Jurnal Informatika Universitas Pamulang **ISSN: 2541-1004**
Penerbit: Program Studi Teknik Informatika Universitas Pamulang **e-ISSN: 2622-4615**
Vol. 5, No. 3, September 2020 (356-363) **10.32493/informatika.v5i3.6571**

Sentimen Analisis Komentar *Toxic* pada Grup Facebook *Game Online* Menggunakan Klasifikasi Naïve Bayes

Renaldy Permana Sidiq[1], Budi Arif Dermawan[2], Yuyun Umaidah[3]

[1]Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang, Jl. HS. Ronggo Waluyo, Kec. Telukjambe Timur, Karawang, Indonesia, 41361

e-mail: [1]renaldy.16173@student.unsika.ac.id, [2]budi.arif@staff.unsika.ac.id, [3]yuyun.umaidah@staff.unsika.ac.id

Submitted Date: August 22nd, 2020 Reviewed Date: September 22nd, 2020
Revised Date: September 24th, 2020 Accepted Date: September 30th, 2020

Abstract

Toxic comments are comments made by social media users that contain expressions of hatred, condescension, threatening, and insulting. Social media users who are on average still teenagers with a nature that still cannot be controlled completely becomes a matter of great concern when they comment, their comments can be studied as text processing. Sentiment analysis can be used as a solution to identifying toxic comments by dividing them into two classifications. Where the data used amounted to 1,500 taken from social media Facebook in the private group Arena of Valor community. The dataset is divided into 2 classes: toxic and non-toxic. This research uses Naive Bayes with TF-IDF transformation and Information Gain feature selection and use distribution ratio 80:20. It will be compared the results of the evaluation where Naive Bayes without transformation, using TF-IDF transformation, and TF-IDF using Information Gain feature selection. The results of the comparison of evaluations from confusion matrix that have been carried out obtained the best classification model is to use the ratio of training and testing data 80:20 with TF-IDF transformation resulting in an accuracy of 75%, precision of 63%, recall of 67%, and F-measure of 64%.

Keywords: Toxic comments; TF-IDF; Information Gain; Sentiment Analysis; Naive Bayes

Abstrak

Komentar Toxic adalah komentar yang dilontarkan oleh pengguna media sosial yang berisi ungkapan kebencian, merendahkan, mengancam, dan menghina. Pengguna media sosial yang rata-rata masih remaja dengan sifat yang masih belum dapat dikontrol sepenuhnya menjadi hal yang sangat perlu diperhatikan ketika mereka berkomentar, komentar mereka dapat dikaji sebagai pemrosesan teks. Sentimen analisis dapat digunakan sebagai solusi mengidentifikasi komentar toxic dengan membaginya menjadi dua kelas klasifikasi. Dimana data yang digunakan berjumlah 1.500 yang diambil dari media sosial Facebook di grup private komunitas Arena of Valor. Dataset tersebut dibagi menjadi 2 kelas yaitu kelas toxic, dan non-toxic. Penelitian ini menggunakan Naive Bayes dengan transformasi TF-IDF dan seleksi fitur Information Gain serta penggunaan rasio pembagian data 80:20. Akan dibandingkan hasil dari evaluasi dimana Naive Bayes tanpa transformasi, menggunakan transformasi TF-IDF, dan TF-IDF menggunakan seleksi fitur Information Gain. Hasil perbandingan dari evaluasi yang telah dilakukan dengan confusion matrix didapatkan model klasifikasi terbaik ialah menggunakan rasio pembagian data training dan data testing 80:20 dengan transformasi TF-IDF menghasilkan akurasi sebesar sebesar 75%, precision sebesar 63%, recall sebesar 67%, dan F-measure sebesar 64%.

Kata kunci: Komentar toxic; TF-IDF; Information Gain; Sentimen Analisis; Naive Bayes

