

PREDIKSI FITUR KEMASAN PRODUK MINYAK KAYU PUTIH DENGAN SUPPORT VECTOR MACHINE (SVM)

W. Latuny^{1,*}, V. O. Lawalata², D. B. Pailin³, R. Ohoirenan⁴

¹Jurusan Teknik Industri Fakultas Teknik Universitas Pattimura, Ambon 97233

*Email: wlatuny@gmail.com

²Jurusan Teknik Industri Fakultas Teknik Universitas Pattimura, Ambon 97233

Email: victor.lawalata@fatek.unpatti.ac.id

³Jurusan Teknik Industri Fakultas Teknik Universitas Pattimura, Ambon 97233

Email: dani.ti.fatek@gmail.com

⁴Jurusan Teknik Industri Fakultas Teknik Universitas Pattimura, Ambon 97233

Email: rahmanohoirenan45@gmail.com

Abstrak. UD Sinar Baru memiliki produk minyak kayu putih dengan berbagai ukuran dari 30 ml sampai 550 ml, dan ukuran 550 ml menjadi produk minyak kayu putih yang paling banyak di konsumsi. Kendati demikian, produk ini dikritik konsumen tentang kemasannya yang belum memenuhi keinginan mereka. Penelitian ini bertujuan memperoleh metode klasifikasi sentimen konsumen yang akurat dan mendapatkan fitur-fitur yang berpengaruh pada desain ulang kemasan produk minyak kayu putih ukuran 550 ml. Pengumpulan data menggunakan metode survei secara *online* dari media sosial facebook untuk mendapatkan komentar konsumen menggunakan *power query*. Analisis data memakai konsep metode *Support Vector Machine* (SVM) dengan dukungan aplikasi WEKA untuk menyajikan analisis sentimen dan akurasi dari komentar konsumen. Hasil penelitian menyajikan kecenderungan komentar terhadap setiap atribut dengan penilaian berupa akurasi untuk seluruh *class* sebesar 83%, untuk *class* positif sebanyak 3% komentar dan 57% komentar untuk *class* negatif. Adapun sentimen yang menunjukkan kemasannya cenderung biasa saja sebesar 20% yang diartikan sebagai netral. Kesimpulan dari hasil penelitian ini SMO mempunyai tingkat prediksi yang sangat akurat untuk menganalisis sentimen konsumen tentang fitur kemasan minyak kayu putih 550 ml, serta perlu dilakukan desain ulang kemasan saat ini dengan memperhatikan fitur bentuk, warna, ukuran, dan efisiensi.

Kata kunci: Akurasi Prediksi, Fitur Kemasan, Minyak Kayu Putih, *Support Vector Machine*

Abstract. UD Sinar Baru has eucalyptus oil products with various sizes from 30 ml to 550 ml, and the size of 550 ml is the most consumed eucalyptus oil product. However, this product has been criticized by consumers for its packaging which has not met their expectations. This study aims to obtain an accurate method of classifying consumer sentiment and obtain features that affect the redesign of the 550 ml eucalyptus oil product packaging. Collecting data using an online survey method from social media facebook to get consumer comments using power queries. Data analysis uses the concept of the *Support Vector Machine* (SVM) method with the support of the WEKA application to provide sentiment analysis and accuracy of consumer comments. The results of the study present the tendency of comments on each attribute with an assessment of 83% accuracy for the entire class, 3% for positive class comments and 57% comments for negative class. The sentiment that shows the packaging tends to be normal at 20% which is interpreted as neutral. The conclusion from the results of this study is that SMO has a very accurate prediction rate to analyze consumer sentiment about the features of the 550 ml eucalyptus oil packaging, and it is necessary to redesign the current packaging by considering the features of shape, color, size, and efficiency.

Keywords: Prediction Accuracy, Packaging Features, Eucalyptus Oil, *Support Vector Machine*

1. PENDAHULUAN

Salah satu sektor industri yang berkontribusi dalam proses pembangunan di Maluku adalah industri kecil penyulingan minyak kayu putih. Keberadaan industri kecil tersebut mempunyai andil dalam memperkokoh struktur industri di Indonesia termasuk di Maluku [1]. Minyak kayu putih menjadi andalan karena khasiatnya untuk memberikan rasa hangat pada tubuh, penghilang gatal dari digigit serangga, meredakan sakit perut, perut kembung, masuk angin, dan aromanya yang wangi dan aromatik untuk kesehatan. Produk ini umumnya dihasilkan oleh industri kecil dan menengah (IKM) di Maluku.

Persaingan produk minyak kayu putih di pasaran cukup ketat, membuat IKM perlu mengantisipasi produk pesaingnya. Salah satu inovasi adalah mendesain kemasan produk yang menarik konsumen untuk membeli produk tersebut. Kemasan merupakan “pemicu” karena ia langsung berhadapan dengan konsumen. Untuk itu, kemasan harus dapat mempengaruhi konsumen untuk memberikan respon positif [2]. Hal ini berarti desain kemasan berkontribusi pada daya tarik produknya.

Tantangan bagi IKM berkaitan dengan komunikasi dengan konsumen untuk mengetahui hal-hal yang diinginkan, kebutuhan atau daya tarik yang melekat pada kemasan produk minyak kayu putih. Keterbatasan akses informasi dan kemampuan mengoperasikan teknologi pendukung menjadikan industri tersebut lebih mengandalkan pengalaman konsumen dan promosi “mouth to mouth” untuk mendapatkan respon dari konsumen akan kualitas produknya. Saat ini masyarakat lebih bergantung pada media sosial seperti facebook, twitter, Instagram dan WhatsApp untuk akses informasi, komunikasi dan perdagangan. Survei pembelian online secara global termasuk Indonesia, sebanyak 71% konsumen melakukan peninjauan terhadap produk sebelum membeli produk tersebut. Sebanyak 43% setuju bahwa media sosial menjadi alat bantu untuk memenuhi kebutuhan pengetahuan berupa review produk dan ulasan forum, guna membantu membuat keputusan pembelian [3]. Review produk maupun ulasan forum disampaikan melalui komentar di sosial media yang berisi keluhan, pujian atau pandangan terhadap produk atau jasa dari suatu toko online. Komentar tersebut mendeskripsikan tanggapan yang berbeda-beda dari setiap konsumen. Komentar-komentar berupa teks tersebut dapat dikumpulkan dan diolah dengan analisis sentimen. Pendekatan ini menganalisa pendapat, sentimen, evaluasi, penilaian, sikap dan emosi publik terhadap entitas seperti produk, jasa,

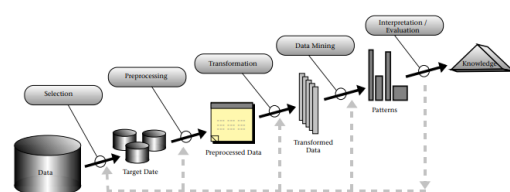
organisasi, individu, masalah, peristiwa, topik, dan atribut mereka [4]. Analisis sentimen mempelajari opini yang mengungkapkan atau mengekspresikan pandangan positif atau negatif [5]. Opini yang dibutuhkan untuk melakukan analisis berasal dari komentar halaman toko online di facebook. Facebook dipilih karena penggunanya saling berinteraksi secara masif, dimana total pengguna facebook sebanyak 1,44 miliar dengan pengguna harian sebanyak 936 juta [6]. Sejak peningkatan penggunaan media sosial dan keterbatasan aktivitas sosial masyarakat, maka survei pendapat akan lebih efektif dan efisien menggunakan media sosial untuk perbaikan kemasan produk minyak kayu putih.

2. BAHAN DAN METODE

Text mining adalah lintas disiplin ilmu yang mengacu pada pencarian informasi, *data mining*, *machine learning*, statistik, dan komputasi *linguistic* [7]. *Text mining* juga dikenal dengan *text data mining* atau pencarian pengetahuan di basis data tekstual adalah proses yang semi otomatis melakukan ekstraksi dari pola data [8].

Tipe pekerjaan *text mining* meliputi kategorisasi, *text clustering*, ekstraksi konsep/entitas, analisis sentimen, *document summarization*, dan *entity-relation modeling* (yaitu hubungan pembelajaran antara entitas) [7]. Sumber data yang digunakan pada *text mining* adalah kumpulan teks yang memiliki format yang tidak terstruktur atau minimal semi terstruktur. Tujuan dari *text mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen.

Text mining merupakan variasi dari *data mining* yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar. Perbedaan terletak pada pola yang digunakan, pola *text mining* diambil dari sekumpulan bahasa alami yang tidak terstruktur sedangkan dalam *data mining* pola diambil dari *database* terstruktur [9]. Beberapa tahapan proses pokok dalam *text mining*, yaitu pemrosesan awal teks (*text preprocessing*), transformasi teks (*text transformation*) atau (*feature generation*), pemilihan fitur (*feature selection*), dan penemuan pola teks atau *data mining* (*pattern discovery*).



Gambar 1.1 Proses *Text Mining* [1]

Text mining memiliki keunggulan dalam analisis data tekstual. Ide utama dari *text mining* adalah untuk menemukan potongan-potongan kecil informasi dari volume data teks yang besar tanpa harus membaca semuanya [10]. Kekuatan pendekatan ini pada ketersediaan data teks melalui Internet [11] berupa dokumen, *email*, media sosial dan web [12]. Kelebihannya adalah analisis yang cepat dan mudah memvisualisasi hasil analisis, sedangkan kelemahannya yaitu sangat subyektif dalam pengukuran sentimen terhadap teks yang dikaji [13].

2.1. Sentimen, Opini dan Analisis Sentimen

Sentimen

Menurut Kamus Besar Bahasa Indonesia (KBBI), *sentiment* berarti pendapat atau pandangan yang didasarkan pada perasaan yang berlebih-lebihan terhadap sesuatu. Sedangkan menurut Merriam-Webster's Online Dictionary, *sentimen* menunjukkan pendapat tetap (terus menerus) yang merefleksikan/ mencerminkan perasaan seseorang.

Opini

Opini dan konsep terkait seperti *sentimen*, *evaluasi*, *tingkah laku*, dan *emosi* merupakan subyek studi dari analisis *sentimen* dan *opinion mining* [4]. Opini atau pendapat merupakan pusat hampir semua aktivitas manusia dan menjadi pengaruh utama dari perilaku. Persepsi terhadap realitas untuk mengevaluasi objek disekitar. Hajmohammadi *et al.* [5] mendefinisikan opini sebagai berikut:

- 1) Pandangan atau penilaian yang terbentuk tentang sesuatu, tidak selalu berdasarkan fakta atau pengetahuan.
- 2) Keyakinan atau pandangan dari sejumlah besar atau mayoritas orang-orang tentang hal tertentu. Secara umum, opini mengacu pada apa yang orang pikirkan tentang sesuatu. Dengan kata lain, opini adalah keyakinan subyektif, dan merupakan hasil emosi atau interpretasi fakta.

Analisis Sentimen

Analisis *sentimen* (*opinion mining*) adalah bidang ilmu yang menganalisa pendapat, *sentimen*, *evaluasi*, *penilaian*, *sikap* dan *emosi publik* terhadap entitas seperti produk, jasa, organisasi, individu, masalah, peristiwa, topik, dan atribut mereka [4]. Analisis *sentimen* berfokus pada *opini-opini* yang mengekspresikan atau mengungkapkan *sentimen positif* atau *negatif*.

Secara umum analisis *sentimen* yang telah diteliti memiliki tiga tingkat (*level*), yaitu:

- 1) *Level dokumen*: mengklasifikasikan apakah seluruh dokumen *opini* mengungkapkan *sentimen positif* atau *negatif*. Analisis mengasumsikan bahwa setiap dokumen mengungkapkan *opini* yang objektif tentang suatu entitas tunggal (misalnya, produk tunggal).
- 2) *Level kalimat*: menentukan apakah setiap kalimat menyatakan *opini positif*, *negatif*, atau *netral*.
- 3) *Level entitas dan aspek*: menemukan *sentimen* pada entitas dan/atau aspeknya. Sebagai contoh, kalimat "kualitas panggilan iPhone baik, tetapi daya tahan baterai pendek". Ada dua aspek *evaluasi*, *kualitas panggilan* dan *baterai kehidupan*, dari iPhone (*entitas*). *Sentimen* pada *kualitas panggilan iPhone* adalah *positif*, tapi *sentimen* pada *hidup baterai* *negatif*. *Kualitas panggilan* dan *daya tahan baterai iPhone* adalah *target pendapat*.

Analisis *sentiment* merupakan salah satu cabang penelitian *text mining* [14]. Analisis *sentimen* hadir untuk menangani kondisi ledakan informasi teks yang tidak terstruktur. Dalam *text mining*, analisis *opini* terhadap ulasan *online* mengekstrak atribut produk dan *opini* terkait, yang oleh analisis *sentimen*, dapat menetapkan skor *sentimen* untuk setiap pasangan aspek *opini* [15]

2.2. Support Vector Machine (SVM)

Salah satu metode statistik yang dapat diterapkan untuk melakukan klasifikasi adalah *Support Vector Machine* (SVM). SVM merupakan suatu teknik untuk menemukan *hyperplane* yang bisa memisahkan dua *set data* dari dua kelas yang berbeda [16]. SVM memiliki kelebihan diantaranya adalah dalam menentukan jarak menggunakan *support vector* sehingga proses komputasi menjadi cepat [17] dan memiliki tingkat akurasi lebih tinggi [18].

SVM adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi [19]. Metode ini memiliki prinsip dasar *linier classifier*, yaitu kasus klasifikasi yang secara linier dapat dipisahkan, namun SVM telah dikembangkan agar dapat bekerja pada problem non-linier dengan memasukkan konsep kernel pada ruang kerja berdimensi tinggi. Pada ruang berdimensi tinggi, akan dicari *hyperplane* yang dapat memaksimalkan jarak (*margin*) antara kelas data.

Algoritma *sequential minimal optimization* (SMO), telah diusulkan untuk memecahkan

masalah regresi menggunakan SVM [20]. Algoritma ini perpanjangan dari algoritma SMO yang diusulkan oleh Platt untuk Desain pengklasifikasi SVM. Kecepatan komputasi dan kemudahan implementasi adalah beberapa fitur penting dari SMO algoritma. Beberapa disarankan perbaikan pada Platt Algoritma SMO untuk desain pengklasifikasi SVM. Secara khusus, ditunjukkan sumber penting dari ketidakefisienan yang disebabkan oleh cara SMO memelihara dan memperbarui satu nilai ambang batas. Mendapatkan petunjuk dari kriteria optimalitas yang terkait dengan kondisi Karush – Kuhn – Tucker (KKT) untuk masalah ganda, kami menyarankan penggunaan dua parameter ambang batas dan merancang dua versi modifikasi dari SMO untuk regresi itu jauh lebih efisien daripada SMO asli.

Komputasi perbandingan pada dataset menunjukkan bahwa modifikasi dilakukan jauh lebih baik daripada SMO asli. Setiap cabang menggambarkan hasil dari atribut yang diuji, dan setiap daun menggambarkan kelas.

1) K-Fold Cross Validation

Cross validation digunakan sebagai metode evaluasi hasil klasifikasi. Pengujian dilakukan untuk memprediksi *error rate*. *Training data* dibagi menjadi *K* buah subset secara acak dengan ukuran yang sama, satu diantara subset acak tersebut digunakan sebagai *testing data*. Setelah itu dilakukan iterasi sebanyak *K* kali dan dilakukan perhitungan *error rate* tiap *subset*. Kemudian hitung *error rate* pada setiap *subset*. Berdasarkan hasil *error rate* tiap *subset*, dihitung rata-ratanya untuk mendapatkan nilai *error rate* keseluruhan.

2) Confusion Matrix

Salah satu metode evaluasi yang digunakan untuk klasifikasi naïve bayes adalah *confusion matrix*. Matriks ini adalah salah satu *tools* penting dalam metode visualisasi yang digunakan pada mesin pembelajaran yang biasanya memuat dua kategori atau lebih [21]. Sebanyak setengah atau dua pertiga dari data keseluruhan digunakan untuk keperluan proses *training* sedangkan sisanya digunakan untuk keperluan *testing* [22].

