

## BAB V PENUTUP

### 5.1 Kesimpulan

Berdasarkan rumusan masalah dan tujuan penelitian yang telah dijabarkan, penelitian ini menggunakan metode Support Vector Machine (SVM) dan Naïve Bayes untuk mengklasifikasi sentimen positif atau negatif pada tweet yang berkaitan dengan *Cyberbullying*. Analisis mendalam menunjukkan bahwa SVM mengungguli Naïve Bayes dalam hal kinerja. SVM mencapai tingkat akurasi sebesar 84%, precision 82%, recall 82%, F1-Score 82%, dan ROC curve sebesar 98%, sementara Naïve Bayes mencatat nilai akurasi 90%, precision 88%, recall 89%, F1-Score 89%, dan ROC curve 96%. Model Naïve Bayes berhasil mengenali dan mengklasifikasikan sentimen dengan tingkat akurasi yang tinggi, yang menegaskan bahwa keunggulannya dalam analisis sentimen terhadap tweet *Cyberbullying* di platform Twitter (X).

Agar Dapat Membedakan hasil Tweet *Cyberbullying* atau tidak Penulis memberikan label secara manual yang penulis berfokus pada hasil stemming Dimana penulis memberikan 2 parameter yaitu Positif (1) untuk tweet yang mengandung *Cyberbullying* dan Negatif (0) untuk tweet yang tidak mengandung *Cyberbullying*, dengan total *Positif* 2927 tweet dan *Negatif* 1337 tweet.

Dari Dataset Sebelumnya Yang berjumlah 6330 tweet hasil crawling data menggunakan *Tweet-Harvest*, setelah melakukan *Pre-Processing* Data dan Membuang data duplikat pada hasil stemming berjumlah 4263 tweet dan setelah memastikan agar dataset Bahasa Indonesia pada dataset penulis dengan jumlah 3592 Dimana kata "*bullying*" dengan total 2530 kata yang sering muncul.

## 5.2 Saran

Berdasarkan Hasil penelitian penulis dapat menjabarkan beberapa saran agar dapat dikembangkan serta Penelitian ini memiliki keterbatasan. Berikut Saran yang dapat penulis sampaikan:

1. Pada Kunci Pencarian dapat lebih spesifik lagi agar tweet yang berisi Cyberbullying lebih bervariasi dan mewakili topik bahasan itu sendiri.
2. Keterbatasan pada penelitian ini pada saat crawling data Dimana penulis menggunakan library *harest-tweet* yang sudah dibuat oleh orang lain.
3. Dapat menambah lebih banyak lagi jumlah dataset agar dapat mengetahui nilai akurasi pada model klasifikasi.
4. Penelitian Analisa sentiment dapat digunakan pada media sosial lainnya seperti Instagram, Facebook, dan Threads atau dari website di internet
5. Pada Proses Stemming memakan waktu yang sangat lama contoh penulis yang menggunakan 6596 hasil crawling membutuhkan waktu lebih dari 1 jam.
6. Pelabelan data dilakukan secara manual oleh penulis ini dilakukan agar hasil pelabelan dapat menjadi model pembelajaran oleh mesin untuk proses analisis dan pelabelan data penulis berfokus pada kolom stemming data.

Demikian Saran dan Kesimpulan yang dapat penulis sampaikan semoga penelitian ini dapat memberikan dampak positif baik sebagai bahan acuan untuk penelitian lain atau sama,