

Analisis Sentimen pada Review Produk Shopee dengan Pendekatan Lexicon dan Metode Support Vector Machine

Tugas Akhir

diajukan untuk memenuhi salah satu syarat

memperoleh gelar sarjana

dari Program Studi Informatika

Fakultas Informatika

Universitas Telkom

1301170278

Muhammad Mukhtar Dwi Putra



Program Studi Sarjana Informatika

Fakultas Informatika

Universitas Telkom

Bandung

2020

LEMBAR PENGESAHAN

**Analisis Sentimen pada Review Produk Shopee dengan Pendekatan Lexicon dan Metode
Support Vector Machine**

*Sentiment Analysis on Review Product Shopee using Lexicon Approach and Support Vector
Machine Method*

NIM: 1301170278

Muhammad Mukhtar Dwi Putra

Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh
gelar pada Program Studi Sarjana Informatika

Fakultas Informatika

Universitas Telkom

Bandung, 8 November 2020

Menyetujui

Pembimbing I

Pembimbing II

Dr. Eng. Wikky Fawwaz Al Maki, M.Eng

NIP: 19820006

Dr. Ade Romadhony, S.T., M.T.

NIP: 06840337-1

Ketua Program Studi
Sarjana Informatika,

Dr. xxx.

NIP: 91930541

LEMBAR PERNYATAAN

Dengan ini saya, Muhammad Mukhtar Dwi Putra, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul "**Analisis Sentimen pada Review Produk Shopee dengan Pendekatan Lexicon dan Metode Support Vector Machine**" beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika dikemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya.

Bandung, 8 November 2020

Yang Menyatakan,

Muhammad Mukhtar Dwi Putra

Analisis Sentimen pada Review Produk Shopee dengan Pendekatan Lexicon dan Metode Support Vector Machine

Muhammad Mukhtar Dwi Putra¹, Pemb 1 tanpa gelar², Pemb 2 tanpa gelar³, Pemb 3 tanpa gelar⁴

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹mmukhtar@students.telkomuniversity.ac.id, ²pembimbing1@telkomuniversity.ac.id,

³pembimbing2@telkomuniversity.ac.id, ⁴pembimbingluar@telkom.co.id

Abstrak

Dokumen ini merupakan panduan penulisan jurnal Tugas Akhir (TA) di lingkungan Fakultas Informatika Universitas Telkom. Meskipun demikian, dimungkinkan/dipersilahkan untuk pembimbing TA menggunakan struktur penulisan yang tidak sama persis dengan yang ada di dokumen ini. Panjang abstrak tidak lebih dari 200 kata dan diketik dalam ukuran huruf 10 pts. TA sebagai salah satu sarana latihan penulisan akademik dan memperjelas tulisan, abstrak dibagi menjadi empat paragraf atau sub-bagian. Setiap sub bagian bisa diberi judul yang digaris bawahi. Abstrak berisi apa, mengapa, bagaimana, dan hasil utama (kesimpulan).

Apa permasalahan pada topik. Yang juga menjelaskan latar belakang permasalahan topik. Sebaiknya tuliskan juga apa masukan dan keluaran secara sangat singkat.

Mengapa topik menarik atau penting. Sebisa mungkin tuliskan contohnya secara sangat singkat. Pada bagian ini sebaiknya ditulis juga apa masalah/kekurangan yang terjadi unt kondisi saat ini (gap antara kondisi sekarang dengan yang diharapkan)?

Bagaimana solusinya. Jelaskan secara garis besar sistem solusi yang telah dilakukan. Biasanya penjelasan solusi ini merupakan yang terpanjang pada abstrak.

Hasil utama. Hasil utama dari eksperimen ditulis singkat dua-tiga kalimat. Akan lebih baik (optional), kalau dituliskan secara eksplisit kontribusi yang telah dihasilkan. Kontribusi bisa dituliskan diantara bagian solusi dan hasil eksperimen.

Pastikan abstrak pada jurnal TA tidak copas dari abstrak proposal TA. Pada abstrak proposal kadang ada kata *akan*, seperti misalnya *yang akan dilakukan*; sedangkan pada abstrak Jurnal TA tidak ada kata *akan* spt itu. Tidak boleh ada sitasi pada abstrak. Pada abstrak tidak menggunakan penamaan, simbol atau istilah yang teknis, misalnya *minsup* untuk menyatakan nilai support minimal.

Kata kunci : merupakan kata-kata kunci yang menjelaskan isi tulisan, biasanya bisa diambil dari judul dan abstrak. Maksimal enam buah dan ditulis dengan huruf kecil, kecuali singkatan

Abstract

The abstract should state briefly the general aspects of the subject and the main conclusions. The length of abstract should be no more than 200 word and should be typed be with 10 pts.

Keywords: keyword should be chosen that they best describe the contents of the paper and should be typed in lower-case, except abbreviation. Keyword should be no more than 6 word

1. Pendahuluan

Latar Belakang

Pada saat ini, perkembangan teknologi terjadi secara pesat sehingga membuat informasi tentang jual beli menjadi lebih mudah didapatkan. Salah satunya adalah *e-commerce* yang menyediakan informasi tentang produk yang dijual ataupun dibeli oleh penggunanya. *E-commerce* menjadi pilihan oleh banyak orang untuk kegiatan jual beli secara daring karena selain menyediakan informasi, *e-commerce* juga menjadi perantara pembayaran sehingga

membuat penggunanya merasa aman dari penipuan. Pada *e-commerce* juga terdapat informasi tentang ulasan produk dari toko daring oleh pembeli yang telah melakukan transaksi sebelumnya. Ini tentunya akan membuat pengguna *e-commerce* lainnya menjadi lebih percaya untuk membeli produk dari toko daring yang mempunyai penilaian tinggi.

