



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil dari analisis deteksi ulasan palsu terhadap toko online menggunakan algoritma *Naïve Bayes*, *Support Vector Machine*, dan *Logistic Regression*, berikut adalah kesimpulan dari penelitian:

1. Data bersih yang diberikan label pada setiap *e-commerce* adalah 500 data. Setelah melewati proses labeling, Bukalapak hanya bisa mendapatkan data sebanyak 491 yang dapat dilabeling. Dikarenakan terlalu banyak data yang bersifat netral. Data yang digunakan untuk proses klasifikasi adalah sebanyak 1490 data. Dengan rincian data *true* sebanyak 1159 *tweet* sedangkan data *fake* sebanyak 331 *tweet*. Dilihat dari hasil tersebut, maka masih terdapat banyak *tweet* palsu yang sengaja dibuat pengguna untuk menaikkan atau menjatuhkan kesan terhadap suatu barang.
2. Dari total 1490 data yang ada, total data *true* Tokopedia sebanyak 426 data, sedangkan total data *fake* Tokopedia sebanyak 76 data. Total data *true* Shopee sebanyak 376 data, sedangkan total data *fake* Shopee sebanyak 126 data. Yang terakhir total data *true* Bukalapak sebanyak 360 data, sedangkan total data *fake* Bukalapak sebanyak 132 data. *E-commerce* yang memiliki data *true* terbanyak adalah Tokopedia dengan 426 data, sedangkan *e-commerce* yang memiliki data *fake* terbanyak adalah Bukalapak dengan 132 data.

3. Hasil perbandingan algoritma dapat dilihat pada tabel berikut:

Tabel 5. 1 Hasil Perbandingan Algoritma

Algoritma	Jenis Data	Akurasi
<i>Support Vector Machine</i>	<i>Data True</i>	70.58%
<i>Support Vector Machine</i>	<i>Data Fake</i>	92.45%
<i>Naive Bayes</i>	<i>Data True</i>	69.89%
<i>Naive Bayes</i>	<i>Data Fake</i>	91.84
<i>Logistic Regression</i>	<i>Data True</i>	70.49%
<i>Logistic Regression</i>	<i>Data Fake</i>	92.45%

Dilihat dari hasil perbandingan algoritma, *Support Vector Machine* mendapatkan hasil akurasi 70.58% untuk data *real* dan 92.45% untuk data *fake*. Algoritma *Logistic Regression* memiliki tingkat akurasi 70.49% untuk data *real* dan 92.45% untuk data *fake*. Algoritma *Naive Bayes* 69.89% untuk data *real* dan 91.84% untuk data *fake*.

4. Algoritma *Support Vector Machine* mendapatkan hasil akurasi tertinggi yaitu 70.58% untuk data *real* dan 92.45% untuk data *fake*. Algoritma *Logistic Regression* memiliki tingkat akurasi tertinggi kedua yaitu 70.49% untuk data *real* dan 92.45% untuk data *fake*. Yang terakhir adalah algoritma *Naive Bayes* dengan akurasi terendah yaitu 69.89% untuk data *real* dan 91.84% untuk data *fake*.

5.2 Saran

Penelitian berupa analisis sentimen ulasan palsu di *Twitter* terhadap toko online menggunakan algoritma *Naive Bayes*, *Support Vector Machine*, dan *Logistic Regression* memberikan saran untuk penelitian selanjutnya yang akan datang, yaitu:

1. Penambahan menggunakan algoritma klasifikasi lainnya untuk pilihan perbandingan akurasi yang lebih banyak.
2. Perluasan deteksi ulasan palsu menggunakan bahasa asing.
3. Penelitian menggunakan *dataset* yang lebih variatif dan berbeda.
4. Proses labeling dilakukan menggunakan *Rapidminer*.



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA