

KLASIFIKASI ULASAN PALSU MENGGUNAKAN BORDERLINE OVER-SAMPLING (BOS) DAN SUPPORT VECTOR MACHINE (SVM) (STUDI KASUS: ULASAN TEMPAT MAKAN)

Aisyah Awalina^{*1}, Fitra A. Bachtiar², Indriati³

^{1,2,3}Universitas Brawijaya, Malang

Email: ¹awalinaa@gmail.com, ²fitra.bachtiar@ub.ac.id, ³indriati.tif@ub.ac.id

*Penulis Korespondensi

(Naskah masuk: 18 Oktober 2021, diterima untuk diterbitkan: 21 Februari 2022)

Abstrak

Kemudahan memperoleh informasi saat ini telah banyak membantu manusia, salah satu mencari ulasan untuk tempat makan baru. Pencarian ulasan ini dipicu karena pengunjung tidak mengetahui layanan dari tempat tersebut. Ulasan juga dapat menguntungkan penjual, karena mereka mengetahui pengalaman yang dimiliki pengunjungnya. Oleh karena itu, ulasan palsu dimanfaatkan banyak orang untuk membuat ulasan palsu. Ulasan palsu bisa secara efektif dibedakan menggunakan *machine learning*. Namun, banyak dari dataset ulasan palsu ini tidak seimbang (*imbalanced dataset*) sehingga dapat mempengaruhi hasil klasifikasi. Oleh karena itu, penelitian ini menggunakan metode BOS untuk mengatasi tidak seimbangnya data dan melakukan klasifikasi dengan metode SVM. Adapun tahapan dari penelitian yaitu preprocessing, lalu pembobotan kata dengan TF-IDF dan fitur sentimen menggunakan *lexicon-based features*, dilanjutkan proses menyeimbangkan dataset dengan BOS, setelah itu proses klasifikasi oleh SVM. Adapun langkah dalam pengujian BOS dan SVM yaitu pembagian data latih dan uji dengan 80%:20%, setelah itu pencarian parameter terbaik pada data latih dengan *5-fold cross validation*, dan dievaluasi dengan data uji. Adapun nilai parameter terbaik pada BOS dan SVM yaitu N dengan nilai 400% dimana hasil evaluasi akurasi dengan nilai 78,6%; *precision* dengan nilai 19,7%; *recall* dengan nilai 17,1%; *f-measure* dengan nilai 14,4%; dan *g-mean* dengan nilai 32%. Oleh karena itu, penggunaan BOS dapat meningkatkan hasil evaluasi dari terhadap klasifikasi ulasan palsu.

Kata kunci: *ulasan palsu, imbalance dataset, SVM, BOS*

SPAM REVIEW CLASSIFICATION USING BORDERLINE OVER-SAMPLING AND SUPPORT VECTOR MACHINE ALGORITHM

Abstract

The convenience of obtaining information nowadays has helped many people such as looking for reviews for new places to eat. The search for reviews was triggered because visitors were not aware of the services of the place. Reviews can also benefit sellers, because they know the experience their visitors have had. Therefore, many people abuse reviews to create spam reviews. Spam reviews can be effectively resolved using machine learning. However, many of these spam review datasets are imbalanced and thus may affect classification results. In this study, BOS algorithm was used to overcome data imbalances, and SVM algorithm for the classification of spam reviews. The stages of the research are preprocessing, then weighting words with TF-IDF and sentiment features using lexicon-based features, followed by the process of balancing the dataset with BOS, and classification process with SVM. Step in testing BOS and SVM are split data of training and test data with 80%:20%, after that the search for the best parameters in the training data with 5-fold cross-validation, and evaluated with test data. The best parameter values for BOS and SVM were N with a value of 400% where the results of the accuracy evaluation were 78.6%; precision with a value of 19.7%; recall with a value of 17.1%; f-measure with a value of 14.4%; and g-mean with a value of 32%. Therefore, use of BOS can improve the evaluation results from the classification of spam reviews.

Keywords: *spam review, imbalance dataset, SVM, BOS*

1. PENDAHULUAN

Perkembangan teknologi saat ini, kita dapat dengan mudah untuk memperoleh informasi seperti

mencari ulasan dari tempat makan baru (Resendes, 2020). Pencarian ulasan tersebut dipicu karena pengunjung tidak mengetahui layanan yang dimiliki

oleh tempat makan tersebut sehingga dengan mencari pengunjung akan mengetahui layanan dari tempat tersebut. Berdasarkan data survei yang dilakukan oleh TripAdvisor (2018) yang dilakukan di 5 negara yaitu Amerika, Inggris, Perancis, Spanyol dan Italia menunjukkan dengan rata-rata nilai 91%, ulasan pada situs *online* mempengaruhi untuk pemilihan tempat makan.