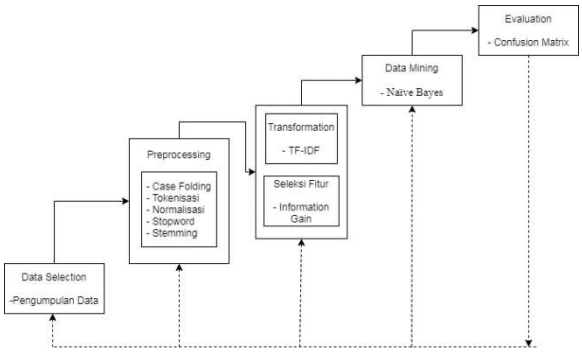
1 Pendahuluan

Penggunaan telepon genggam (*smartphone*) merupakan barang yang umum digunakan oleh umat manusia di seluruh dunia. Smartphone memiliki sistem operasi mobile yang lebih baik dari telepon genggam generasi sebelumnya, kemampuan komputasi yang lebih cepat, dan resolusi kamera yang lebih tinggi (Daeng et al., 2017). Perkembangan smartphone diiringi dengan penggunaan internet yang berkembang setiap tahunnya. Internet memberikan fasilitas yang memudahkan masyarakat dalam memperoleh informasi dari dalam maupun luar negeri dengan sangat cepat hingga hitungan detik dan biaya yang murah (Novianti & Riyanto, 2018). Berdasarkan hasil survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), pengguna internet di Indonesia bertambah 10,12% pada tahun 2018 dibandingkan tahun 2017 dengan total mencapai 171,17 juta pengguna dari total populasi sebanyak 264,16 juta jiwa (Rif'an, 2019). Berkaitan dengan pengguna internet yang semakin meningkat, diikuti oleh pengguna media sosial yang berkembang pula. Media sosial merupakan layanan dengan cakupan yang luas dimana didalamnya terdapat pertukaran informasi dan topik secara berkelanjutan (Schrape, 2016). Rentang usia remaja erdasarkan Badan Kependudukan dan Keluarga Berencana (BKKBN) ialah 10-24 tahun dan belum menikah. Analisis menunjukkan remaja yang menggunakan Facebook di Indonesia dengan rentang umur 13-17 tahun sebanyak 12,7% dan rentang umur 18-24 tahun sebanyak 30,7% (NapoleonCat.com, 2020). Dalam komentar pada sosial media, seringkali ditemukan sekelompok orang jahat dimana ia menghalangi diskusi yang saling menghormati dengan komentarnya yang beracun (*toxic comment*) yang didominasi oleh remaja. Komentar beracun *(toxic comment*) didefinisikan sebagai komentar yang kasar, tidak sopan, tidak masuk akal, atau bahkan sampai mempermalukan seseorang di media sosial yang cenderung membuat pengguna lain merasa tidak nyaman (Risch & Krestel, 2020). Dalam pendeteksian komentar dapat menggunakan pendekatan machine learning yaitu analisis sentimen yang sangat diperlukan dalam menyaring komentar di media sosial. Fauzi, Akbar & Asmawan (2019) menyebutkan bahwa dalam mengukur suatu sentimen komentar dari suatu media sosial memiliki kendala berupa dibutuhkan analisis yang dalam agar opini dapat diartikan. Sehingga diperlukan pendekatan machine learning yang dapat memisahkan komentar yang mengandung *toxic* dan tidak mengandung *toxic*.

Penelitian sebelumnya yang dilakukan oleh Sharma & Patel (2018) adalah melakukan analisis sentimen komentar beracun (*toxic comment*). Penelitian yang dilakukan untuk klasifikasi keberadaan berbagai komentar beracun pada platform online, seperti media sosial. Kekurangan dari penelitian tersebut ialah tidak adanya pembersihan komentar pada dataset sehingga menghasilkan klasifikasi yang kurang akurat dan kurang menjanjikan. Pembersihan dataset merupakan proses mengurangi atribut yang tidak cukup berpengaruh dalam proses klasifikasi. Berdasarkan penelitian yang dilakukan oleh Hakimi (2018), pembersihan dataset memang diperlukan sebelum proses klasifikasi agar informasi di dalam dataset dapat diproses. Proses pembersihan dataset dikenal dengan preprocessing, selain preprocessing dalam klasifikasi komentar dibutuhkan pula pembobotan kata dan algoritme klasifikasi. Dalam penelitian ini digunakan pembobotan kata TF-IDF dan algoritme Naive Bayes karena keduanya memiliki metode yang dalam implementasinya sederhana, memiiki performa cepat, dan efektif. Namun algoritme Naive Bayes memiliki permasalahan pada dimensi tinggi dari fitur, maka diperlukan seleksi fitur dimana akan menyeleksi fitur yang diperlukan untuk proses klasifikasi. Penelitian yang dilakukan oleh Negara, Muhardi, & Putri (2020), mengatasi dimensi yang tinggi dari fitur pada algoritme Naive Bayes dimana digunakan seleksi fitur Information Gain yang dapat digunakan untuk klasifikasi sentimen analisis. Berdasarkan hasil penelitian sebelumnya, penelitian ini akan melakukan sentimen analisis pada komentar beracun (*toxic comment*) pada grup komunitas di Facebook menggunakan algoritme Naive Bayes dengan menggunakan seleksi fitur TF-IDF dan Information Gain.

