

## BAB 5

### SARAN DAN KESIMPULAN

#### 5.1 Kesimpulan

Berdasarkan penelitian dan analisis yang telah dilakukan, dapat disimpulkan bahwa :

1. Terdapat *buzzer* yang melakukan *tweet* yang sama berulang-ulang terhadap masing-masing bakal calon presiden dengan tujuan mendongkrak citra tiap calon. Hasil dari *pra-processing cleaning* dan filterasi data terhadap *buzzer* tiap bakal calon presiden, awalnya terdapat masing-masing 5000 data *tweet* hasil *crawling*, namun hasil akhir data *tweet* masing-masing bakal calon presiden setelah dilakukan proses filterasi yaitu, Anies Baswedan (1211 data *tweet*), Ganjar Pranowo (1101 data *tweet*), Prabowo Subianto (1481 data *tweet*). Hasil akhir tersebut menunjukkan bahwa Ganjar Pranowo memiliki *tweet buzzer* paling banyak, disusul dengan Anies Baswedan, dan Prabowo Subianto.
2. Hasil pelabelan emosi dari tiap bakal calon presiden menunjukkan nilai yang berbeda satu sama lain, adapun jumlah pelabelan emosi masing-masing calon yaitu : Anies Baswedan (senang : 484, marah : 381, netral : 232, jijik : 41, sedih : 36, takut : 32, terkejut : 4), Ganjar Pranowo (senang : 668, netral : 274, marah : 70, terkejut : 40, takut : 21, jijik : 18, sedih : 9), Prabowo Subianto (senang : 638, netral : 468, marah : 126, terkejut : 111, sedih : 72, takut : 54, jijik : 11).
3. Performa model algoritma Naïve Bayes tanpa menggunakan label netral menunjukkan performa yang lebih baik, dengan nilai akurasi model sebesar 58% pada data Anies Baswedan, 58% pada data Prabowo Subianto, dan 76% pada data Ganjar Pranowo, serta gabungan 69%.
4. Performa model menggunakan label netral menunjukkan adanya pengaruh label netral terhadap akurasi dari model, hal ini terjadi karena label netral mendominasi kelas emosi lainnya. Klasifikasi pada skenario pengujian menggunakan label netral menunjukkan akurasi sebesar 55% pada data Anies Baswedan, 60% pada data Ganjar Pranowo, dan 53% pada data Prabowo Subianto, sedangkan untuk gabungan semuanya nilai akurasi sebesar 51%.

5. Tahap *random sampling* dilakukan bertujuan untuk menyeimbangkan seluruh data dalam label, sehingga model dapat memberikan hasil yang lebih baik. Hasil klasifikasi *random sampling* menunjukkan nilai akurasi 44% untuk data Anies menggunakan label netral, dan 44% tanpa label netral. Ganjar Pranowo menunjukkan nilai akurasi 71% menggunakan label netral, dan 71% tanpa label netral. Prabowo Subianto memberikan nilai akurasi model 43% menggunakan label netral, dan 52% tanpa label netral.

6. Terdapat beberapa kelas emosi yang tidak dapat diklasifikasikan dengan *fscore* di bawah satu persen, karena sedikitnya jumlah data terkait kelas tersebut.

## 5.2 Saran

Terdapat beberapa evaluasi terkait penelitian ini, adapun saran yang dapat dilakukan untuk penelitian ke depannya antara lain:

1. Dapat mempertimbangkan penggunaan algoritma klasifikasi lainnya, seperti *K-Nearest Neighbors*, *Decision Tree*, *Random Forest*, dll, untuk mengecek performa dan dapat menentukan algoritma paling baik dalam membuat model.
2. Perlu dilakukan pencarian dan penambahan data *tweet* untuk kelas emosi jijik, takut, dan terkejut agar tidak ada ketidakseimbangan data antara kelas emosi.
3. Ketika proses pengumpulan data dapat dilakukan filterasi dan menyeleksi secara langsung data *tweet* yang sama untuk menghindari adanya data homogen karena *buzzer*.