, tokenisasi, stopword removal, stemming* dan pembobotan. Proses selanjutnya melakukan klasifikasi berdasarkan pelabelan data bullying dan non bullying bertujuan mempermudah proses pencarian nilai akurasi klasifikasi dataset dengan metode *Support Vector Machine* dan naïve bayes. Hasil yang didapat dengan menggunakan metode *Support Vector Machine* sebesar 82.,29% lebih baik dari metode *Naïve Bayes* denga hasil sebesar 80,84%.

Kata Kunci: Naïve Bayes, Support Vector Machine, Cyberbullying, Bullying, Klasifikasi

PENDAHULUAN

Twitter merupakan salah satu media sosial terpopuler yang berperan sebagai wadah komunikasi di masyarakat. Dengan menggunakan Twitter, seluruh orang di dunia dapat terhubung dengan keluarga, teman, dan kerabat melalui komputer atau ponsel mereka. Salah satu layanan yang disediakan oleh Twitter kepada penggunanya adalah pembuatan pesan status (disebut "*tweets*") yang dapat dibaca oleh pengguna Twitter lainnya dan biasanya berisi ungkapan pendapat pengguna dalam berbagai topik dengan batasan sebanyak 140 karakter, sehingga twitter menjadi salah satu situs yang menyediakan kumpulan data opini dari masyarakat di seluruh dunia.

Media sosial dapat menunjukkan perkembangan yang sangat besar pada masa ini. Hampir sebagian besar orang dari berbagai kalangan menggunakan media sosial dalam kehidupan sehari–hari (Chazawi and A. Ferdian, 2015), seperti media promosi, aktivitas politik, serta tindakan kriminalitas dapat dilakukan melalui media sosial twitter (de Semiocast, 2012). Penyebaran informasi yang sangat cepat diiringi banyaknya pengguna akun

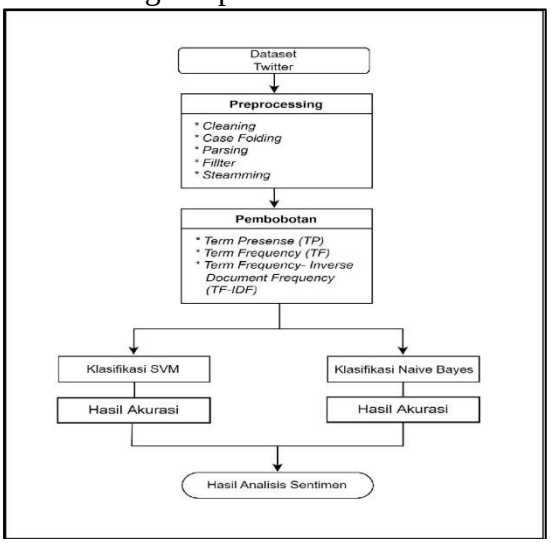
dalam pembuatannya bisa menggunakan identitas asli, hal ini menjadi sulitnya tindak kejahatan sulit ditemukan. Aktivitas kriminal atau tindak kejahatan siber yang sering dilakukan melalui Twitter pada umumnya berupa *cyber crime* di antaranya yaitu ancaman keamanan cyber seperti rekayasa sosial, eksploitasi kerentanan perangkat lunak, dan serangan jaringan.

Analisis sentimen merupakan proses memahami, mengekstrak, dan mengolah data tekstual secara otomatis untuk mendapatkan informasi analisis ini untuk mendeteksi sebuah informasi emosional seseorang yang terdapat dalam pesan para pengguna jejaring sosial twitter terhadap topik yang dibahas penggunanya, seperti dalam mengklasifiksai *review* film dengan *machine learning*, hal ini didukung dengan hasil penelitian oleh Pang yang menemukan adanya klasifikasi positif dan negatif(M. Badri, 2011), selain itu terdapat pendekatan klasifikasi aktivitas krimial lainnya seperti pendekatan dengan *Naïve Bayes (NB)* dan *Support Vector Machine (SVM)(* B. Pang and L. Lee, 2008), *metode algoritma Naïve Bayes* (G. Balakrishnan, P. Priyanthan, 2012) dan masih terdapat kekurangan dalam mendeteksi perubahan emosional yang signifikan dalam mengekstrak informasi tentang polaritas sentimen pengguna meliputi positif, netral atau negatif. Analisis sentimen mengelompokkan polaritas dari teks yang ada dalam kalimat atau dokumen untuk mengetahui pendapat yang dikemukakan dalam kalimat atau dokumen tersebut apakah bersifat positif atau negatif (A. Ortigosa, J. M. Martín, 2014) . Penelitian ini yang di usulkan untuk mengetahui kinerja metode *Support Vector Machine* dalam mengklasifikasikan tweets masyarakat mengenai *Opinion Cyberbullying* di media sosial Twitter.

