

# Klasifikasi Komentar *Bullying* pada Instagram Menggunakan Metode *K-Nearest Neighbor*

Reski Mai Candra<sup>1</sup>, Anindya Nanda Rozana<sup>2</sup>

Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau<sup>1,2</sup>

reski.candra@uin-suska.ac.id<sup>1</sup>, anindya.nanda.rozana@students.uin-suska.ac.id<sup>2</sup>

## Article Info

### History:

Dikirim 5 Mei 2020

Direvisi 2 Juni 2020

Diterima 14 Juli 2020

### Kata Kunci:

Bullying

Klasifikasi

KNN

K-Nearest Neighbor

Komentar Instagram

## Abstrak

Bullying salah satu tindakan negatif yang dilakukan oleh manusia secara terus menerus atau berulang kali. Tindakan ini sering menyebabkan korban menjadi tidak berdaya, terluka secara fisik dan mental. Tindakan bullying kini telah memasuki dunia sosial media. Instagram menjadi salah satu media sosial yang menjadi tempat terjadinya bullying melalui komentar pada foto/video yang diunggah oleh artis/selebgram. Salah satu cara mewaspadaai tindakan bullying adalah dengan cara mengklasifikasi komentar yang bersangkutan di media sosial (dalam hal ini Instagram) menggunakan metode klasifikasi K-Nearest Neighbor (KNN). Jumlah data yang digunakan sebanyak 1000 data, dimana 500 merupakan bullying dan 500 tidak bully. Dengan adanya penelitian ini diharapkan dapat membantu seseorang untuk mengetahui perbedaan komentar bullying dan yang bukan bullying. Metode KNN telah diuji menggunakan 5 nilai k yang berbeda (7, 9, 11, 13, 15) dengan pembagian data 70:30, 80:20 dan 90:10 menggunakan confusion matrix, dengan masing-masing fold yang dibentuk 4, 5 dan 10 sehingga menghasilkan total 95-fold pengujian, pengujian ini menghasilkan akurasi rata-rata sebesar 58,83%. Terdapat akurasi tertinggi pada k 13 dengan perbandingan 90:10 data pada fold ke 6 yaitu sebesar 77%.

© This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

## Koresponden:

Reski Mai Candra

Program Studi Teknik Informatika, Fakultas Sains dan Teknologi

Universitas Islam Negeri Sultan Syarif Kasim Riau

Jl.H.R. Soebrantas No.155 KM 18 Simpang Baru Panam, Pekanbaru 28293

Email: reski.candra@uin-suska.ac.id

## 1. PENDAHULUAN

Bullying merupakan salah satu tindakan negatif yang dilakukan oleh manusia secara terus menerus atau berulang-ulang. Tindakan ini kerap kali menyebabkan korban tidak berdaya, terluka secara fisik maupun mental [1]. Bullying mengakibatkan dampak yang dapat mengancam setiap pihak yang terlibat, baik pihak yang di-bully, pihak yang mem-bully dan pihak yang menyaksikan bullying [2]. Turunnya kesehatan fisik dan mental merupakan pengaruh buruk dari bullying. Dalam kasus yang berat, bullying dapat menjadi penyebab tindakan yang fatal, seperti bunuh diri dan lainnya.

Media sosial Instagram merupakan sebuah media internet yang memungkinkan pengguna merepresentasikan dirinya maupun berinteraksi, bekerja sama, berbagi, berkomunikasi dengan