Beberapa *e-commerce* yang terkenal adalah Shopee, Tokopedia, Bukalapak, OLX dan lainnya. Pada tahun 2019, jumlah pengguna *e-commerce* di Indonesia diproyeksikan akan mencapai 168,3 juta pengguna dan terus meningkat mencapai 212,2 juta pada 2023 [1]. Semua ulasan produk pada *e-commerce* ini mempunyai pola yang sama, yaitu mengulas tentang produk ataupun toko dengan kalimat yang mempunyai sentimen positif atau negatif. Tetapi pada penelitian ini, data yang diambil hanya berasal dari situs *e-commerce* Shopee sebagai sampel dari *e-commerce*.

Pada analisis sentimen, terdapat dua pendekatan algoritma yang bisa digunakan untuk menyelesaikan permasalahan ini. Pendekatan algoritma tersebut adalah pendekatan *lexicon* dan pendekatan *learning*. Pendekatan menggunakan *lexicon* bergantung pada kamus sentimen yang berupa data latih berisi kumpulan kalimat yang mempunyai nilai sentimen [2]. Pada setiap kalimat pada data latih ini terdapat kata opini yang mempunyai nilai sentimen yang akan digunakan untuk klasifikasi sentimen pada kalimat yang baru jika kata opini pada kalimat yang baru terdapat pada kamus. Sedangkan pendekatan menggunakan *learning* pendekatan yang menerapkan algoritma *machine learning* dengan menggunakan fitur linguistik yang ada [2]. Kedua pendekatan ini mempunyai kelemahan, maka untuk mengatasi kelemahan tersebut dibuatlah sistem dengan kombinasi dua pendekatan ini.

Topik dan Batasannya

Pendekatan *lexicon* bisa dipakai untuk memberi label pada data agar pelabelan bisa dilakukan secara otomatis atau untuk mengklasifikasi langsung suatu kalimat yang berisi sentimen. Tetapi, pendekatan *lexicon* terkadang tidak bisa mendeteksi skor dari beberapa kata opini karena kata tersebut tidak terdapat di dalam kamus opini yang dipakai. Skor positif dan skor negatif yang dihasilkan oleh pendekatan *lexicon* juga bisa sama besarnya yang membuat kebingungan dalam pemberian label. Maka dari itu, untuk menutupi kelemahan tersebut digunakan juga pendekatan *learning* yang sifatnya bisa melabeli data tanpa perlu kamus opini atau sistem skor.

Data yang digunakan pada penelitian ini berasal dari ulasan pada situs *e-commerce* Shopee. Data tersebut berupa kalimat tentang ulasan suatu produk yang dibeli dari suatu toko. Data akan diberi label positif (1) dan negatif (0) sesuai dengan sentimennya secara manual. Sistem analisis sentimen yang dibuat untuk mendeteksi data tersebut tidak bersifat *real-time online* sehingga jika ada data baru maka model dengan pendekatan *learning* harus dilatih ulang.

Tujuan

Untuk menganalisa apakah metode kombinasi yang digunakan lebih baik daripada hanya menggunakan pendekatan *lexicon*. Mempelajari proses metode kombinasi yang dipilih agar dapat meningkatkan kinerja *lexicon* dalam permasalahan analisis sentimen. Melihat dampak penggunaan metode kombinasi dibanding hanya menggunakan pendekatan *lexicon*.

Organisasi Tulisan

Pada bagian 1 dijelaskan mengenai latar belakang masalah, perumusan masalah dan tujuan dari penelitian. Pada bagian 2 dijelaskan landasan teori yang terkait dengan penelitian ini. Pada bagian 3 dijelaskan gambaran sistem yang di bangun, pada bagian 4 dijelaskan mengenai hasil dan analisis dari metode tersebut. Pada bagian 5 merupakan kesimpulan yang menjawab permasalahan pada penelitian ini serta saran dari penulis untuk penelitian selanjutnya.

2. Studi Terkait

2.1 Analisis Sentimen

Menurut paper *Techniques and Applications For Sentiment Analysis* [3], sentimen dapat didefinisikan sebagai perasaan positif atau negatif individu dan analisis sentimen dapat didefinisikan sebagai kegiatan mencari pendapat penulis terhadap suatu entitas tertentu. Sehingga dapat dikatakan bahwa analisis sentimen adalah kegiatan yang bertujuan untuk mengklasifikasikan data tekstual menjadi kelas. Pada kasus masalah penelitian ini terdapat dua kelas, yaitu positif dan negatif. Dalam implementasinya, data teks akan memiliki sentimen positif jika data mengandung opini positif lalu akan diklasifikasikan sebagai kelas positif, begitu juga dengan kelas negatif.

2.2 Pendekatan Lexicon

Metode berbasis *lexicon* bergantung pada kata-kata dalam kalimat dari ulasan produk yang akan menghasilkan sentimen positif atau negatif. Kata pada kalimat ini biasa disebut sebagai kata opini atau kata sentimen.