ISSN: 2541-1004
e-ISSN: 2622-4615
10.32493/informatika.v5i3.6571

2 Metodologi Penelitian



Gambar 1 Metode KDD

Metode penelitian yang digunakan dalam penelitian ini adalah dengan *Knowledge Discovery in Database* (KDD). Terdapat 5 tahap dalam proses KDD, yaitu:

1. *Data Selection*
   Data selection atau seleksi data merupakan proses untuk menyeleksi data yang ada agar dapat digunakan sebelum tahap penggalian informasi dalam *Knowledge Discovery in Database* (KDD) dimulai.
2. *Preprocessing*
   Proses *preprocessing* merupakan proses dimana *data* yang memiliki duplikasi akan dihapus, selain itu proses ini akan memeriksa *data* yang inkonsisten dan *data* yang ada akan diperbaiki apabila memiliki kesalahan tulisan.
3. *Transformation*
   Selanjutnya *transformation* atau transformasi dimana dilakukan proses tranformasi pada data yang telah dipilih menjadi bentuk vector agar dapat dlilakukan proses data mining.
4. *Data Mining*
   Data mining merupakan proses untuk mencari informasi yang menarik didalam *data* terpilih dengan menggunakan algoritme tertentu. Pemilihan algoritme yang tepat akan sangat bergantung pada tujuan dan proses dari metodologi penelitian secara keseluruhan.
5. *Evalution*
   Evalution atau evaluasi merupakan proses untuk menampilkan pengetahuan atau informasi yang dihasilkan dari proses klasifikasi yang ada pada *data mining*, dimana dalam menampilkan pengetahuan tersebut perlunya penyajian yang sederhana, menarik, namun mudah untuk dipahami.

3 Hasil dan Pembahasan

Hasil penelitian yang telah dilakukan ialah bagaimana melakukan sentimen analisis yang merupakan klasifikasi dari komentar beracun (*toxic comment*) di grup Facebook komunitas Arena of Valor (AOV). Sentimen analisis tersebut menggunakan algoritme Naive Bayes serta seleksi fitur TF-IDF dan *Information Gain*.

3.1 *Data selection* (pemilihan data)

Data komentar dari Facebook pada grup komunitas AOV telah diambil melalui proses *scrape* maka selanjutnya masuk ke tahap *data selection*. Tahap *data selection* adalah tahap dimana dilakukannya pelabelan secara *manual* lalu dilakukan verifikasi oleh seorang ahli bahasa yang merupakan dosen Bahasa Indonesia. Tabel 1 menunjukkan rincian dari komentar Facebook yang telah diverifikasi oleh ahli bahasa.

Tabel 1 Jumlah Komentar yang Telah Diverifikasi

Jenis Komentar	Jumah Komentar
Non-toxic	1237
Toxic	263
Total	1500

Tabel 1 menunjukkan jumlah komentar Facebook yang telah diseleksi oleh ahli bahasa yaitu 1500 komentar, terdiri dari 1237 komentar *non-toxic* dan 263 komentar *toxic*.

3.2 Preprocessing

Setelah dilakukan seleksi *data* dan verifikasi oleh ahli bahasa maka selanjutnya akan dilakukan *preprocessing*. *Processing* merupakan tahapan yang digunakan untuk menghilangkan kata yang tidak diperlukan atau dianggap tidak penting dalam proses klasifikasi. Terdapat 5 tahap dalam proses *preprocessing*, yaitu:

1. *Case Folding*

```
In [83]:  1  data.com_cleaning[8:15]

Out[83]: 8    bug fragment  solusi nya reinstal garena emang...
         9    jangan menyerah walau tower udah runtuh semua ...
         10   baru sebulan pkl jarang bgt buka nih gim  malo...
         11   kalian mau tahu apa yang lebih lucu dari angka...
         12   bang kalo ngepick itu liat musuh dulu dong ban...
         13   halo yg punya ini akun tolong amankan akun lu ...
         14   kenapa lah todolist  maloch  jarang dipake  p...
```

Gambar 1 Hasil *Case Folding*

Berdasarkan Gambar 1 seluruh huruf kapital di data komentar sudah tidak ada. Huruf kapital tersebut diubah menjadi huruf kecil, hal tersebut dilakukan agar dalam proses pembacaan dalam mesin terhadap corpus akan lebih mudah serta tidak memakan waktu yang banyak.

2. Tokenisasi

Telah dilakukan tokenisasi seperti yang ditunjukkan di Gambar 2. Data komentar yang sebelumnya berbentuk kalimat sekarang dipecah menjadi bentuk kata perkata. Tokenisasi ini dilakukan agar memudahkan dalam tahap transformasi sehingga dalam prosesnya tidak memproses berdasarkan kalimat tapi memproses kata demi kata.

```
In [10]: 1 data.com_tokenisasi[8:15]

Out[10]: 8     [bug, fragment, solusi, nya, reinstal, garena,...
         9     [jangan, menyerah, walau, tower, udah, runtuh,...
         10    [baru, sebulan, pkl, jarang, bgt, buka, nih, g...
         11    [kalian, mau, tahu, apa, yang, lebih, lucu, da...
         12    [bang, kalo, ngepick, itu, liat, musuh, dulu, ...
         13    [halo, yg, punya, ini, akun, tolong, amankan, ...
         14    [kenapa, lah, todolist, maloch, jarang, dipake...
```