Ulasan palsu sendiri memberikan hasil sangat efektif dalam jangka pendek bagi penjual karena dalam 2 minggu saja penjual dapat meningkatkan *rating* tokonya dengan rata-rata 0,16 dengan kenaikan berlipat ganda dari 5 sampai 10 ulasan per minggu (Proserpio, et al., 2020). Dari penjelasan tersebut, diperlukan deteksi ulasan palsu sehingga tidak merugikan pihak manapun. Salah satu cara untuk membedakan ulasan palsu dapat menggunakan *machine learning*. Akan tetapi, ulasan palsu juga memiliki tantangan tersendiri, karena kebanyakan dataset ulasan palsu sangat sedikit dibandingkan ulasan asli sehingga menyebabkan ketidakseimbangan data (*imbalanced dataset*).

Penelitian oleh Ott, et. al (2013) menggunakan 4 pendekatan, dimana pendekatan *n-grams* memiliki nilai tertinggi dibandingkan dengan pendekatan lain. Metode SVM digunakan untuk melakukan proses evaluasi dengan hasil *precision* tertinggi adalah 89,1%. Penelitian oleh Mukherjee, et. al (2013) melakukan menggunakan metode *Support Vector Machine* (SVM) dengan *5-fold cross validation* pada dataset Yelp dan didapatkan hasil akurasi tertinggi yaitu 86,1%. Penelitian oleh Elsharif & Gherbi (2017) melakukan klasifikasi ulasan palsu menggunakan metode *machine learning* dan didapatkan metode SVM menjadi metode yang terbaik dengan hasil akurasi tertinggi yaitu 81,35%.

Pada penelitian oleh Elsharif & Gherbi (2017) melakukan klasifikasi ulasan palsu menggunakan berbagai metode *machine learning* dengan pendekatan metode pada analisis sentimen. Disebutkan bahwa dasar dari analisis sentimen dapat dihitung dari polaritas pada teks di level dokumen, kalimat dan aspek dengan nilai positif atau negatif. Namun, dalam penelitian Elmungrhi & Gherbi memiliki kekurangan karena kurang merepresentasikan dari sentimen yaitu nilai positif atau negatif. Penelitian dari Melleng et. al (2019) melakukan deteksi ulasan palsu dengan pendekatan sentimen dan emosi. Adapun analisa sentimen dan emosi menggunakan 3 *lexicon* dan 1 API analisis. Hasil dari penelitian ini menemukan bahwa representasi emosi dan sentimen bekerja secara efektif dalam klasifikasi ulasan palsu.

Penelitian terhadap untuk mengatasi ketidakseimbangan data telah banyak dilakukan salah satunya penelitian Nguyen et. al (2011) menggunakan metode *Borderline Over-sampling* (BOS). Dimana metode ini menggunakan pendekatan metode *over-sampling* dengan metode SVM yang menjadi dasar *classifier*. Hasil yang

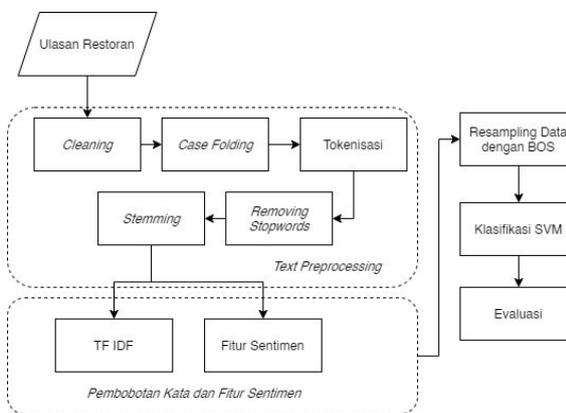
diperoleh, metode ini cukup efektif dibandingkan metode *over-sampling* lainnya.

Dalam studi literatur diatas terlihat bahwa metode SVM memiliki kinerja yang baik dalam klasifikasi ulasan palsu. Namun, *dataset* ulasan palsu sendiri kebanyakan *imbalance*. Sehingga dalam penelitian mengkombinasikan SVM dan BOS untuk penyelesaian masalah penelitian yang diangkat. Metode BOS sendiri terbukti efektif dalam penyelesaian *imbalance dataset* di berbagai dataset. Sehingga dalam penelitian ini, metode BOS digunakan untuk menyelesaikan masalah *imbalance dataset* pada ulasan palsu, dan SVM untuk proses klasifikasinya.

Berdasarkan uraian permasalahan latar belakang di atas serta penelitian-penelitian yang dilakukan sebelumnya, maka penelitian ini akan melakukan klasifikasi ulasan palsu dengan metode BOS untuk mengatasi ketidakseimbangan data, dan metode SVM untuk proses klasifikasinya. Adapun ekstraksi fitur yang digunakan adalah TF-IDF dan fitur sentimen.

2. METODE PENELITIAN

Adapun gambaran umum pada penelitian yang dapat dilihat pada Gambar 1. Adapun beberapa tahapan yang dilakukan yaitu proses *text preprocessing*, pembobotan kata dengan TF-IDF, fitur sentimen, *resampling data* dengan metode BOS, klasifikasi dengan metode SVM dan evaluasi.



