



# Analisis Sentimen Pengguna Aplikasi *Marketplace* Tokopedia Pada Situs Google Play Menggunakan Metode *Support Vector Machine* (SVM), *Naïve Bayes*, dan Logistic Regression

Meishita Inelza Putri<sup>a,\*</sup>, Iqbal Kharisudin<sup>b</sup>

<sup>a,b</sup> Universitas Negeri Semarang, Semarang 50299, Indonesia

\* Alamat Surel: [meishitainelzap@students.unnes.ac.id](mailto:meishitainelzap@students.unnes.ac.id)

## Abstrak

Pertumbuhan *e-commerce* di Indonesia sangat pesat, terbukti dengan data pengguna internet yang terus meningkat dari tahun ke tahun, serta adanya pandemi virus COVID-19 yang mengharuskan masyarakat untuk melakukan *social distancing*. Tokopedia merupakan salah satu penyedia *e-commerce* yang umum digunakan masyarakat Indonesia saat ini, khususnya generasi milenial. Data *review* hanya akan dilabel sebagai dua jenis sentimen, yaitu sentimen positif dan sentimen negatif kemudian metode klasifikasi sentimen positif dan sentimen negatif dilakukan dengan menggunakan beberapa metode, yakni *Support Vector Machine*, *Naïve Bayes*, dan *Logistic Regression*. Hasil pelabelan kelas sentimen pada data *review* Tokopedia diperoleh 3125 *review* dengan jumlah *review* positif sebanyak 2598 *review* dan *review* negatif sebanyak 527 *review*. Pada penelitian ini akan menggunakan dua jenis model pembagian data, yaitu dengan split data 80:20 dan *k-fold cross validation*. Ketidakseimbangan banyak data pada tiap kelas sentimen ditangan menggunakan teknik SMOTE. Penerapan SMOTE terbukti dapat meningkatkan kinerja pada data tidak seimbang, sehingga hal tersebut menunjukkan bahwa SMOTE lebih efektif dalam meningkatkan ketepatan akurasi klasifikasi daripada non-SMOTE. Metode klasifikasi yang tepat digunakan untuk klasifikasi *review* pengguna Tokopedia adalah *k-fold cross validation Support Vector Machine-SMOTE* karena memiliki nilai AUC paling tinggi.

## Kata kunci:

SMOTE, Analisis Sentimen, Supervised Learning, Text Mining.

© 2022 Dipublikasikan oleh Jurusan Matematika, Universitas Negeri Semarang

## 1. Pendahuluan

Meningkatnya prevalensi penyebaran virus COVID-19 (*Coronavirus Disease 2019*) di Indonesia mendorong pemerintah untuk mengeluarkan kebijakan pengendalian penyebaran virus tersebut dengan memberlakukan pembatasan berskala besar yang sering dikenal dengan istilah Pembatasan Sosial Berskala Besar (PSBB). Kebijakan tersebut pada akhirnya menyebabkan banyak bisnis, sekolah, dan aktivitas yang mengharuskan interaksi tatap muka untuk menerapkan *work form home* dan *online-learning* dalam aktivitas apa pun. Namun, dengan diberlakukannya kebijakan tersebut untuk mengantisipasi penularan virus COVID-19 di Indonesia telah membawa perubahan yang menguntungkan (positif) di bidang *e-commerce*, di mana *e-commerce* menjadi salah satu opsi yang tepat untuk mencegah penyebaran virus COVID-19 dan sangat berperan besar dalam menunjang peningkatan pengguna.

Saat ini, semakin banyak layanan *e-commerce* yang bermunculan dan dengan cepat mendapatkan popularitas karena pangsa pasar yang besar, banyak penyedia layanan *e-commerce* berlomba-lomba untuk menjadi yang terdepan. Persaingan antar penyedia layanan *e-commerce* secara tidak langsung semakin ketat saat ini. Pertumbuhan *e-commerce* pun sangat pesat di Indonesia dilihat dari data pengguna internet semakin tahun terus meningkat. Berdasarkan pengamatan yang dilakukan oleh Statista (Statista 2020), Tokopedia merupakan salah satu penyedia *e-commerce* yang umum digunakan masyarakat Indonesia saat ini, khususnya generasi milenial.

