

Analisis Sentimen Menggunakan Metode Neural Network

Team members:

1. Arief Budiman
2. Muhidin Jaariya
3. Rommy



C O N T E N T S

01.

Pendahuluan

02.

Metode Penelitian

03.

Kesimpulan

PENDAHULUAN





Latar belakang

Analisis sentimen merupakan proses penting dalam menganalisis opini, perasaan, dan sikap yang terkandung dalam teks. Dalam era digital saat ini, jumlah data teks yang dihasilkan melalui media sosial, ulasan pelanggan, dan sumber lainnya semakin meningkat. Untuk mengatasi kompleksitas dan volume besar data tersebut, pendekatan yang digunakan adalah jaringan saraf tiruan (neural network) telah terbukti berhasil dalam analisis sentimen.



Rumusan masalah

Dalam konteks analisis sentimen dengan menggunakan neural network, terdapat beberapa rumusan masalah yang dapat diajukan, antara lain:

1. Bagaimana cara menerapkan analisis sentimen menggunakan Neural Network untuk mengklasifikasikan teks ke dalam tiga kategori sentimen: positif, negatif, dan netral?
2. Seberapa akurat dan efektifkah metode analisis sentimen berbasis Neural Network dalam mengklasifikasikan sentimen pada teks



Tujuan penelitian

1. Menerapkan metode analisis sentimen berbasis Neural Network untuk mengklasifikasikan teks ke dalam tiga kategori sentimen: positif, negatif, dan netral.
2. Mengevaluasi akurasi dan efektivitas metode analisis sentimen berbasis Neural Network dalam memprediksi sentimen pada teks,
3. Memberikan pemahaman yang lebih baik tentang penggunaan Neural Network dalam analisis sentimen dan potensinya untuk mendukung pengambilan keputusan berbasis data.

Metode Penelitian



Data Asli

	Text	Sentimen
0	warung ini dimiliki oleh pengusaha pabrik tahu...	positive
1	mohon ulama lurus dan k212 mmbri hujjah partai...	neutral
2	lokasi strategis di jalan sumatera bandung . t...	positive
3	betapa bahagia nya diri ini saat unboxing pake...	positive
4	duh . jadi mahasiswa jangan sombong dong . kas...	negative
...
10995	tidak kecewa	positive
10996	enak rasa masakan nya apalagi kepiting yang me...	positive
10997	hormati partai-partai yang telah berkoalisi	neutral
10998	pagi pagi di tol pasteur sudah macet parah , b...	negative
10999	meskipun sering belanja ke yogya di riau junct...	positive
11000 rows × 2 columns		

Stopwords NLTK

	Text	Sentimen	text_clean	text_filter
0	warung ini dimiliki oleh pengusaha pabrik tahu...	positive	warung ini dimiliki oleh pengusaha pabrik tahu...	warung dimiliki pengusaha pabrik puluhan terke...
1	mohon ulama lurus dan k212 mmbri hujjah partai...	neutral	mohon ulama lurus dan k212 mmbri hujjah partai...	mohon ulama lurus k212 mmbri hujjah partai diw...
2	lokasi strategis di jalan sumatera bandung . t...	positive	lokasi strategis di jalan sumatera bandung t...	lokasi strategis jalan sumatera bandung nya ny...
3	betapa bahagia nya diri ini saat unboxing pake...	positive	betapa bahagia nya diri ini saat unboxing pake...	betapa bahagia nya unboxing paket barang nya b...
4	duh . jadi mahasiswa jangan sombong dong . kas...	negative	duh jadi mahasiswa jangan sombong dong kas...	duh mahasiswa sombong kasih kartu kuning belaj...
...
10995	tidak kecewa	positive	tidak kecewa	kecewa
10996	enak rasa masakan nya apalagi kepiting yang me...	positive	enak rasa masakan nya apalagi kepiting yang me...	enak masakan nya kepiting menyenangkan memilih...
10997	hormati partai-partai yang telah berkoalisi	neutral	hormati partai partai yang telah berkoalisi	hormati partai partai berkoalisi
10998	pagi pagi di tol pasteur sudah macet parah , b...	negative	pagi pagi di tol pasteur sudah macet parah b...	pagi pagi tol pasteur macet parah bikin jengkel
10999	meskipun sering belanja ke yogya di riau junct...	positive	meskipun sering belanja ke yogya di riau junct...	belanja yogya riau junction kali lihat foodlif...

