

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

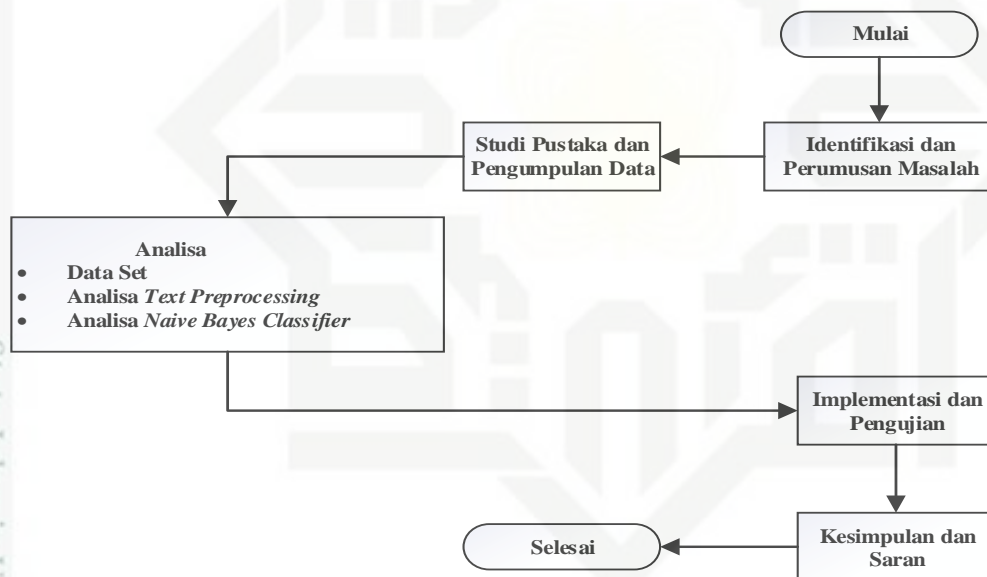
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB III METODOLOGI PENELITIAN

1.1 Tahapan Penelitian

Suatu penelitian yang akan dilakukan tentunya membutuhkan sebuah metode pendekatan sebagai bentuk penyelesaian yang berfungsi untuk melaksanakan penelitian agar berjalan lancar dan sesuai dengan tujuan yang diharapkan. Berikut ini adalah tahapan-tahapan penelitian yang dijadikan penulis sebagai panduan dalam menyelesaikan tugas akhir ini yaitu:



Gambar 3.1 Tahapan Penelitian

1.2 Identifikasi dan Perumusan Masalah

Tahapan awal yang dilakukan pertama kali untuk menemukan identifikasi dan permasalahan apa yang akan diteliti. Pada tahap ini mempelajari tujuan, ruang lingkup dan metodologi penelitian. Beberapa kegiatan yang dilakukan pada tahap ini yaitu studi pustaka terkait penelitian yang sudah ada, memilih topik permasalahan yang di anggap menarik untuk penelitian dan mencari ide-ide baru yang sudah ada untuk penelitian baru.

1.3 Studi Pustaka dan Pengumpulan Data

Studi pustaka ini dipelajari dari buku, situs internet, media elektronik dan media cetak. Berikutnya latar belakang masalah dan tujuan dari penelitian ini menjadi landasan mengapa penelitian ini dilakukan.

Tahap selanjutnya yaitu tahap pengumpulan data, pada tahap pengumpulan data berupa *tweet* pelecehan *online*. Adapun pengambilan data ini bertujuan untuk mendapatkan informasi tentang analisa klasifikasi *tweet* pelecehan *online* di Indonesia dengan menggunakan metode *Naïve Bayes*.

Adapun sumber data dan jumlah data sebagai berikut:

1. Sumber data berasal dari Twitter berbahasa Indonesia metode yang digunakan dalam pengumpulan data adalah mengunduh data *tweet* menggunakan API Twitter.
2. Data yang digunakan adalah yang mengacu berdasarkan dari beberapa kasus pelecehan *online* yang populer dibicarakan di Twitter.
3. Setelah data *tweet* dikumpulkan selanjutnya memberi label pada seluruh data *tweet* sesuai dengan kelas yang telah ditentukan yaitu *racist*, *embarrass*, *physically threatened*, *sexually harassed* dan *neutral*.