Gambar 1. Gambaran Umum Penelitian

Berdasarkan Gambar 1 dapat dilihat gambaran umum yang akan dilakukan dalam penelitian ini. Dataset yang merupakan ulasan restoran masuk kedalam proses *text preprocessing*. Setelah di *text preprocessing* kemudian dilakukan pembobotan kata dan fitur sentimen. Setelah fiturnya telah didapatkan, dilanjutkan proses *resampling* data dengan metode BOS. Kemudian dilakukan klasifikasi dengan metode SVM dan proses evaluasi.

2.1. Data yang digunakan

Dalam penelitian ini menggunakan ulasan-ulasan yang telah digunakan dalam penelitian

Rayana (2015). Data ulasan tersebut diambil dari laman Yelp yang telah dilakukan pelabelan dengan label kelas. Yelp sendiri merupakan perusahaan publik yang didirikan pada tahun 2004 oleh Russel Simmons dan Jeremy Stoppelman. Area yang digunakan adalah restoran Johny's Luncheonette di kota New York. Data yang digunakan yaitu 500 data dengan proporsi kelas ulasan asli sebanyak 430 dan kelas ulasan palsu sebanyak 70.

2.2. Alur Metode

Adapun alur metode dalam penelitian ini, yaitu *text preprocessing*, ekstraksi fitur dengan TF-IDF dan fitur sentimen, kemudian proses *resampling data* dilanjutkan dengan proses klasifikasi dengan SVM. Proses terakhir adalah evaluasi.

2.2.1 Text Preprocessing

Proses pertama yaitu data ulasan restoran yang masuk ke proses *text preprocessing* yang bertujuan mereduksi serta mengurangi *noise* pada data. Adapun tahapan dari *text preprocessing* melakukan proses *cleaning* dimana melakukan penghapusan terhadap komponen-komponen selain huruf (kata). Proses selanjutnya yaitu *case folding* dimana dalam penelitian dilakukan proses merubah semua huruf menjadi huruf kecil (*lower case*). Dilanjutkan dengan proses tokenisasi dimana struktur teks sebelumnya diubah menjadi token. Lalu dilakukan proses penghapusan *stopword* dengan *stopwords* dari Savoy (2005) yang dapat diakses dalam laman resmi *Université de Neuchâtel*. Selanjutnya proses *stemming* dengan *Snowball Stemmer*.

2.2.2 Pembobotan kata dan fitur sentimen

Setelah proses *text preprocessing*, dilakukan proses pencarian fitur. Pada penelitian ini menggunakan dua pendekatan yaitu menggunakan *bag-of-words* untuk pembobotan kata menggunakan TF-IDF dan fitur sentimen. Adapun proses pembobotan kata yang digunakan adalah normalisasi TF-IDF. Kemudian fitur sentimen sendiri mempunyai 2 fitur, yaitu jumlah kata positif dan jumlah kata negatif dengan menggunakan *lexicon* Bing Liu. *Lexicon* Bing Liu sendiri memiliki 6789 kata yang dibagi menjadi 2006 kata positif dan 4783 kata negatif. Jumlah kata dari setiap kelas positif dan negatif akan dilakukan normalisasi *min-max*.

2.2.3 Resampling data

Setelah dilakukan proses pencarian fitur, dilakukan proses *resampling*. Dataset akan dibagi menjadi data latih dan data uji dengan proporsi 80%:20%. Setelah itu data latih akan dilakukan *resampling* sebanyak N dengan metode BOS.

2.2.4 Klasifikasi

Setelah proses *resampling data* dilakukan, maka selanjutnya data latih yang telah di-*resampling* akan dilatih menggunakan metode *Sequential Training*

SVM. Setelah itu dilakukan proses klasifikasi dengan metode SVM ulasan palsu dan asli. Kernel yang digunakan dalam metode SVM ini adalah *Gaussian RBF*.

2.2.5 Rancangan Pengujian

Adapun perancangan pengujian yang dilakukan yaitu menguji parameter *sigma*, *lambda*, *learning rate*, konstanta C, jumlah iterasi maksimum dan *epsilon* pada metode SVM serta parameter N pada metode BOS. Adapun pengujian dari parameter metode SVM dilakukan dengan tanpa *resampling* data latih dimana pengujian dijalankan sebanyak 5 kali. Sedangkan pengujian parameter metode BOS dilakukan sebanyak 5 kali, dimana parameter terbaiknya didapatkan dari proses *k-fold cross validation* pada data latih. Adapun parameter terbaiknya didapatkan dengan cara mencari nilai *g-mean* tertinggi pada setiap hasil parameter SVM. Urutan pencarian parameter terbaik dimulai dari *sigma*, *lambda*, *learning rate*, konstanta C, jumlah iterasi maksimum kemudian *epsilon*.

2.2.6 Evaluasi

Setelah hasil dari pengujian metode BOS dan SVM didapatkan dilakukan evaluasi. Evaluasi yang digunakan dalam metode ini menggunakan *confusion matrix*, dimana nilai-nilai dari hasil *confusion matrix* nantinya akan digunakan dalam penilaian akurasi, *recall*, *precision*, *F-Measure* dan *G-mean*. Adapun contoh *confusion matrix* dapat dilihat pada Tabel 1.