Sentimen positif ditunjukkan oleh kata-kata yang bermakna positif, seperti bagus, hebat dan murah. Sedangkan sentimen negatif ditunjukkan oleh kata-kata yang bermakna negatif, seperti jelek, tidak berguna dan mahal. Biasanya kata dari sentimen diikuti dengan kata keterangan seperti sangat bagus, jelek sekali atau sangat bagus sekali. Kata benda juga terkadang mempunyai nilai sentimennya sendiri. Untuk menentukan nilai sentimen positif dan negatif diperlukan kamus yang disebut dengan kamus opini. Jika kata opini tidak terdapat dalam kamus opini, maka akan terklasifikasi netral. Dalam penelitian ini, kamus opini yang dipakai adalah SentiWordNet 3.0. Karena bahasa pada kamus ini berupa Bahasa Inggris, maka data yang sudah didapat harus diterjemahkan terlebih dahulu agar kata opini bisa terdeteksi oleh kamus ini.

2.3 Pendekatan Learning

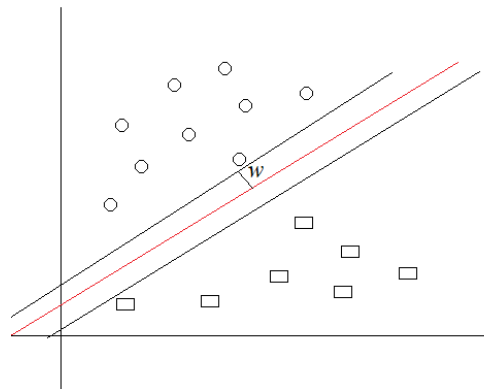
Metode berbasis learning adalah metode pengklasifikasian suatu objek dengan memperhatikan fitur yang ada. Dalam sentimen analisis, objek yang diklasifikasikan adalah suatu kalimat atau dokumen dengan kata-kata yang terdapat dalam kalimat sebagai fiturnya. Untuk dapat mengklasifikasikan kalimat, metode ini akan melatih pengklasifikasi sentimen dengan menggunakan metode dari *machine learning* untuk menentukan suatu kalimat memiliki sentimen positif atau sentimen negatif dengan mempertimbangkan nilai sentimen yang dihasilkan oleh kalimat itu sendiri.

2.4 Support Vector Machine

Support Vector Machine (SVM) adalah algoritma komputer yang belajar melalui sample untuk memberikan label pada objek-objek [4]. Tujuannya adalah untuk mendapatkan model yang telah dilatih dari fitur-fitur yang telah diekstrak dari data latih agar dapat memprediksi label dari data pengujian. Model ini akan memisahkan data menjadi dua kelas atau lebih. Dalam membuat model klasifikasi, *SVM* akan membuat garis linier yang dinamakan *hyperplane*. Namun, dalam implementasinya tidak banyak data yang dapat dipisahkan secara linier sehingga *SVM* akan mengubah dimensi dari data latih agar model dapat memisahkan data secara linier.

Hyperplane akan memisahkan data berdasarkan nilai batas yang paling besar. Pada kasus ini, kelas akan dibagi menjadi dua yaitu positif dan negatif. *Hyperplane* dicari berdasarkan persamaan 1.

$$f : \vec{w} \cdot \vec{x} + b = 0 \quad (1)$$



Gambar 1. Hyperplane dengan optimasi

Dengan \vec{w} adalah *weight*, \vec{x} sebagai nilai vektor fitur dan b adalah nilai bias. Vektor w dan b merupakan parameter pada SVM yang akan memaksimalkan nilai batas. Sehingga *hyperplane* yang dicari akan seperti pada gambar 1.

2.5 Kombinasi Pendekatan Lexicon dan Pendekatan Learning

Karena ketidakmampuan pendekatan *lexicon* dalam mengatasi kata singkatan, simbol, *emoticon* dan pendekatan *learning* yang harus mempunyai data berlabel secara manual, maka dibentuklah metode kombinasi pendekatan *lexicon* dan *learning*. Data yang dapat diklasifikasi oleh pendekatan *lexicon* akan menjadi data latih pada model SVM. Model SVM akan menjadi pengoptimasi model *lexicon* karena menggunakan data latih yang didapat dari pendekatan *lexicon* (yang berhasil dilabeli). Model hasil analisis sentimen akhir ini akan berupa pengklasifikasi

dengan pendekatan *lexicon* lalu jika tidak terdeteksi atau terklasifikasi netral maka akan dibantu oleh pendekatan *learning* dengan model SVM.

2.6 TF-IDF

TF-IDF adalah perhitungan bobot dari frekuensi kemunculan kata pada suatu dokumen. Pada dasarnya, TF-IDF bekerja dengan menentukan frekuensi relatif kata-kata secara spesifik pada dokumen dengan membandingkan proporsi kebalikan dari kata itu terhadap seluruh kata dalam dokumen [5]. Setelah dapat frekuensinya, dihitunglah bobot dari kata terhadap dokumen tersebut yang dinamakan *term weight*. Persamaan yang digunakan untuk menghitung *term weight* dari TF-IDF seperti pada 2 atau 3.

$$w_{ij} = t f_i \cdot i d f_j \quad (2)$$

$$w_{ij} = t f_i \cdot \log(D/d f_j) \quad (3)$$

Dimana w_{ij} adalah *term weight* terhadap dokumen. Sedangkan $t f_i$ adalah jumlah kemunculan *term* dalam dokumen. D adalah jumlah semua dokumen yang ada dan $d f_j$ adalah jumlah dokumen yang mengandung *term* (kata) t_j .