Gambar 2 Hasil Tokenisasi

3. Normalisasi

```
In [61]: 1 data.com_normalisasi[8:18]

Out[61]: 8     [bug, fragment, solusi, nya, reinstal, garena,...
         9     [jangan, menyerah, walau, tower, sudah, runtuh...
         10    [baru, sebulan, pkl, jarang, banget, buka, nih...
         11    [kalian, ingin, tahu, apa, yang, lebih, lucu, ...
         12    [abang, kalau, ngepick, itu, lihat, musuh, dah...
         13    [halo, yang, punya, ini, akun, tolong, amankan...
         14    [kenapa, lah, todolist, maloch, jarang, dipake...
         15    [coi, ini, kenapa, ya, setiap, ingin, masuk, m...
         16                  [saya, akan, cary, kamu, kontol]
         17              [saran, lur, beli, budy, atau, binx]
```

Gambar 3 Hasil Normalisasi

Data komentar sudah diubah menjadi kata baku berdasarkan KBBI (Kamus Besar Bahasa Indonesia) seperti yang ditunjukkan pada Gambar 3. Pengubahan kata tidak baku menjadi kata baku tersebut bertujuan agar kata yang diproses memiliki makna seperti yang ada dalam KBBI.

4. *Stopword*

```
In [64]: 1 data.com_stopword[8:18]

Out[64]: 8     [bug, fragment, solusi, reinstal, garena, eman...
         9                   [menyerah, tower, runtuh, old]
         10    [sebulan, pkl, jarang, banget, buka, nih, gim,...
         11    [lucu, angka, obs, miskin, wajar, omen, miskin...
         12    [abang, ngepick, musuh, abang, mentang, top, s...
         13    [akun, tolong, amankan, akun, penyusup, barusa...
         14    [todolist, maloch, jarang, dipake, bet, sup, d...
         15    [coi, match, delay, gitu, jaringan, bagus, hps...
         16                                  [cary, kontol]
         17                  [saran, lur, beli, budy, binx]
```

Gambar 4 Hasil *Stopword*

Kata sambung seperti “yang”, “dan”, “di”, “dari” sudah tidak ada ketika proses *stopword* dilakukan, terlihat di Gambar 4 kata-kata tersebut tidak ada disana. Proses *stopword* ini merupakan proses untuk menyaring kata-kata yang tidak diperlukan dalam klasifikasi, seperti kata penghubung “yang”, “dan”, “di”, dan lain sebagainya. Dalam proses *stopword* ini menggunakan bantuan *library* Sastrawi yang ada pada bahasa pemogramman Python.

5. *Stemming*

Gambar 5 menunjukkan hasil dari *stemming* telah berhasil menghilangkan imbuhan pada kata. *Stemming* merupakan tahapan akhir dalam *preprocessing* yang akan menghilangkan semua imbuhan yang ada pada awal, akhir, dan kombinasi dari keduanya pada kata atau kalimat. Proses *stemming* ini juga menggunakan bantuan *library* Sastrawi.

```
In [23]: 1 data.com_stemming[8:18]

Out[23]: 8     [bug, fragment, solusi, reinstal, garena, eman...
         9                     [serah, tower, runtuh, old]
         10    [bulan, pkl, jarang, banget, buka, nih, gim, m...
         11    [lucu, angka, obs, miskin, wajar, omen, miskin...
         12    [abang, ngepick, musuh, abang, mentang, top, s...
         13    [akun, tolong, aman, akun, susup, barusan, cek...
         14    [todolist, maloch, jarang, dipake, bet, sup, d...
         15    [coi, match, delay, gitu, jaring, bagus, hpskr...
         16                                  [cary, kontol]
         17                  [saran, lur, beli, budy, binx]
```

Gambar 5 Hasil *Stemming*

3.3 *Transformation* (Transformasi)

Show TFIDF sample ke-206

saran upi ladenin menang telak lapor biar admin urus

	TF	IDF	TF-IDF	Term
array position 4	0.142857	6.010635	0.858662	admin
array position 124	0.142857	4.649659	0.664237	biar
array position 471	0.142857	7.620073	1.088582	ladenin
array position 478	0.142857	6.010635	0.858662	lapor
array position 581	0.142857	6.233779	0.890540	menang
array position 799	0.142857	6.115996	0.873714	saran
array position 946	0.142857	7.620073	1.088582	urus

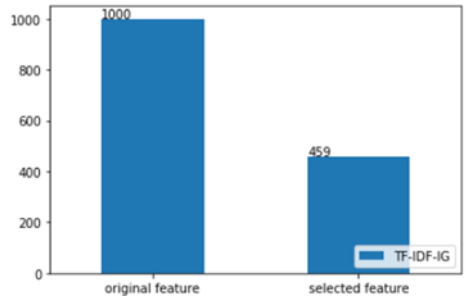
Gambar 6 Hasil Penerapan TF-IDF

Seperti yang ditunjukkan Gambar 6 terlihat kalimat nilai dari TF-IDF pada kalimat tersebut pada setiap *term* atau kata. Pada Gambar 6 dilakukan penentuan nilai dari TF (*Term Frequency)* terlebih dahulu, lalu dilakukan penentuan nilai dari IDF (*Inverse Document Frequency*), dan terakhir mengalikan hasil dari nilai TF dan IDF.

