

ANALISIS SENTIMEN REVIEW INSTITUSI PENDIDIKAN TINGGI DARI FACEBOOK MENGGUNAKAN JARINGAN SARAF TIRUAN

Habibullah Akbar

Fakultas Ilmu Komputer, Universitas Esa Unggul
Jalan Arjuna Utara No. 9, Kebon Jeruk, Jakarta, 11510E-mail:
habibullah.akbar@esaunggul.ac.id

Abstract

Nowadays, the development of digital technology allows the higher institutions such as universities to observe public opinion/sentiment through social media. Unfortunately, the analysis of the data is still done manually. In this study, we develop an automated sentiment analysis framework to predict the polarity of public opinion on several higher education institutions in Indonesia. The data used comes from the comments that are available on the official page of the institution's Facebook. We make comparisons of several combined models between classifiers (Logistic Regression, Support Vector Machine, Random Forest Classifier and Neural Network) with feature extraction methods (Count vectorizer and TF-IDF). The experimental results show that the best sentiment analysis results are given by the Neural Network model which uses the TF-IDF weighting technique with an accuracy of more than 86%. These results indicate that the proposed sentiment analysis framework is potential to be used to find out public opinion on higher education institutions in Indonesia.

Keywords: *Data mining, sentiment analysis, machine learning.*

Abstrak

Di masa sekarang ini, perkembangan teknologi digital memungkinkan institusi pendidikan tinggi seperti universitas untuk mengetahui opini/sentimen masyarakat lewat media sosial. Sayangnya, analisis terhadap data tersebut masih dilakukan secara manual. Pada penelitian ini, kami membangun sebuah kerangka analisis sentimen otomatis untuk memprediksi polaritas dari opini masyarakat terhadap beberapa institusi pendidikan tinggi di Indonesia. Data yang digunakan berasal dari komentar-komentar yang berada pada halaman resmi Facebook institusi yang digunakan untuk penelitian. Kami melakukan perbandingan dari beberapa model gabungan antara klasifikator (*Logistic Regression, Support Vector Machine, Random Forest Classifier* dan *Neural Network*) dengan metode ekstraksi fitur (*Count vectorizer* dan *TF-IDF*). Hasil eksperimen menunjukkan bahwasanya hasil analisis sentimen terbaik diberikan oleh model *Neural Network* yang menggunakan teknik pembobotan *TF-IDF* dengan akurasi lebih dari 86%. Hasil ini menunjukkan bahwa kerangka analisis sentimen yang diusulkan berpotensi untuk digunakan untuk mengetahui opini masyarakat terhadap institusi pendidikan tinggi yang ada di Indonesia.

Kata kunci: *Data mining, analisis sentimen, machine learning.*

Pendahuluan

Pada situasi saat ini, internet telah memungkinkan institusi pendidikan tinggi seperti universitas dan sekolah tinggi untuk mengumpulkan sejumlah besar data yang berasal dari *stakeholder* seperti siswa, teman, keluarga, pengusaha, dan pihak otoritas pendidikan lewat media sosial, blog, surat kabar *online*, bahkan forum *online* (Abbasi, Chen, & Salem, 2008).

Karena jumlah data tekstual yang tersedia pada media sosial sangatlah besar, maka hampir tidaklah mungkin bagi institusi pendidikan tinggi untuk menganalisis jenis data ini secara manual. Untuk melakukan analisis yang lebih praktis, maka mereka seyogyanya dapat menggunakan pendekatan komputasi terhadap data tekstual dari postingan pada media sosial untuk mengidentifikasi polaritas sentimennya. Polaritas disini terbagi menjadi positif (komentar yang

baik), netral dan negatif (komentar yang tidak baik). Pendekatan komputasi ini akan memungkinkan institusi untuk melakukan analisis opini secara otomatis.

Diantara manfaat utama dari analisis sentimen akan memungkinkan lembaga pendidikan tinggi untuk meningkatkan kegiatan pengajaran mereka (Altrabsheh, Gaber, & Cocea, 2013). Dengan melakukan analisis terhadap data ini, institusi-institusi dapat mengetahui sentimen para *stakeholder* sehingga dapat menjadi masukan untuk penyusunan strategi dan taktik yang lebih baik bagi masyarakat. Sekiranya institusi-institusi pendidikan tinggi dapat merespon sentimen dari para *stakeholder* secara cepat maka hal ini mungkin akan berdampak positif dan signifikan terhadap pertumbuhan institusi itu sendiri.

Diantara media sosial yang berpotensi untuk digunakan dalam kegiatan akademik adalah *Twitter* (Mollett, Moran, & Dunleavy, 2011). *Twitter* telah digunakan untuk mengevaluasi 9 universitas di Jerman (Abdelrazeq, Janßen, Tummel, Jeschke, & Richert, 2016). Namun selain *Twitter*, media sosial lain seperti *Facebook* dan *Whatsapp* juga dapat digunakan untuk mengetahui pendapat masyarakat terhadap dunia akademik. Akan tetapi, masih jarang penelitian yang menggunakan kedua media tersebut untuk menganalisis sentimen masyarakat terhadap universitas. Pada penelitian ini, kami akan membahas mengenai teknik-teknik komputasi yang dapat secara otomatis melakukan analisis sentimen terhadap komentar pada situs Facebook. Kami akan membahas proses mulai dari teknik-teknik pemrosesan data tekstual hingga klasifikasi kata berdasarkan pendekatan kecerdasan buatan.

Untuk itu, kami mendefinisikan analisis sentimen sebagai sebuah proses eksplorasi dan eksploitasi data yang berasal dari media sosial untuk mengetahui opini, perasaan ataupun emosi para *stakeholders* (semua pihak terkait baik mahasiswa, keluarga ataupun masyarakat secara umum) terhadap institusi pendidikan tinggi. Sentimen tersebut dapat berasal dari upaya pemasaran langsung dari institusi atau pengalaman yang dialami *stakeholders* itu sendiri.

Berbagai studi analisis sentimen sebenarnya telah dilakukan dalam berbagai bentuk seperti evaluasi program pendidikan tinggi untuk kegiatan akademis. (Koehler, Greenhalgh, & Zellner, 2015) menunjukkan potensi penggunaan analisis sentimen di lingkungan pendidikan untuk menganalisis bagaimana masyarakat menanggapi kegiatan akademik. Mereka mengevaluasi keramahan dua konferensi teknologi pendidikan. Mereka menemukan bahwasanya sebagian besar respon bersifat netral. Sedangkan 34% memberikan tanggapan yang positif dan 2% untuk tanggapan negatif. Contoh lain dari penerapan analisis sentimen telah diusulkan oleh (Cummins, Burd, & Hatch, 2010). Mereka menggunakan *tag* umpan balik untuk menghasilkan sumber materi pembelajaran untuk kursus pemrograman.

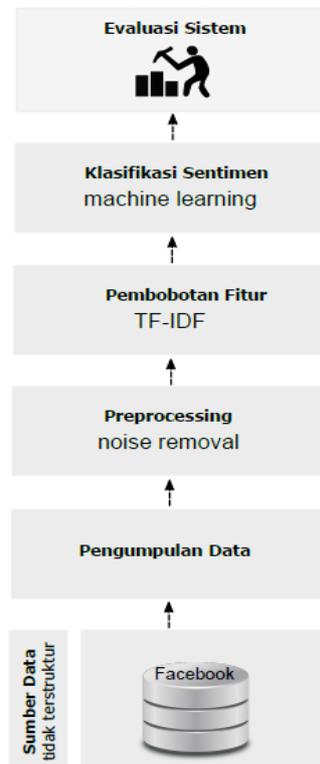
Selain kegiatan pembelajaran secara tradisional di ruang kelas, analisis sentimen juga telah diterapkan untuk meningkatkan kegiatan *e-learning*. (Kechaou, Ammar, & Alimi, 2011) menggunakan analisis sentimen untuk mengevaluasi respon siswa berkaitan dengan blog dan sistem manajemen pembelajaran. Mereka menggunakan model pembelajaran gabungan berdasarkan *hidden markov model* dan *support vector machine (SVM)* untuk mengklasifikasikan sentimen dari postingan media sosial. Namun, mereka menemukan hasil akurasi yang kurang memuaskan karena penggunaan gaya bahasa dan kata-kata yang tidak baku pada konten postingan.

Studi lain dilakukan oleh (Wen, Yang, & Rose, 2014). Mereka mengukur dampak sentimen pada postingan di forum *massive open online course* terhadap tingkat atrisi siswa (ukuran yang digunakan untuk mengukur kemampuan institusi untuk mempertahankan siswa agar tidak keluar atau berpindah ke tempat lain). Di luar dugaan, mereka tidak menemukan adanya korelasi yang signifikan antara sentimen yang diekspresikan dari siswa terhadap tingkat atrisi. Kendatipun demikian, pengetahuan sentimen *stakeholders* tetap diperlukan pihak institusi untuk mengetahui tingkat kepuasan ataupun respon masyarakat terhadap kebijakan-kebijakan yang diambil.

Namun seperti yang telah disebutkan sebelumnya, kebanyakan dari penelitian-penelitian diatas masih bergantung pada platform media sosial *Twitter*. Belum banyak penelitian yang membahas secara khusus perbandingan model-model prediksi berdasarkan data komentar dari Facebook. Oleh sebab itu, diperlukan sebuah penelitian tersendiri untuk membahas masalah ini.

Metode Penelitian

Pada bagian ini, kami mempresentasikan pendekatan kami untuk menemukan pendapat *stakeholders* terkait dengan lembaga pendidikan tinggi melalui analisis sentimen. Pendekatan ini terdiri dari lima tahap utama: pengumpulan data, *preprocessing*, ekstraksi fitur, klasifikasi sentimen dan evaluasi sistem seperti yang diilustrasikan pada gambar 1.



Gambar 1

Lima tahapan utama dalam metode analisis sentimen yang digunakan untuk klasifikasi komentar *Facebook*.

Tahapan pengumpulan data

Kami mengumpulkan data dari halaman *Facebook* resmi universitas-universitas di Indonesia. Pemilihan universitas dilakukan berdasarkan situs Sinta (*Science and Technology Index*) yang berasal dari Kementerian Riset, Teknologi, Dan Pendidikan Tinggi Republik Indonesia yang dapat diakses di <http://sinta.ristekdikti.go.id/>.