3. Sistem yang Dibangun

3.1 Pengambilan Data

Data yang dipakai dalam penelitian ini adalah kalimat ulasan produk dalam suatu *e-commerce*. Kalimat ini akan berisikan tentang ulasan terhadap suatu macam-macam produk seperti kamera, ponsel pintar, sepatu dan sebagainya. Data yang diambil berasal dari komentar ulasan pada *e-commerce* Shopee.

Teknik *web scraping* adalah teknik untuk meminta data, membuat *query* dari *web*, lalu menguraikan kalimat dengan mengekstrak informasi yang didapatkan secara otomatis [6]. Pengambilan data pada penelitian ini dilakukan semi-manual dengan menggunakan teknik *web scraping* dengan mengambil data menggunakan api pada situs Shopee.

3.2 Preprocessing

Sebelum melatih *sentiment classifier* dengan menggunakan *lexicon* dan SVM, fitur berupa *term* dalam data latih akan diekstrak terlebih dahulu. Data latih harus dilakukan *preprocessing* agar model dapat mengklasifikasi kata opini secara maksimal. Berikut merupakan tahap *preprocessing* :

1. Case Folding : Pada tahap ini, semua karakter huruf pada kalimat data latih akan diubah menjadi huruf kecil (*lowercase*), huruf yang diubah hanya dari 'a' sampai dengan 'z' (bukan simbol dan huruf yang tidak terdapat pada huruf abjad).
2. Tokenization : Setelah huruf pada kalimat diubah menjadi huruf kecil, kalimat akan dibagi menjadi per-kata dengan cara dipisah, tahap ini disebut dengan *tokenization*. Kata akan dibagi menjadi unit-unit dasar yang tidak perlu diuraikan dalam pemrosesan selanjutnya [7]. Kata yang dipisah ini disebut dengan token, huruf pada kata ini dipisah berdasarkan *delimiter*.
3. POS Tagging : *Part of Speech tagging* dilakukan untuk mengkategorikan kelas kata dalam suatu kalimat. Contoh POS Tag adalah kata sifat (JJ), kata benda (NN), kata kerja (VB) dan lainnya.
4. Stop Word Removal : *Stop word* adalah kata yang sering digunakan untuk menyambungkan kalimat seperti 'dan', 'karena', 'bahwa' dan lainnya. *Stop word* tidak terlalu berpengaruh sebagai fitur dalam analisis sentimen sehingga untuk meningkatkan kinerja sistem, *stop word* akan dihilangkan pada tahap ini [8].

Terdapat perbedaan proses dalam mengelola kalimat untuk dimasukkan ke dalam pendekatan yang dipilih ini. Pada pendekatan *lexicon*, proses yang dilakukan adalah proses *case folding* (1), *tokenization* (2) dan *POS Tagging* (3). Sedangkan pendekatan SVM menggunakan proses *case folding* (1), *tokenization* (2) dan *stop word removal* (4). Semua proses pada *preprocessing* ini akan dibantu dengan menggunakan *library* nltk.

Proses *stop word removal* tidak dilakukan pada pendekatan *lexicon*, karena pada proses pengklasifikasi nanti kalimat harus utuh untuk mencari POS tag per katanya. Sedangkan pada model SVM, proses *POS tagging* tidak ada karena pada proses *learning* nanti tidak perlu menggunakan POS tag.

3.3 Klasifikasi dan Pembagian Data dengan Pendekatan Lexicon

Klasifikasi menggunakan pendekatan *lexicon* akan menghitung skor pada suatu kalimat per kata dengan POS *tag* tertentu. POS *tag* yang akan diambil untuk dihitung skornya adalah JJR, VBG, JJ, VBN, VB, dan VBD. Ini dilakukan untuk mempercepat proses perhitungan skor sehingga proses akan menjadi lebih efektif daripada menghitung semua skor kata di kalimat.

Pada sistem dengan pendekatan *lexicon* ini mempunyai penanganan kata negasi yang jika ada kata dengan awalan "not", maka skor positif dan skor negatif dari kata akan ditukar (karena maknanya terbalik dari aslinya). Contoh kata negasi: "works well experimented with it by placing it right under the aircon and the numbers will change accordingly **not** 100 sure if it's accurate so far so good fast delivery"

Sistem ini juga mempunyai penanganan kata perumpamaan seperti jika ada kata "unless", "if" atau "will", maka kata POSTag tertentu yang terdeteksi selanjutnya akan dihitung menjadi nol. Lalu kata "but" dan "unfortunately" juga akan membuat skor kata sebelumnya dari POSTag tertentu menjadi nol. Hal ini dilakukan karena kata itu hanya perumpamaan yang berarti sentimen dari kata tersebut tidak ada. Contoh kata perumpamaan: "Good deal and will continue to support XiaoMi **if** this products is good"

Nilai sentimen akan dihitung berdasarkan kelas positif dan negatif lalu kalimat akan diklasifikasi berdasarkan nilai sentimen yang terbesar [9]. Untuk menentukan klasifikasi sentimen, persamaan yang dipakai adalah 4 dan 5.

$$S_{positif} = \sum_{i \in t}^n nilaipositif_i \quad (4)$$

$$S_{negatif} = \sum_{i \in t}^n nilainegatif_i \quad (5)$$

Nilai sentimen positif dan negatif akan dihitung untuk setiap kata dalam kalimat seperti 4 dan 5. Nilai sentimen keseluruhan untuk setiap kalimat diperoleh dengan membandingkan nilai skor total positif dan skor negatif seperti pada persamaan 6.

$$Sentence_{sentimen} \{ positif \text{ jika } S_{positif} > S_{negatif}, \text{netral jika } S_{positif} = S_{negatif}, \text{negatif jika } S_{positif} < S_{negatif} \quad (6)$$

Nilai sentimen ini akan menentukan pengklasifikasi sentimen nantinya. Lalu kalimat akan diklasifikasi positif atau negatif berdasarkan hasil dari sentimen pada persamaan 6. Jika tidak terklasifikasi atau netral, maka data akan masuk ke data uji. Jika terklasifikasi positif atau negatif, maka data akan masuk ke data latih.

3.4 Klasifikasi dengan Support Vector Machine

Klasifikasi yang dilakukan dengan pendekatan *lexicon* tidak terlalu efektif karena jika kata pada kalimat tidak terdapat pada kamus opini, maka pendekatan *lexicon* tidak bisa mengklasifikasinya. Untuk mengatasi hal ini pendekatan *lexicon* akan dioptimasi dengan pendekatan *machine learning* menggunakan metode *Support Vector Machine*.

Sebelum dijadikan input model dan setelah proses preprocessing, fitur dari data latih akan diekstrak dengan menggunakan TF-IDF terlebih dahulu. Fitur yang diekstrak sebagai input dari model SVM adalah *term weight* dari TF-IDF. Data yang dipakai untuk melatih SVM adalah data latih yang didapat dari hasil kata yang terklasifikasi pada pendekatan *lexicon*. Pada penelitian ini, *library scikit-learn* akan dipakai untuk membantu proses pembuatan model SVM dan pembuatan TF-IDF.

3.5 Teknik Evaluasi

Nilai Prediksi	Kelas Asli	
	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
<i>Negative</i>	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

Tabel 1. Confusion Matrix

Confusion matrix adalah alat untuk mengevaluasi seberapa bagus model untuk bisa memprediksi berbagai kelas yang ada [10]. Data latih yang telah dihasilkan dari proses pendekatan *lexicon* akan dievaluasi dengan *confusion matrix*. Data latih ini akan dibandingkan dan dievaluasi dengan data set yang sudah ada dilabeli secara manual

yang ditentukan dari bintang yang diberikan oleh pengulas (Dengan bintang 1 - 2 dilabeli 0, 4 - 5 dilabeli 1 dan 3 tergantung dengan sentimennya). Dari *confusion matrix* akan didapatkan informasi untuk mencari *accuracy*, *precision* dan *recall*. Tabel *confusion matrix* dapat dilihat pada 1.

Dengan *TP* adalah model benar dalam mengklasifikasi sentimen positif. *TN* adalah model benar dalam mengklasifikasi sentimen negatif. *FP* adalah model salah dalam mengklasifikasi sentimen positif. *FN* adalah model salah dalam mengklasifikasi sentimen negatif.

Accuracy merupakan nilai prediksi yang mempunyai tingkat kedekatan dengan nilai aslinya. Persamaannya bisa dilihat pada persamaan 7.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \quad (7)$$

Recall adalah nilai keberhasilan model dalam memprediksi data uji. Persamaannya seperti pada persamaan 8.

$$Recall = \frac{TP}{TP + FN} * 100\% \quad (8)$$

Precision merupakan nilai ketepatan model dalam memprediksi data uji. Persamaannya bisa dilihat pada persamaan 9.

$$Precision = \frac{TP}{TP + FP} * 100\% \quad (9)$$

F1-Score merupakan nilai yang mengindikasikan bahwa nilai *precision* dan nilai *recall* suatu model itu baik. Persamaannya bisa dilihat pada persamaan 10.

$$F1-Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (10)$$

Cross validation adalah teknik evaluasi dengan pendekatan statistik. Metode *cross validation* akan dipakai untuk mencari *hyperparameter* dari model SVM dengan teknik *k-fold cross validation*. Pada *k-fold cross validation* data latih akan dibagi menjadi *k* bagian untuk dicari nilai rata-rata akurasi. Jumlah *k* dapat ditentukan secara bebas. Namun, semakin besar nilai *k* maka akan semakin baik [11]. Pada penelitian ini, jumlah *k* yang digunakan adalah 10.

Sisten dengan pendekatan *lexicon* dan pendekatan learning (SVM) akan dievaluasi menggunakan hasil akurasi, *precision*, *recall* dan *f-measure* (*f1-score*). Data yang dibandingkan untuk evaluasi adalah data yang sudah diberi label secara manual saat pengumpulan data.