11000 rows × 4 columns

Feature Extraction

```
df_vektor = df.text_clean.tolist()
```

```
count_vect = CountVectorizer()
```

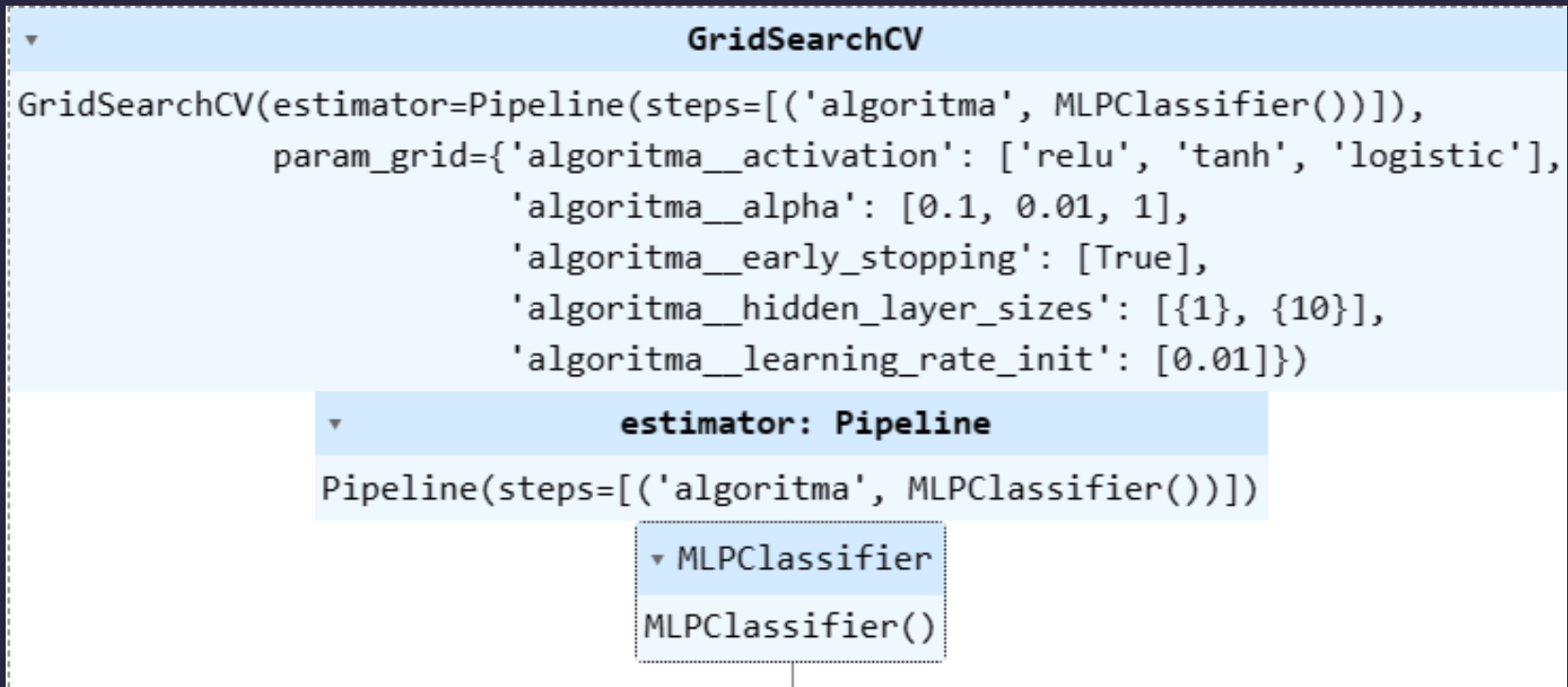
```
# melakukan fitting dan transformasi pada dokumen  
count_vect.fit(df_vektor)
```

```
# melihat hasil representasi bag of words  
X = count_vect.fit_transform(df_vektor)
```