	Term	Jumlah TF-IDF	TF-IDF-IG
717	peduli	0.544291	0.000000
387	judul	0.544291	0.000000
786	rotasi	0.593801	0.120122
713	parman	0.639478	0.078458
953	ulang	0.639478	0.092739
417	kelongaran	0.639478	0.040927
788	rule	0.653149	0.165346
937	toro	0.680364	0.023519
734	pikir	0.692768	0.100412
66	auto	0.692768	0.017224

Gambar 7 Hasil Penerapan TF-IDF_IG

Terlihat di Gambar 7 terdapat hasil dari penerapan seleksi fitur Information Gain ke TF-IDF dimana penerapan tersebut mereduksi fitur atau tidak memberikan bobot terhadap kata-kata yang tidak saling terhubung, terlihat pada Gambar 7 adanya kata yang tidak memiliki bobot. Hasil dari seleksi fitur ini pun dapat terlihat jumlah fitur yang telah tereduksi yang ditunjukkan pada grafik di Gambar 8.



Gambar 8 Grafik Pengurangan Fitur TF-IDF-IG

Grafik di Gambar 8 menunjukkan jumlah reduksi fitur dari penerapan Information Gain pada TF-IDF. Pada penerapan tersebut dapat terlihat di Gambar 8 yang sebelumnya fitur berjumlah 1000 lalu ketika diterapkan Information Gain jumlahnya menjadi 459, hal tersebut menunjukkan sebanyak 541 fitur telah direduksi oleh Information Gain.

3.4 *Data Mining*

Transformasi telah dilakukan ditahap sebelumnya, lalu kini akan dilakukan tahap *data mining*. Tahap *data mining* merupakan tahap klasifikasi data menggunakan Naive Bayes akan ditampilkan dalam bentuk tabel *Confusion Matrix*. Penelitian ini akan menggunakan rasio pembagian *data* 80:20 yang merupakan implementasi dari hokum pareto, lalu akan dibandingkan model yang tidak menerapkan TF-IDF, menerapkan TF-IDF, dan menerapkan Information Gain pada TF-IDF.

Tabel 2 Hasil Klasifikasi Rasio 80:20 tanpa TF-IDF

Prediksi	*Actual*	
	Toxic	*Non-toxic*
Toxic	35	20
Non-toxic	61	184

Berdasarkan Tabel 2 terlihat terdapat 35 *data* dalam kelas *toxic* diprediksi benar sebagai kelas *toxic*, sedangkan 61 *data* kelas *toxic* diprediksi bukan kelas *toxic*. Lalu sebanyak 184 *data* kelas *non-toxic* benar diprediksi sebagai kelas *non-toxic*, sedangkan 20 *data* kelas *non-toxic* diprediksi sebagai kelas *toxic*.

Tabel 3 Hasil Klasifikasi Rasio 80:20 dengan TF-IDF

Prediksi	*Actual*	
	Toxic	*Non-toxic*
Toxic	30	25
Non-toxic	50	195

Berdasarkan Tabel 3 terlihat terdapat 30 data dalam kelas *toxic* diprediksi benar sebagai kelas *toxic*, sedangkan 50 *data* kelas *toxic* diprediksi bukan kelas *toxic*. Lalu sebanyak 195 *data* kelas *non-toxic* benar diprediksi sebagai kelas *non-toxic*, sedangkan 25 *data* kelas *non-toxic* diprediksi sebagai kelas *toxic*.

Tabel 4 Hasil Klasifikasi Rasio 80:20 dengan TF-IDF dan Information Gain

Prediksi	*Actual*	
	Toxic	*Non-toxic*
Toxic	32	23
Non-toxic	114	131

Berdasarkan Tabel 4 terlihat terdapat 32 *data* dalam kelas *toxic* diprediksi benar sebagai kelas *toxic*, sedangkan 114 *data* kelas *toxic* diprediksi bukan kelas *toxic*. Lalu sebanyak 131 *data* kelas *non-toxic* benar diprediksi sebagai kelas *non-toxic*, sedangkan 23 *data* kelas *non-toxic* diprediksi sebagai kelas *toxic*.

