

# Klasifikasi Ujaran Kebencian Dengan Naïve Bayes Classification pada Page Meme Comic Indonesia

Fadly Febriya<sup>1</sup>, Regina Sukma Citra<sup>2</sup>, Tarsinah Sumarni<sup>3</sup>, Muhamad Malik Mutoffar<sup>4</sup>

<sup>1234</sup>Teknik Informatika

Sekolah Tinggi Teknologi Bandung

Bandung, Indonesia

e-mail: <sup>1</sup>febriyastudy@gmail.com, <sup>2</sup>reginas.citra@gmail.com, <sup>3</sup>tarsinahsumarni@sttbandung.ac.id,

<sup>4</sup>malik@sttbandung.ac.id

## Abstrak

Hadirnya teknologi di lingkungan masyarakat disambut dengan baik hal ini disebabkan dengan mudahnya mengakses informasi di sosial media tanpa melihat batas ruang dan waktu. Mudahnya mengakses informasi bisa mendatangkan dampak negatif seperti penyalahgunaan sosial media dengan menebarkan ujaran kebencian. Ujaran kebencian sering dilakukan pada postingan-postingan khususnya di postingan yang membagikan berbagai informasi kepada pengikutnya (followers). Berdasarkan penuturan dari Polri bahwa angka kasus ujaran kebencian sebanyak 3.325 kasus dengan presentasi kenaikan sebanyak 44,99% dengan ujaran kebencian terbanyak adalah kasus penghinaan. Kasus nyata dari ujaran kebencian yang sering terjadi ini maka perlu adanya pengetahuan baru terhadap kata atau komentar yang termasuk dalam ujaran kebencian. Pengetahuan baru ini akan dilakukan dengan pengklasifikasian ujaran kebencian yang ada dalam komentar di postingan page dengan menggunakan Algoritma Naïve Bayes Classification. Diharapkan dengan adanya pengklasifikasian ini dapat membantu masyarakat untuk berhati-hati dalam menulis postingan ataupun komentar di sosial media.

**Kata kunci:** Klasifikasi, Ujaran Kebencian, Naïve Bayes Classification

## Abstract

The presence of technology in the community is welcomed, this is due to the ease of accessing information on social media without looking at the boundaries of time and space. Easy access to information can have negative impacts such as misuse of social media by spreading hate speech. Hate speech is often carried out in posts, especially in posts that share various information with their followers. Based on the narrative from the National Police, the number of hate speech cases was 3,325 cases with an increase of 44.99% with the most hate speech being contempt cases. The real case of hate speech that often occurs, there is a need for new knowledge of the words or comments that are included in hate speech. This new knowledge will be done by classifying the hate speech in the comments on the page post using the Naïve Bayes Classification Algorithm. It is hoped that this classification can help the public to be careful in writing posts or comments on social media.

**Keywords:** Classification, Hate Speech, Naïve Bayes Classification

## 1. Pendahuluan

Dewasa ini munculnya perkembangan teknologi dan informasi yang berkembang sangat pesat, begitupula dengan perkembangan data. Perkembangan dari data ini akan menghasilkan sejumlah data besar yang digunakan oleh manusia untuk menghasilkan informasi dan pengetahuan (*knowledge*). Penyebaran informasi di internet berkembang sangat cepat khususnya di *platform* sosial media seperti Facebook, hal ini berdasarkan data dari Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) [1] dan *We Are Social dan Hootsuite*, bahwa pengguna internet di Indonesia sebanyak 143,26 juta jiwa dan pengguna Facebook di Indonesia sebanyak 130 juta akun [2].

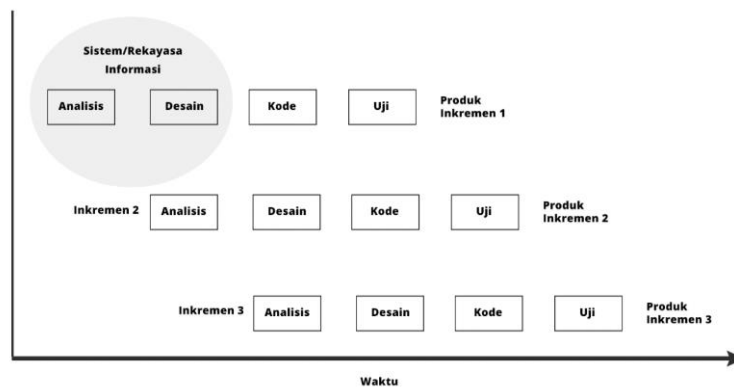
Melihat dari banyaknya pengguna sosial media khususnya pengguna Facebook di Indonesia dan penyebaran informasi yang sangat cepat maka akan ada dampak negatif dari hal ini. Contoh dampak negatif dari penggunaan sosial media yaitu penyalahgunaan sosial media dengan menebarkan ujaran kebencian. Ujaran kebencian adalah komunikasi yang dilakukan oleh individu atau kelompok dalam bentuk provokasi, hasutan maupun hinaan kepada individu atau kelompok lainnya [3]. Menurut Polri, angka kasus ujaran