Tabel 1. Confusion Matrix

		Aktual	
		True	False
Prediksi	True	True Positive (TP)	False Positive (FP)
	False	False Negative (FN)	True Negative (TN)

Dari *confusion matrix* kita dapat menghitung penilaian evaluasi yang digunakan. Adapun persamaan dari penilaian evaluasi tersebut dapat dilihat pada Persamaan (5), (6), (7), (8) dan (9).

$$Akurasi = \frac{TP+TN}{(TP+TN+FP+FN)} \quad (5)$$

$$Precision = \frac{TP}{(FP+TP)} \quad (6)$$

$$Recall = \frac{TP}{(FN+TP)} \quad (7)$$

$$F - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (8)$$

$$G - Mean = \sqrt{\left(\frac{TP}{TP+FN}\right) \times \left(\frac{TN}{TN+FP}\right)} \quad (9)$$

3. DASAR TEORI

3.1. Ulasan Palsu

Opinion Spam atau Opini palsu merupakan kegiatan ilegal yang sengaja untuk menggiring sentimen dengan memberikan opini positif untuk

mempromosikan atau opini negatif untuk merusak reputasi (Jindal & Liu, 2008). Opini palsu dibagi 3 menurut penelitian Jindal & Liu (2008), yaitu:

1. *Untruthful opinion*: penulis ulasan menyimpangkan ulasan sebenarnya dengan tujuan mempromosikan produknya atau memfitnah saingan mereka. Tipe ini sulit dibedakan bahkan melalui pakar sekalipun.
2. *Reviews on brand only*: penulis ulasan hanya berkomentar tentang merek atau produsen nya, bukan produk yang dijual.
3. *Non-reviews*: ulasan berisikan iklan atau ulasan yang tidak relevan.

3.2. Support Vector Machine

Support Vector Machine (SVM) dikembangkan pada tahun 1992 oleh Boser pada acara *Annual Workshop on Computational Learning Theory*. Berbeda dengan konsep neural network yang berusaha mencari *hyperplane* (garis) pemisah antar *class*, konsep SVM lebih berusaha untuk menemukan *hyperplane* yang terbaik pada *input space*. Prinsip dasar dari SVM adalah *linear classifier* dengan memaksimalkan *margin* yang merupakan jarak pemisah antara *class*. SVM sendiri hanya menggunakan *Support Vector*, yaitu titik data yang terpilih karena berkontribusi, untuk membentuk model yang akan digunakan pada proses klasifikasi. Salah satu alternatif untuk melatih dan mencari *support vector* dari SVM dapat menggunakan metode sekuensial dari penelitian Vijayakumar (1999).

3.3. Borderline Over-sampling

Adapun tahapan dari algoritme *Borderline Over-sampling* yang dikembangkan oleh Nguyen et. al (2011):

1. Inisialisasi nilai X, N, k, m , dimana X adalah data latih yang digunakan, N adalah banyaknya *sampling level* yang digunakan, k adalah nilai *nearest neighbor* seperti metode SMOTE, dan m adalah nilai *nearest neighbor* untuk memutuskan tipe *sampling* yang nanti akan diproses.
2. Ekstraksi X^+ terhadap X , dimana X^+ adalah data latih pada kelas minoritas.
3. Menghitung variabel T dapat dilihat dari Persamaan (1).

$$T = \left(\frac{N}{100}\right) \times |X^+| \quad (1)$$

Dimana T adalah jumlah dari *artificial instance* yang dibuat pada kelas minoritas.

4. Menghitung variabel SV^+ dengan melatih memakai metode SVM terhadap variabel X . SV^+

sendiri merupakan hasil *support vector* pada kelas minoritas.

5. Menginisialisasi *array amount* dengan mendistribusikan variabel T secara merata pada variabel SV^+ . *Array amount* adalah sebuah *array* yang berisikan jumlah *oversample* pada setiap *support vector* kelas minoritas.
6. Menghitung nn pada X^+ dengan metode *K-Nearest Neighbor* (KNN) dengan nilai sebanyak k .
7. Dilakukan perulangan pada setiap *support vector* positif ke- i di variabel SV^+ ($svpositif_i \in SV^+$).
 - a. Melakukan perhitungan variabel m *nearest neighbors* terhadap variabel X .
 - b. Jika jumlah dari m *nearest neighbors* dari kelas minoritas lebih kecil dari jumlah setengah dari m *nearest neighbors*, maka sepanjang garis setiap $svpositif_i$ bergabung dengan k positif *nearest neighbor*, membuat $amount[i]$ untuk *artificial instance* menggunakan ekstrapolasi (memperluas area kelas positif) dengan Persamaan (2).

$$xpositif = svpositif_i + \rho(svpositif_i - nn[i][j]) \quad (2)$$

- c. Jika tidak, maka menggunakan *interpolasi* (seperti di metode SMOTE) untuk mengkonsolidasikan area *boundary* dari kelas positif saat ini. Perhitungan tersebut dapat dilihat pada Persamaan (3).