pengguna lain dan membentuk ikatan sosial secara virtual [1]. Media sosial yang semakin diminati masyarakat khususnya kalangan kaum milenial saat ini adalah Instagram. Instagram mempunyai suatu alat penyampaian pesan (aplikasi) untuk bisa berkomunikasi dengan khalayak secara luas dengan saling berbagi foto atau video, yang di dalamnya juga terdapat fitur-fitur lain seperti DM (Direct Message), comment, love, dan lain-lain. Sistem sosial pada Instagram menghubungkan antara satu akun dengan akunnya lainnya dengan menjadi follower atau following. Hubungan komunikasi antara sesama pengguna Instagram dapat terjalin dengan memberikan tanda love dan komentar pada foto atau video yang telah diunggah oleh pengguna lainnya. Namun pemanfaatan komunikasi melalui komentar Instagram ini banyak disalahgunakan oleh pengguna yang belum memahami etika-etika dalam bersosialisasi pada dunia maya, sering terlihat ejekan, caci, maki, dan bully-an melalui komentar pada Instagram dan apabila berkelanjutan akan menjadi tindakan Cyberbullying.

Penelitian sebelumnya membahas tentang penggunaan metode K-Nearest Neighbor untuk menganalisis sentimen data komentar komentar sosial media Facebook [3]. Penelitian tentang penggunaan algoritma Naive Bayes Classifier untuk deteksi komentar negatif Instagram. Hasil akurasi menggunakan metode Naive Bayes sebesar 76,7% sedangkan hasil akurasi menggunakan metode K-Nearest Neighbor sebesar 79,21%. Maka, dalam penelitian ini metode K-Nearest Neighbor dianggap lebih baik digunakan untuk mengklasifikasi data komentar [4]. Penelitian yang lain menganalisis sentimen *cyberbullying* pada komentar instagram dengan metode klasifikasi *Support Vector Machine*, penelitian ini mencapai hasil akurasi terbaik sebesar 90%, *precisio* sebesar 94,44%, *recall* sebesar 85% dan *f-measure* sebesar 89,47% dengan komposisi data latih 50% dan data uji 50% tanpa mengimplementasikan algoritma *Lexicon Based Features* [5].

Algoritma K-Nearest Neighbor merupakan metode klasifikasi yang mana data baru dikelompokkan berdasarkan jarak data tersebut ke tetangga terdekatnya [4]. Hasil klasifikasi didapatkan dari kelas yang paling banyak muncul. K-NN termasuk algoritma yang sering digunakan untuk klasifikasi teks dan data. Metode K-Nearest Neighbor adalah salah satu metode berbasis NN yang paling tua dan populer. Metode ini menggunakan nilai k untuk menyatakan jumlah tetangga terdekat yang terlibat dalam penentuan prediksilabel kelas pada data uji. Setelah terpilih kemudian dilakukan voting kelas dari K tetangga terdekat tersebut. Kelas dengan jumlah suara tetangga terbanyaklah yang diberikan sebagai label kelas hasil prediksi pada data uji tersebut [6].

Algoritma ini bertujuan untuk mengklasifikasikan objek menurut atribut dan sampel data latih [7]. Seperti halnya K-means, metode ini cukup sederhana, tidak ada asumsi mengenai distribusi data, mudah diaplikasikan dan sering dipakai dalam kasus nyata [7]. Di bawah ini adalah tahapan langkah algoritma K-NN:

1. Tentukan k
2. Hitung jarak data baru dengan semua data training.
3. Urutkan jarak tersebut dari yang terdekat
4. Periksa kelas k tetangga terdekat
5. Kelas data baru = Kelas mayoritas k tetangga terdekatnya

## 2. METODE PENELITIAN

Metode penelitian merupakan pedoman atau tahapan dalam melakukan penelitian. Metode penelitian ini dilakukan bertujuan untuk mendapatkan hasil yang sesuai dengan yang diinginkan. Berikut tahapan-tahapan dalam penelitian:

### 2.1 Pengumpulan Data

Data yang dibutuhkan pada penelitian ini adalah nama akun instagram artis atau selebgram, pemberi komentar, dan teks komentar. Komentar Instagram yang digunakan sebagai data latih dan data uji dikumpulkan melalui akun Instagram artis atau selebgram Indonesia yang memiliki jumlah follower di atas 500 ribu.