## To cite this article:

Putri, M. I., & Kharisudin, I. (2022). Penerapan *Synthetic Minority Oversampling Technique* (SMOTE) Terhadap Analisis Sentimen Data *Review* Pengguna Aplikasi *Marketplace* Tokopedia. *PRISMA, Prosiding Seminar Nasional Matematika 5*, 759-766

Pada situs Google Play menawarkan fitur komentar sebagai forum di mana pengguna dapat mengungkapkan pendapat mereka tentang aplikasi yang digunakan. *Review* yang diberikan juga terdiri dari berbagai jenis *review*, baik negatif maupun positif, yang nantinya akan mempengaruhi calon pengguna lainnya. Pengumpulan dan penyortiran data *review* tidaklah hal yang mudah karena banyaknya *review* yang tersedia pada situs *website* biasanya sangat banyak. Menurut Moraes et al. (2013), metode yang cocok untuk mengumpulkan data informasi tersebut adalah penggunaan metode *web scraping*. Namun, *review* tersebut masih berupa teks tanpa makna yang konkrit. Oleh karena itu, diperlukan suatu analisis yang mampu mengklasifikasikan *review* pengguna tersebut menjadi sentimen positif dan negatif berdasarkan isi *review* yang diberikan dengan menggunakan *text mining*.

Permasalahan yang sering muncul pada sebagian besar studi analisis sentimen, yaitu kebanyakan data *review* cenderung tidak seimbang (*imbalanced datasets*) dari segi jumlah tiap kelas individu, misalnya cenderung ke arah kelas positif atau sebaliknya. Secara umum, algoritma *machine learning* akan menghasilkan suatu model dengan tingkat sensitivitas yang rendah terhadap kelas minoritas ketika menerima set data yang tidak seimbang karena hal tersebut akan menyebabkan performa klasifikasi analisis sentimen yang dilakukan menjadi buruk. Terdapat beberapa pendekatan yang telah diusulkan untuk mengatasi permasalahan dataset tidak seimbang di antaranya, yaitu *one-class classification* (Juszczak and Duin 2003), *re-sampling* (Chawla et al. 2002), dan *cost-sensitive learning* (Zhou and Liu 2006). Pada penelitian ini, metode *Synthetic Minority Oversampling Technique* (SMOTE) digunakan untuk menangani kasus data yang tidak seimbang.

Berdasarkan penjabaran tersebut, pada penelitian ini akan menganalisis sentimen *review* tentang aplikasi *marketplace* yang semakin kompetitif dengan menggunakan metode klasifikasi *supervised learning* di antaranya *Support Vector Machine* (SVM), *Naïve Bayes*, dan *Logistic Regression* yang mana metode klasifikasi tersebut kemudian dibandingkan kinerja performa klasifikasinya untuk mendapatkan metode terbaik untuk mengklasifikasikan kelas sentimen dan penelitian ini akan menggunakan dua jenis model pembagian data, yaitu dengan split data 80:20 dan *k-fold cross validation*. Berdasarkan uraian di atas, tujuan dari penelitian ini untuk memperoleh hasil klasifikasi sentimen menggunakan metode klasifikasi *supervised learning* kemudian membandingkan kinerja performanya dengan metode klasifikasi yang disertai teknik SMOTE untuk menemukan metode klasifikasi terbaik dalam pengklasifikasian analisis sentimen. Selain itu, penelitian ini juga bertujuan untuk memperoleh kata-kata yang sering muncul dalam sentimen positif maupun negatif, sehingga nantinya informasi yang diperoleh dapat berguna bagi pihak *developer* aplikasi Tokopedia maupun pengguna yang ingin mendownload aplikasi maupun pihak lain yang membutuhkan untuk pengembangan teori berikutnya.

Penelitian mengenai analisis sentimen dengan menggunakan algoritma SVM dan KNN pernah dilakukan terhadap data *review* wisatawan Candi Borobudur pada situs TripAdvisor oleh (Saputri, Winahju, and Fithriasari 2019). Penelitian tersebut memberikan kesimpulan bahwa algoritma SVM dengan menggunakan metode SMOTE menghasilkan kinerja klasifikasi yang sangat baik.