$$xpositif = svpositif_i + \rho(svpositif_i - nn[i][j]) \quad (3)$$

Dimana $svpositif_i$ adalah *support vector* ke- i pada kelas minoritas. Kemudian $nn[i][j]$ adalah *nearest neighbor* positif ke- j dari $svpositif_i$ dimana nilai j adalah $1 \leq j \leq k$. Dan ρ adalah nilai random dengan rentang nilai $[0, 1]$

8. Menggabungkan $\{xpositif\}$ sebagai set baru dari train set dengan dataset sebenarnya. Disimpan ke dalam variabel baru X_{new} seperti pada Persamaan (4).

$$X_{new} = X \cup xpositif \quad (4)$$

4. HASIL DAN PEMBAHASAN

Dalam subbab ini akan dijelaskan beberapa hasil pengujian yang telah dilakukan dalam klasifikasi ulasan palsu. Adapun hasil pengujian pengaruh parameter yang dilakukan pada SVM yaitu *Sigma*, *Lambda*, *Learning Rate*, Konstanta C , iterasi maksimum dan *epsilon*. Sedangkan hasil pengujian pengaruh yang dilakukan pada BOS yaitu parameter N .

4.1. Pengujian Nilai Parameter Sigma

Hasil pengujian pengaruh nilai parameter Sigma pada *Sequential Training SVM* ditunjukkan pada Tabel 2. Adapun parameter lain yang digunakan adalah $\lambda=0,5$; $\gamma=0,001$; $C=1$ dan iterasi maksimum=20.

Tabel 2. Hasil Pengujian Nilai Parameter Sigma

Sigma	Akura- rasi	Preci- sion	Recall	F- Measure	G- mean
0,5	86,2%	1,5%	20%	2,9%	5,5%
1	14,2%	98,5%	13,9%	24,3%	4,3%
1,5	14,2%	98,5%	13,9%	24,3%	4,3%
2	14,2%	98,5%	13,9%	24,3%	4,3%
2,5	14,2%	98,5%	13,9%	24,3%	4,3%
4	14,4%	100%	14,1%	24,3%	4,3%
8	14,4%	100%	14,1%	24,6%	4,3%

Berdasarkan hasil pengujian dari Tabel 2 menunjukkan hasil akurasi dan *gmean* tertinggi pada saat nilai *sigma* (σ) = 0,5 akan tetapi hasil nilai *precision* dan *f-measure* memiliki nilai terendah dibandingkan nilai *sigma* yang lain. Hal tersebut dikarenakan banyaknya kelas ulasan palsu yang tidak teridentifikasi yang disebabkan tidak seimbangnya kelas ulasan asli dan palsu pada data latih. Saat *sigma* bernilai 1 mengalami penurunan akurasi yang sangat signifikan pada sistem, namun mengalami kenaikan drastis pada *precision*. Hal tersebut dikarenakan data uji pada kelas ulasan asli terdeteksi sebagai ulasan palsu. Nilai-nilai *sigma* setelah bertambah mengalami hasil yang tetap konstan.

4.2. Pengujian Nilai Parameter Lambda

Hasil pengujian pengaruh nilai parameter Lambda pada *Sequential Training SVM* ditunjukkan pada Tabel 3. Adapun parameter lain yang digunakan adalah $\sigma=0,5$; $\gamma=0,001$; $C=1$ dan iterasi maksimum=20.

Tabel 3. Hasil Pengujian Nilai Parameter Lambda

Lam- bda	Akurasi	Preci- sion	Recall	F- Measur- e	G-mean
0,01	85,4%	1,5%	20%	2,9%	5,5%
0,1	85,6%	1,5%	20%	2,9%	5,5%
0,5	86,2%	1,5%	20%	2,9%	5,5%
1	80,8%	11,1%	11,9%	11,1%	23,4%
2	50,6%	42,9%	12,2%	18,2%	42,8%

Berdasarkan hasil pengujian dari Tabel 3 menunjukkan hasil nilai akurasi yang tertinggi adalah ketika nilai *lambda* = 0,5 yaitu dengan akurasi = 86,2%, akan tetapi nilai *precision* yang dihasilkan relatif kecil yaitu sebesar 1,5%. Hal tersebut dikarenakan tidak seimbangnya data latih antara ulasan asli dan palsu. Ketika nilai *lambda* = 2, nilai evaluasi dari *g-mean* tertinggi dengan nilai sebesar 42,8%. Akan tetapi terjadi *trade off* pada hasil nilai akurasi sistem. Nilai *lambda* berpengaruh dalam mendapatkan perpotongan *hyperplane* yang lebih baik dikarenakan saat nilai *lambda* mengecil,

maka jarak *margin* akan mengecil dan titik dari dalam *margin* akan bergerak ke luar *margin* (Hasanah, et al., 2016).

4.3. Pengujian Nilai Parameter Learning Rate

Hasil pengujian pengaruh nilai parameter *learning rate* pada *Sequential Training SVM* ditunjukkan pada Tabel 4. Adapun parameter lain yang digunakan adalah $\sigma=0,5$; $\lambda=2$; $C=1$ dan iterasi maksimum=20.