Sebenarnya, *Facebook* merilis antarmuka pemrograman aplikasi untuk membagikan data secara otomatis dengan menggunakan *REST API* ataupun *web service*. Namun demikian sejak kasus skandal *Cambridge Analytica*, maka *Facebook* sangat berhati-hati dalam membuka akses data melalui API. Oleh karena data dikumpulkan manual langsung dari halaman resmi universitas. Data yang sudah dikumpulkan dari *Facebook* akan diberikan label secara manual yaitu label positif, negatif, atau netral.

Tahapan *preprocessing*

Teks yang diarsipkan akan diubah menjadi daftar kata menggunakan tokenisasi. Tokenisasi merupakan proses yang sangat penting dari setiap eksplorasi data teks. Proses ini akan membagi atau mensegmentasi sebuah rangkaian kata dalam sebuah kalimat menjadi potongan kata per kata yang disebut juga sebagai token. Selain kata, token juga dapat berupa suatu frase, suatu angka, ataupun suatu tanda baca.

Untuk menghapus kata-kata yang tidak berhubungan langsung dengan opini atau sentimen, kami menggunakan teknik *preprocessing* lainnya seperti penghilangan *username*, alamat situs, tanda baca, emotikon dan menjadikan seluruh teks menjadi *lower case*.

Tahap pembobotan fitur

Data yang sudah diproses tentunya masih mengandung kata-kata yang sulit untuk diproses oleh kebanyakan algoritma klasifikasi *machine learning*. Hal ini disebabkan sebagian besar algoritma klasifikasi membutuhkan data yang direpresentasikan sebagai nilai numerik. Oleh karena itu, kami akan mengubah data tekstual yang telah melalui tahapan preprocessing kedalam bentuk numerik menggunakan teknik *term-frequency-inverse document frequency (TF-IDF)*.

Metode ini akan menghitung nilai *term frequency (TF)* dan *inverse document frequency (IDF)* pada setiap kata yang telah ditokenisasi. Rumus yang digunakan untuk melakukan pembobotan adalah:

(1)

$$W_{t,d} = TF_{t,d} \times \left(\frac{IDF}{DF} \right) \quad (2)$$

dimana:

TF = *Term Frequency*

IDF = *Inverse Document Frequency*

DF = *Document Frequency*

N = jumlah dokumen

t = kata

d = komentar

Selain TF-IDF, kami juga menggunakan *Count vectorizer (Cvec)* dimana setiap kata yang telah ditokenisasi akan dihitung frekuensi kemunculannya.

Tahap klasifikasi sentimen

Tahap ini akan mengukur polaritas sentimen. Kami akan menggunakan metode-metode *machine learning* seperti *Logistic Regression*, *Support Vector Machine (SVM)*, *Random Forest Classifier (RFC)* dan *Neural Network (NN)* untuk mengklasifikasikan komentar Facebook sebagai kategori positif, negatif atau netral.

Logistic Regression

LR adalah model statistik yang menggunakan fungsi *logistic*. Model ini hanya memodelkan probabilitas output dan sebenarnya bukan ditujukan untuk melakukan klasifikasi. Namun, *LR* dapat digunakan untuk membangun model klasifikator dengan memilih nilai *cutoff*, kemudian mengklasifikasikan input yang lebih besar dari *cutoff* sebagai satu kelas, sedangkan nilai di bawah *cutoff* sebagai kelas lainnya. Selain itu, *LR* juga telah digunakan untuk analisis sentimen terhadap data perhotelan (Hamdan, Bellot, & Bechet, 2015).

Pada kasus analisis sentimen yang lebih dari 2 kelas, maka kami menggunakan multinomial *LR*. Multinomial *LR* menggunakan skema *one-vs-rest* (misal kelas positif atau bukan, dst). Parameter yang diatur adalah kekuatan regularisasi *C* untuk mencegah terjadinya *over-fitting* yaitu model yang terlalu kaku terhadap data ataupun *under-fitting* yaitu model yang terlalu sederhana sehingga tidak sesuai dengan data. Pada nilai *C* yang besar, kekuatan regularisasi akan menurun yang memungkinkan model untuk meningkatkan kerumitannya, dan sebaliknya. Pada penelitian ini, nilai *C* yang diuji adalah 0.01, 0.1, 1, 10 dan 100.

Support Vector Machine

Serupa dengan *LR*, *SVM* biasanya digunakan untuk memisahkan 2 kelas data yang berbeda. Namun *SVM* juga dapat digunakan untuk memisahkan kelas yang lebih dari 2 yang disebut juga sebagai *multi-class SVM*. *SVM* yang digunakan pada penelitian ini mengimplementasikan pendekatan *one-against-one* (Knerr, Personnaz, & Dreyfus, 1990). Detil komputasi dapat dilihat pada paper tersebut. Jika n_{class} adalah jumlah kelas, maka model klasifikator $n_{class} \cdot (n_{class} - 1) / 2$ akan dibangun untuk masing-masing model dengan melatih data dari dua kelas.

