

Aplikasi Pendeteksi *Cyberbullying* Terhadap Komentar Postingan Media Sosial Instagram dengan Metode *Naïve Bayes Classifier* Berbasis Website

Alfan Syahada Hutagalung^{a1}, Arif Bijaksana Putra Negara^{a2}, Enda Esyudha Pratama^{a3}

^aProgram Studi Informatika, Fakultas Teknik, Universitas Tanjungpura

Jl. Prof. Dr. H. Hadari Nawawi, Pontianak 78124

¹alfansh@student.untan.ac.id

²arifbpn@informatika.untan.ac.id

³enda@informatika.untan.ac.id

Abstrak

Berdasarkan survey data perusahaan analisis Sosial Media Marketing yang berbasis di Warsawa, Polandia, NapoleonCat mencatat jumlah pengguna Instagram di Indonesia tahun 2019 mencapai 61.610.000. Instagram memungkinkan penggunanya untuk berbagi momen lewat foto dan video. Instagram juga memiliki fitur untuk mengomentari postingan. Namun dengan adanya fitur tersebut dapat memberikan dampak yang positif dan negatif bagi masyarakat. Salah satu dampak negatifnya adalah cyberbullying. Pada kenyataannya sosial media instagram tidak dapat menyaring komentar sentimen negatif yang mengandung unsur cyberbullying. Dampak dari tindakan cyberbullying dapat mengakibatkan depresi berat hingga bunuh diri. Karenanya, aplikasi pendeteksi cyberbullying di Instagram adalah solusi untuk menemukan komentar yang mengandung kata atau kalimat negatif. Aplikasi yang telah dibuat menerapkan sebuah pemodelan text mining dengan mengklasifikasikan komentar ke dalam sentimen negatif (bullying dan bukan bullying) dan sentimen positif. Pemodelan text mining yang diterapkan menggunakan metode Naïve Bayes. Dengan adanya Aplikasi Pendeteksi Cyberbullying untuk mendeteksi komentar negatif dengan menggunakan metode Naïve Bayes Classifier diharapkan dapat membantu pengguna instagram dalam mendeteksi dan menghapus komentar negatif. Hasil pengujian confusion matrix menunjukkan nilai akurasi sebesar 98,5%, nilai kerusakan (error rate) sebesar 1,5%, nilai presisi sebesar 96%, dan nilai recall sebesar 96%.

Kata kunci: *Cyberbullying*, Instagram, *Naïve Bayes Classifier*

Cyberbullying Detection Application for Social Media Instagram Post Comments with *Naïve Bayes Classifier* Method Website Based

Abstract

Based on a survey of data from a Social Media Marketing analyst company based in Warsaw, Poland, NapoleonCat noted that the number of Instagram users in Indonesia in 2019 reached 61,610,000. Instagram allows users to share moments through photos and videos. Instagram also has a feature for commenting on posts. However, these features can have both positive and negative impacts on society. One of the negative effects is cyberbullying. In fact, Instagram social media cannot filter negative sentiment comments that contain elements of cyberbullying. The impact of cyberbullying can lead to severe depression and suicide. Therefore, the cyberbullying detection application on Instagram is a solution for finding comments that contain negative words or sentences. The application that has been made applies a text mining modeling by classifying comments into negative sentiments (bullying and not bullying) and positive sentiments. The applied text mining modeling uses the Naïve Bayes method. With the Cyberbullying Detection Application to detect negative comments using the Naïve Bayes Classifier method, it is hoped that it can help Instagram users detect and delete negative comments. The results of the Confusion matrix test showed an accuracy value of 98.5%, an error rate of 1.5%, a precision value of 96%, and a recall value of 96%.

I. PENDAHULUAN

Di era teknologi saat ini, perkembangan teknologi informasi dan komunikasi sangatlah pesat. Hal ini dibuktikan dengan mudahnya dalam mengakses internet. Salah satu fenomena dari perkembangan teknologi informasi dan komunikasi adalah penggunaan internet sebagai media komunikasi.

Berdasarkan survei data Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) terakhir di tahun 2018, pengguna internet di Indonesia mencapai 171,17 juta. Dari segmen umur, ternyata remaja dari usia 15-19 tahun mempunyai penetrasi paling tinggi hingga mencapai 91%. Para remaja mengakses internet untuk memperoleh informasi, menambah pengetahuan, dan memanfaatkan internet sebagai media komunikasi dan media menambah teman baru dengan menggunakan media sosial [1].

Salah satu media sosial yang sangat populer saat ini adalah Instagram. Berdasarkan survey data perusahaan analisis Sosial Media Marketing yang berbasis di Warsawa, Polandia, NapoleonCat mencatat jumlah pengguna Instagram di Indonesia tahun 2019 mencapai 61.610.000 juta atau setara dengan 35,9% dari pengguna internet di Indonesia, bahkan jumlah pengguna Instagram meningkat setiap tahunnya[2].

Instagram memungkinkan penggunaannya untuk berbagi momen lewat foto dan video. Momen yang dibagikan biasa disebut postingan. Bukan hanya untuk berbagi, Instagram juga memiliki fitur untuk mengomentari postingan. Namun dengan adanya fitur tersebut dapat memberikan dampak yang positif dan negatif bagi masyarakat. Salah satu dampak negatifnya adalah *Cyberbullying*.

Cyberbullying merupakan kejadian manakala seseorang diejek, dihina, diintimidasi, dan dipermalukan oleh orang lain melalui internet, teknologi digital, atau telepon seluler. *Cyberbullying* yang dimaksud dalam hal ini mencakup komentar sentimen negatif yang mengandung unsur *cyberbullying* pada postingan tertentu dengan pesan personal tak bersahabat di media sosial instagram. *Cyberbullying* dapat berdampak pada mental korbannya, cenderung mengalami depresi sedang hingga berat, emosional, dan merasa tidak percaya diri yang dapat mengakibatkan depresi, percobaan pembunuhan bahkan bunuh diri [3].

Jika pendeteksi tindakan *cyberbullying* dilakukan secara konvensional maka akan memakan waktu yang lama karena harus mendeteksi ribuan komentar. Maka diperlukan suatu aplikasi yang dapat membantu masyarakat dalam mendeteksi tindakan *cyberbullying* tersebut secara massal.

Berdasarkan penjabaran di atas solusi dari permasalahan tersebut adalah dibuatnya suatu "Aplikasi Pendeteksi *Cyberbullying* Terhadap Komentar Postingan Media Sosial Instagram Dengan Metode *Naive Bayes Classifier*". Instagram dipilih karena sosial media ini adalah salah satu sosial media teraktif di Indonesia. Instagram juga dipilih karena instagram menyediakan *Application Programming Interface* (API) yang

digunakan untuk mengambil data pengguna dan komentar [4].

Pada penulisan skripsi ini, peneliti menggali informasi dari penelitian-penelitian sebelumnya sebagai bahan perbandingan, baik mengenai metode yang digunakan dan kesimpulan yang ada. Selain itu, peneliti juga menggali dari buku, skripsi dan jurnal dalam rangka mendapatkan suatu informasi yang ada sebelumnya tentang teori tentang judul yang digunakan untuk memperoleh landasan teori ilmiah.

Penelitian tentang Aplikasi Pendeteksi *Cyberbullying* menggunakan metode *Naive Bayes Classifier* sebelumnya sudah pernah dilakukan oleh peneliti lain, salah satunya adalah Restio Afrinza pada tahun 2019, yang merupakan seorang mahasiswa Program Studi Ilmu Komputasi Fakultas Informatika Universitas Telkom Bandung dengan judul "Deteksi *Cyberbullying* Pada Twitter Menggunakan *Naive Bayes Classifier*". Peneliti menggunakan metode *Naive Bayes Classifier* untuk menentukan mengklasifikasikan data yang dikumpulkan dari *tweets* pada *twitter* menjadi dua berupa sentimen positif dan sentimen negatif. Aplikasi yang dibuat oleh peneliti berupa Aplikasi Pendeteksi *Cyberbullying* berbasis website yang dibuat menggunakan bahasa pemrograman PHP, HTML, CSS, dan MySQL. Penggunaan aplikasi tersebut terbatas untuk 9750 data latih dan 3250 data uji. Untuk menentukan akurasi dari algoritma *Naive Bayes*, yaitu menggunakan metode confusion matrix. Hasil penelitian yang dibuat oleh peneliti menghasilkan sistem yang dapat mengklasifikasikan sentimen dengan hasil pengujian untuk 3250 tweet data uji yaitu accuracy 82,12%, precision 85,28%, recall 95,22% dan error rate 15,70% [5].

Penelitian lain yang serupa dilakukan oleh Ika Yunida Angraini pada tahun 2018, peneliti merupakan seorang mahasiswi Program Studi Sistem Informasi Fakultas Teknik Universitas Nusantara PGRI Kediri dengan judul "Pemodelan Deteksi *Cyberbullying* Pada Jejaring Sosial Twitter". Peneliti membuat sebuah pemodelan *text mining* dengan mengklasifikasikan *tweets* pada jejaring sosial Twitter kedalam kelas *bullying* dan kelas bukan *bullying* menggunakan metode *Naive Bayes Classifier*. Peneliti menggunakan 5 tahap dalam proses *preprocessing* yaitu *replace tokens*, *transform case*, *tokenization*, *filter stopwords*, dan *n-grams*. Untuk memvalidasi model peneliti melakukan proses 10-Fold Validation dan untuk menguji kinerja model pada data real maka peneliti melakukan proses testing. Pemodelan tersebut pada proses validasi menggunakan 10-Fold Validation memiliki hasil accuracy sebesar 82,50% dan standard deviation kurang lebih 5,12%. Sementara pada proses testing model bekerja dengan baik dengan accuracy sebesar 90%. Berdasarkan hasil penelitian yang dilakukan peneliti, maka dapat ditarik kesimpulan bahwa deteksi cyber bullying pada jejaring sosial Twitter dapat dilakukan dengan beberapa teknik. Pertama, data berupa *tweets* dikumpulkan melalui proses *data crawling*. Kedua, dilakukan proses seleksi data, *cleaning* data dan *preprocessing* untuk mempersiapkan data pada proses

mining. Ketiga, dilakukan klasifikasi menggunakan *Naive Bayes Classifier* [6].