3. HASIL DAN PEMBAHASAN

Data set di ambil dari halaman facebook peneliti yang telah di posting selama 6 bulan dan setiap 1 bulan sekali dilakukan pemostingan kembali agar mendapatkan jumlah *data set* yang diinginkan peneliti untuk menganalisa sentimen komentar publik.

Komentar yang telah didapatkan kemudian di-*crowling text* ke Ms. Excel Power Query untuk selanjutnya dilakukan tahap *Text Preprocessing*.

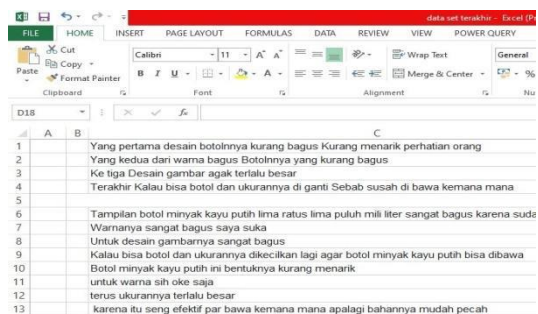
Data yang diolah berjumlah 100 komentar dari tiap orang yang berbeda. Data ini dibersihkan yang belum teratur untuk menjadi satu data set yang baik untuk digunakan pada WEKA. Pembersihan ini terdiri dari pemrosesan awal teks (*text preprocessing*), transformasi teks (*text transformation* atau *feature generation*), pemilihan fitur (*feature selection*), penemuan pola *text* atau *data mining (pattern discovery)*, dan *interpretation/ evaluation*.

3.1. Text Preprocessing

Preprocessing disini menggunakan Ms. Excel dengan memperhatikan data yang telah dipisahkan seperti saran pada komentar yang tidak berkaitan dengan pengolahan untuk metode klasifikasi untuk sentimen komentar dengan atribut yang telah ditentukan. Agar pada tahap klasifikasi lebih optimal dalam perhitungannya. Tahap *preprocessing* pada penelitian ini diantaranya:

1) Tokenizing

Pada tahap ini akan dilakukan penghapusan delimiter yaitu karakter angka dan karakter simbol kecuali karakter huruf dengan cara membuat daftar kode karakter yang diperlukan.



Gambar 2. Tokenizing

2) Case Folding

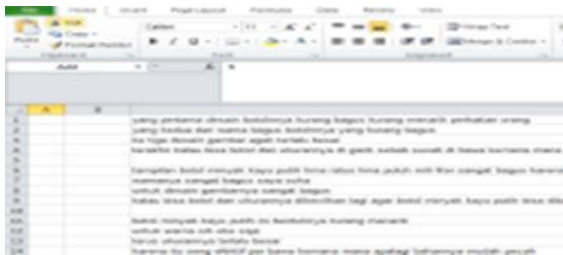
Pada tahap ini data pada komentar diubah menjadi huruf kecil semua dan karakter emotikon yang lain akan dihilangkan.



Gambar 3. Tahap Case Folding

Transformasi teks (*text transformation* atau *feature generation*). Pada tahap ini peneliti

membagi dan memilih tiap komentar sesuai dengan 4 atribut untuk metode klasifikasi yang akan di proses. Data dibagi terdiri atas data Bentuk, Ukuran, Warna dan Efisiensi.



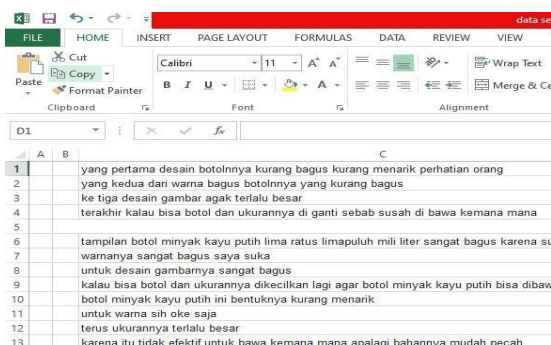
Gambar 4. Feature Generation

Featured Selection

Tahap ini merupakan tahap lanjut dari pengurangan dimensi pada proses transformasi teks yang terbagi atas:

1) Tahap *Filtering*

Pada tahap ini akan di hapus kata-kata yang sering muncul dan bersifat umum.