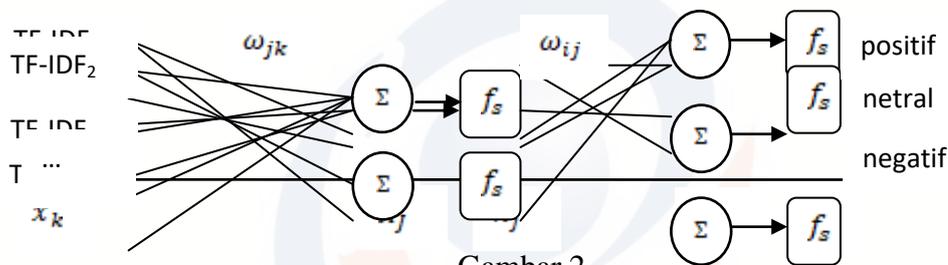
Akurasi model yang akan dihasilkan dari proses training SVM bergantung pada parameter yang digunakan seperti fungsi kernel. Fungsi kernel yang akan digunakan adalah kernel *linear*, *polynomial* dan *radial basis function*. Adapun parameter gamma sebagai konstan 1.

Random Forest Classifier

RFC adalah estimator yang menyesuaikan jumlah dari gabungan dari beberapa pengklasifikasi *decision tree* pada berbagai sub-sampel dataset dan menggunakan rata-rata untuk meningkatkan akurasi prediktif dan kontrol *over-fitting*. Ukuran sub-sampel selalu sama dengan ukuran sampel input asli. Parameter yang akan untuk diuji pada penelitian ini adalah *number of trees in the forest* (jumlah pohon). Nilai yang diuji pada penelitian adalah 100, 200, 300, 400 dan 500. Adapun nilai dari parameter jumlah level maksimal pada setiap percabangan diset 20.

Neural Network

Arsitektur jaringan saraf yang digunakan dalam penelitian ini adalah jaringan saraf *feedforward* dua lapis yang biasanya dilatih dengan algoritma *Levenberg-Marquardt* ataupun *gradient-based* (Akbar, Suryana, & Sahib, 2011). Namun pada penelitian ini, kami menggunakan algoritma *L-BFGS*. Jaringan ini memiliki dua lapisan yaitu lapisan tersembunyi dan lapisan luaran. Arsitektur jaringan saraf ditunjukkan pada Gambar 2.



Gambar 2
Arsitektur Neural Network

Pertama, input jaringan adalah hasil transformasi teks yang sudah dibersihkan menjadi angka numerik menggunakan *TF-IDF* ataupun *count vectorizer*. Untuk setiap neuron input x_k , maka koneksi ke layer tersembunyi membobotkan input tersebut dengan bobot w_{jk} dimana $k = 1, 2, \dots, K$ mewakili jumlah total fitur input.

Dengan demikian model matematika umpan-maju umum diberikan sebagai berikut

$$h_j = f_s(X_j) \quad (3)$$

dimana

$$X_j = \sum_{k=1}^N \omega_{jk} x_k \quad (4)$$

x_k , h_j dan y_i masing-masing mewakili input, *hidden* dan *neuron output*. Persamaan di atas adalah model untuk menghitung neuron tersembunyi h_j dari input neuron x_k . Oleh karena itu, output neuron y_i dapat diperoleh dengan menggunakan

$$y_i = f_s(H_j) \quad (5)$$

$$H_j = \sum_{j=1}^M \omega_{ij} h_j \quad (6)$$

Fungsi aktivasi f_s yang digunakan adalah sigmoid $f(x)$ dipilih dalam penelitian ini dihitung seperti di bawah ini

$$f(x) = \frac{1}{1 + e^{-\tau x}} \quad (7)$$

di mana parameter τ mengontrol kecuraman kurva dan x adalah masukkan dari luaran *neuron* pada lapisan sebelumnya. Parameter yang diuji pada penelitian ini adalah jumlah neuron M pada *hidden*

layer. Nilai yang diuji yaitu 20, 40, 60, 80 dan 100. Adapun jumlah iterasi diset maksimum 1000 dan *learning rate* = 0.01.

Evaluasi Model Prediksi Sentimen

Metrik evaluasi yang akan digunakan untuk mengukur tingkat akurasi adalah *confusion matrix* 3 x 3 seperti yang diberikan pada tabel 1. Pada eksperimen, evaluasi analisis sentimen akan menggunakan *k-fold cross validation* (CV), dengan nilai *k* = 5. Adapun kami telah mencoba secara garis besar nilai *k* lain, namun tidak memberikan perbedaan hasil yang signifikan.

Tabel 1
Confusion Matrix yang digunakan untuk mengevaluasi model.