Train - Test

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.20, random_state = 40)
```

Model NN



Model Neural Network

```
{'algoritma__activation': 'logistic',  
 'algoritma__alpha': 0.01,  
 'algoritma__early_stopping': True,  
 'algoritma__hidden_layer_sizes': {10},  
 'algoritma__learning_rate_init': 0.01}
```

Evaluasi Model (Data Training)

	precision	recall	f1-score	support
negative	0.91	0.96	0.94	2775
neutral	0.95	0.89	0.92	915
positive	0.97	0.96	0.97	5110
accuracy			0.95	8800
macro avg	0.95	0.94	0.94	8800
weighted avg	0.95	0.95	0.95	8800

Menggunakan data training

Evaluasi Model (Data Testing)

	precision	recall	f1-score	support
negative	0.79	0.85	0.82	661
neutral	0.82	0.70	0.76	233
positive	0.92	0.91	0.91	1306
accuracy			0.87	2200
macro avg	0.84	0.82	0.83	2200
weighted avg	0.87	0.87	0.87	2200

Menggunakan data testing

Kesimpulan



Accuracy

Akurasi keseluruhan dari model Anda adalah 95%. Ini menunjukkan bahwa model memiliki tingkat keakuratan yang tinggi dalam melakukan klasifikasi.

Precision

Precision mengukur sejauh mana hasil positif yang diprediksi oleh model adalah benar. Untuk kelas "negative", precision adalah 0.91, yang berarti sekitar 91% dari prediksi yang diklasifikasikan sebagai "negative" oleh model adalah benar. Untuk kelas "neutral", precision adalah 0.95, yang berarti sekitar 95% prediksi "neutral" adalah benar. Untuk kelas "positive", precision adalah 0.97, yang berarti sekitar 97% prediksi "positive" adalah benar.

Recall

Recall mengukur sejauh mana model dapat mengidentifikasi dengan benar contoh-contoh positif. Untuk kelas "negative", recall adalah 0.96, yang berarti sekitar 96% contoh "negative" yang ada telah diidentifikasi dengan benar oleh model. Untuk kelas "neutral", recall adalah 0.89, yang berarti sekitar 89% contoh "neutral" telah diidentifikasi dengan benar. Untuk kelas "positive", recall adalah 0.96, yang berarti sekitar 96% contoh "positive" telah diidentifikasi dengan benar.

F1-score

F1-score adalah ukuran gabungan dari precision dan recall yang memberikan informasi tentang keseimbangan antara keduanya. Untuk kelas "negative", f1-score adalah 0.94. Untuk kelas "neutral", f1-score adalah 0.92. Untuk kelas "positive", f1-score adalah 0.97. Semua nilai f1-score menunjukkan tingkat klasifikasi yang baik untuk masing-masing kelas..

Macro Average

Macro average menghitung rata-rata dari precision, recall, dan f1-score untuk semua kelas dengan bobot yang sama. Macro average precision, recall, dan f1-score semua adalah sekitar 0.95, menunjukkan kinerja yang seimbang di antara semua kelas.

Weighted Average

Weighted average juga menghitung rata-rata dari precision, recall, dan f1-score untuk semua kelas, tetapi dengan bobot yang berdasarkan jumlah sampel dalam setiap kelas. Weighted average precision, recall, dan f1-score semua adalah sekitar 0.95, menunjukkan kinerja yang seimbang dengan mempertimbangkan jumlah sampel dalam setiap kelas.

THANKS