Dari nama-nama akun instagram artis atau selebgram, dipilih. Kemudian dilakukan pelabelan secara manual. Proses pengambilan data komentar Instagram dilakukan secara manual dengan mengunjungi profil akun Instagram artis/selebgram dan memilih foto/video yang akan menjadi objek penelitian dengan *dataset* sebanyak 1.000 komentar yang menggunakan bahasa Indonesia. Kemudian

komentar dari foto/video tersebut akan di-copy dan disusun ke dalam Microsoft Excel dan dilakukan Pelabelan manual oleh seorang dosen pakar Bahasa Indonesia Jurusan Pendidikan Bahasa Indonesia UIN Suska Riau bernama Dra. Murny, M. Pd.

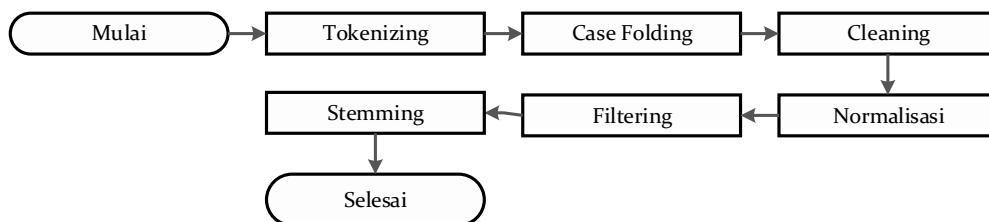
Tabel 1 dibawah ini adalah contoh data komentar yang telah dikumpulkan dan diberi label kelas *bullying* dan tidak *bullying*.

Tabel 1. Contoh Data Komentar.

No	Instagram	Komentar	Komentator	Kelas
D1	rosameldianti	DASAR MUKA COMBERAN!!	feby_rahma9	<i>Bully</i>
D2	rosameldianti	Penyanyi dgn suara terbagus dan gak jelek yaitu kakak	ro.ni6172	Tidak <i>Bully</i>
D3	rosameldianti	Jelek ya wajah di tv aneh ih sama di ig edit jago pake aplikasi apa kak	lid2737	<i>Bully</i>
D4	rosameldianti	horeeeee nggak pernah manggung2	adityazuka	Tidak <i>Bully</i>
D5	rosameldianti	Cieeeee yg di usir Penonton..... Wkkwkwkwk	amel_yeyen	Tidak <i>Bully</i>
D6	rosameldianti	Lupa ngedit ya sayyyaang muka lu buluk tuh 🤔	evie_fc	<i>Bully</i>

## 2.2 Analisa Preprocessing

Praproses (*Preprocessing*) data merupakan langkah penting dalam melakukan analisa klasifikasi dengan tujuan untuk membersihkan data dari unsur-unsur yang tidak dibutuhkan untuk mempercepat proses klasifikasi. Gambar 1 di bawah ini, merupakan *flowchart* tahapan praproses data yang digunakan.



Gambar 1. *Flowchart* Tahapan *Preprocessing*.

Melakukan penjelasan tahapan *Preprocessing* Komentar pada Data nomor D1,

D1	rosameldianti	DASAR MUKA COMBERAN!!	feby_rahma9	<i>Bully</i>
----	---------------	-----------------------	-------------	--------------

### a. *Tokenizing*

Pada proses *Tokenizing* ini, memisahkan setiap kata dari komentar di dalam dokumen yang mulanya berupa kalimat-kalimat menjadi kata-kata dan menghilangkan tanda baca.

Tabel 2. *Tokenizing*

<b>D1</b>
dasar
muka
comberan

b. *Case Folding*

Pada proses *Case Folding*, menyeragamkan seluruh kata dari komentar di dalam dokumen menjadi huruf kecil semuanya.

Tabel 3. *Case Folding*

No	Hasil <i>Cleaning</i>	Hasil <i>Case Folding</i>
D1	DASAR MUKA COMBERAN	dasar muka comberan

c. *Cleaning*

Pada proses *Cleaning*, membersihkan data komentar dari komponen-komponen yang tidak memiliki hubungan dengan informasi yang ada pada dokumen, seperti karakter atau simbol (@#\$%^&\*()\_+";:}{<>.,?!~/[]), angka, *link url* (http://link.com), *hashtag* (#), dan *mention* (@username).

Tabel 4. *Cleaning*

No	Komentar	Hasil <i>Cleaning</i>
D1	DASAR MUKA COMBERAN!!	DASAR MUKA COMBERAN

d. Normalisasi

Pada proses Normalisasi, kata komentar dikembalikan ke kata yang tidak baku kedalam kata baku dalam kamus bahasa Indonesia.

Tabel 5. Normalisasi

<b>D1</b>
dasar
muka
comberan

e. *Filtering*

Pada proses *Filtering*, membuang kata komentar yang tidak penting dan tidak berpengaruh untuk proses klasifikasi. Dengan menggunakan penghapusan *stopword removal*.

Tabel 6. *Filtering*

<b>D1</b>
muka
comberan

f. *Stemming*

Pada proses *Stemming*, pengubahan kata komentar yang berimbuhan menjadi kata dasarnya. Algoritma *stemming* yang digunakan pada penelitian ini adalah ECS (*Enhanced Confix Stripping*).

Tabel 7. *Stemming*

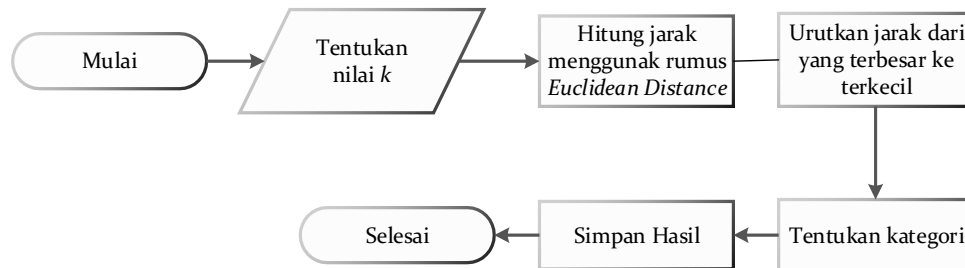
<b>D1</b>
muka
comberan

### 2.3 Pembobotan *TF-IDF*

Merupakan tahapan terakhir setelah *preprocessing*. Pada tahapan *indexing* ini menggunakan *tf* dan *idf*, proses pembobotan yang dilakukan yaitu mengubah kata komentar menjadi bentuk vektor. Tahapan-tahapan yang dilakukan pada proses ini yaitu menghitung *document frequency (df)*, *term frequency (tf)*, dan *inverse document frequency (idf)* dari setiap kata pada komentar.

### 2.4 Analisa Metode *K-Nearest Neighbor*

Pada tahap ini, dilakukan analisa metode klasifikasi yang diperoleh dengan data latih. *Flowchart* proses *K-Nearest Neighbor* dalam klasifikasi penelitian ini akan diperlihatkan pada gambar 2 berikut.



Gambar 2. *Flowchart* Proses Klasifikasi

Adapun langkah-langkah metode *K-Nearest Neighbor* adalah:

- Tentukan  $k$
- Hitung jarak data baru dengan semua data training. Ukuran jarak yang digunakan: Euclidean distance

$$\begin{aligned}
 d(p, q) = d(q, p) &= \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} \\
 &= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}
 \end{aligned} \quad (1)$$

- Urutkan jarak tersebut dari yang terdekat
- Periksa kelas  $k$  tetangga terdekat
- Kelas data baru = Kelas mayoritas  $k$  tetangga terdekatnya

### 3. HASIL DAN PEMBAHASAN

Klasifikasi terdiri dari 2 kelas: kelas *bullying* dan tidak *bullying*, dengan 500 komentar *bullying* dan 500 komentar tidak *bullying*. Perbandingan *dataset* yang digunakan adalah 70:30, 80:20 dan 90:10. Tabel 8 berisikan rincian *dataset* yang digunakan:

Data Komentar	Data Latih	900 (90%)		
	Data Uji	100 (10%)		
		Latih	Uji	Jumlah
Kelas	Bully	450	50	500
	Tidak Bully	450	50	500
Jumlah		900 komentar	100 komentar	1000 komentar

Pengujian yang dilakukan menggunakan *confusion matrix*, pengujian *confusion matrix* dilakukan dengan cara menguji data latih menjadi data uji sebanyak 3 kali pengujian. Menggunakan 1000 data komentar menggunakan 5 nilai  $k$  yang berbeda (7, 9, 11, 13, 15) dengan mekanisme pengujian data latih dan data uji menggunakan perbandingan 70:30, 80:20, 90:10. Pada pengujian 70:30 menggunakan  $k$ -fold sebanyak 4 kali, sedangkan perbandingan 80:20 sebanyak 5 kali dan

perbandingan 90:10 sebanyak 10 k-fold. Penentuan fold pada tiap perbandingan didapatkan dari hasil jumlah keseluruhan data dibagi dengan data uji. Tujuan pengujian ini digunakan untuk mencari data yang memiliki nilai optimal atau nilai yang sama, antara kelas awal dengan hasil menggunakan metode K-Nearest Neighbor.

a. Pengujian akurasi perbandingan 70:30

Pengujian pertama dilakukan perbandingan 70:30 dari 1000 data, sehingga terbentuk data uji sebanyak 300 data dan data latih sebanyak 700 data. Agar pengujian ini seimbang, maka pengujian pada perbandingan ini dibagi menjadi 4-fold.

Tabel 9. Pengujian k-13 Perbandingan 70:30

Fold	Label Benar	Label Salah	Recall	Akurasi
1	145	155	90,7%	48,33%
2	158	142	94,22%	52,84%
3	168	132	92,73%	56%
4	165	135	92,81%	55%

Pada tabel 3 telah dilakukan pengujian terhadap k-7, k-9, k-11, k-13, k-15 dengan perbandingan 70:30, setelah dilakukan pengujian diambil pengujian k-13 yang mempunyai akurasi tinggi pada fold ke-3 dengan akurasi 56%.

b. Pengujian akurasi perbandingan 80:20

Pengujian kedua dilakukan dengan perbandingan 80:20 dari 1000 data, sehingga terbentuk data uji sebanyak 200 data dan data latih sebanyak 800 data. Agar pengujian seimbang, maka pengujian pada perbandingan ini dibagi menjadi 5-fold.

Tabel 10. Pengujian k-13 Perbandingan 80:20

Fold	Label Benar	Label Salah	Recall	Akurasi
1	104	96	91,94%	52%
2	71	129	92,815%	35,5%
3	118	82	95,095%	59,3%
4	98	102	95,40%	49%
5	123	77	93,715%	61,5%

Pada tabel 4 telah dilakukan pengujian terhadap k-7, k-9, k-11, k-13, k-15 dengan perbandingan 80:20, setelah dilakukan pengujian diambil pengujian k-13 yang mempunyai akurasi tinggi pada fold ke-5 dengan akurasi 61,5%.

c. Pengujian akurasi perbandingan 90:10

Pengujian ketiga dilakukan dengan perbandingan 90:10 dari 1000 data, sehingga terbentuk data uji sebanyak 100 data dan data latih sebanyak 900 data. Agar pengujian seimbang, maka pengujian pada perbandingan ini dibagi menjadi 10-fold.