kebencian sebanyak 3.325 kasus dengan presentasi kenaikan sebanyak 44,99% dari tahun sebelumnya. Kasus ujaran kebencian paling banyak terjadi pada kasus penghinaan sebanyak 1.657 kasus dengan presentasi kenaikan sebanyak 73,14%.

Dalam penelitian terdahulu dengan judul “Penerapan Data Mining Untuk Memprediksi Klasifikasi Jumlah Pembaca Sebuah Artikel Pada Situs Beranda.co.id Menggunakan Algoritma *Bayesian Classification*” [4] dijelaskan bahwa aplikasi *Data Mining* dengan menggunakan Algoritma *Bayesian Classification* telah berhasil diterapkan ke dalam perancangan dan pengembangan dengan dibuktikan hasil pada perhitungan manual dan perhitungan aplikasi. Klasifikasi dapat didefinisikan sebagai proses menyatakan suatu objek data sebagai kategori atau sebuah kelas yang telah didefinisikan sebelumnya [5]. Pengklasifikasian menurut *Naïve Bayes* ini digunakan untuk memprediksi probabilitas keanggotaan pada suatu *class* [6]. Pengertian lainnya *Naïve Bayes* berasumsi bahwa nilai antar variabel saling independent pada nilai outputnya [7]. Berdasarkan dari permasalahan di atas dan penelitian terdahulu maka dalam penelitian ini akan dilakukan pengklasifikasian ujaran kebencian dengan menggunakan Algoritma *Naïve Bayes Classification* dan metode pengembangan aplikasi yang akan digunakan adalah Model Iterasi.

## 2. Metode Penelitian

Metode penelitian yang dipilih adalah model Iteratif, model ini akan menghasilkan versi pada perangkat lunak yang sudah mengalami penambahan fungsi untuk setiap pertambahannya [8]. Pada Gambar 1 menjelaskan tahapan pada model pengembangan Iteratif.



Gambar 1. Model Pengembangan Iteratif [8]

Metode pengumpulan data dilakukan secara wawancara, wawancara dilakukan untuk menemukan permasalahan yang harus diteliti dan mengetahui hal-hal yang mendalam [9]. Hasil wawancara dilakukan untuk mengetahui *page* yang akan diambil, *page* diambil berdasarkan jumlah like menurut Socialbakers.

Metode algoritma yang dipakai pada penelitian ini adalah Algoritma *Naïve Bayes Classification*. Analisis dalam pengklasifikasian teksnya menggunakan *Text Mining*, *text mining* dilakukan untuk memperoleh informasi berkualitas tinggi dengan memperhatikan pola dan tren [10]. Tahapan dalam *text mining* adalah *text-preprocessing*, *text-preprocessing* adalah tahap ini mencakup semua rutinitas dan proses untuk mempersiapkan data yang akan digunakan operasi *knowledge discovery* sistem *text mining* [11]. Kriteria ujaran kebencian pada penelitian ini adalah penghinaan, penistaan, perbuatan tidak menyenangkan, *body shamming*, pencemaran nama baik dan bukan ujaran kebencian. Hasil pengujian dari pengklasifikasian ini adalah menentukan tingkat akurasi dari *page* Meme Comic Indonesia. Tingkat akurasi adalah Salah satu aspek yang menjadi parameter kehandalan dari algoritma klasifikasi. Sistem akan melakukan klasifikasi yang diharapkan dapat mengklasifikasikan semua set dengan benar, tetapi tidak semua sistem bekerja dengan baik atau tidak bisa 100% akurat [12]. Berikut adalah rumus formula untuk menghitung tingkat akurasi.

$$\text{Akurasi} = \frac{\text{Jumlah data yang diprediksi secara benar}}{\text{Jumlah prediksi yang dilakukan}} \quad (1)$$

