

LAPORAN TUGAS PEKAN 10 NLP : TEXT CLASSIFICATION

CLARISA HASYA YUTIKA | 1301174256 | IF 41 GAB01

1. Deskripsi Masalah

Implementasi metode Support vector machines (SVM) untuk klasifikasi analisis sentimen. Menggunakan dataset review femaledaily yang berjumlah 3080. Untuk evaluasi hasil klasifikasi menggunakan Accuracy, Precision, Recall, dan F-Measure. Kemudian akan di analisis pengaruh persentase data latih dan data uji, waktu pelatihan yang diperlukan, dan data yang salah diklasifikasikan.

2. Perancangan Sistem

• PreProcessing

Sebelum masuk tahap klasifikasi, data akan di cleaning terlebih dahulu dari simbol, angka, url, dan double space. Kemudian semua huruf di lowercase.

• TF IDF

Dilakukan pembobotan TF-IDF untuk menghitung bobot setiap kata yang umum digunakan. Pada eksperimen kali ini menggunakan 'max_features=1000', yang artinya hanya mengambil top 1000 terms.

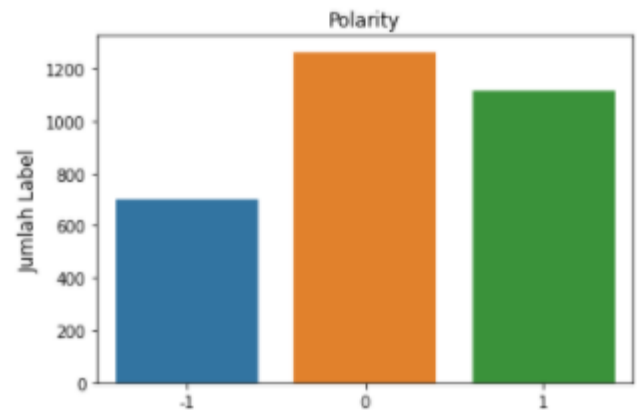
• Klasifikasi

Untuk klasifikasi analisis sentiment menggunakan metode algoritma SVM, dengan kernel linear.

3. Hasil dan Analisis

Dari 3080 data text review, dengan 3 kelas yang masing-masing berjumlah sebagai berikut :

0	1264	0 = Netral
1	1118	1 = Positive
-1	698	-1 = Negative



Dilakukan 2 eksperimen, yaitu perbedaan perbandingan split data train dan data test. Pada eksperimen pertama perbandingan data train : data test adalah 80 : 20, pada eksperimen kedua adalah 70 : 30.

Pada eksperimen pertama, untuk proses training dan predict dibutuhkan waktu sebagai berikut :

```
CPU times: user 4 µs, sys: 0 ns, total: 4 µs
Wall time: 8.82 µs
```

Kemudian hasil predict dari 616 data testing dihasilkan :

```
[[ 85 28 27]
 [ 12 184 64]
 [ 19 54 143]]
```

	precision	recall	f1-score	support
-1	0.73	0.61	0.66	140
0	0.69	0.71	0.70	260
1	0.61	0.66	0.64	216
accuracy			0.67	616
macro avg	0.68	0.66	0.67	616
weighted avg	0.67	0.67	0.67	616

LAPORAN TUGAS PEKAN 10 NLP : TEXT CLASSIFICATION

CLARISA HASYA YUTIKA | 1301174256 | IF 41 GAB01

Pada eksperimen kedua, untuk proses training dan predict dibutuhkan waktu sebagai berikut :

```
CPU times: user 2 µs, sys: 1e+03 ns, total: 3 µs
Wall time: 5.01 µs
```

Kemudian hasil predict dari 924 data testing dihasilkan :

```
[[125 45 34]
 [ 24 281 69]
 [ 38 94 214]]
```

	precision	recall	f1-score	support
-1	0.67	0.61	0.64	204
0	0.67	0.75	0.71	374
1	0.68	0.62	0.65	346
accuracy			0.67	924
macro avg	0.67	0.66	0.66	924
weighted avg	0.67	0.67	0.67	924

Dari kedua hasil eksperimen tidak terlalu menunjukkan perbedaan yang signifikan. Hanya proses training dan predict yang membutuhkan waktu yang berbeda dikarenakan jumlah data pada eksperimen pertama lebih banyak sehingga membutuhkan waktu lebih banyak. Kemudian nilai akurasi pun sama, hanya ada perbedaan tipis di nilai precision, recall, dan f1-measure.