Selain itu ada juga penelitian serupa yang berhubungan dengan Aplikasi Pendeteksi Cyberbullying menggunakan logika *Naive Bayes Classifier*. Penelitian ini dilakukan oleh Alfa Saleh pada tahun 2014 yang merupakan seorang mahasiswa Teknik Informatika STMIK Potensi Utama Medan dengan judul “Klasifikasi Metode Naive Bayes Dalam Data Mining Untuk Menentukan Konsentrasi Siswa (Studi Kasus di MAS PAB 2 Medan)”. Peneliti menerapkan metode *Naive Bayes Classifier* untuk menentukan konsentrasi siswa berdasarkan data latih dengan kriteria jenis kelamin, nilai mata pelajaran, angket, rekomendasi dan konsentrasi. Hasil penelitian yang dibuat oleh peneliti untuk mengetahui konsentrasi siswa dengan metode Naive Bayes menunjukkan keakuratan hasil klasifikasi yang diperoleh sesuai dengan data yang diujikan sebesar 90,8333%, dimana data 120 siswa yang diuji, terdapat sebanyak 109 siswa data berhasil diklasifikasikan dan diprediksi dengan benar. Berdasarkan hasil penelitian yang dilakukan peneliti, maka dapat ditarik kesimpulan bahwa metode *Naive Bayes Classifier* dapat membantu proses *data mining* untuk menentukan konsentrasi siswa. Pertama, data berupa *tweets* dikumpulkan melalui proses *data crawling*. Metode *Naive Bayes Classifier* memanfaatkan data latih untuk menghasilkan probabilitas setiap kriteria untuk *class* yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi konsentrasi siswa [7].

Berdasarkan beberapa penelitian tersebut, penelitian yang akan dilakukan yaitu berfokus pada aplikasi pendeteksi komentar *cyberbullying* di postingan media sosial Instagram. Tujuan dari penelitian ini adalah menghasilkan suatu aplikasi yang dapat mendeteksi komentar negatif yang mengandung unsur *cyberbullying*.

Penelitian ini menjadi penting karena dapat menjadi solusi dari permasalahan *Cyberbullying* Terhadap Komentar Postingan Media Sosial Instagram. Berdasarkan latarbelakang yang diuraikan maka perlu dibangun sebuah Aplikasi Pendeteksi *Cyberbullying* Terhadap Komentar Postingan Media Sosial Instagram Dengan Metode *Naive Bayes Classifier* ini diharapkan dapat mempermudah dalam mendeteksi dan menghapus komentar *cyberbullying*.

II. METODOLOGI PENELITIAN

A. Metode Penelitian

Metode atau langkah-langkah penelitian yang akan dilakukan adalah sebagai berikut:

1. Analisis Kebutuhan

Pada tahapan ini dilakukan analisis mengenai kebutuhan fungsional dan kebutuhan non fungsional aplikasi.

2. Pengumpulan Data

Pengumpulan data dilakukan dengan teknik studi literatur dan *web scrapping*.

3. Perancangan Aplikasi

Perancangan aplikasi terdiri dari perancangan arsitektur aplikasi, perancangan DFD (*Data Flow Diagram*), perancangan basis data, *flowchart* aplikasi, preprocessing, perhitungan *naive bayes classifier* dan perancangan antarmuka aplikasi.

4. Pembuatan Aplikasi

Pembuatan aplikasi merupakan tahap pembuatan aplikasi berdasarkan hasil perancangan aplikasi.

5. Pengujian Aplikasi

Pengujian dilakukan untuk mengetahui kelayakan aplikasi yang telah dibuat terhadap tujuan pembuatan aplikasi. Pengujian menggunakan metode *confussion matrix* untuk menguji fungsionalitas aplikasi.

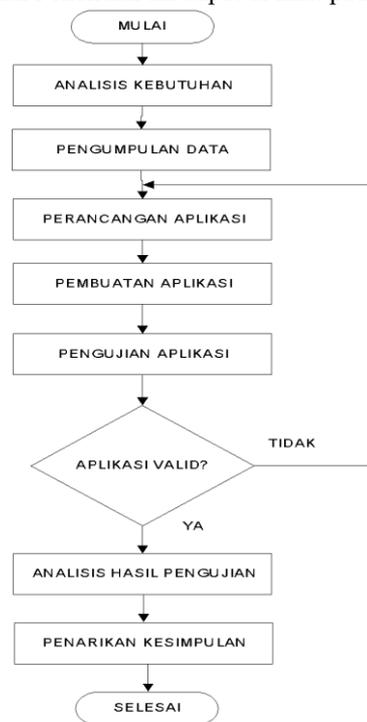
6. Analisis Hasil Pengujian

Analisis terhadap hasil pengujian dilakukan untuk mengetahui karakteristik aplikasi yang telah dikembangkan.

7. Penarikan Kesimpulan

Kesimpulan dirumuskan berdasarkan pengujian yang telah dilakukan.

. Langkah Penelitian ini dapat di lihat pada gambar 1.