Gambar 5. Hasil Tahap *Filtering*

2) Tahap *Stemming*

Tahap ini kata-kata pada komentar akan di kembalikan ke bentuk dasarnya dengan menghilangkan imbuhan pada setiap kata.



Gambar 6. Hasil Tahap *Stemming*

3) *Convert Negation*

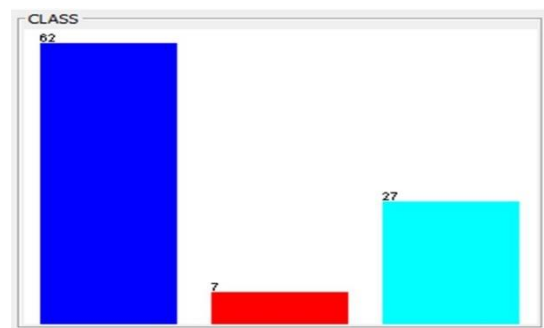
Pada tahap ini kata yang bersifat negasi akan dilakukan konversi seperti, “tidak”, “enggak”, “nga”, “seng”, dan lain-lain.



Gambar 7. Hasil Tahap *Convert Negation*

3.2. Pattern Discovery

Klasifikasi *data mining* dan pengujian data untuk klasifikasi data yang telah di preprocessing di atas, peneliti menggunakan software WEKA untuk mengklasifikasi data dimana dalam WEKA terdapat *naïve bayes*, *k-NN* dan *decision tree*.



Gambar 8. Visualize untuk Atribut Class

Dari gambar histogram di atas, diidentifikasi bahwa balok yang berwarna biru adalah jumlah keseluruhan untuk data negatif pada data *set* yang di peroleh dengan jumlah 62. Balok berwarna merah merupakan jumlah keseluruhan data positif yang di peroleh dengan jumlah sebanyak 7 data. Balok yang berwarna biru muda merupakan total keseluruhan data netral yang berjumlah 27 data.

1) *Classifier Zero-R*

Klasifikasi data ini adalah sebagai *baseline* untuk melihat data akurat untuk membandingkan tiga metode klasifikasi tersebut, dimana klasifikasi ini digunakan untuk standar penilaian tingkat keberhasilan akurasi data untuk tiga metode yang nantinya akan di uji.

Tabel 1. *Stratified Cross-Validation*

Stratified cross-validation		
Correctly Classified Instances	62	64,58%
Incorrectly Classified Instances	34	35,42%

Tabel 2. Confusion Matrix

Confusion Matrix				
a	b	c	<--	classified as
62	0	0		A = Negatif
7	0	0		B = Positif
27	0	0		C = Netral

Pada klasifikasi Zero-R dapat diketahui bahwa data yang di peroleh pada Classifier output yaitu *Run information*, *Stratified cross-validation*, *Detailed Accuracy by Class*, dan *Confusion Matrix*. Dari hasil ini diketahui nilai standar yang akan menjadi *baseline* untuk 3 metode yang dipakai untuk *Correctly Classified Instances* sebesar 64,58% dan untuk *Confusion Matrix* yaitu 62 data *instances* untuk class negatif, 7 *instances* untuk class positif dan 27 *instances* untuk class netral.

2) Support Vector Machine (SVM)

Pada pengolahan data *training* menggunakan Classifier SMO dapat diketahui hasil *Run information*, *Classifier model (full training set)*, *Time taken to build model*: 0,18 seconds, *Stratified cross-validation*, *Detailed Accuracy by Class*, dan *Confusion Matrix*. Hasil ini menunjukkan untuk *Correctly Classified Instances* berada pada standar *baseline* dengan jumlah 81,25% dan *Confusion Matrix* yang terklasifikasi keseluruhan sebanyak 54 *instances* untuk negatif, 3 *instances* untuk positif dan 21 *instances* untuk netral.