Dari table confusion matrix cukup lumayan banyak data yang tidak sesuai dengan actual labelnya. Sehingga nilai precision, recall, f1-measure tidak terlalu bagus. Nilai precision rendah maka model tidak cukup baik untuk memprediksi data yang positif, sedangkan nilai recall rendah maka model tidak cukup mengenali data yang positif. Kemudian nilai f1-measure adalah harmonic mean dari precision

dan recall dikarenakan dapat terjadi tradeoff antara precision dan recall.

Dikarenakan jumlah data antara masing-masing kelas tidak terlalu jauh jumlahnya, maka nilai precision recall dan f1-measure masing-masing kelas tidak terlalu jauh

Cukup lumayan banyak data yang salah prediksi. Hal ini dapat terjadi karena terdapat beberapa review yang ambigu. Dalam satu review membahas kelebihan dan kekurangan suatu produk sehingga dapat terjadi salah pelabelan. Ada beberapa terms yang termasuk positif atau negatif, namun dalam beberapa kalimat terdapat penambahan negasi di depan terms tersebut, misalkan 'tidak berjerawat', 'tidak cocok', 'tidak bagus', dll.

Berikut beberapa contoh data yang salah klasifikasi :

saya tidak pernah cocok menggunakan toner apa pun padahal toner dikatakan penting untuk perawatan kulit terutama kulit berminyak dan pori pori besar saya sudah mencoba berbagai merek toner tapi tetap tidak cocok jika tidak kulit terasa sangat kering jerawat mengelupas dll setiap kali menggunakan toner ini pertama kalinya saya bertemu toner yang benar benar cocok dan kulit saya tidak mengeluh sama sekali toner ini juga bisa dikatakan sebagai penguat fte lho sehingga akan bekerja lebih baik jika menggunakan toner ini sebelum menggunakan fte membeli kembali iya	1	-1
ini adalah sejenis serum anti penuaan untuknya di atas an saya penasaran dengan merk biokos ini karena ada di salon kecantikan saya mencoba serumnya secara langsung hasilnya tidak memuaskan saya dan akhirnya mencoba yang lain	-1	1
menyukai tekstur yang ringan dan cepat menyerap cukup melembabkan untuk kulit kering tapi untuk memutihkan tidak terlalu menjadi masalah mungkin hanya untuk mempertahankannya bukan untuk memperbaiki harga yang terjangkau sangat suka menggunakan banyak ha ha	0	1
krim bb yang saya coba untuk pertama kalinya saya tidak merekomendasikan menggunakan ini untuk banyak bekas jerawat bagi saya cakupannya juga kurang warnanya memang pas dengan kulit saya tapi kulit saya jadi sangat berminyak baunya oke menurut saya tidak terlalu harum pas produk ini sebenarnya cocok karena tidak membuat breakout cuma sayang bikin wajahmu berminyak	0	-1