Gambar 1. Diagram Alir Penelitian

B. *Naive Bayes Classifier*

Naive Bayes Classifier merupakan suatu algoritma yang digunakan untuk mencari nilai probabilitas tertinggi untuk mengklasifikasikan data uji pada kategori yang paling tepat. Pada penelitian ini yang menjadi data uji adalah dokumen komentar postingan sosial media Instagram. Ada dua tingkat klasifikasi terhadap komentar postingan media sosial Instagram. Klasifikasi pertama, aplikasi mengklasifikasikan kriteria sentimen positif dan negatif. Jika klasifikasi pertama menunjukkan komentar

adalah sentimen positif, maka tidak dilanjutkan untuk ke tingkat selanjutnya. Jika klasifikasi pertama menunjukkan komentar adalah sentimen negatif, maka dilanjutkan ke tingkat klasifikasi yang kedua, yaitu klasifikasi kelas *bullying* dan kelas bukan *bullying* [8]. Kelebihan dari Naïve Bayes adalah proses klasifikasi data dapat disesuaikan dengan sifat dan kebutuhan masing-masing [9].

Dalam algoritma *Naïve Bayes Classifier* setiap dokumen akan direpresentasikan dengan pasangan atribut “x1, x2, x3, ... xn” dimana x1 adalah kata pertama, x2 adalah kata kedua dan seterusnya. Kemudian himpunan kategori komentar direpresentasikan dengan atribut V. Pada saat klasifikasi, algoritma akan mencari nilai probabilitas kemunculan tertinggi dari semua kategori yang diujikan. Nilai probabilitas ini akan direpresentasikan dengan atribut VMAP, maka persamaan 2.1 adalah sebagai berikut :

$$V_{MAP} = \underset{V_j \in V}{\text{arg max}} \frac{P(x_1, x_2, x_3, \dots, x_n | V_j) P(V_j)}{P(x_1, x_2, x_3, \dots, x_n)} \quad (2.1)$$

C. Alat Bantu Perancangan Sistem

Perancangan dilakukan dengan membuat konsep perancangan berupa rancangan awal aplikasi yaitu dengan *Flowchart*, *Data Flow Diagram*, dan *Entity Relationship Diagram*.

Flowchart adalah suatu bagan dengan simbol-simbol tertentu yang menggambarkan urutan proses secara mendetail dan hubungan antara suatu proses (instruksi) dengan proses lainnya dalam suatu program. [10].

Data flow diagram (DFD) adalah diagram yang menggunakan notasi simbol untuk menggambarkan arus data sistem [11].

Entity Relationship Diagram (ERD) adalah model teknik pendekatan yang menyatakan atau menggambarkan hubungan suatu model [12].

D. MySQL

MySQL adalah *relational database management system* (RDBMS) yang didistribusikan secara gratis dibawah licensi GPL (*General Public License*) [13]. MySQL sebenarnya merupakan turunan salah satu konsep utama dalam database sejak lama yaitu SQL (*Structured Query Language*). SQL adalah sebuah konsep pengoperasian database terutama untuk pemilihan/Seleksi dan pemasukan data yng memungkinkan pengoperasian data dikerjakan dengan mudah dan secara otomatis.

Kehandalan suatu sistem basis data dapat diketahui dari cara kerja *optimizer*-nya dalam melakukan proses perintah-perintah SQL, yang dibuat oleh *user* maupun program-program aplikasinya. Sebagai *database server*, MySQL dapat dikatakan lebih unggul dibandingkan database server lainnya dalam query data. Hal ini terbukti untuk *query* yang dilakukan oleh *single user*, kecepatan *query* My SQL dapat sepuluh kali lebih cepat dari

PostgreSQL dan lima kali lebih cepat dibandingkan Interbase.

E. Pengujian Pengujian Confussion Matrix

Confusion matrix merupakan alat yang digunakan untuk menganalisa seberapa baik klasifikasi mengenali tuple dari kelas yang berbeda. *True Positif* dan *True Negatif* memberikan informasi ketika klasifikasi benar, sedangkan *False Positif* dan *False Negatif* memberikan informasi ketika klasifikasi salah [14]. Confusion matrix memuat informasi tentang klasifikasi yang diprediksi dengan benar oleh sebuah sistem klasifikasi [15].

Tahapan pengujian menggunakan confusion matrix dilakukan dengan beberapa parameter pengujian yakni *accuracy*, *recall*, dan *error rate*. Kemudian hasil dari pengujian tersebut selanjutnya akan dilakukan analisis lebih lanjut untuk dapat ditarik kesimpulan [16].

TABEL I

PENGUJIAN CONFUSION MATRIX

| | | |
|----------------------|----|-------|
| Klasifikasi / Aktual | YA | TIDAK |
| YA | TP | FN |
| TIDAK | FP | TN |

Tabel I. menunjukkan bahwa hasil klasifikasi benar sesuai dengan keputusan sistem (TP) dan Klasifikasi salah menurut sistem (FP). Sedangkan dokumen yang tidak termasuk dalam hasil klasifikasi kategori, adakalanya memang bukan anggota dari kategori itu (TN) dan adakalanya ternyata seharusnya menjadi anggota kategori tersebut (FN).

Keempat parameter pada Tabel I. digunakan untuk menghitung 4 metode evaluasi yakni :

1. Accuracy

Merupakan perbandingan jumlah seluruh dokumen relevan dengan jumlah seluruh dokumen. Accuracy memiliki rumusan sebagai berikut :

$$Accuracy = (TP + TN)$$

1. Recall

Merupakan perbandingan jumlah dokumen yang relevan terkenal dengan jumlah seluruh dokumen relevan. Recall memiliki rumusan sebagai berikut :

$$Recall = TP / (TP + FP)$$

2. Precision

Merupakan perbandingan jumlah dokumen yang relevan terkenal dengan jumlah dokumen yang terkenal. Precision memiliki persamaan sebagai berikut :

$$Precision = TP / (TP + FN)$$

3. Error Rate

Merupakan nilai benar (1) dikurangi dengan nilai akurasi. *Error rate* memiliki persamaan yang mewakili seluruh kinerja sistem yang merupakan penggabungan nilai Recall dan Presicion. *Error rate* memiliki persamaan sebagai berikut :

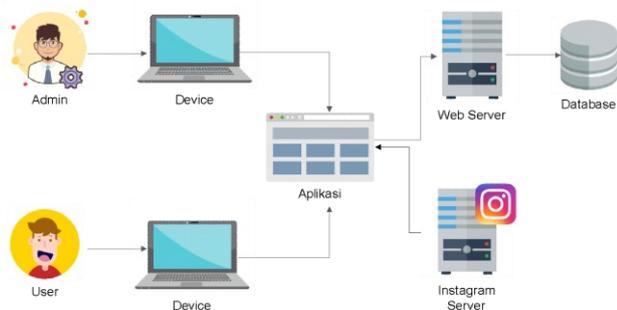
$$Error\ rate = 1 - Accuracy$$

| Nilai Asli | Nilai Prediksi | |
|-------------------------|----------------|---------------------|
| | True Bullying | True Bukan Bullying |
| Prediksi Bullying | 70 | 0 |
| Prediksi Bukan Bullying | 5 | 25 |

Nilai *accuracy*, *recall*, dan *precision* dinyatakan dalam persen. Semakin tinggi persentase ketiga nilai tersebut menunjukkan semakin baik kinerja sistem klasifikasi teks otomatis. Sedangkan untuk *error rate* berlaku sebaliknya.

F. Perancangan Arsitektur Sistem

Arsitektur sistem merupakan cara yang digunakan untuk menggambarkan sistem secara umum, gambaran arsitektur tersebut dibangun berdasarkan informasi yang didapatkan dari analisis kebutuhan dan pengumpulan data. Komponen aplikasi terdiri dari data dan informasi atau proses, seperti pada Gambar 2, aplikasi yang akan dibangun ini berbasis *web*, yang dapat diakses oleh pengguna instagram dan admin. Pada aplikasi yang dibuat, admin memanajemen data yang ada pada aplikasi melalui *web browser*, kemudian dikirim ke server dan disimpan dalam *database*. Pengguna atau user membuka aplikasi melalui *web browser* dan memilih postingan yang akan dideteksi atau diklasifikasi *cyberbullying*.



Gambar 2. Arsitektur Sistem

III. HASIL DAN PENGUJIAN

Pada bagian ini dibahas mengenai perhitungan dan pengujian dari aplikasi pendeteksi *cyberbullying* terhadap komentar instagram dengan metode *Naïve Bayes Classifier*. Pengujian perhitungan dilakukan menggunakan pengujian *confussion matrix*. Pengujian *confussion matrix* dilakukan untuk memeriksa apakah aplikasi telah berjalan dengan baik dan benar sesuai dengan tujuan yang diharapkan. *Confussion matrix* akan diujikan pada 100 data sampel komentar, yang terdiri dari 70 komentar bullying dan 30 komentar bukan bullying diuji dengan membandingkan dengan

data latih. Hasil pengujian *confussion matrix* disajikan dalam sebuah tabel seperti pada Tabel II.

TABEL II
HASIL PENGUJIAN *CONFUSION MATRIX*

A. Hasil Penelitian

Perancangan yang telah dilakukan terdiri dari perancangan arsitektur sistem, diagram alir sistem, dan basis data. Perancangan tersebut menghasilkan sebuah Aplikasi Pendeteksi *Cyberbullying* terhadap komentar postingan media sosial Instagram. Antarmuka dari sistem diantaranya adalah antarmuka admin dan user.

B. Antarmuka Login

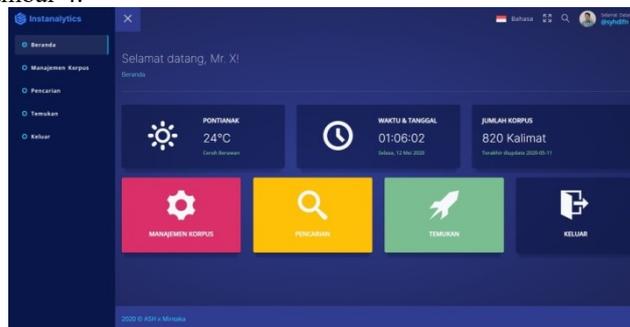
Halaman ini menampilkan *form login*. Pengguna atau admin akan mengisi kolom username, password, dan kemudian klik tombol *submit*. Tampilan halaman *login* tahunan dapat dilihat pada Gambar 3.



Gambar 3. Tampilan Antarmuka Login

C. Antarmuka Admin

Halaman beranda admin merupakan halaman yang berisi informasi cuaca, waktu, tanggal, jumlah korpus dan menu-menu yang terdiri dari menu manajemen korpus, menu pencarian, menu temukan, dan menu *logout* atau keluar. Tampilan lokasi beranda admin dapat dilihat pada Gambar 4.