3.6 Alur Sistem Kombinasi lexicon dan SVM

Sistem kombinasi *sentiment classifier* ini akan mengklasifikasi menggunakan pendekatan *lexicon* terlebih dahulu, Lalu jika tidak terklasifikasi atau terklasifikasi netral, maka akan diklasifikasi dengan pendekatan *learning*.

4. Evaluasi

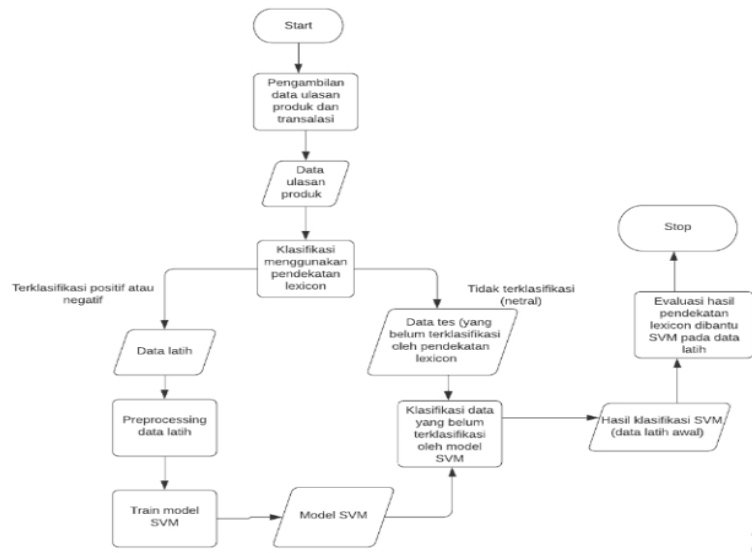
4.1 Dataset dan Proses Pemilihan Data Train

Pada penelitian ini digunakan *training set* yang berjumlah 3617 data dan diambil dari ulasan komentar suatu produk di situs *e-commerce* Shopee. Data pada *training set* akan dibagi menjadi data latih dan data tes. Pemilihan data latih dilakukan dengan cara mengambil data pada *training set* yang berhasil diklasifikasi oleh pendekatan *lexicon*. Data yang tidak berhasil diklasifikasi oleh pendekatan *lexicon* akan dimasukan ke data tes dan akan dilanjutkan klasifikasinya dengan model SVM.

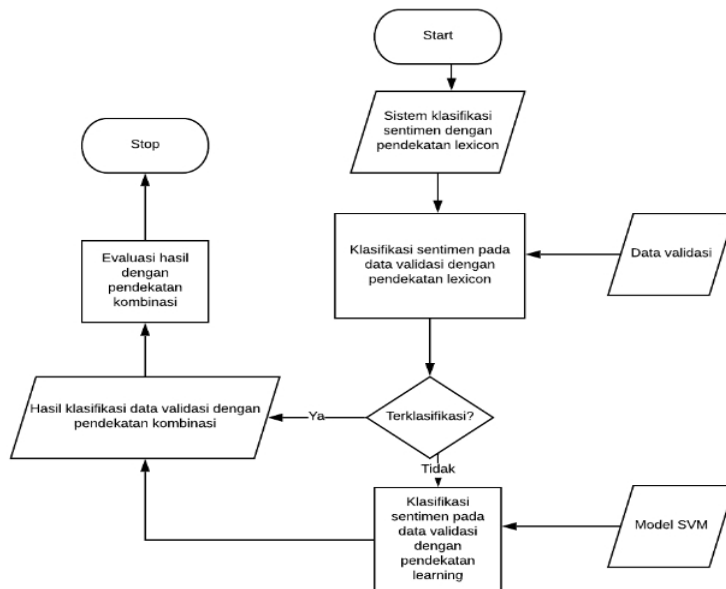
Dataset mempunyai tiga jenis kategori barang pembelian, yaitu *acc* (aksesoris), *fashion* dan *tech* (teknologi). Jumlah data pada data latih dan data validasi yang berhasil dikumpulkan dapat dilihat pada Tabel 2.

Data	Jumlah Data per Kelas		Jumlah Data
	Positif	Negatif	
Data Latih	1808	1808	3616
Data Validasi	251	198	449

Tabel 2. Jumlah data yang berhasil dikumpulkan



Gambar 2. Alur sistem pengklasifikasi pada data latih



Gambar 3. Alur sistem pengklasifikasi pada data validasi

4.2 Hasil Pengujian

Setelah melakukan proses seperti yang disebutkan sebelumnya, sistem mendapatkan pembagian data seperti pada Tabel 3. Data latih yang ada pada tabel tersebut adalah data yang terklasifikasi benar oleh pendekatan *lexicon*. Ini dilakukan karena jika data yang terklasifikasi salah juga dimasukkan ke data latih, maka akan mengacaukan model dengan pendekatan *learning* nantinya. Sedangkan data tes akan dilanjutkan klasifikasinya oleh pendekatan *learning* dengan model SVM. Hasil evaluasi dari pendekatan *lexicon* dan pendekatan *learning* pada data latih bisa dilihat pada Tabel 4.