Pada penelitian ini, dokumen tweet yang diambil diklasifikasikan menjadi kategori *crime indicated opinion*. *Crime indicated opinion* yang akan diambil dalam penelitian ini adalah tweet atau pesan pada Twitter yang mengandung unsur aktivitas kriminal. Selanjutnya peneliti akan mencari nilai akurasi dari metode *Naïve Bayes* dan *Support Vector Machine*.

METODE

Pada bagian ini akan dijelaskan mengenai perancangan sistem dalam penelitian ini. Metode dalam penelitian ini dilakukan melalui beberapa tahap. Secara garis besar, alur penelitian dapat dijelaskan melalui diagram pada Gambar 1.



Gambar 1. Proses Penelitian

2.1 Dataset

Data yang digunakan dalam penelitian ini adalah tweets yang mengandung kata "*Cyberbullying*", "Cyber" atau "Bullying". Data diperoleh dengan menggunakan API (*application programming interface*) dalam proses ini kata kunci penarikan dataset yaitu dengan menulis kata *Cyberbullying* dimana aplikasi ini dapat membantu menghubungkan dua atau lebih perangkat. Kemudian komentar tersebut akan diberikan label nilai positif dan negatif.

2.2 Pra-Pengolahan Data / Preprocessing Data

Pre-processing adalah proses pembersihan dokumen yang bertujuan untung menghilangkan noise. Proses pembersihan dokumen yang dilakukan pada penelitian ini antara lain:

2.2.1 *Cleaning*

Data cleaning adalah suatu prosedur untuk memastikan kebenaran, konsistensi, dan kegunaan suatu data yang ada dalam *dataset*. Caranya adalah dengan mendeteksi adanya *error* atau *corrupt* pada data, kemudian memperbaiki atau menghapus data jika memang diperlukan.

2.2.2 *Case Folding*

Case folding merupakan proses mengubah seluruh huruf menjadi huruf kecil. Pada proses ini karakter-karakter 'A'-'Z' yang terdapat pada data diubah kedalam karakter 'a'-'z'.

2.2.3 *Parsing*

Parsing data merupakan proses pengambilan data dalam satu format kemudian mengubahnya ke format yang lain.

2.2.4 *Filter*

Fitur *Filter* atau sering disebut juga sebagai *autofilter* adalah salah satu fitur yang terdapat pada microsoft office Excel yang digunakan menampilkan baris data tertentu sesuai filter yang kita terapkan dan menyembunyikan sisa baris data lainnya.

2.2.5 *Stemming*

Tahap *stemming* adalah tahapan yang juga diperlukan untuk memperkecil jumlah indeks yang berbeda dari satu data sehingga sebuah kata yang memiliki *suffix* maupun *prefix* akan kembali ke bentuk dasarnya.

2.3 Klasifikasi *Support Vector Machine*

SVM adalah algoritma *Machine Learning* yang menerapkan fungsi *hyperplane* pada data sehingga terbentuk daerah-daerah tiap kelas. *Hyperplane* sendiri merupakan sebuah fungsi yang digunakan sebagai pemisah antar kelas yang ada. Dalam memprediksi suatu kelas dari data, SVM akan melabelinya berdasarkan daerah kelas mana yang merupakan tempat dari data tersebut. SVM biasanya digunakan pada dataset besar yang diambil dari situs online dan menjadi popular karena penerapannya dalam klasifikasi teks (S. Hanggara, T. Akhriza, 2017). Prinsip SVM adalah membangun hyperplane yang memiliki ukuran margin yang sama dan tidak cenderung mendekati daerah dari salah satu kelas. Hal tersebut dapat dilakukan dengan pengukuran margin kemudian dilakukan pencarian titik maksimalnya. Usaha pada pencarian hyperplane terbaik sebagai pemisah antar kelas merupakan inti dari metode SVM, yaitu:

1. Proses *training*: pada proses *training* digunakan data *training* set yang telah diketahui label-labelnya untuk membangun model atau fungsi.
2. Proses *testing* : untuk mengetahui keakuratan model atau fungsi yang akan dibangun pada proses *training*, maka digunakan data yang disebut dengan testing set untuk memprediksi label-labelnya [8]

2.4 Klasifikasi *Naïve Bayes*

Algoritma Naïve Bayes merupakan salah satu algoritma yang terdapat dalam teknik klasifikasi. Naïve Bayes merupakan pengklasifikasian dengan bentuk model probabilistik dan statistik yang disederhanakan dengan berdasar pada teorema Bayes dengan asumsi bahwa setiap atribut bersifat bebas (independence).

HASIL DAN PEMBAHASAN PENELITIAN

Dataset yang dikumpulkan berjumlah 13442 data setelah dilakukan pembersihan secara manual. Kemudian dilakukan proses labelisasi. Pembagian label dataset twitter dapat dilihat pada Tabel 1 dibawah:

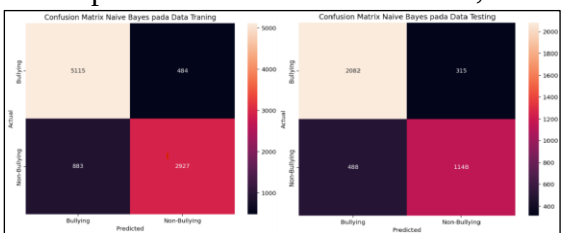
Tabel 1. Dataset *Cyberbullying*

Dataset (Sentimen)	**Jumlah**
Bullying (Positif)	7996
Non-Bullying (Negatif)	5446
Jumlah	**13442**

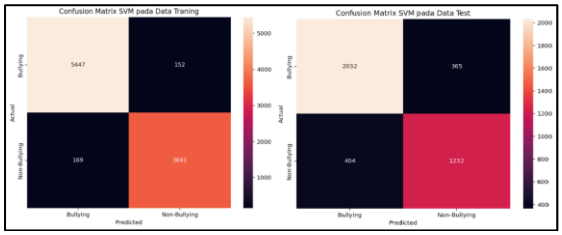
Dataset Cyber-Bullying merupakan data mentah yang masih perlu di proses lagi pada Preprocessing data. Berikut merupakan langkah dalam pengambilan dataset opini umum dari twitter sebagai berikut; *Twitter API Credential* - Developer Consumer API *Keys*- Proses Pearikan data library Tweepy - Dataset *Cyberbullying* - dataset dalam google-colab - Dataset *Cyberbullying* - Kategori Tweet *Cyberbullying*

3.1. Pengujian Metode *Naïve Bayes* dan *Support Vector Machine*

Setelah dilakukanya semua tahapan preprocessing data tahap selanjutnya yaitu pengujian *Confusion Matrix* pada klasifikasi *Naive Bayes* dan *Support Vector Machine*, yang dilakukan dalam 5 kali percobaan. **Percobaan pertama** menggunakan rasio 70 persen data training dan 30 persen data testing. Hasil dapat dilihat pada Gambar 2 dan Gambar 3;

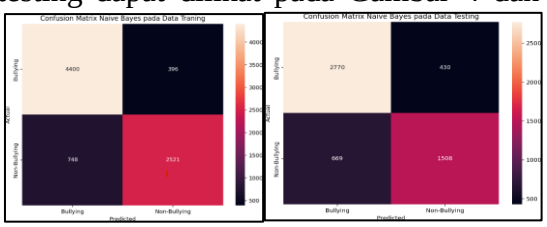


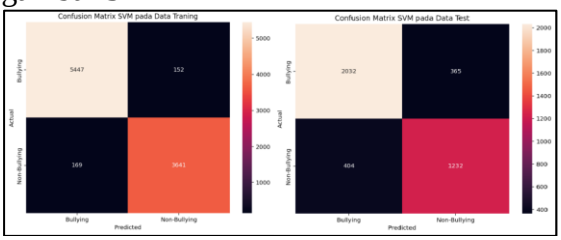
Gambar 2 *Confusion Matrix Naive Bayes* pada data Training