Tabel 4. Hasil Pengujian Nilai Parameter Learning Rate

Learnin- g Rate	Akuras- i	Precisi- on	Recall	F- Measu- re	G- mean
0,0001	85,6%	1,53%	10%	2,67%	5,52%
0,001	50,6%	42,9%	12,2%	18,2%	42,8%
0,01	86%	0%	0%	0%	0%
0,1	86%	0%	0%	0%	0%
1	86%	0%	0%	0%	0%

Berdasarkan hasil pengujian dari Tabel 4 menunjukkan hasil nilai akurasi yang tertinggi adalah ketika nilai *learning rate* = 0,01, 0,1 dan 1 yaitu dengan hasil akurasi sebesar 86%, akan tetapi nilai *precision* yang dihasilkan sebesar 0% dimana data uji dari kelas ulasan palsu tidak dapat diidentifikasi yang disebabkan oleh tidak seimbangnya data latih pada kelas ulasan asli dan palsu. Nilai dari *g-mean* paling tinggi pada nilai *learning rate* = 0,001 namun mengalami *trade off* pada hasil akurasi sistem. *Learning rate* digunakan untuk mengontrol kecepatan proses *learning* yang dibutuhkan untuk mencapai konvergensi (Pratama, et al., 2018)

4.4. Pengujian Nilai Parameter Konstanta C

Hasil pengujian pengaruh nilai parameter konstanta *C* pada *Sequential Training SVM* ditunjukkan pada Tabel 5. Adapun parameter lain yang digunakan adalah $\sigma=0,5$; $\lambda=2$; $\gamma=0,001$ dan iterasi maksimum=20.

Tabel 5. Hasil Pengujian Nilai Parameter Konstanta C

Konsta- nta C	Akuras- i	Precisi- on	Recall	F- Measur- e	G- mean
0,0001	17,2%	98,5%	14,3%	24,9%	15,5%
0,001	85,8%	1,54%	20%	2,86%	5,6%
0,01	85,6%	1,54%	10%	3,67%	5,6%
0,1	58,6%	35,8%	13,2%	18%	42,1%
0,5	51%	42,9%	12,3%	18,3%	43,2%
1	50,6%	42,9%	12,2%	18,2%	42,8%
5	49,8%	44,1%	12,3%	18,6%	43,1%

Berdasarkan hasil pengujian dari Tabel 5 menunjukkan hasil nilai akurasi yang tertinggi adalah ketika konstanta *C* = 0,001 yaitu dengan hasil akurasi sebesar 85,4%, namun hasil dari nilai *precision* relatif kecil sebesar 1,538%. Hal tersebut disebabkan tidak seimbangnya data latih dari data ulasan asli dan palsu. Pada nilai konstanta *C* = 0,5 menghasilkan nilai *g-mean* tertinggi sebesar 43,15%,

nilai dari *precision* juga cukup bagus yaitu 42,89%, akan tetapi mengalami *trade off* pada nilai akurasi. Tujuan dari parameter kompleksitas sendiri adalah mengontrol *trade off* dari *margin* dan meminimalkan nilai dari error (Vijayakumar & Wu, 1999). Proses meminimalkan error ini yang digunakan pada proses *training* yang menghitung nilai bobot dan bias (Parapat, et al., 2018).

4.5. Pengujian Nilai Parameter Iterasi Maksimum

Hasil pengujian pengaruh nilai parameter iterasi maksimum pada *Sequential Training SVM* ditunjukkan pada Tabel 6. Adapun parameter lain yang digunakan adalah $\sigma=0,5$; $\lambda=2$; $\gamma=0,001$ dan $C=0,5$.

Tabel 6. Hasil Pengujian Nilai Parameter Iterasi Maksimum

Iterasi Maksimum	Akurasi	Precisi on	Recall	F-Measure	G-mean
10	51%	42,9%	12,3%	18,3%	43,2%
20	51%	42,9%	12,3%	18,3%	43,2%
30	51%	42,9%	12,3%	18,3%	43,2%
40	51%	42,9%	12,3%	18,3%	43,2%
50	51%	42,9%	12,3%	18,3%	43,2%

Berdasarkan hasil pengujian dari Tabel 6 menunjukkan nilai evaluasi keseluruhan yang konstan pada seluruh jumlah iterasi maksimum yaitu akurasi = 51%, *precision* = 42,89%, *recall* = 12,25%, *f-measure* = 18,3%, dan *g-mean* = 43,15%. Sehingga jumlah iterasi maksimum pada penelitian tidak terlalu berpengaruh terhadap tingkat akurasi.

4.6. Pengujian Nilai Parameter Epsilon

Hasil pengujian pengaruh nilai parameter epsilon pada *Sequential Training SVM* ditunjukkan pada Tabel 7. Adapun parameter lain yang digunakan adalah $\sigma=0,5$; $\lambda=2$; $\gamma=0,001$, $C=0,5$ dan iterasi maksimum=20.

