

## BAB 7

### KESIMPULAN DAN SARAN

#### 7.1 Kesimpulan

Berdasarkan hasil penelitian yang sudah dilakukan pada tugas akhir ini, berikut ini adalah beberapa kesimpulan yang dapat diambil:

1. Data Twitter berhasil dikumpulkan menggunakan Kafka dengan format CSV setiap 30 menit sekali dalam rentang waktu antara November 2022 hingga Februari 2023, dengan menggunakan kata kunci yang didapatkan melalui hasil penelusuran secara nyata. Jumlah *tweet* yang didapatkan sebanyak 3.805.936 *tweet* yang terdiri dari 245.438 akun *unique*.
2. Dari data Twitter yang sudah dikumpulkan dilakukan eksplorasi dan penyiapan data secara berulang dan sudah didapatkan hasil yang terbaik.
3. Akun *buzzer* berhasil dideteksi atau ditemukan dengan menggunakan 5 tahap yang sudah dirancang berdasarkan hasil studi literatur dan eksplorasi secara nyata. Dari analisis terhadap akun *buzzer* yang berhasil dideteksi, dapat dilihat bahwa hasil deteksi yang didapatkan sudah bagus, sebesar 86.4% akun merupakan *buzzer* dan juga didapatkan akun *buzzer* dari hasil eksplorasi secara nyata, yaitu *amelkecil46*.
4. Prediksi sentimen akun *buzzer* hasil deteksi berhasil dilakukan dengan menggunakan metode *lexicon based* dan metode *machine learning* (*decision tree* dan *random forest*). Dilakukan 2 eksperimen, dimana eksperimen 1 menggunakan 51 baris data sedangkan eksperimen 2 menggunakan 311 baris data (berbeda pada jumlah perwakilan *tweet* yang diambil). Hasil terbaik eksperimen 1 dimiliki oleh model *random forest* 1 dan hasil terbaik eksperimen 2 dimiliki oleh model *decision tree* 2.
5. Perangkat lunak deteksi *buzzer* berbasis *website* berhasil dibuat untuk mendeteksi *buzzer*, memprediksi sentimen *buzzer*, dan menampilkan hasil dari deteksi dan prediksi yang telah dilakukan sebelumnya.

#### 7.2 Saran

Berikut ini adalah beberapa saran yang dapat dilakukan untuk pengembangan selanjutnya:

1. Menggunakan sistem pengumpul data lain yang menyediakan status *quote* dan *reply* agar hasil yang didapatkan dapat dianalisis lebih lanjut menggunakan graf.
2. Membuat sendiri korpus kata-kata positif dan negatif dari teks yang didapatkan atau berdasarkan referensi seperti Kamus Besar Bahasa Indonesia (KBBI) agar hasil yang didapatkan lebih bagus lagi.

## DAFTAR REFERENSI

- [1] Ishwarappa dan J, A. (2015) A brief introduction on big data 5vs characteristics and hadoop technology. *Procedia Computer Science*, **48**, 319–324.
- [2] Firmansyah, M. A., Karlinah, S., dan Sumartias, S. (2017) Kampanye pilpres 2014 dalam konstruksi akun twitter pendukung capres. *The Messenger*, **9**, 79–90.
- [3] houh Chen, C., Härdle, W., dan Unwin, A. (2008) *Handbook of Data Visualization*. Springer Berlin, Heidelberg, Berlin.
- [4] Karau, H., Konwinski, A., Wendell, P., dan Zaharia, M. (2015) *Learning Spark: Lightning-Fast Data Analysis*. O'Reilly Media, Inc., USA.
- [5] Ravi, M. dan Moertini, V. S. (2020) Pengumpulan data twitter dengan teknologi big data. Bagian dari Moertini, V. S. dan Adithia, M. T. (ed.), *Pengantar Data Science dan Aplikasinya bagi Pemula*. Unpar Press, Bandung, Indonesia.
- [6] Supriyanto, E. E., Bakti, I. S., dan Furqon, M. (2021) The role of big data in the implementation of distance learning. *Paedagoria*, **12**, 61–68.
- [7] Gupta, V. dan Hewett, R. (2020) Real-time tweet analytics using hybrid hashtags on twitter big data streams. *Multidisciplinary Digital Publishing Institute*, **11**, 1–23.
- [8] Karya, G. (2020) Teknologi big data. Bagian dari Moertini, V. S. dan Adithia, M. T. (ed.), *Pengantar Data Science dan Aplikasinya bagi Pemula*. Unpar Press, Bandung, Indonesia.
- [9] Chambers, B. dan Zaharia, M. (2018) *Spark: The Definitive Guide Big Data Processing Made Simple*. O'Reilly Media, Inc., USA.
- [10] Felicia dan Loisa, R. (2018) Peran buzzer politik dalam aktivitas kampanye di media sosial twitter. *Koneksi*, **2**, 352–359.
- [11] Gustomy, R. (2020) Pandemi ke infodemi: Polarisasi politik dalam wacana covid-19 pengguna twitter. *Jurnal Ilmiah Ilmu Pemerintahan*, **5**, 190–205.
- [12] Arianto, B. (2020) Peran buzzer media sosial dalam memperkuat ekosistem pemasaran digital. *Jurnal Optimum*, **10**, 52–65.
- [13] Ibrahim, M., Abdillah, O., Wicaksono, A. F., dan Adriani, M. (2015) Buzzer detection and sentiment analysis for predicting presidential election results in a twitter nation. *2015 IEEE 15th International Conference on Data Mining Workshops*, Chicago, USA, 14-17 November, pp. 1348–1353. IEEE, USA.
- [14] Hannan, S. A. (2016) An overview on big data and hadoop. *International Journal of Computer Applications*, **154**, 29–35.
- [15] Kolajo, T., Daramola, O., dan Adebisi, A. (2019) Big data stream analysis: a systematic literature review. *Journal of Big Data*, **6**, 1–30.

- [16] Narkhede, N., Shapira, G., dan Palino, T. (2017) *Kafka: The Definitive Guide*. O'Reilly Media, Inc., USA.
- [17] Coletta, L. F. S., da Silva, N. F. F., Hruschka, E. R., dan Jr., E. R. H. (2014) Combining classification and clustering for tweet sentiment analysis. *2014 Brazilian Conference on Intelligent Systems*, Sao Paulo, Brazil, 18-22 October 2014, pp. 210–215. IEEE, Brazil.
- [18] Cvijikj, I. P. dan Michahelles, F. (2011) Understanding social media marketing: A case study on topics, categories and sentiment on a facebook brand page. *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, Tampere, Finland, 28-30 September 2011, pp. 175–182. ACM, New York.
- [19] Han, J., Kamber, M., dan Pei, J. (2012) *Data Mining Concepts and Techniques*, 3rd edition. Morgan Kaufmann, Burlington.
- [20] James, G., Witten, D., Hastie, T., dan Tibshirani, R. (2013) *An Introduction to Statistical Learning with Applications in R*. Springer, New York.
- [21] Ravi, M. (2020) Studi dan implementasi spark streaming untuk mengumpulkan big data stream. Skripsi. Universitas Katolik Parahyangan, Indonesia.