Label	Prediksi		
	Terdeteksi Positif	Terdeteksi Netral	Terdeteksi Negatif
Positif	TP_P	FA _P	FN _P
Netral	FP _A	TA_A	FN _A
Negatif	FP _N	FA _N	TN_N

Berdasarkan tabel 1, metrik yang akan digunakan untuk penelitian ini adalah akurasi, yaitu Jumlah prediksi yang benar (**TP_P** + **TA_A** + **TN_N**) dibagi dengan jumlah total prediksi yang dibuat (**TP_P** + FA_P + FN_P + FP_A + **TA_A** + FN_A + FP_N + FA_N + **TN_N**) dimana

- **TP_P** = Hasil prediksi positif dan data memang berlabel sentimen positif.
- FA_P = Hasil prediksi netral padahal data berlabel sentimen positif.
- FN_P = Hasil prediksi negatif padahal data berlabel sentimen positif.
- FP_A = Hasil prediksi positif padahal data berlabel sentimen netral.
- **TA_A** = Hasil prediksi netral dan data memang berlabel sentimen netral.
- FN_A = Hasil prediksi negatif padahal data berlabel sentimen netral.
- FP_N = Hasil prediksi positif padahal data berlabel sentimen negatif.
- FA_N = Hasil prediksi netral padahal data berlabel sentimen negatif.
- **TN_N** = Hasil prediksi negatif dan data memang berlabel sentimen negatif.

Hasil dan Pembahasan

Pada bagian ini, hasil dari tahapan-tahapan yang telah dijelaskan pada bagian sebelumnya akan dipaparkan sebagai berikut.

Hasil Pengumpulan Data

Pengambilan data dilakukan dengan situs Facebook yang diambil sebanyak 4 universitas.

- Universitas Indonesia (UI) - <https://www.facebook.com/ui.ac.id/>
- Universitas Sebelas Maret (UNS) - <https://www.facebook.com/UNSOOfficial/>
- Universitas Telkom (Tel-U) - <https://www.facebook.com/telkomuniversity/>
- Universitas Islam Indonesia (UII) - <https://www.facebook.com/Universitas.Islam.Indonesia/>

Kampus dipilih berdasarkan situs Sinta (Science and Technology Index) yang berasal dari Kementerian Riset, Teknologi, Dan Pendidikan Tinggi Republik Indonesia yang dapat diakses di <http://sinta.ristekdikti.go.id/>. Berdasarkan skor 3 tahun terakhir, Universitas Indonesia, Universitas Sebelas Maret, Universitas Telkom, dan Universitas Islam Indonesia masing-masing berada pada ranking 1, 11, 19, dan 35.

Awalnya kami memilih universitas dalam rentang ranking 10. Universitas Indonesia dan Universitas Sebelas Maret merupakan kampus publik. Namun Universitas Negeri Semarang yang berada pada ranking 21 juga merupakan kampus publik, sehingga kami memilih Universitas Telkom sebagai representasi kampus swasta. Selain itu, kampus ini juga merupakan kampus swasta terbaik di Indonesia versi skor Sinta.

Kampus berikutnya harusnya adalah Universitas Lampung. Namun, kampus ini tidak memiliki halaman facebook resmi. Demikian pula dengan universitas-universitas berikutnya (Universitas Lambung Mangkurat, Universitas Negeri Padang, dan Universitas Halu Oleo). Akhirnya, kami baru mendapatkan kampus berikutnya yang memiliki halaman facebook aktif beserta ulasan dari akademisi pada Universitas Islam Indonesia yang berada pada ranking 35. Selain itu, kampus ini juga merupakan kampus swasta sehingga pada dataset yang kami dapatkan adalah 2 kampus publik dan 2 kampus swasta. Selain kampus-kampus tersebut kami juga telah mengumpulkan komentar dari kampus lainnya, namun jumlah komentar tidak sampai 100 sehingga diabaikan pada penelitian ini. Adapun untuk keempat kampus tersebut kami telah mendapatkan setidaknya 100 data untuk masing-masing kampus tersebut. Setelah seluruh data berhasil didapatkan, kami memberikan label pada komentar-komentar yang didapatkan dari situs Facebook tersebut secara manual. Pemberian label tersebut membagi data menjadi tiga kategori yaitu sentimen positif, netral dan negatif.

Hasil *Preprocessing* Data

Data yang didapatkan dari Facebook merupakan data mentah yang masih mengandung kata ataupun karakter yang tidak berhubungan dengan sentimen. Kata ataupun karakter tersebut mencakup garis miring, @, link, emoticon, simbol dan lain sebagainya. *Preprocessing* data bertujuan untuk menghilangkan kata-kata ataupun karakter-karakter tersebut sebagaimana yang telah dijelaskan pada bagian 2.2. Perbandingan antara komentar original dan yang telah diproses dapat dilihat di tabel 2. Seperti yang diperlihatkan, hasil preproses akan menghilangkan karakter-karakter tertentu seperti emotikon, tanda kutip, dan tanda tanya. Selain itu, hasil preproses juga telah membuat semua huruf menjadi huruf kecil.

Tabel 2
Hasil *preprocessing*

Original	Preprocessed
Peringkat 133 di asia...ini ok atau tidak ok?	peringkat 133 di asia ini ok atau tidak ok
Sebagai mantan fe ka es gw mah kaga heran 😊	sebagai mantan fe ka es gw mah kaga heran
Naaahhh looo mulai mencurigakan???? Dulu atau...	naaahhh looo mulai mencurigakan dulu ataupun s...