Tabel 7. Hasil Pengujian Nilai Parameter Epsilon

Epsilon	Akurasi	Precisi on	Recall	F-Measure	G-mean
10 ⁻⁴	51%	42,9%	12,3%	18,3%	43,2%
10 ⁻⁵	51%	42,9%	12,3%	18,3%	43,2%
10 ⁻⁶	51%	42,9%	12,3%	18,3%	43,2%

Berdasarkan hasil pengujian dari Tabel 7 menunjukkan nilai evaluasi keseluruhan yang konstan pada seluruh nilai epsilon yang diujikan yaitu akurasi = 51%, *precision* = 42,89%, *recall* = 12,25%, *f-measure* = 18,3%, dan *g-mean* = 43,15%. Sehingga nilai epsilon pada penelitian ini tidak terlalu berpengaruh terhadap tingkat akurasi.

4.7. Pengujian Pengaruh pada BOS

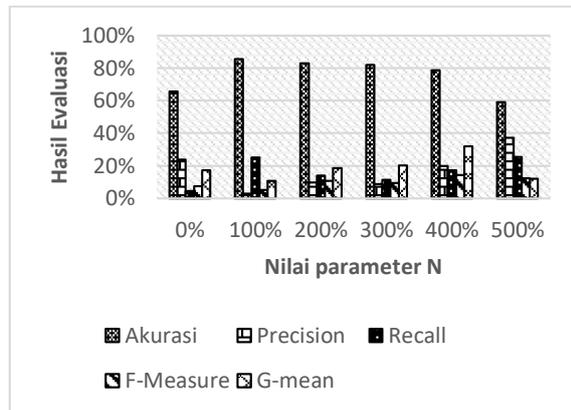
Hasil pengujian pengaruh nilai parameter N pada BOS ditunjukkan pada Tabel 8. Adapun parameter *Sequential training SVM* yang digunakan

dalam metode BOS adalah $\sigma=1$; $\lambda=0,5$; $\gamma=0,001$, $C=1$ dan iterasi maksimum=20. Adapun nilai parameter lain dalam BOS adalah $k=5$ dan $m=5$ mengikuti penelitian yang dilakukan Nguyen et. al (2011).

Tabel 8. Hasil Pengujian Pengaruh pada BOS

N	Akurasi	Precisi on	Recall	F-Measure	G-mean
0%	65,6%	23,8%	4,4%	7,35%	17,3%
100%	85,6%	2,97%	25%	5,08%	10,8%
200%	83%	9,74%	13,9%	10,7%	18,4%
300%	81,8%	8,76%	11,3%	9,37%	20,1%
400%	78,6%	19,7%	17,1%	14,4%	32%
500%	59%	37,3%	25,6%	12,5%	11,9%

Berdasarkan hasil pengujian dari Tabel 8 menunjukkan nilai nilai akurasi yang bermacam-macam. Nilai akurasi pada N=0%-100% mengalami peningkatan, namun setelah itu mengalami penurunan. Meski mengalami penurunan pada N=100%-300%, nilai akurasi yang dihasilkan relatif tinggi. Nilai *g-mean* juga mengalami peningkatan dari N=100%-400% dengan nilai *g-mean* tertinggi pada N=400%. Pada N=500%, nilai akurasi yang dihasilkan pada *precision* cukup tinggi dibandingkan yang lain, namun terjadi *trade off* pada nilai akurasi. Hal tersebut disebabkan karena data latih pada ulasan palsu sedikit lebih dominan dibandingkan dengan data latih pada ulasan asli, sehingga menyebabkan nilai *precision* juga naik. Grafik hasil dari pengujian pengaruh parameter N dapat dilihat pada Gambar 2.



Gambar 2. Grafik hasil pengujian parameter N

Berdasarkan hasil dari Gambar 2 menunjukkan ketika hasil *g-mean* tertinggi, nilai dari *precision* tidak terlalu rendah dan nilai akurasinya juga bukan yang tertinggi. Hal tersebut dikarenakan hasil evaluasi klasifikasi ulasan palsu dan asli hampir seimbang. Evaluasi menggunakan *g-mean* juga dapat menghindari *overfitting* dari mayoritas kelas (Bekkar, et al., 2013). Adapun contoh kesalahan klasifikasi ulasan palsu yaitu dari kalimat ulasan "I ordered two eggs on a roll with ham and cheese, and it was hard to beat for \$4. Coffee came with my sandwich also" yang dikenali sebagai ulasan asli

pada sistem. Hal tersebut disebabkan frekuensi pada setiap *term* pada kalimat tersebut mayoritas berada kelas ulasan asli, sehingga sistem gagal untuk mengenali.

5. KESIMPULAN

Berdasarkan hasil pengujian serta analisis metode SVM pada klasifikasi ulasan palsu mendapatkan parameter terbaiknya yaitu $\sigma=0,5$; $\lambda=2$; $\text{learning rate}=0,001$; konstanta $C=0,5$; jumlah iterasi maksimum=20 dan $\text{epsilon}=0,0001$ yang memiliki hasil evaluasi akurasi 51%; *precision* 42,9%; *recall* 12%; *f-measure* 18,3%; dan *g-mean* 43,1%. Hasil dari nilai akurasi, *precision* dan *g-mean* cukup baik, akan tetapi nilai *recall* dan *f-measure* cukup rendah. Hal tersebut disebabkan adanya data yang tidak seimbang.