Tabel 3. *Stratified Cross-Validation*

<i>Stratified cross-validation</i>		
<i>Correctly Classified Instances</i>	81	81,25%
<i>Incorrectly Classified Instances</i>	18	18,75%

Tabel 4. *Confusion Matrix*

<i>Confusion Matrix</i>				
a	b	c	<--	classified as
54	0	8		A = Negatif
0	3	4		B = Positif
3	3	21		C = Netral

Interpretation/Evaluation

Berdasarkan *cross-validation*, metode yang di uji (SMO) memiliki *Correctly Classified Instances* lebih besar (81,25%) dan *Incorrectly Classified Instances* lebih kecil (18,75%). *Confusion Matrix* menunjukkan metode ini memperoleh akurasi yang baik (*class* negatif = 54 *instances* terklasifikasi dengan benar, 21 data *instances* terklasifikasi pada *class* netral, dan *class* positif = 3 *instances*). Artinya metode SMO dapat menganalisis sentimen komentar publik untuk penelitian ini.

3.3. Class Balancing

Class Balancing dilakukan untuk membuat jumlah *data set* untuk tiap atribut menjadi

seimbang. Hasil analisis menggunakan *cost sensitive classifier* disajikan pada tabel 5 dan 6.

Tabel 5. *Stratified Cross-Validation*

<i>Stratified cross-validation</i>		
<i>Correctly Classified Instances</i>	80	83,33%
<i>Incorrectly Classified Instances</i>	16	16,67%

Tabel 6. *Confusion Matrix*

<i>Confusion Matrix</i>				
a	b	c	<--	classified as
57	0	5		a = Negatif
1	3	3		b = Positif
6	1	20		c = Netral

3.4. Testing dan Training

Testing menggunakan 10 data dengan label atau *class* yang belum diketahui, sedangkan *training* memakai data yang sebelumnya telah di proses menggunakan metode klasifikasi SMO. Hasil *testing* dan *training* disajikan pada tabel 7.

Tabel 7. Hasil *Predicted Class*

Predicted CLASS	Prediction Margin	
	Statistic	value
Netral = 5 komentar	Minimum	-0,995
Negatif = 3 komentar	Maximum	0,993
Positif = 2 komentar	Maen	0,582
	StdDev	0,627

3.5. Diskusi

Teknik *text mining* klasifikasi SMO mempunyai tingkat prediksi yang sangat akurat. Klasifikasi ini dapat menganalisa *data instances* dengan baik untuk tiap atribut *class* dengan tingkat akurasi 81%. Peneliti ini mendapatkan hasil prediksi sebesar 83,33% untuk *Correctly Classified Instances* dan 18,75% *Incorrectly Classified Instances*. Hasil *class* prediksi untuk positif 2 komentar, negatif 3 komentar dan netral 5 komentar berarti presentasi prediksi menggunakan klasifikasi SMO mendapatkan hasil yang sangat baik.

Hasil diatas menunjukkan fitur kemasan minyak kayu putih 550 ml seperti Bentuk, Warna, Ukuran, dan Efisiensi berpengaruh pada desain ulang kemasan botol minyak kayu putih 550 ml. Komentar pengguna pada sosial media untuk fitur kemasan seperti bentuk yang memiliki nilai *class* negatif sebanyak 47 komentar, sedangkan *class* positif sebanyak 36 komentar, serta untuk *class* netral sebanyak 13 komentar. Oleh karena *class* negatif lebih besar dari *class* lainnya, maka kemasan baru harus memiliki daya tarik yang lebih

tinggi. Beberapa usulan yang perlu diperhatikan berupa bentuk yang sederhana, bentuk regular, bentuk lebih kecil dari desain awal tanpa mengurangi volumenya, bentuk yang lebih nyaman saat digunakan atau dibawa, dan bentuk cembung dari pada bentuk cekung. Atribut ukuran kemasan dapat mempengaruhi desain ulang kemasan dengan nilai *class* negatif 82 komentar, *class* positif 11 komentar, sedangkan *class* netral hanya 3 komentar. Untuk ukuran masih harus disesuaikan dengan kondisi penggunaan, seperti jika untuk dibawa sebagai oleh-oleh dapat dibuat lebih kecil tanpa menurunkan volume kemasan. Efisiensi mempengaruhi desain ulang pada kemasan yang sangat harus diperhatikan 3 hal seperti, penambahan pegangan atau genggaman agar mudah digunakan atau dibawa, penggunaan bahan selain kaca untuk mengurangi beban pada kemasan tanpa menghilangkan khasiat dari minyak kayu putih tersebut, dan lekukan pada kepala kemasan botol dibuat menlekung agar mudah digunakan. Hal ini dikarenakan untuk atribut efisiensi memiliki nilai *class* negatif 94 komentar, dan *class* positif 2 komentar. Untuk atribut warna, kemasan diusulkan lebih bening atau transparan agar konsumen dapat melihat langsung warna dari minyak kayu putih. Usulan ini dikarenakan atribut warna memiliki nilai *class* negatif 46 komentar, *class* positif 32 komentar, dan untuk *class* netral 18 komentar.