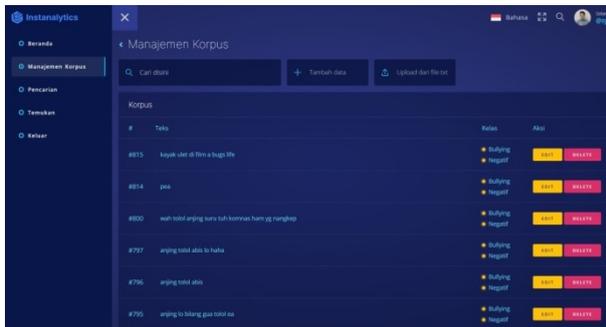


Gambar 4. Tampilan Antarmuka Admin

D. Antarmuka Manajemen Korpus

Halaman manajemen korpus merupakan halaman yang digunakan oleh admin untuk mengolah data korpus. Pada manajemen korpus admin dapat melakukan tambah, ubah dan hapus data korpus. Data korpus merupakan data yang menjadi acuan aplikasi untuk mendeteksi komentar

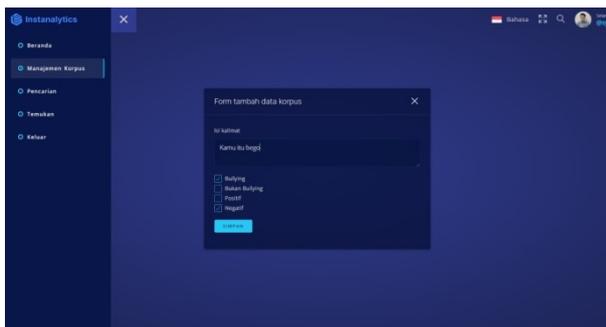
cyberbullying. Data korpus berisi kumpulan kalimat yang sudah dilabeli kelasnya yang terdiri dari positif, negatif, bullying, dan bukan bullying. Halaman manajemen korpus dapat dilihat pada Gambar 5.



Gambar 5. Tampilan Antarmuka Manajemen Korpus

E. Antarmuka Input Data Korpus

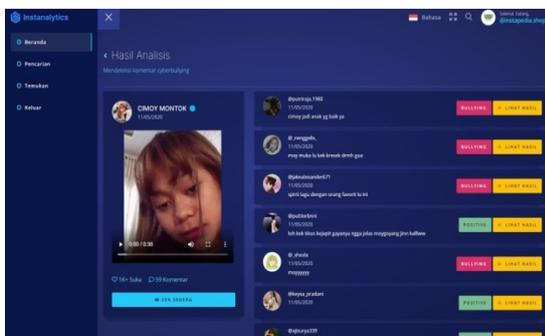
Tampilan tambah data korpus menampilkan form yang berisi kolom isi kalimat, label kelas kalimat, dan tombol simpan. Tampilan tambah data korpus dapat dilihat pada Gambar 6.



Gambar 6. Antarmuka Halaman Input Data Korpus

F. Antarmuka Hasil Analisis Komentar

Halaman ini menampilkan postingan instagram yang sedang dianalisis cyberbullying komentarnya. Halaman ini juga menampilkan isi komentar beserta hasil analisisnya berupa hasil klasifikasi menggunakan metode Naïve Bayes Classifier. Tampilan halaman hasil analisis komentar dapat dilihat pada Gambar 7.



Gambar 7. Tampilan Antarmuka Hasil Analisis Komentar

G. Pengujian

Pada penelitian dilakukan dua jenis pengujian yaitu pengujian black box dan pengujian confusion matrix.

1. Pengujian Black Box

Black Box pada perangkat lunak dilakukan untuk menguji kesesuaian antara masukan dengan hasil yang ditampilkan pada aplikasi. Pengujian ini berfokus pada spesifikasi fungsional dari software, menguji kondisi input program berdasarkan apa yang dilihat, hanya fokus terhadap fungsionalitas dan output. Pengujian ini perlu dilakukan untuk melihat respon yang diberikan oleh aplikasi saat melakukan proses input data.

a) Pengujian Login

Pengujian input data dilakukan pada halaman login. Input data yang diuji adalah saat memasukkan username dan password. Lihat Tabel III.

TABEL III
PENGUJIAN HASIL UJI INPUT LOGIN

| Skenario Pengujian | Hasil Uji | Hasil Yang Diharapkan (Keterangan) |
|---|----------------|--|
| 1 Input data kosong | Tidak Berhasil | Sistem akan menolak untuk diproses, dan menampilkan pesan kesalahan: "All fields is required!" |
| 2 Input data salah satu kosong | Tidak Berhasil | Sistem akan menolak untuk diproses, dan menampilkan pesan kesalahan: "All fields is required!" |
| 3 Input semua data diisi | Berhasil | Berhasil login dan masuk ke aplikasi |
| 4 Input username & password tidak benar | Tidak Berhasil | Sistem akan menolak untuk diproses, dan menampilkan pesan kesalahan: "User credentials are wrong!" |

b) Pengujian Hapus Komentar Cyberbullying

Pengujian berikut dilakukan pada halaman hasil analisis dalam fungsi hapus komentar cyberbullying. Fungsi ini hanya dapat dijalankan jika postingan yang dianalisis merupakan postingan milik pengguna itu sendiri. Lihat Tabel IV.