Hasil Ekstraksi Fitur

Setelah melewati tahapan *preprocessing*, maka dibentuk *bag-of-words* agar terbentuk tabel yang berisi daftar kata-per-kata. Daftar ini diperlukan karena klasifikator tidak dapat melakukan proses *training* secara langsung terhadap data tekstual. Dari daftar *bag-of-words* akan dilakukan perhitungan jumlah kata yang muncul (*count vectorizer*) maupun perhitungan bobot TF-IDF. Contoh hasil dari *bag-of-words* yang terbentuk diberikan pada gambar 3.

```

1 print( cvec.fit_transform(x_train).todense() )
2 print( cvec.vocabulary_ )

[[0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]
 ...
 [0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]]

{'saya': 1906, 'sangat': 1897, 'bersyukur': 387, 'belajar': 315, 'di': 591, 'uii': 2322, 'dari': 560, 'awal': 260, 'mau': 1323, 'menyerah': 1428, 'sampai': 1894, 'semester': 1981, 'smester': 2051, 'akhir': 151, 'semangat': 1977, 'menuntut': 1422, 'ilmu': 900, 'ingat': 926, 'anak2': 197, 'pada': 1584, 'best': 397, 'one': 1567, 'ever': 740, 'take': 2138, 'the': 2250, 'science': 1917, 'trmkasih': 2296, 'shabat': 2022, 'atas': 248, 'knpirmasi': 1151, 'nya': 1549, 'tman2shabat': 2272, 'ui': 2321, 'met': 1444, 'sre': 2086, 'tman2': 2271, 'smua': 2059, 'my': 1494, 'favorite': 756, 'university': 2353, 'but': 469, 'can': 476, 'not': 1539, 'reach': 1836, 'bangunan': 287, 'yang': 2430, 'luas': 1255, 'juneto': 1021, 'tempat': 2185, 'segudang': 1947, 'halo': 843, 'delia': 574, 'iyaa': 969, 'aku': 159, 'dulunya': 703, 'mahasiswa': 1272, 'fkb': 767, 'monggoo': 1469, 'welcome': 2408, 'top': 2284, 'class': 508, 'support': 2112, 'kpd': 1171, 'suami': 2101, 'tercinta': 2207, 'tuk': 2310, 'dokter': 679, 'alhamdulillah': 170, 'smg': 2052, 'ilmunya': 902, 'dpt': 692, 'berguna': 341, 'bagi': 270, 'maasyarakat': 1316, 'semua': 1989, 'amin': 194, 'yra': 2442, 'netral': 1514, 'kata2': 1068, 'yg': 2435, 'selalu': 1966, 'terucap': 2234, 'ketika': 1129, 'mengingat': 1406, 'kembali': 1104, 'dosen': 686, 'dekan': 571, 'pegawai': 1623, 'dan': 554, 'lainnya': 1197, 'semuanya': 1990, 'ramah': 1824, '13th': 24, 'lalu': 1200, 'sekarang': 1959, 'rasanya': 1834, 'kok': 115

```

Gambar 3

Kata-kata yang diekstrak dari komentar-komentar Facebook dan telah dihitung frekuensi kemunculannya pada dataset yang digunakan.

K-Fold Cross Validation

Setelah data melewati tahapan *preprocessing* dan ekstraksi fitur, maka data dibagi menjadi dua yaitu data *training* dan data *testing* (80% data *training*, 20% data *testing*). Data *testing* akan disimpan hingga proses *training* selesai dilakukan. Lebih lanjut lagi, data *training* dibagi menjadi beberapa lipatan berdasarkan metode *K-fold cross validation*. Pada eksperimen, kami menggunakan nilai $k = 5$. note: Kami juga telah mencoba nilai $k = 4$ dan $k = 5$, namun hasilnya tidak menunjukkan perbedaan yang signifikan.

Hasil Klasifikasi Logistic Regression

Tabel 3 berisi hasil akurasi dari analisis sentimen dengan menggunakan metode LR. Training dari LR dilakukan dengan beberapa varian nilai dari parameter C yaitu 0.01, 0.1, 1, 10 dan 100.

Tabel 3
Perbandingan akurasi LR terhadap fitur Cvec dan TF-IDF.

Universitas	Akurasi-Cvec					Akurasi-TF-IDF				
	Parameter: C									
	0.01	0.1	1	10	100	0.01	0.1	1	10	100
UI	0.7	0.75	0.8	0.8	0.75	0.7	0.7	0.75	0.75	0.85
UNS	0.81	0.81	0.857	0.81	0.857	0.81	0.81	0.81	0.81	0.857
Tel-U	0.8	0.8	0.8	0.85	0.85	0.7	0.7	0.75	0.9	0.9
UII	0.9	0.9	0.9	0.9	0.9	0.90	0.9	0.9	0.9	0.9
semua universitas	0.642	0.642	0.827	0.852	0.84	0.642	0.642	0.827	0.852	0.84
rata-rata	0.77	0.78	0.837	0.842	0.839	0.75	0.75	0.807	0.842	0.869

Dari tabel 3 dapat dilihat dengan menggunakan parameter $C = 10$, akurasi rata-rata dari LR terhadap seluruh percobaan dengan fitur Cvec menghasilkan akurasi sebesar 0.842 atau 84.2%. Sedangkan hasil akurasi untuk model gabungan LR dan fitur TF-IDF lebih baik yaitu sebesar 86.9% dengan parameter $C=100$.