Berdasarkan hasil pengujian serta analisis metode BOS dan SVM mendapatkan parameter terbaiknya yaitu 400% dengan hasil evaluasi akurasi 78,6%; *precision* 19,7%; *recall* 17%; *f-measure* 14,4%; dan *g-mean* 32%. Hasil ini menunjukkan bahwa terdapat efek dari proses *resampling* data dibandingkan tanpa proses *resampling data* yaitu kenaikan pada hasil evaluasi.

Dalam penelitian selanjutnya dapat dilakukan analisis efek dari setiap fitur yang digunakan pada klasifikasi ulasan palsu karena ekstraksi fitur merupakan salah satu proses yang paling penting. Pemilihan ekstraksi fitur yang efisien juga dapat meningkatkan hasil evaluasi (Hasan, et al., 2020). Lalu dapat menambahkan algoritme yang bisa mengoptimalkan pemilihan parameter yang akan digunakan seperti algoritme genetika. Tujuan penambahan algoritme agar waktu komputasi lebih cepat dan dapat meningkatkan hasil dari evaluasi. Dalam penelitian ini pencarian parameter masih memakan banyak waktu yaitu 10 jam dengan satu kali pengujian.

DAFTAR PUSTAKA

- BEKKAR, M., DJEMAA, H. K. & ALITOUCE, T. A., 2013. Evaluation Measures for Models Assesment over Imbalanced Data Sets. *Journal of Information Engineering and Applications*, 10(3), pp. 27-38.
- ELSHRIF, E. & GHERBI, A., 2017. Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques. *DATA ANALYTICS 2017 : The Sixth International Conference on Data Analytics Detecting*, pp. 65-72.
- HASANA, U., M., L. R., PRATAMA, A. & CHOLISSODIN, I., 2016. Perbandingan Metode SVM, Fuzzy-KNN, dan BDT-SVM untuk Klasifikasi Detak Jantung Hasil Elektrokardiografi. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 3(3), pp. 201-207.
- HASAN, T. ET AL., 2020. A Comparative Analysis of Feature Extraction Methods for Human Opinion Grouping Using Several Machine Learning Techniques. *IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering*, pp. 272-275.
- JINDAL, N. & LIU, B., 2008. Opinion spam and analysis. *Proceedings of the 2008 International Conference on Web Search and Data Mining*, pp. 219-230.
- MELLENG, A., LOUGHREY, A.-J. & P, D., 2019. Sentiment and Emotion Based Text Representation for Fake Reviews Detection. *Proceedings of Recent Advances in Natural Language Processing*, pp. 750-757.
- MUKERJEE, A. VENKATARAMAN, V. LIU, B. & GLANCE, N., 2011. What Yelp Fake Review Filter Might Be Doing?. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, pp. 409-418.
- NGUYEN, H. M., COOPER, E. W. & KAMEI, K., 2011. Borderline Over-sampling for Imbalanced Data Classification. *Int. J. Knowl. Eng. Soft Data Paradigms*, Volume 3, pp. 4-21.
- OTT, M., CARDIE, C. & HANCOCK, J. T., 2013. Negative Deceptive Opinion Spam. *NAACL HLT 2013 - 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Main Conference*, pp. 497-501.
- PARAPAT, I. M., FURQON, M. T. & S., 2018. Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 10(2), pp. 3163-3169.
- PRATAMA, A., WIHANDIKA, R. C. & RATNAWATI, D. E., 2018. Implementasi Algoritme Support Vector Machine (SVM) untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 4(2), pp. 1704-1708.
- PROSERPIO, D., HOLLENBECK, B. & HE, S., 2020. *How Fake Customer Reviews Do-and Don't Work*. [Online] Available at: <https://hbr.org/2020/11/how-fake-customer-reviews-do-and-dont-work#> [Diakses 12 04 2021].
- RAYANA, S. & AKOGLU, L., 2015. Collective Opinion Spam Detection: Bridging Review Networks and Metadata. *Proceeding of the 21st ACM SIGKDD international conference on Knowledge discovery and data mining*.
- RESENDES, S., 2020. *50+ Restaurant Industry Statistics Restaurateurs Should Know in 2020*. [Online]

Available at: <https://upserve.com/restaurant-insider/industry-statistics/>
[Diakses 12 04 2021].

SAVOY, J., 2005. *IR Multilingual Resources at UniNE*. [Online]
Available at:
<http://members.unine.ch/jacques.savoy/clef/>
[Diakses 13 09 2020].

TRIPADVISOR, 2018. *Influences on Diner Decision-Making*. [Online]
Available at:
<https://www.tripadvisor.com/ForRestaurants/r3227>
[Diakses 25 12 2020].

VIJAYAKUMAR, S. & WU, S., 1999. Sequential Support Vector Classifiers and Regression.