3.3 *Evaluation* (Evaluasi)

Setelah dilakukan seluruh pengujian maka akan dilakukan perbandingan untuk mencari model terbaik dalam proses klasifikasi. Gambar 9 menunjukkan hasil dari evaluasi dengan rasio 80:20 tanpa TF-IDF.

```
                    actual:toxic  actual:non-toxic
predicted:toxic               35                20
predicted:non-toxic           61               184

Accuracy: 0.73
Precision (mean): 0.6332720588235294
Recall (mean): 0.6936920222634508
F-measure (mean): 0.6415876340359002
```

Gambar 9 Evaluasi Klasifikasi Rasio 80:20 tanpa TF-IDF

Telah ditunjukkan di Gambar 9 bahwa hasil akurasi menggunakan Naive Bayes tanpa menerapkan TF-IDF menghasilkan akurasi sebesar 73%, *precision* sebesar 63%, *recall* sebesar 69%, dan *F-measure* sebesar 64%. Berdasarkan Gambar 9 terlihat terdapat 35 data dalam kelas *toxic* diprediksi benar sebagai kelas *toxic*, sedangkan 61 data kelas *toxic* diprediksi bukan kelas *toxic*. Lalu sebanyak 184 data kelas *non-toxic* benar diprediksi sebagai kelas *non-toxic*, sedangkan 20 data kelas *non-toxic* diprediksi sebagai kelas *toxic*. Hasil akurasi dengan menerapkan TF-IDF ditunjukkan di Gambar 10.

```
                    actual:toxic  actual:non-toxic
predicted:toxic               30                25
predicted:non-toxic           50               195

Accuracy: 0.75
Precision (mean): 0.6306818181818181
Recall (mean): 0.6706864564007421
F-measure (mean): 0.6415770609318996
```

Gambar 10 Evaluasi Klasifikasi Rasio 80:20 dengan TF-IDF

Gambar 10 telah menampilkan hasil akurasi dengan menerapkan TF-IDF dengan akurasi sebesar 75%, *precision* sebesar 63%, *recall* sebesar 67%, dan *F-measure* sebesar 64%. Berdasarkan Gambar 10 terlihat terdapat 30 data dalam kelas toxic diprediksi benar sebagai kelas toxic, sedangkan 50 data kelas toxic diprediksi bukan kelas toxic. Lalu sebanyak 195 data kelas non-toxic benar diprediksi sebagai kelas non-toxic, sedangkan 25 data kelas non-toxic diprediksi sebagai kelas toxic.

```
                    actual:toxic  actual:non-toxic
predicted:toxic               32                23
predicted:non-toxic          114               131

Accuracy: 0.5433333333333333
Precision (mean): 0.5349137164205657
Recall (mean): 0.5582560296846011
F-measure (mean): 0.487524782104515
```

Gambar 11 Evaluasi Klasifikasi Rasio 80:20 dengan TF-IDF dan Information Gain

Sedangkan untuk hasil akurasi penerapan TF-IDF dan Information Gain terdapat pada Gambar 11 dengan menghasilkan akurasi 54,3%, *precision* sebesar 53%, *recall* sebesar 55,8%, dan *F-measure* sebesar 48,7%. Gambar 11 terlihat terdapat terdapat 32 data dalam kelas toxic diprediksi benar sebagai kelas toxic, sedangkan 114 data kelas toxic diprediksi bukan kelas toxic. Lalu sebanyak 131 data kelas non-toxic benar diprediksi sebagai kelas non-toxic, sedangkan 23 data kelas non-toxic diprediksi sebagai kelas toxic.

3.4 Visualisasi

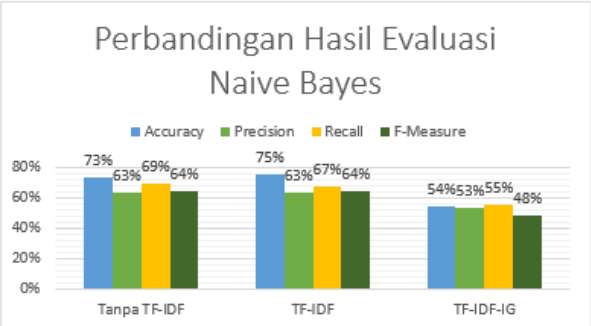
A. *Wordcloud*

Gambar 12 *Wordcloud* Kata Sering Muncul

Terlihat di Gambar 12 masih terdapat kata *toxic* seperti kata "ngentot" yang berukuran ukuran teksnya *medium* dan kata "anjing" yang ukuran teksnya kecil, hal tersebut dapat berarti masih munculnya beberapa komentar/kata yang bersifat *toxic*. Namun sentimen lebih cenderung ke arah *non-toxic* dikarenakan jumlah kata pada sentimen *non-toxic* yang muncul pada *wordcloud* di Gambar 12 lebih banyak dibandingkan sentimen *toxic* serta ukuran teksnya lebih besar.