Setelah itu, dilakukan juga proses evaluasi *sentiment classifier* terhadap data validasi (diluar dari data latih awal) yang hasilnya bisa dilihat pada tabel 5. Untuk evaluasi menggunakan data validasi, hasil dengan pendekatan *lexicon* dibagi menjadi dua, dengan netral dan tanpa netral. Maksud dari dengan dan tanpa netral disini adalah perhitungan evaluasi yang menyertakan dan tidak menyertakan banyaknya hasil data yang tidak dapat diklasifikasi oleh pendekatan *lexicon* pada akurasi, *precision*, *recall* dan *f1-score*. Data yang berhasil terklasifikasi oleh pendekatan *lexicon* hanya berjumlah 229 data dan yang tidak bisa terklasifikasi ada 220 data.

Data		Jumlah Data per Kelas		Jumlah
		Positif	Negatif	
Aksesoris	Data latih	331	197	528
	Data tes	260	357	617
Fashion	Data latih	293	234	527
	Data tes	240	274	514
Teknologi	Data latih	229	184	413
	Data tes	316	309	625
Semua data	Data latih	853	615	1468
	Data tes	816	940	1756

Tabel 3. Jumlah data latih dan data tes yang telah dibagi dari data latih

Sentiment Classifier		Accuracy	Precision	Recall	F1-Score	Jumlah Data Terklasifikasi
Semua Data	Pendekatan <i>Lexicon</i>	78.92 %	77.12 %	85.98 %	81.31 %	1860
	Pendekatan <i>Learning</i> (SVM)	82.85 %	78.45 %	87.00 %	82.51 %	1756
	Hasil gabungan	80.83 %	77.72 %	86.44 %	81.85 %	3616
Teknologi	Pendekatan <i>Lexicon</i>	75.09 %	70.89 %	84.19 %	76.97 %	650
	Pendekatan <i>Learning</i> (SVM)	79.52 %	78.65 %	81.64 %	80.12 %	625
	Hasil gabungan	77.44 %	74.80 %	82.82 %	78.61 %	1175
Aksesoris	Pendekatan <i>Lexicon</i>	80.24 %	79.37 %	88.26 %	83.58 %	658
	Pendekatan <i>Learning</i> (SVM)	79.25 %	72.60 %	81.53 %	76.81 %	617
	Hasil gabungan	79.76 %	76.58 %	85.51 %	80.80 %	1275
Fashion	Pendekatan <i>Lexicon</i>	80.82 %	80.05 %	84.92 %	82.41 %	652
	Pendekatan <i>Learning</i> (SVM)	78.21 %	73.70 %	82.91 %	78.03 %	514
	Hasil gabungan	79.67 %	77.35 %	84.10 %	80.58 %	1166

Tabel 4. Hasil evaluasi sentimen *classifier* pada data latih

4.3 Analisis Hasil Pengujian

Pada pengujian data latih (Tabel 4) maupun data validasi (Tabel 5), terlihat bahwa pendekatan *lexicon* tidak mampu mendapatkan hasil akurasi yang sangat tinggi. Seharusnya, pengklasifikasi dengan pendekatan berbasis kamus akan mendapatkan hasil yang maksimal karena pendekatan ini mengklasifikasi dengan menggunakan skor kata yang ada pada kalimat. Saat ditelusuri lebih lanjut, terdapat skor yang salah karena kata mempunyai dua makna seperti kata "*cheap*" yang bisa berarti murahan atau terjangkau. Jika pada kalimat seharusnya bersentimen positif karena maksudnya terjangkau, tetapi pada kamus *lexicon* yang dipakai (Sentiwordnet 3.0) kata "*cheap*" mempunyai skor yang negatif karena pada kamus Sentiwordnet 3.0 "*cheap*" berarti murahan.

Pada Table 5, terlihat bahwa *precision* dari pendekatan *lexicon* (evaluasi tanpa netral) selalu lebih tinggi daripada *precision* pendekatan *learning* (SVM). Dan terlihat juga bahwa *recall* dari pendekatan *learning* (SVM) lebih tinggi daripada *recall* dengan pendekatan *lexicon* kecuali pada kategori barang aksesoris dan *fashion*. Ini berarti, jika kedua pendekatan tersebut digabung bisa membuat pendekatan kombinasi yang menghasilkan performa sistem pengklasifikasi sentimen yang lebih baik.

Lalu jika dilihat pada Tabel 5, hasil evaluasi "semua data" tanpa membagi per kategori barang mendapat *precision* dari pendekatan *lexicon* yang lebih baik daripada pendekatan *learning* (SVM) dan *recall* dari pendekatan *learning* lebih tinggi daripada pendekatan *lexicon*. Berarti jika data tidak dibagi per kategori pun, model pengklasifikasi sentimen dengan pendekatan *learning* tidak akan terpengaruh untuk menjadi model yang lebih buruk.

Saat model kombinasi dari sistem yang dibuat dievaluasi yang hasilnya bisa dilihat pada Tabel 5, hasil akurasi-nya selalu lebih baik daripada pengklasifikasi sentimen dengan pendekatan *lexicon* (yang hanya bisa mengklasifikasi beberapa data) ataupun pendekatan *learning*. Hasil *precision*, *recall* maupun *f1-score* model kombinasi selalu menghasilkan hasil yang lebih baik dari salah satu pendekatan *lexicon* atau pendekatan *learning*. Dengan begitu, model dengan pendekatan kombinasi dapat dikatakan lebih baik daripada model pendekatan *lexicon* atau pendekatan *learning*.