Hasil Klasifikasi *Support Vector Machine*

Tabel 4 berisi hasil akurasi dari analisis sentimen dengan menggunakan metode SVM. Training dari SVM dilakukan dengan beberapa varian kernel yaitu kernel *linear*, *polynomial* dan *radial basis function*.

Tabel 4
Perbandingan akurasi SVM terhadap fitur Cvec dan TF-IDF

Universitas	Akurasi-Cvec			Akurasi-TF-IDF		
	Parameter: kernel					
	Linear	RBF	Polynomial	Linear	RBF	Polynomial
UI	0.75	0.65	0.35	0.8	0.75	0.75
UNS	0.857	0.81	0.762	0.857	0.81	0.81
Tel-U	0.85	0.7	0.8	0.85	0.75	0.7
UII	0.9	0.9	0.9	0.9	0.9	0.9
semua universitas	0.802	0.679	0.728	0.84	0.827	0.691
rata-rata	0.832	0.748	0.708	0.849	0.807	0.770

Dari tabel 4 dapat dilihat dengan menggunakan parameter kernel *linear*, SVM dapat menghasilkan akurasi rata-rata masing-masing 83.2% untuk fitur Cvec dan 84.9% untuk fitur TF-IDF. Adapun nilai terendah dicapai oleh kernel *polynomial* yang hanya dapat memberikan akurasi 70.8% untuk fitur dari Cvec dan 77% untuk hasil pembobotan TF-IDF.

Hasil Klasifikasi *Random Forest Classifier*

Tabel 5 berisi hasil akurasi dari analisis sentimen dengan menggunakan metode RFC. Training dari RFC telah dilakukan dengan beberapa varian nilai dari parameter *number of trees* yaitu 100, 200, 300, 400 dan 500.

Tabel 5
Perbandingan akurasi RFC terhadap fitur Cvec dan TF-IDF.

Universitas	Akurasi-Cvec					Akurasi-TF-IDF				
	Parameter: <i>number of trees</i>									
	100	200	300	400	500	100	200	300	400	500
UI	0.8	0.75	0.8	0.75	0.75	0.8	0.75	0.75	0.75	0.75
UNS	0.81	0.857	0.81	0.81	0.81	0.857	0.81	0.81	0.81	0.81
Tel-U	0.75	0.8	0.8	0.75	0.75	0.75	0.75	0.8	0.8	0.7
UII	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9

Universitas	Akurasi-Cvec					Akurasi-TF-IDF				
	Parameter: <i>number of trees</i>									
	100	200	300	400	500	100	200	300	400	500
semua universitas	0.728	0.753	0.728	0.753	0.728	0.728	0.741	0.753	0.728	0.691
rata-rata	0.798	0.812	0.808	0.793	0.788	0.807	0.790	0.803	0.798	0.770

Dari tabel 5 dapat dilihat dengan menggunakan parameter *number of trees* 200, RFC dapat menghasilkan akurasi rata-rata 81.2% untuk fitur Cvec. Sedangkan untuk fitur TF-IDF, akurasi 80.7% dicapai oleh parameter *number of trees* 100. Adapun nilai terendah dicapai oleh parameter *number of trees* 500 baik untuk ekstraksi fitur Cvec maupun TF-IDF.

Hasil Klasifikasi *Neural Network*

Tabel 6 berisi hasil akurasi dari analisis sentimen dengan menggunakan metode NN. Training dari NN dilakukan dengan beberapa varian nilai dari parameter jumlah *hidden neurons* yaitu 20, 40, 60, 80 dan 100.

Tabel 6
Perbandingan akurasi NN terhadap fitur Cvec dan TF-IDF.

Universitas	Akurasi-Cvec					Akurasi-TF-IDF				
	Parameter: <i>number of neurons</i>									
	20	40	60	80	100	20	40	60	80	100
UI	0.75	0.7	0.75	0.75	0.75	0.85	0.85	0.85	0.85	0.85
UNS	0.81	0.857	0.857	0.81	0.81	0.81	0.81	0.81	0.81	0.857
Tel-U	0.8	0.85	0.85	0.85	0.85	0.9	0.9	0.9	0.9	0.9
UII	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
semua universitas	0.827	0.84	0.815	0.778	0.84	0.84	0.84	0.852	0.877	0.802
rata-rata	0.817	0.829	0.834	0.818	0.830	0.860	0.860	0.862	0.867	0.862

Dari tabel 6 dapat dilihat dengan menggunakan parameter *number of neurons* 60, NN dapat menghasilkan akurasi rata-rata 83.4% untuk fitur Cvec. Sedangkan fitur TF-IDF dan parameter *number of neurons* 80 berhasil memberikan tingkat akurasi yang lebih tinggi yaitu 86.7%. Adapun nilai terendah dicapai oleh parameter *number of neurons* 20, baik untuk ekstraksi fitur Cvec maupun TF-IDF.