ISSN: 2541-1004
e-ISSN: 2622-4615
10.32493/informatika.v5i3.6571

Hasil pengujian yang telah dilakukan dalam tahap *transformation* dan *data mining* menghasilkan hasil evaluasi yang akan dilihat perbandingannya agar mendapatkan model klasifikasi yang terbaik, perbandingan tersebut ditampilkan dengan grafik di Gambar 13.



Gambar 13 Grafik Perbandingan Hasil Evaluasi

Grafik di Gambar 13 menampilkan presentase dari keseluruhan hasil evaluasi terhadap hasil pengujian, dimana pada TF-IDF-IG memiliki presentase akurasi yang paling rendah sebesar 54%. Sedangkan presentase antara penggunaan TF-IDF dan tanpa TF-IDF berbeda sedikit, dimana dengan tanpa TF-IDF memiliki presentase akurasi sebesar 73% dan dengan menggunakan TF-IDF memiliki presentase akurasi sebesar 75%. Presentase terbesar dimiliki oleh penggunaan TF-IDF dengan akurasi sebesar 75%, didapatkan 30 kelas *toxic* yang diprediksi dengan benar dan 195 kelas *non-toxic* yang diprediksi dengan benar dari 300 *data testing*.

Selain akurasi dapat dilihat pula pada grafik di Gambar 4.19 terdapat nilai presentase *precision*, *recall*, dan *F-Measure*. *Precision* antara tanpa TF-IDF dan dengan TF-IDF memiliki persamaan, yaitu sebesar 63% sedangkan *precision* yang terendah berada di TF-IDF-IG sebesar 53%. Berdasarkan nilai *precision* dengan menggunakan TF-IDF didapatkan 30 kelas *toxic* yang diprediksi dengan benar dari 300 *data testing*. Berbeda dari *precision* dimana nilai presentase *recall* tidak memiliki persamaan antara tanpa TF-IDF dan menggunakan TF-IDF. Dengan nilai *precision* tertinggi sebesar 69% pada tanpa TF-IDF, lalu nilai presentase *precision* menggunakan TF-IDF sebesar 67%, dan nilai presentase *precision* terendah berada di TF-IDF-IG sebesar 55%. Berdasarkan nilai *recall* tertinggi dimana tanpa menggunakan TF-IDF didapatkan 35 kelas *toxic* yang diprediksi dengan benar dari 300 *data testing*. *F-measure* memiliki persamaan nilai presentase pada tanpa TF-IDF dan menggunakan TF-IDF sebesar 64%, sedangkan nilai presentasi *F-measure* terendah berada pada TF-IDF-IG sebesar 48%.

Dapat disimpulkan berdasarkan hasil perbandingan tersebut model yang terbaik untuk model klasifikasi Naïve Bayes pada rasio pembagian data 80:20 berada apada TF-IDF dengan tingkat presentase akurasi tertinggi sebesar 75% dengan *precision* sebesar 63%, *recall* sebesar 67%, dan *F-measure* sebesar 64%. Pemilihan model tersebut bukan hanya dilihat dari nilai akurasi yang tertinggi namun presentase dari *precision*, *recall*, dan *F-measure* yang baik dibandingkan model yang lainnya. TF-IDF pun tidak hanya menghitung bobot dari seringnya *term*/kata muncul namun juga menghitung bobot dari kata yang unik atau tidak sering muncul, hal tersebut yang membuat TF-IDF dalam pembobotannya *balance* atau seimbang dengan menghitung setiap *term*/kata pada *dataset* tidak terkecuali. Berbeda dengan pembobotan tanpa TF-IDF yang hanya memberikan bobot *term*/kata yang sering muncul dan pembobotan TF-IDF-IG yang hanya memberikan bobot untuk kata-kata tertentu saja.

4 Kesimpulan

Berdasarkan dari penelitian yang telah dilakukan kini dapat disimpulkan proses dari klasifikasi sentimen komentar *toxic* di grup komunitas AOV dari komentar yang divalidasi sejumlah 1500 dengan sentimen *non-toxic* sejumlah 1237 komentar, dan 263 sentimen *toxic*. Menghasilkan sentimen cenderung ke *non-*toxic dikarenakan jumlah kata *non-toxic* yang sering muncul pada *wordcloud* serta ukuran kata *non-toxic* yang lebih besar dibandingkan kata *toxic*. Pengujian terhadap model klasifikasi dengan rasio pembagian *data* 80:20 dimana model tanpa TF-IDF, menggunakan TF-IDF, dan penggunaan Information Gain pada TF-IDF menghasilkan nilai akurasi tertinggi berada pada penggunaan TF-IDF dengan akurasi sebesar 75%, *precision* sebesar 63%, *recall* sebesar 67%, dan *F-measure* sebesar 64%. Hasil tersebut dikarenakan pada TF-IDF tidak hanya memberikan bobot pada kata yang sering muncul namun kata yang unik atau tidak sering muncul pada *dataset*, hal inilah yang membuat pembobotan TF-IDF menjadi *balance* atau seimbang. Dengan hasil tersebut dapat disimpulkan penggunaan seleksi fitur belum tentu dapat mempengaruhi kenaikan akurasi klasifikasi. Tidak hanya memilih seleksi fitur saja namun jumlah *dataset* yang lebih banyak dan proses

preprocessing yang baik dapat membantu kenaikan akurasi klasifikasi.