<i>Sentiment Classifier</i>		<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	Data terklasifikasi
Semua Data	Pendekatan <i>Lexicon</i> (tanpa netral)	79.47 %	82.97 %	83.57 %	83.27 %	229
	Pendekatan <i>Lexicon</i> (dengan netral)	40.53 %	32.40 %	32.5 %	32.45 %	449
	Pendekatan <i>Learning</i> (SVM)	80.84 %	78.08 %	86.34 %	82.00 %	449
	Pendekatan Kombinasi	81.06 %	82.47 %	83.46 %	82.96 %	449
Teknologi	Pendekatan <i>Lexicon</i> (tanpa netral)	71.62 %	83.78 %	67.39 %	74.69 %	74
	Pendekatan <i>Lexicon</i> (dengan netral)	31.92 %	24.03 %	22.46 %	23.22 %	166
	Pendekatan <i>Learning</i> (SVM)	75.30 %	66.66 %	73.84 %	70.07 %	166
	Pendekatan Kombinasi	75.90 %	75.0 %	71.05 %	72.97 %	166
Aksesoris	Pendekatan <i>Lexicon</i> (tanpa netral)	86.36 %	87.93 %	96.22 %	91.89 %	66
	Pendekatan <i>Lexicon</i> (dengan netral)	41.00 %	38.93 %	40.47 %	39.68 %	139
	Pendekatan <i>Learning</i> (SVM)	72.66 %	72.07 %	91.95 %	80.80 %	139
	Pendekatan Kombinasi	80.57 %	81.98 %	92.85 %	87.08 %	139
Fashion	Pendekatan <i>Lexicon</i> (tanpa netral)	80.89 %	76.08 %	85.36 %	80.45 %	89
	Pendekatan <i>Lexicon</i> (dengan netral)	50.0 %	34.65 %	36.45 %	35.53 %	144
	Pendekatan <i>Learning</i> (SVM)	71.52 %	69.11 %	70.14 %	69.62 %	144
	Pendekatan Kombinasi	75.69 %	73.52 %	74.62 %	74.07 %	144

Tabel 5. Hasil evaluasi terhadap data validasi

5. Kesimpulan

Setelah melakukan penelitian dan mendapatkan hasil pengujian, dapat diambil kesimpulan bahwa penggunaan pendekatan *learning* dapat membantu pendekatan *lexicon* yang bisa disebut dengan pendekatan kombinasi. Pendekatan kombinasi ini menghasilkan akurasi yang lebih baik daripada pendekatan *lexicon* ataupun pendekatan *learning*. Hasil *precision*, *recall* dan *f1-score* dari pendekatan kombinasi juga menunjukkan bahwa sistem dapat memperbaiki kinerja dari salah satu pendekatan *lexicon* atau *learning*.

Dari hasil pengujian, dapat dilihat bahwa pendekatan *lexicon* tidak mampu menghasilkan akurasi yang tinggi. Tidak semua data juga dapat diklasifikasi oleh pendekatan *lexicon* karena keterbatasan kata di kamus. Lalu skor kata juga masih ada yang salah. Penggunaan library nltk juga terkadang masih salah mengklasifikasi POS *tag*. Sehingga tidak semua kata yang terklasifikasi dengan tag tertentu ada pada kamus.

Saran yang bisa digunakan untuk mengembangkan Tugas Akhir ini adalah sebagai berikut :

1. Menggunakan kamus *lexicon* yang berkaitan dengan subjek penelitian dan tidak ada skor yang salah.
2. Menggunakan pengklasifikasi POS *tag* lain yang lebih baik.
3. Menggunakan model dari pendekatan *learning* yang lain (mungkin saja ada model yang lebih baik dari SVM).

Daftar Pustaka

- [1] Mediaindonesia.com Developer. Pengguna e-commerce di tanah air akan capai 168,3 juta, Oct 2019.

- [2] Walaa Medhat, Ahmed Hassan, and Hoda Korashy. Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4):1093–1113, 2014.
- [3] Ronen Feldman. Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4):82–89, 2013.
- [4] William S Noble. What is a support vector machine? *Nature biotechnology*, 24(12):1565–1567, 2006.
- [5] Juan Ramos et al. Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning*, volume 242, pages 133–142. Piscataway, NJ, 2003.
- [6] Ryan Mitchell. *Web scraping with Python: Collecting more data from the modern web*. ” O’Reilly Media, Inc.”, 2018.
- [7] Jonathan J Webster and Chunyu Kit. Tokenization as the initial phase in nlp. In *COLING 1992 Volume 4: The 15th International Conference on Computational Linguistics*, 1992.
- [8] Subbu Kannan and Vairaprakash Gurusamy. Preprocessing techniques for text mining. *International Journal of Computer Science & Communication Networks*, 5(1):7–16, 2014.
- [9] Endang Wahyu Pamungkas and Divi Galih Prasetyo Putri. An experimental study of lexicon-based sentiment analysis on bahasa indonesia. In *2016 6th International Annual Engineering Seminar (InAES)*, pages 28–31. IEEE, 2016.
- [10] Jiawei Han, Jian Pei, and Micheline Kamber. *Data mining: concepts and techniques*. Elsevier, 2011.
- [11] Tadayoshi Fushiki. Estimation of prediction error by using k-fold cross-validation. *Statistics and Computing*, 21(2):137–146, 2011.

Lampiran