

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil penelitian ini dapat diketahui bahwa daerah provinsi dengan jumlah penduduk yang padat seperti DKI Jakarta, Jawa Barat, dan Bali cenderung menolak adanya kenaikan harga BBM. Mayoritas masyarakat pada setiap provinsi menggunakan kata “harga” dalam memberikan pendapatnya mengenai kenaikan harga BBM baik sentimen negatif, sentimen netral, ataupun sentimen positif. Pada kategori sentimen negatif dan sentimen netral, kata “harga” merupakan kata paling sering digunakan yaitu sebanyak 25 provinsi dari total 34 provinsi di Indonesia yang mayoritas menggunakan kata tersebut. Pada kategori sentimen positif, kata yang paling sering digunakan cukup beragam, diantaranya adalah “harga”, “dukung”, “subsidi”, dan lain-lain. Beberapa provinsi seperti DKI Jakarta, Kalimantan Timur, Kepulauan Riau, Maluku Utara, Nusa Tenggara Timur, Papua, Papua Barat, Sulawesi Tenggara, Sumatera Selatan, dan Sumatera Utara memiliki perbedaan frekuensi kata yang paling sering digunakan pada setiap kategorinya dan tidak ada kata yang mendominasi untuk semua kategori.

Akurasi terendah terdapat pada provinsi Sulawesi Utara yaitu sebesar 0.65. Nilai akurasi tersebut rendah disebabkan oleh fitur kata yang terdapat pada data *testing* tidak terdapat dalam data *training*, sehingga model kesulitan untuk mempelajarinya dan menghasilkan akurasi yang buruk. Akurasi tertinggi terdapat pada provinsi DIY, Jambi, Jawa Timur, Kalimantan Selatan, Papua Barat, dan Sumatera Barat dengan nilai akurasi sebesar 0.95. Nilai akurasi yang tinggi tersebut disebabkan karena data *testing* yang digunakan memiliki kosa kata yang terdapat pada data *training*, sehingga model dapat mempelajarinya berdasarkan data *training* dan memproses data *testing* dengan baik. Model yang dibuat telah berhasil memprediksi setiap *tweet* pada data *testing* dengan akurasi rata-rata sebesar 0,882.

## 5.2 Saran

Model yang baik dapat dibangun berdasarkan pelabelan data yang baik. Pada penelitian ini, pelabelan data dilakukan manual oleh satu orang sehingga hasil dari pelabelan masih bersifat subjektif, perlu dilakukan proses pelabelan dengan melibatkan ahli bahasa agar pelabelan data lebih objektif. Hal lain yang menjadi dasar dalam pembuatan model adalah *feature extraction*. Salah satu kekurangan metode TF-IDF yaitu tidak memperhatikan urutan kata sehingga metode tersebut tidak memiliki kemampuan untuk menangkap makna. Perlu dilakukan *feature extraction* dengan metode yang lebih baik seperti metode Word2Vec, FastText, dan Glove.