Gambar 3 *Confusion Matrix Support Vector Machine*

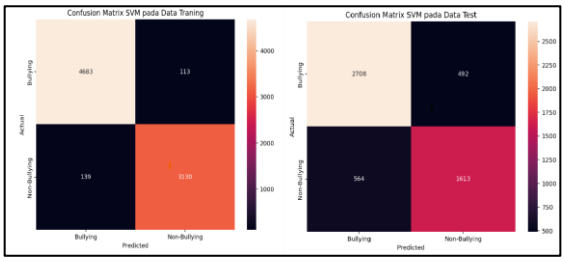
Percobaan Kedua menggunakan rasio 60 persen data training dan 40 persen data testing. Hasil *Confusion Matrix Naive Bayes* pada dataset *Cyberbullying* pada data training dan testing dapat dilihat pada Gambar 4 dan gambar 5





Gambar 4. *Confusion Matrix Naïve Bayes*

Percobaan Ketiga menggunakan rasio 75 persen data training dan 25 persen percobaan Keempat menggunakan rasio 80 persen data training dan 20 persen dan percobaan Kelima menggunakan rasio 65 persen data training dan 35 persen pada data testing. Hasil Confusion Matrix Naive Bayes pada dataset Cyberbullying pada data training dan testing dapat dilihat pada Gambar dibawah



Gambar 5. *Confusion Matrix Support Vector Machine*

Dataset Cyber-Bullying merupakan data mentah yang masih perlu di proses lagi pada Preprocessing data. Berikut merupakan langkah dalam pengambilan dataset opini umum dari twitter sebagai berikut; *Twitter API Credential* - Developer Consumer API *Keys*- Proses Pearikan data library Tweepy - Dataset *Cyberbullying* - dataset dalam google-colab - Dataset *Cyberbullying* - Kategori Tweet *Cyberbullying*.

Tabel 2. Perbandingan Nilai Akurasi metode *Naive Bayes* dan *Support Vector Machine*

N0	Data Test	Data Training	Naïve Bayes		SVM	
			Acu Traing	Acu Test	Acu Traing	Acu Test
1	30%	70%	85.47%	80.08%	96.58%	80.93%
2	40%	60%	85.81%	79.56%	96.87%	80.36%
3	25%	75%	85.49%	80.18%	95.86%	79.76%

4	20%	80%	85.21%	**80.84%**	95.66%	**82.29%**
5	35%	65%	85.61%	79.57%	96.62%	80.93%

Berdasarkan hasil dari Tabel 2 diatas, penggunaan metode *Support Vector Machine* untuk klasifikasi teks dengan menggunakan dataset *Cyberbullying* mengungguli dibandingkan dengan menggunakan *Naïve Bayes* dengan hasil nilai akurasi 82.29% dengan menggunakan metode *Support Vector Machine* dan 80.84% dengan menggunakan metode *Naïve Bayes*. Secara keseluruhan hasil dari akurasi, *precision, recall,* dan F1-*Score Support Vector Machine* lebih diunggulkan bila dibandingkan metode *Naive Bayes*.

SIMPULAN

Provide Pada penelitian ini, ada beberapa poin yang bisa ditarik untuk dijadikan kesimpulan dengan penggunaan metode *Support Vector Machine* dan antara lain sebagai berikut:

1. Metode *Support Vector Machine* yang memiliki nilai akurasi sebesar 82,29% lebih unggul dari *Naïve Bayes* yang memiliki nilai sebesar 80,84% untuk klasifikasi teks *bullying* atau *non-bullying* dengan menggunakan *dataset twitter* pada ukuran data test sebesar 0,2.
2. Berdasarkan pengujian yang telah dilakukan dengan menggunakan *dataset* yang sama, hasil yang didapatkan dengan menggunakan metode *Support Vector Machine* mengungguli bila dibandingkan dengan menggunakan metode *Naïve Bayes*. Metode *Support Vector Machine* mampu menghasilkan nilai akurasi sebesar 82,29%, sedangkan nilaiakurasi dengan menggunakan metode *Naïve Bayes* sebesar 80,84%. Ini membuktikan bahwa penggunaan metode *Support Vector Machine* mampu meningkatkan nilai akurasi klasifikasi teks bila dibandingkan dengan menggunaan metode *Naïve Bayes*.
3. Dari kelima percobaan dengan ukuran data test yang berbeda, didapatkan bahwa *Support Vector Machine* memiliki hasil akurasi yang lebih baik dibandingkan Naïve Bayes. Sehingga metode *Support Vector Machine* sangat disarankan penggunaannya pada klasifikasi teks dengan penggunaan dataset twitter.