### 3. Hasil dan Pembahasan

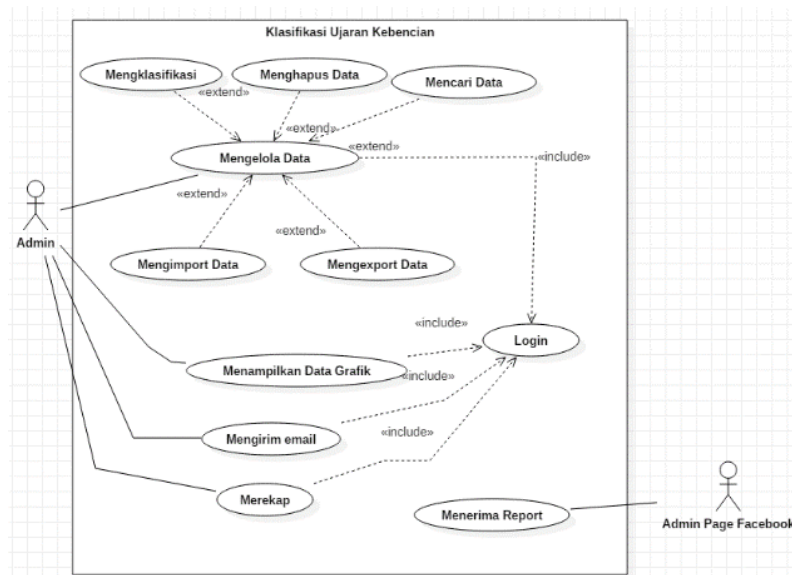
#### 3.1 Analisis Penelitian

Hasil analisis yang dilakukan berdasarkan data yang diperoleh oleh dari page Facebook Meme Comic Indonesia pada Tanggal 24-25 Januari 2019 dan 29 Januari – 04 Februari 2019 menggunakan *API Key Facebook*. Tahap kedua adalah mengolah data dengan melakukan tahapan *text-preprocessing*. Pada tahapan *text-preprocessing* ini mencakup sebagai berikut.

1. *Tokenizer*, memotong urutan karakter untuk disesuaikan dengan kebutuhan sistem.
2. *Stopword*, menghilangkan karakter, tanda baca serta kata-kata umum yang tidak memiliki makna.
3. *Case folding*, merubah semua huruf dalam dokumen menjadi [^a-zA-Z 0-9].
4. *Data training*, menentukan labeling pada data training secara manual dengan enam label yaitu *body shamming*, penghinaan, penistaan, pencemaran nama baik, perbuatan tidak menyenangkan dan bukan ujaran kebencian.
5. *Testing*, proses uji dari data latih yang akan diimport ke dalam sistem.

#### 3.2 Desain

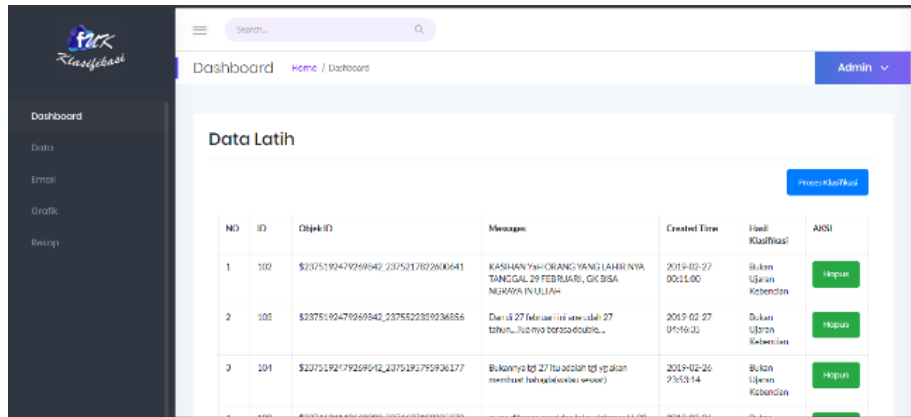
Berikut adalah rancangan *use case diagram*, dalam *use case diagram* ini dijelaskan alur sistem yang akan dibuat dalam penelitian ini, yang akan dijelaskan pada Gambar 1 di bawah ini.



Gambar 1. Use Case Diagram

#### 3.3 Pengujian

Hasil dari implementasi sistem ini adalah dengan melakukan pengujian proses data latih menggunakan metode *Naïve Bayes Classification*. Data latih yang akan diuji sudah melakukan tahapan proses pre-processing (*tokenizer*, *stopword* dan *case folding*) terlebih dahulu dan pembuatan indeks *keyword (stopword)* data latih. Data latih yang digunakan dalam proses ini adalah sebanyak 210 data. Pengujian klasifikasi dilakukan dari data yang sudah diambil dengan proses *import data excel* ke sistem, ketika data berhasil di *import* ke sistem dan mengklik tombol klasifikasi maka sistem secara otomatis akan melakukan proses pengklasifikasian komentar sesuai dengan kategori yang sudah ditentukan. Pada Gambar 2 berikut adalah hasil klasifikasi yang sudah sesuai dengan kategori.