4. KESIMPULAN

Berdasarkan hasil penelitian, disimpulkan bahwa SMO mempunyai tingkat prediksi yang sangat akurat untuk menganalisis sentimen konsumen tentang fitur kemasan minyak kayu putih 550 ml. Fitur kemasan yang berpengaruh dalam perubahan desain kemasan produk tersebut adalah bentuk, warna, ukuran, dan efisiensi.

DAFTAR PUSTAKA

- [1] Smith, H. dan Idrus, S. (2016). Estimated Production of Cajuput Oil on Small.
- [2] Dwiningih, N. (2002). Desain Produk dan Manajemen Kualitas.
- [3] Utomo, B. (2014). Analisis Sentimen pada Twitter Menggunakan Text Mining.
- [4] Liu, B. (2012). Sentiment Analysis and Opinion Mining.
- [5] Ibrahim, R. dan Othman, Z. A. (2012). Opinion Mining and Sentiment Analysis.
- [6] Hapsari, Y., Hidayattullah, M. F., & Khambali, M. (2018). Opinion Mining Terhadap Toko Online Di Media Sosial Menggunakan Algoritma Naïve Bayes.
- [7] Han, J. (2012). Data Mining Concepts and Techniques.
- [8] Turban (2005). Decision Support Systems and Intelligent Systems.
- [9] Kamber, M. (2006). Classification and Prediction.
- [10] Mashreghi, Z. dan Esmaeili, M. (2017). A Combination Method for Improving Text Summarization.
- [11] Rose, J. dan Lennerholt, C. (2017). Low Cost Text Mining as a Strategy for Qualitative Researchers.
- [12] Hassani, H., Beneki, C., Unger, S. and Mazinani, M. J. (2020). Text Mining in Big Data Analytics.
- [13] Lukitowati, H. L. dan Paryatno, L. (2020). Evaluasi Pelaksanaan Webinar Pusdiklat Perdagangan Dengan Menggunakan Text Mining: Analisis Saran Peserta Webinar.
- [14] Sudalma, S. (2015). The Effect of SO₂ and NO₂ from Transportation and Stationary Emissions Sources to SO₄²⁻ and NO₃⁻ in Rain Water in Semarang.
- [15] Gao, S., Tang, O., Wang, H. dan Yin, P. (2018). Identifying Competitors through Comparative Relation Mining of Online Reviews in the Restaurant Industry.
- [16] Vapnik, V. N. (1999). An Overview of Statistical Learning Theory, 10(5), 988–999.
- [17] Cortes, C., & Vapnik, V. (1995). Support-Vector Network. Machine Learning 20, 273-297.
- [18] Rachman, F. dan Purnami, S.W. (2012). Klasifikasi Tingkat Keganasan Breast Cancer dengan Menggunakan Regresi Logistik Ordinal dan Support Vector Machine (SVM). Jurnal Sains dan Seni ITS, Vol. 1, No. 1 ISSN: 2301-928X.
- [19] Santosa, B. (2007). Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. Yogyakarta: Graha Ilmu.
- [20] Schölkopf, B. dan Smola, A. J. (2002). Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond. Massachusetts Institute of Technology. ISBN 0-262-19475-9.
- [21] Horn, C. dan Horn, C. (2010). Analysis and Classification of Twitter messages Analyse und Kategorisierung von Twitter Nachrichten.
- [22] Kantardzic, M. (2003). Data Mining: Concepts, Models, Methods, and Algorithms.