TABEL IV
PENGUJIAN HASIL UJI HAPUS KOMENTAR CYBERBULLYING

| Skenario | Hasil | Hasil Yang Diharapkan |
|----------|-------|-----------------------|
|----------|-------|-----------------------|

| Pengujian | Uji | (Keterangan) |
|---------------------------------|----------------|---|
| 1 <i>Input</i> data kosong | Tidak Berhasil | Sistem akan menolak untuk diproses, dan menampilkan pesan kesalahan: “ <i>No comments are selected!</i> ” |
| 2 <i>Input</i> semua data diisi | Berhasil | Berhasil menghapus komentar dan menampilkan pesan: “ <i>Success delete comments!</i> ” |

c) Pengujian Input Data Korpus

Pengujian *input* data dilakukan pada halaman tambah data korpus. *Input* data yang diuji adalah saat memasukkan kelas dan isi kalimat. Lihat Tabel V.

TABEL V
PENGUJIAN HASIL UJI INPUT DATA KORPUS

| Skenario Pengujian | Hasil Uji | Hasil Yang Diharapkan (Keterangan) |
|---------------------------------------|----------------|---|
| 1 <i>Input</i> data kosong | Tidak Berhasil | Sistem akan menolak untuk diproses, dan menampilkan pesan kesalahan: “ <i>Failed Insert Data!</i> ” |
| 2 <i>Input</i> data salah satu kosong | Tidak Berhasil | Sistem akan menolak untuk diproses, dan menampilkan pesan kesalahan: “ <i>Failed Insert Data!</i> ” |
| 3 <i>Input</i> semua data diisi | Berhasil | Berhasil menambah data korpus |

2. Pengujian Confussion Matrix

Pengujian confussion matrix dilakukan untuk memeriksa apakah aplikasi telah berjalan dengan baik dan benar sesuai dengan tujuan yang diharapkan. Pengujian berikut dilakukan pada suatu postingan yang diposting oleh akun dengan *username @cimoyluv*. Dari hasil deteksi melalui aplikasi didapatkan 90 data komentar yang sudah diketahui klasifikasi kelasnya. Dari 90 data komentar, terdapat 21 komentar *bullying* dan 69 komentar bukan *bullying*. Hasil klasifikasi data komentar diuji melalui aplikasi dengan membandingkan dengan data latih secara otomatis menggunakan metode *Naïve Bayes Classifier*. Hasil klasifikasi data komentar yang dilakukan dengan aplikasi akan dibandingkan dengan hasil klasifikasi yang seharusnya. Perbandingan hasil klasifikasi data komentar tersebut dapat dilihat pada Tabel VI.

TABEL VI.
PERBANDINGAN HASIL KLASIFIKASI DATA KOMENTAR

| No. | Komentar | Hasil Klasifikasi Aplikasi | Hasil Klasifikasi Seharusnya |
|-----|--|----------------------------|------------------------------|
| 1. | lu kaya gembel najis banget | Bullying | Bullying |
| 2. | kenapa muka lu uda kaya jamet kuli bangunan | Bullying | Bullying |
| 3. | manis banget bang | Bukan Bullying | Bukan Bullying |
| 4. | sekarang udah bersih ya bang putih lagi | Bukan Bullying | Bukan Bullying |
| 5. | suaranya sih udah bagus tapi masih perlu banyak latihan lagi | Bukan Bullying | Bukan Bullying |
| 6. | glowup banget ya sist aku jadi makin sayang bangsat | Bukan Bullying | Bukan Bullying |
| 7. | imut banget si kamu gemes deh liatnya | Bukan Bullying | Bukan Bullying |
| 8. | semoga anakku kelak mirip lo ya bang | Bukan Bullying | Bukan Bullying |
| 9. | banyak tingkah banget lo | Bukan Bullying | Bukan Bullying |
| 10. | bajingan kau ya bisa bisanya bertingkah di keadaan yang seperti ini | Bullying | Bullying |
| 11. | kenapa harus ada dia disamping lu gaterima gua muka dia kaya kontol basi | Bullying | Bullying |

Dari table Table VI. dapat disimpulkan bahwa hasil klasifikasi yang seharusnya menunjukkan 75 komentar *bullying* dan 125 komentar bukan *bullying*. Selanjutnya perbandingan hasil klasifikasi akan diujikan dengan metode *confussion matrix*. Hasil pengujian *confussion matrix* dapat dilihat pada Tabel VII.

TABEL VII
HASIL PENGUJIAN CONFUSION MATRIX

| Nilai Asli | Nilai Prediksi | |
|-------------------------|----------------|---------------------|
| | True Bullying | True Bukan Bullying |
| Prediksi Bullying | 75 | 3 |
| Prediksi Bukan Bullying | 3 | 122 |

Berdasarkan table hasil pengujian diatas, didapatkan hasil sebagai berikut:

1. $Accuracy = (TP + TN) / N = (75 + 122) / 200 = 98,5\%$
2. $Error Rate = 1 - Accuracy = 1 - 98,5\% = 1,5\%$
3. $Precision = TP / (TP + FP) = 75 / (75 + 3) = 96\%$
4. $Recall = TP / (TP + FN) = 75 / (75 + 3) = 96\%$

Hasil dari proses klasifikasi komentar *cyberbullying* melalui aplikasi yang didapat memiliki akurasi yang sangat tinggi, yaitu 98,5% dengan nilai rating *error* 1,5%, nilai presisi 96% dan nilai *recall* 96%.