5 Saran

Berdasarkan hasil penelitian kesimpulan yang diperoleh dalam penelitian ini masih dapat dikembangkan dalam segi model klasifikasi, maka peneliti menyarankan, yaitu sebagai berikut:

1) Penelitian selanjutnya dapat menambahkan jumlah dataset yang lebih banyak dibandingkan penelitian ini.
2) Penelitian selanjutnya dapat menerapkan pembobotan dan seleksi fitur yang lain, seperti Chi-Square, Particle Swarm Optimazion, Stable Mutation Jump Strategy, dan lain sebagainya.

Referensi

Brassard-Gourdeau, E., & Khoury, R. (2019). Subversive toxicity detection using sentiment information. 1–10. https://doi.org/10.18653/v1/w19-3501\

Fauzi, A., Akbar, M. F., & Asmawan, Y. F. A. (2019). Sentimen Analisis Berinternet Pada Media Sosial dengan Menggunakan Algoritma Bayes. *Jurnal Informatika*, 6(1), 77–83. https://doi.org/10.31311/ji.v6i1.5437

Hilman, M., Nurjaman, A., & Mubarok, M. S. (2017). Analisis sentimen pada ulasan buku berbahasa Inggris Menggunakan Information Gain dan Support Vector Machine. *E-Proceeding of Engineering*: Vol.4, No.3 Desember 2017, 4(3), 4900–4906.

Ibrahim, M., Torki, M., & El-Makky, N. (2019). Imbalanced toxic comments classification using data augmentation and deep learning. *Proceedings-17th IEEE International Conference on Machine Learning and Applications*, ICMLA 2018, 875–878. https://doi.org/10.1109/ICMLA.2018.00141

Lu, C., Wang, D., Liu, X., & Gan, K. (2018). A mining and visualizing system for large-scale Chinese technical standards. P*roceedings - IEEE 4th International Conference on Big Data Computing Service and Applications, BigDataService* 2018, 1–8. https://doi.org/10.1109/BigDataService.2018.00010

Manning, C. D., Raghavan, P., & Schütze, H. (2009). An Introduction to Information Retrieval. Library Review, 53(9), 462–463. https://doi.org/10.1108/00242530410565256

Maron, M. E., & Kuhns, J. L. (1960). On relevance, probabilistic indexing and information retrieval. *Journal of the ACM* (JACM), 7(3), 216–244. https://doi.org/10.1145/321033.321035

Pradikdo, A. C., & Ristyawan, A. (2018). Model klasifikasi abstrak skripsi menggunakan text mining untuk pengkategorian skripsi sesuai bidang kajian. Simetris: *Jurnal Teknik Mesin, Elektro Dan Ilmu Komput*er, 9(2), 1091–1098.

Ridwansyah, & Aji, S. (2017). Sentimen etika posting pada media sosial menggunakan performa terbaik. *Information Management For Educators And Professionals*, 2(1), 67–76.

Risch, J., & Krestel, R. (2020). Toxic comment detection in online discussions. https://doi.org/10.1007/978-981-15-1216-2_4

Saeed, H. H., Shahzad, K., & Kamiran, F. (2019). Overlapping toxic sentiment classification using deep neural architectures. *IEEE International Conference on Data Mining Workshops*, ICDMW, 2018-Novem, 1361–1366. https://doi.org/10.1109/ICDMW.2018.00193

Sharma, R., & Patel, M. (2018). Toxic comment classification using neural networks and machine learning. *IARJSET*, 5(9), 47–52. https://doi.org/10.17148/iarjset.2018.597

Sudiantoro, A. V., & Zuliarso, E. (2018). Analisis sentimen twitter menggunakan text mining dengan algoritma Naïve Bayes Classifier. *Prosiding SINTAK* 2018, 398–401.

Suryani, N. P. S. M., Linawati, & Saputra, K. O. (2019). Penggunaan metode Naive Bayes Classifier pada analisis sentimen Facebook berbahasa Indonesia. *Majalah Ilmiah Teknologi Elektro*, 18(1), 145-148. https://doi.org/https://doi.org/10.24843/MITE.2019.v18i01.P22