Tabel 11. Pengujian k-13 Perbandingan 90:10

Fold	Label Benar	Label Salah	Recall	Akurasi
1	67	33	91,78%	67%
2	51	49	94,31%	51%
3	39	61	92,26%	39%
4	36	64	94,41%	36%
5	61	39	96,36%	61%

6	77	23	96,56%	77%
7	58	42	97,21%	58%
8	59	41	94,57%	59%
9	63	37	91,16%	63%
10	59	41	97,75%	59%

Pada tabel 5 telah dilakukan pengujian terhadap k-7, k- 9, k-11, k-13, k-15 dengan perbandingan 90:10, setelah dilakukan pengujian diambil pengujian k-13 yang mempunyai akurasi tinggi pada fold ke-6 dengan akurasi 77%.

#### 4. KESIMPULAN

Berdasarkan tahapan-tahapan metodologi penelitian yang telah dilaksanakan, maka dapat ditarik beberapa kesimpulan sebagai berikut:

- a. Metode K-Nearest Neighbor dapat diterapkan untuk klasifikasi bully dan tidak bully pada komentar Instagram.
- b. Klasifikasi bully dan tidak bully pada komentar Instagram menggunakan K-Nearest Neighbor ini telah berhasil mengklasifikasikan komentar bully dan tidak bully, dengan akurasi tertinggi terdapat pada k 13 dengan perbandingan 90:10 data pada fold ke 6 yaitu sebesar 77%, sedangkan akurasi terendah terdapat pada k 11 dengan perbandingan data 90:10 pada fold ke 1 yaitu sebesar 35%.



#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada (pihak pemberi dana) yang telah memberi dukungan financial terhadap penelitian ini. (Penjelasan ini dibuat jika penelitian ini mendapatkan pendanaan dari pihak lain)

#### DAFTAR PUSTAKA

- [1] Nasrullah, R. (2017). *Media Sosial: Perspektif Komunikasi, Budaya, dan Sosioteknologi*. Bandung: Simbiosis Rekatama Media.
- [2] Priyatna, A. (2010). *Let's End Bullying: Memahami, Mencegah & Mengatasi Bullying*. Jakarta: Elex Media Komputindo
- [3] Salam, A., Zeniarja, J., & Uswatun Khasanah, R. S. (2018). Analisis Sentimen Data Komentar Sosial Media Facebook Dengan K-Nearest Neighbor (Studi Kasus: Akun Jasa Ekspedisi Barang J&T Ekspres Indonesia). *Prosiding SINTAK*, 480-486.
- [4] Ismiati, M. B. (2018). Deteksi Komentar Negatif Di Instagram Menggunakan Algoritma Naives Bayes Classifier.
- [5] Luqyana, W. A., Cholissodin, I., & Perdana, R. S. (2018). Analisa Sentimen Cyberbullying Pada Komentar Instagram Dengan Metode Klasifikasi Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.
- [6] Prasetyo, E. (2014). *Data Mining: Mengolah Data menjadi Informasi Menggunakan Matlab*. Yogyakarta: ANDI.
- [7] Nugroho, M., & Santoso, H. (2017). Klasifikasi Dokumen Komentar pada Situs Youtube menggunakan Algoritma K-Nearest Neighbor (K-NN).

**BIOGRAFI PENULIS**

	<p><b>Reski Mai Candra</b> obtained Bachelor Degree in Informatics Engineering from Universitas Islam Negeri Sultan Syarif Kasim Riau in 2009, obtained Master Degree in Computer Science from Universiti Teknologi Malaysia (UTM) in 2013. he has been a Lecturer with the Department of Informatics Engineering, Universitas Islam Negeri Sultan Syarif Kasim Riau, since 2010. His current research interests include Human Computer Interaction, UI/UX Design, Deep learning and Natural Language Processing (NLP).</p>
	<p><b>Anindya Nanda Rozana</b> obtained Bachelor Degree Informatics Engineering from Universitas Islam Negeri Sultan Syarif Kasim Riau in 2019. His current research interests include natural language processing and machine learning.</p>