Hasil dari penelitian menunjukkan bahwa penelitian peningkatan akurasi klasifikasi teks dengan menggunakan metode *Support Vector Machine* mampu mendapatkan nilai akurasi yang diharapkan bila dibandingkan dengan metode *Naïve Bayes*. Perlunya penggunaan *dataset* selain data *twitter* atau media sosial seperti *dataset* deteksi email spam dan non spam, kategori artikel berita, sms spam, dan lain-lain untuk melihat apakah penggunaan metode *Support Vector Machine* mampu memberikan nilai akurasi yang tinggi jika menggunakan *dataset* tersebut.

DAFTAR PUSTAKA

A. Chazawi and A. Ferdian, *Tindak Pidana Informasi & Transaksi Elektronik Penyerangan Terhadap Kepentingan Hukum Pemanfaatan Teknologi Informasi dan Transaksi Elektronik*, 364.099 Ch. Malang: Media Nusa Creative, 2015.

Semiocast, “Twitter reaches half a billion accounts. More than 140 million in the U.S.,” *de Semiocast*, 2012.

M. J. Soler, F. Cuartero, and M. Roblizo, “Twitter As A Tool For Predicting Elections Results,” *Proc. 2012 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2012*, pp. 1194–1200, 2012.

M. Badri, “Komunikasi Pemasaran UMKM Di Era Media Sosial. Corporate and Marketing Communication,” no. January, p. Jakarta : Pusat Studi Komunikasi dan Bisnis Progra, 2011.

H. Saif, Y. He, and H. Alani, “Semantic Sentiment Analysis of Twitter BT - The Semantic Web – ISWC 2012,” 2012, pp. 508–524.

B. Pang, L. Lee, and S. Vaithyanathan, "Techniques, Thumbs up? Sentiment Classification using Machine Learning," *Proc. Conf. Empir. Methods Nat. Lang. Process.*, vol. 31, no. 9, pp. 481–482, 2002.

B. Pang and L. Lee, "Opinion Mining And Sentiment Analysis," *World J. Gastroenterol.*, vol. 2, no. 1–2, pp. 1–135, 2008.

G. Balakrishnan, P. Priyanthan, R. Thiruchittampalam, N. Prasath, and A. Perera, "Opinion mining and sentiment analysis on a Twitter data stream," in *International Conference on Advances in ICT for Emerging Regions, ICTer 2012 - Conference Proceedings*, 2012, pp. 182–188.

A. Ortigosa, J. M. Martín, and R. M. Carro, "Sentiment analysis in Facebook and its application to e-learning," *Comput. Human Behav.*, vol. 31, pp. 527–541, 2014.

S. Hanggara, T. Akhriza, and M. Husni, "Aplikasi Web untuk Analisis Sentimen pada Opini Produk dengan Metode *Naive Bayes* Classifier," 2017.

S. Ravikumar, M. Fredimoses, and R. Gokulakrishnan, "Biodiversity of actinomycetes in Manakkudi mangrove ecosystem, Southwest coast of India," *Ann. Biol. Res.*, vol. 2, no. 1, pp. 76–82, 2011.

G. S. Solakidis, K. N. Vavliakis, and P. A. Mitkas, "Multilingual Sentiment Analysis Using Emoticons and Keywords," *2014 IEEE/WIC/ACM Int. Jt. Conf. Web Intell. Intell. Agent Technol.*, vol. 2, pp. 102–109, 2014

V. Kharde and S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques," *Int. J. Comput. Appl.*, vol. 139, pp. 5–15, 2016,

B. Gupta, M. Negi, K. Vishwakarma, G. Rawat, and P. Badhani, "Study of Twitter Sentiment Analysis using Machine Learning Algorithms on Python," *Int. J. Comput. Appl.*, vol. 165, pp. 29–34, 2017.

C. Brogan, *Social Media 101: Tactics and Tips to Develop Your Business Online*, 1st ed. United Stated: Wiley. Retrieved From John Wiley & Son, 2010.