Gambar 2. Data Latih Yang Sudah Di Klasifikasi

3.1.1. 3.3.1 Tingkat Akurasi

Tujuan dari pengujian klasifikasi ini adalah mengetahui seberapa besar kemampuan dari sistem dalam melakukan klasifikasi dari data yang sudah diambil yaitu Meme Comic Indonesia. Contoh proses pada pengujian sistem yang telah berhasil dilakukan dapat dilihat pada Tabel 1 sebagai berikut.

Tabel 1. Contoh Klasifikasi Pada Sistem

Data	Messages	Hasil Pre-processing	Hasil Klasifikasi
Meme Comic Indonesia	Mantap bro ini baru tamatan SMA yg bener" tamat total bukan kek yg lain tamat sekolah tinggi" makin tolot	Mantap baru tamatan SMA bener" tamat total bukan lain tamat sekolah tinggi" makin tolot	Penghinaan
	Wahhh saya baru tau. Ternyata saya juga bisa gerakan genjutsu :v	Wahhh saya baru tau. Ternyata bisa gerakan genjutsu :v	Bukan Ujaran Kebencian

Menentukan tingkat akurasi ini dilakukan dengan melihat jumlah prediksi data uji yang dilakukan pada page Meme Comic Indonesia dengan banyaknya data uji 1491 dari jumlah data yang diambil sebanyak 1530 data. Hasil dari nilai akurasi pada Page Meme Comic Indonesia ini adalah sebagai berikut.

$$Akurasi = \frac{\text{Jumlah data yang diprediksi secara benar}}{\text{Jumlah prediksi yang dilakukan}}$$

$$Akurasi = \frac{1445}{1491} \times 100\%$$

$$Akurasi = 96,91\%$$

4. Kesimpulan

- Berdasarkan dari hasil penelitian yang sudah dilakukan maka dapat disimpulkan sebagai berikut.
1. Hasil klasifikasi komentar yang mengandung ujaran kebencian di page Meme Comic Indonesia sudah sesuai berdasarkan dengan kriterianya.
  2. Pengetahuan baru terhadap kata-kata atau komentar yang termasuk dalam ujaran kebencian beserta kriteria ujaran kebenciannya.
  3. Hasil dari tingkat akurasi dari Meme Comic Indonesia ini adalah 96,91%.

---

**Daftar Isi**

- [1] N. T. Hariyanti, A. Wirapraja “Pengaruh Influencer Marketing sebagai Strategi Pemasaran Digital Era Moderen (Sebuah Studi Literatur),” *Jurnal EKSEKUTIF*, vol. 15, no. 1, pp. 133-146, 2018.
- [2] Y. Fitriani, “Pemanfaatan Media Sosial sebagai Media Penyajian Konten Edukasi atau Pembelajaran Digital,” *Journal of Information System, Applied, Management, Accounting and Research*, vol. 5, no. 4, pp. 1006-1013, 2021.
- [3] S. Mawarti, “Toleransi: Media Komunikasi Umat Beragama,” vol. 10, no. 1, pp. 86-87, 2018.
- [4] Y. Pratama, Penerapan Data Mining untuk Memprediksi Klasifikasi Jumlah Pembaca sebuah Artikel pada Situs Beranda.co.id Menggunakan Algoritma Bayesian Classification, *Media.neliti.com*, 2016.
- [5] F. A. Hermawati, *Data Mining*, Yogyakarta: ANDI, 2013.
- [6] K. and E. T. Luthfi, *Algoritma Data Mining*, Yogyakarta: ANDI, 2009.
- [7] I. W. Saputro and B. W. Sari, “Uji Performa Algoritma Naïve Bayes Untuk Prediksi Masa Studi,” *Creative Information Technology Journal*, vol. 6, no. 1, pp. 1-11, 2020.
- [8] R. A. S and M. Shalahuddin, *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek Edisi Revisi*, Bandung: Informatika Bandung, 2018.
- [9] P. D. S. *Metode Penelitian Kuantitatif, Kualitatif, Dan R&D*, Bandung: Alfabeta, 2017.
- [10] O. W. Purbo, *Text Mining Analisis Medsos, Kekuatan Brand & Intelejen di Internet*, Yogyakarta: ANDI, 2019.
- [11] J. Sanger and R. Feldman, *The Text Mining Handbook : Advanced Approaches in Analyzing*, Newyork: Cambridge University, 2007.
- [12] E. Prasetyo, *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*, 1st ed, Yogyakarta: Data Mining Konsep dan Aplikasi Menggunakan MATLAB, 1st ed, 2012.