Untuk melihat performa setiap model kombinasi secara keseluruhan, kita dapat melihatnya pada tabel 7. Nilai yang diberikan pada tabel 7 merupakan hasil rata-rata dari setiap model kombinasi (paduan antara setiap metode ekstraksi fitur Cvec dan TF-IDF dengan setiap metode klasifikasi LR, SVM, RFC dan NN beserta nilai parameter yang telah digunakan). Hasil rata-rata tersebut juga sudah merupakan hasil terhadap seluruh dataset universitas (UI, UNS, Tel-U dan UII). Dari hasil tersebut, eksperimen menunjukkan bahwa hasil akurasi rata-rata terbaik secara keseluruhan bernilai 86.2% yang diberikan oleh model kombinasi *neural network* dan TF-IDF. Sedangkan model kombinasi yang paling lemah diberikan oleh SVM dan Cvec yang menghasilkan hanya 76.3%.

Tabel 7
Perbandingan akurasi rata-rata untuk seluruh model kombinasi.

Model Kombinasi	Akurasi
LR Cvec	0.814
LR TF-IDF	0.799
SVM Cvec	0.763
SVM TF-IDF	0.809
RFC Cvec	0.799
RFC TF-IDF	0.794
NN Cvec	0.826
NN TF-IDF	0.862

Kesimpulan

Pada penelitian ini, kami telah melakukan analisis sentimen untuk memprediksi polaritas dari opini masyarakat terhadap universitas. Data yang digunakan merupakan komentar-komentar yang berada pada halaman resmi Facebook dari beberapa universitas terkemuka di Indonesia. Kami melakukan perbandingan dari beberapa klasifikator (yaitu *Logistic Regression*, *Support Vector Machine*, *Random Forest Classifier* dan *Neural Network*) dan teknik ekstraksi fitur (*count vectorizer* dan TF-IDF). Hasil eksperimen menunjukkan bahwasanya hasil akurasi analisis sentimen terbaik diberikan oleh *Neural Network* yang menggunakan pembobotan TF-IDF dimana hasil akurasi prediksi sentimen mencapai lebih dari 86%.

Namun demikian, penelitian ini masih dapat diperbaiki lagi agar menjadi lebih berguna dan bermanfaat bagi pihak-pihak yang berkepentingan dalam menganalisis opini masyarakat terhadap kampus-kampus di tanah air. Saat ini terdapat hampir 5000 kampus yang tersebar di seluruh tanah air. Oleh sebab itu, akan lebih baik lagi jika analisis sentimen dapat dilakukan pada lebih banyak kampus agar model prediksi dapat lebih dipercaya. Selain itu, dibutuhkan penelitian lanjutan yang dapat menggunakan algoritma yang lebih terkini dengan jumlah data yang lebih besar.

Daftar Pustaka

- Abbasi, A., Chen, H., & Salem, A. (2008). Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums. *ACM Transactions on Information Systems (TOIS)*, 26(3), 12.
- Abdelrazeq, A., Janßen, D., Tummel, C., Jeschke, S., & Richert, A. (2016). Sentiment analysis of social media for evaluating universities. In *Automation, Communication and Cybernetics in Science and Engineering 2015/2016* (pp. 233–251). Springer.
- Akbar, H., Suryana, N., & Sahib, S. (2011). Training neural networks using Clonal Selection Algorithm and Particle Swarm Optimization: A comparisons for 3D object recognition. In *2011 11th International conference on hybrid intelligent systems (HIS)* (pp. 692–697).
- Altrabsheh, N., Gaber, M. M., & Cocea, M. (2013). SA-E: sentiment analysis for education. In *International Conference on Intelligent Decision Technologies* (Vol. 255, pp. 353–362).
- Cummins, S., Burd, L., & Hatch, A. (2010). Using feedback tags and sentiment analysis to generate

sharable learning resources investigating automated sentiment analysis of feedback tags in a programming course. In *2010 10th IEEE International Conference on Advanced Learning Technologies* (pp. 653–657).

Hamdan, H., Bellot, P., & Bechet, F. (2015). Lsislif: Crf and logistic regression for opinion target extraction and sentiment polarity analysis. In *Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)* (pp. 753–758).

Kechaou, Z., Ammar, M. Ben, & Alimi, A. M. (2011). Improving e-learning with sentiment analysis of users' opinions. In *2011 IEEE global engineering education conference (EDUCON)* (pp. 1032–1038).

Knerr, S., Personnaz, L., & Dreyfus, G. (1990). Single-layer learning revisited: a stepwise procedure for building and training a neural network. In *Neurocomputing* (pp. 41–50). Springer.

Koehler, M., Greenhalgh, S., & Zellner, A. (2015). Potential Applications of Sentiment Analysis in Educational Research and Practice--Is SITE the Friendliest Conference? In *Society for Information Technology & Teacher Education International Conference* (pp. 1348–1354).

Mollett, A., Moran, D., & Dunleavy, P. (2011). Using Twitter in university research, teaching and impact activities. LSE Public Policy Group, London School of Economics and Political Science.

Wen, M., Yang, D., & Rose, C. (2014). Sentiment Analysis in MOOC Discussion Forums: What does it tell us? In *Educational data mining 2014*.