IV. KESIMPULAN

Kesimpulan yang didapatkan dari hasil metodologi penelitian, analisis sistem, perancangan dan pengujian untuk aplikasi pendeteksi *cyberbullying* terhadap komentar postingan media sosial instagram dengan metode *naive bayes classifier* membuktikan bahwa aplikasi yang telah dibuat dalam penelitian ini dapat mendeteksi dan menghapus ribuan komentar *cyberbullying* secara massal. Dari hasil deteksi *cyberbullying* melalui aplikasi terhadap 200 data komentar dari suatu postingan media sosial instagram menunjukkan bahwa terdapat 78 komentar *bullying* dan 122 komentar bukan *bullying*. Hasil deteksi tersebut kemudian diujikan menggunakan metode *confussion matrix*. Dari hasil pengujian *confussion matrix* menunjukkan bahwa perhitungan yang dilakukan oleh aplikasi memiliki nilai akurasi sebesar 98,5%, nilai kerusakan (*error rate*) sebesar 1,5%, nilai presisi sebesar 96%, dan nilai *recall* sebesar 96%. Berdasarkan hasil penelitian, diketahui bahwa kata kunci "tolol" merupakan kata dengan bobot paling tinggi yang paling sering digunakan untuk melakukan *cyberbullying* yaitu sebesar 13,82 %. Pola kalimat yang berpotensi paling besar merupakan kalimat *cyberbullying* adalah "Subyek + kata bully", dengan subyek "lu" berpotensi paling besar digunakan untuk melakukan aktifitas *cyberbullying* yaitu sebesar 15,46%..

REFERENSI

[1] Asosiasi Penyelenggara Jasa Internet Indonesia [APJII]. 2014. Profil Pengguna Internet Indonesia. Jakarta : Asosiasi Penyelenggara Jasa Internet Indonesia.

[2] Napoleoncat. (2019). *Social media users in Indonesia*. [online]. Available

<https://napoleoncat.com/stats/social-media-users-in-indonesia/2019/12>

[3] Oktaviani, Yolanda. "Perundungan dunia maya (*cyberbullying*) menurut undang-undang RI NO.19 Tahun 2016 tentang informasi transaksi elektronik dan hukum islam Universitas Islam Negeri Radenfatah Palembang". Palembang : Universitas Islam Negeri Radenfatah. 2017.

[4] Pratama, Ripanti. "Analysis of Student Academic Performance and Social Media Activities by Using Data Mining Approach". Jurnal. Pontianak : Fakultas Informatika Universitas Tanjungpura. 2020

[5] Afrinza, Restio. "Deteksi Cyber Bullying pada Twitter Menggunakan Naive Bayes Classifier. Skripsi. Bandung" : Fakultas Informatika Universitas Telkom. 2019.

[6] Anggraini, Ika Yunida. "Pemodelan Deteksi Cyberbullying Pada Jejaring Sosial Twitter. Skripsi. Kediri" : Fakultas Teknik Universitas Nusantara PGRI Kediri. 2018.

[7] Saleh, Alfa. "Klasifikasi Metode Naive Bayes Dalam Data Mining Untuk Menentukan Konsentrasi Siswa (Studi Kasus di MAS PAB 2 Medan)". Skripsi. Medan : Fakultas Teknik Informatika Universitas STMIK Potensi Utama Medan. 2014.

[8] Feldman, R., & Sanger. J. "Text Mining Handbook : Advanced Approaches in Analyzing Unstructured Data". New York : Cambridge University Press. 2007.

[9] Gunawan, Billy. "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes". Pontianak : Jurnal Edukasi dan Penelitian Informatika. 2018.

[10] M. Suyanto. "Analisis & desain aplikasi multimedia untuk pemasaran". Yogyakarta : Penerbit Andi. 2004.

[11] S. Yuswanto. "Mengolah Database dengan SQL Server 2000". Surabaya : Prestasi Pustaka Publ. 2005.

[12] Jogiyanto, H.M. "Analisa dan Desain Sistem: Pendekatan Teori Dan Praktek Aplikasi Bisnis. Cetakan Ketiga". Yogyakarta : Andi Offset. 2005.

[13] Dwi Prasetyo, Didik.. "Tips dan Trik Kolaborasi PHP dan MySQL untuk membuat web database yang interaktif". Elex Media Komputindo : Jakarta. 2003.

[14] Han, J., Kamber, M., & Pei, J. "Data Mining Concepts and Techniques Third Edition". Waltham : Elsevier Inc. 2011.

[15] Hamilton, Howard. (2017). *Confusion Matrix*. http://www2.cs.uregina.ca/~hamilton/courses/831/notes/confusion_matrix/confusion_matrix.html, 24 Maret 2018.

[16] Juniarsih, Ripanti, Pratama. "Implementasi Naive Bayes Classifier pada Opinion Mining Berdasarkan Tweets Masyarakat Terkait Kinerja Presiden". Jurnal. Pontianak : Fakultas Informatika Universitas Tanjungpura. 2020.