LAPORAN TUGAS PEKAN 10 NLP : TEXT CLASSIFICATION

CLARISA HASYA YUTIKA | 1301174256 | IF 41 GAB01

sebenarnya saya selalu mencari toner aha yang bagus untuk pemakaian sehari hari dan saya sudah lama mendengar tentang pixi glow tonic tapi baru setahun sejak saya berpikir untuk membelinya kemasannya bagus aman hanya sangat besar jadi apos tidak oke untuk bepergian mungkin pixi juga banyak mendapatkan keluhan tersebut dari pelanggan akhirnya mereka merilis versi ml sangat lezat tidak menyengat dan benar benar segar setelah digunakan tertulis dalam deskripsi produk untuk am amp pm gunakan saya pikir kulit saya terlalu kering untuk menggunakan ini dua kali hari akhirnya gunakan saja untuk malam kurangi jerawat kecil saya u tapi saya juga bingung dengan jerawat di dahi yang semakin membesar setelah menggunakan ini kulit saya kombinasi sangat kering di hidung dan pipi menggunakan kulit kering ini itu menjadi jauh lebih tidak terlihat tapi ya ada jerawat baru yang muncul kesal sepertinya akan membeli kembali tapi botol besar ini juga belum jadi meski kita pakai secara rutin saya berharap rupiah menguat terhadap dolar ya menyedihkan bahwa anda merasa seperti dirampok membeli ini jauh lebih mahal	1	-1
saya rasa kulit saya tidak cocok dengan produk sk ii essence ini membuat kulit saya yang awalnya hanya jerawat sesekali menjadi breakout terus menerus setelah pemakaian dihentikan kulit aman setelah saya menggunakan produk perawatan jerawat dan karena masih penasaran saya kembali mencoba essence sk ii dan hasilnya breakout lagi dan biasanya muncul di area sekitar dahi dan samping hidung	-1	1
suka lipcolor balm ini karena benar benar berpigmen dan melembabkan bibir warnanya paling rose karena cocok untuk pemakaian sehari hari walaupun cepat hilang tapi noda di bibir jadi walaupun sudah hilang masih ada warna di bibir yang tidak membuatnya pucat kemasan bagus handy juga mudah dibawa dan tidak memakan tempat pengaplikasiannya juga mudah karena ujungnya lancip sehingga gampang untuk sudut bibir recommended	1	0
air mawar paporit saya jika misalnya saya ingin masker yang merupakan masker bubuk saya pasti akan menggunakan ini harum tapi jujur saya tidak merasakan apa efek air mawar apakah anda mau dr viva atau sari anda tidak merasakan efeknya hanya baunya lebih bau dari air biasa	1	0
oke sunscreen ini murah simak ulasannya banyak guys yang suka tapi maaf kurang keren pas bikin kilang minyak saya tidak tahu ide saya yang terlalu sensitif untuk produk ini atau apa tapi tetap saja baunya tidak enak membuat saya muak apalagi kalau baru dipakai jadi saya sangat mencium baunya maaf batuk	0	-1
banyak orang pake jasaeng essence tapi saya tidak tidak sesuai dengan kulit saya yang berminyak karena bikin muka saya makin berminyak dan wajah saya kelihatan kusam bikin berjerawat padahal tidak terlalu heboh buat yang mau pake jasaeng essence tapi kurang cocok bisa coba yang herblinic sulwhasoo saya lebih cocok pake daripada jasaeng minusnya juga harga tinggi karena memang di korea sendiri whoo merek skincare premium yang pasti apakah kemasannya cantik	0	-1
sambutan dari blogger sangat bagus jadi saya membeli s el estee lauder advance night repair sehingga tidak benar benar terhapus jika tidak sesuai hehehe pake tiap minggu nggak masalah pokoknya lembab aja tidak seperti review blogger ya yang penting saya sudah mencoba produk ini	-1	1
parah banget kecewa banget sama sunblock dari vaseline ini gue pake pas lagi snorkeling kan masih terbakar sinar matahari kulitnya juga gosong semua tidak ada efeknya sama sekali seperti pakai sunblock lengket banget di kulit saya pokoknya tidak oke sama sekali saya sama sekali tidak ada niat untuk membelinya kembali	-1	0
perbaiki sangat suka pixi glow tonic awalnya saya takut karena banyak orang purging dulu dan kesemutan dulu tapi ternyata wajah saya pas sekali dengan pixi belum ada kesemutan purging dulu setelah digunakan selama bulan hasilnya sudah tidak terlihat lagi wajah setelah menggunakan pixi lebih cerah halus glowing juga tapi memang bikin wajah saya kering kulit saya normal kombinasi setelah pakai pixi glow cenderung kering terus apakah itu menyakiti matamu atau tidak meski area mata sudah dihindari juga namun tetap saja menyengat ke mata kemasannya tidak tahu kenapa sering bocor jadi tidak ramah perjalanan	-1	1
saat wajah bermasalah dengan jerawat h ir semua skincare berubah menjadi skincare khusus untuk kulit berjerawat dan pilihan saya jatuh ke tabir surya dari jerawat yaitu acnes uv tint selain berfungsi sebagai tabir surya produk ini juga bisa difungsikan sebagai alas riasan harganya juga murah dan sangat mudah didapatkan namun ternyata produk ini kurang cocok untuk kulit kering saat saya menggunakan produk ini kulit saya semakin kering dan mengelupas di beberapa bagian selain itu teksturnya membuat tabir surya ini sulit membaaur di kulit meski diklaim cocok untuk semua jenis kulit sepertinya tabir surya ini akan lebih cocok untuk kulit berminyak	0	-1
saya suka tekstur dan aromanya tapi setelah dibilas produk ini meninggalkan residu berminyak seperti itu dibilas lagi tidak lepas kandungan minyaknya juga membuat lantai kamar mandi licin jadi agak merepotkan tidak akan membeli kembali	-1	0