Jumlah *dataset* yang digunakan lebih kurang 1500 *tweet* dan perbandingan data latih dan data uji akan dilakukan klasifikasi dari 70:30, 80:20, 90:10.

1.4 Analisa

Setelah menentukan tahapan pengumpulan data selesai maka tahap selanjutnya adalah penganalisaan *machine learning*. Dalam analisa *machine learning* terdapat beberapa tahapan, diantaranya adalah sebagai berikut:

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Analisa Klasifikasi *Tweet* Pelecehan *Online*

Pada tahap ini penulis menganalisa kelas pelecehan *online* yang akan diklasifikasi. Kelas tersebut dibagi menjadi lima yaitu *racist*, *embarrass*, *physically threatened*, *sexually harassed* dan *neutral*. Selanjutnya data yang ada akan diberi label manual sesuai dengan kategori yang telah ditetapkan.

2. Analisa Data

Pada tahap ini penulis menyiapkan langkah-langkah data berupa seleksi manual, pelabelan manual, *preprocessing* yaitu *cleaning* yaitu penghapusan *retweet*, URL, emotikon, dan selanjutnya *cleaning*, *case folding*, *tokenizing*, *normalization*, *filtering* dan *stemming* dan melakukan pembobotan.

3. Analisa Algoritma *Naïve Bayes*

Tahap ini adalah bagian pembelajaran pola dengan metode *Naïve Bayes* dan mengklasifikasikan *tweet* yang merupakan *tweet* pelecehan *online* kedalam masing-masing kelas berdasarkan data latih.

1. Proses pelatihan *Naïve Bayes*. Untuk proses pembelajaran terbagi lima klasifikasi berdasarkan *tweet* pelecehan *online* pada Twitter. Kemudian didapatkan nilai probabilitas dari setiap kelas terhadap *tweet* melalui persamaan 2.5 di landasan teori model probabilitas. Setelah mendapatkan nilai probabilitas setiap kategori. Kemudian setiap kelas digabungkan dicocokkan dengan ke lima kelas.
2. Proses klasifikasi dilakukan didasarkan pada data latih, dan dicari nilai data yang diuji. Kemudian setelah dapat nilai data yang diuji, data uji tersebut dicocokkan nilai lima kelas pelecehan *online*. Setelah itu dapat disimpulkan data uji tersebut masuk kedalam ke dua kelas *racist*, *embarrass*, *physically threatened*, *sexually harassed* atau *neutral*.

1.5 Implementasi dan Pengujian

Tahap implementasi dan pengujian merupakan tahapan yang dilakukan setelah tahap analisa telah selesai. Pada tahap ini akan menjelaskan tentang penerapan aplikasi yang telah dianalisa sebelumnya.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Penyiapan *dataset*

Pada tahap ini merupakan implementasi analisa yang telah dilakukan sebelumnya terhadap *dataset*. Mulai dari proses unduh data, *preprocessing* data, kemudian pembobotan *dataset*, hingga data siap untuk digunakan dalam proses klasifikasi metode.

2. Rencana pengujian

Setelah dilakukan implementasi, maka dilakukan pengujian sebagai ukuran bahwa dapat dijalankan sesuai dengan tujuan. Pada pengujian ini ada beberapa hal yang dijadikan acuan, yaitu:

- a. Pengujian pembagian data latih dan data uji dengan beberapa perbandingan data untuk mencari tingkat akurasi klasifikasi.
- b. Pengujian perangkat lunak, pengujian yang dilakukan adalah pengujian *white box* yaitu dengan menguji pada keseluruhan proses, dan memegang perhitungan mekanisme internal perangkat lunak dan komponen.

3. Evaluasi

Pada tahap evaluasi, ukuran atau parameter yang digunakan adalah *confusion matrix*. Tujuan dari tahapan evaluasi adalah untuk melihat efektivitas metode pengujian yang dilakukan dengan beberapa model percobaan dalam mengklasifikasi *tweet* pelecehan *online*.

1.6 Kesimpulan dan Saran

Tahapan akhir dari penelitian adalah penarikan kesimpulan berdasarkan hasil yang telah diperoleh dari tahapan sebelumnya, serta memberikan saran-saran kepada pembaca serta untuk menyempurnakan dan mengembangkan penelitian selanjutnya