Selain itu, Sari & Irhamah, 2019 melakukan penelitian tentang analisis sentimen nasabah terhadap layanan perbankan menggunakan metode Regresi Logistik Biner, *Naïve Bayes*, dan SVM. Penelitian ini menunjukkan bahwa metode SMOTE-SVM kernel RBF merupakan metode terbaik untuk mengklasifikasikan sentimen pada layanan BRI, sedangkan untuk Bank Mandiri adalah SMOTE-NBC karena memiliki nilai AUC tertinggi.

---

## 2. Metode

Pada penelitian ini, jenis data yang digunakan adalah data primer. Data dikumpulkan melalui proses *web scraping* pada situs web Google Play Store dan data dikumpulkan dari bulan April 2020 sampai dengan Februari 2021 sebanyak 3125 *review* pengguna Tokopedia. Metode *web scraping* adalah metode yang digunakan untuk mengumpulkan atau mengekstraksi informasi atau data semi-terstruktur dari website, biasanya dalam bentuk halaman web dalam bahasa *markup*, seperti HTML atau XHTML kemudian mengsktak informasi atau data tertentu dari halaman tersebut untuk dianalisis (Turland 2010). Selain itu, penggunaan metode *web scraping* dalam proses mengumpulkan maupun menyortir *review* dapat mempermudah untuk memperoleh informasi dari data yang berjumlah sangat banyak.

*Review* pengguna dalam penelitian ini akan diklasifikasikan menjadi sentimen positif dan negatif. Namun, *review* tersebut masih berupa teks tanpa makna yang konkrit sehingga diperlukan suatu analisis

yang mampu mengklasifikasikan *review* pengguna tersebut menjadi sentimen positif dan negatif berdasarkan isi *review* yang diberikan dengan menggunakan *text mining*. *Text mining* adalah proses menganalisis dan mengeksplorasi data teks tidak terstruktur menjadi terstruktur untuk mengidentifikasi pola berupa pengetahuan baru dan informasi yang bermakna atau berguna dengan berbantuan sebuah teknologi (Hung and Zhang 2012). Saat melakukan *text mining*, menyiapkan terlebih dahulu teks dokumen yang akan digunakan untuk proses analisis utama. Proses menyiapkan teks dokumen atau kumpulan data mentah disebut sebagai proses *text preprocessing*. Tujuan *text preprocessing* adalah mengubah data teks yang awalnya tidak terstruktur menjadi data yang terstruktur. Secara umum proses tahapan *text preprocessing* dapat dilakukan sebagai berikut (Feldman and Sanger 2006):

- (1) *Spelling normalization* adalah proses penyegaraman ejaan kata.
- (2) *Case folding* adalah proses standarisasi bentuk huruf agar tidak ada perbedaan makna.
- (3) *Tokenizing* adalah proses pemisahan kata per kata yang tidak saling mempengaruhi dari teks dokumen.
- (4) *Filtering* adalah proses penyaringan atau pemilihan kata dalam dokumen.

Setelah melewati tahapan *text preprocessing*, data tersebut perlu dimodelkan agar data yang masih berbentuk kata dapat diproses dan dihitung dengan cara mengubah data menjadi vektor, yang kemudian diberi nilai dan bobot untuk setiap kata. Penelitian ini menggunakan pembobotan TF-IDF, pembobotan TF-IDF adalah salah satu model yang digunakan untuk menghitung bobot kata pada suatu dokumen dengan menggunakan model *term frequency* untuk menghitung jumlah keseluruhan kata yang ada pada dokumen dan *Inverse Document Frequency* (IDF) untuk menghitung kata yang muncul di setiap dokumen (Nurjanah, Perdana, and Fauzi 2017).

Analisis sentimen adalah proses mengolah dan memahami data tekstual secara otomatis untuk mendapatkan informasi tentang sentimen yang terkandung dalam suatu kalimat opini, analisis dilakukan untuk melihat *review* atau tren opini terhadap suatu isu atau objek oleh seseorang, terlepas dari kecenderungannya apakah memiliki pandangan/pendapat negatif atau positif (Berliana, Shaufiah, and Sa'adah 2018). *VADER Lexicon* digunakan untuk melakukan proses pelabelan secara otomatis kelas sentimen pada data *review* berbahasa Inggris.

Data yang telah dilakukan pelabelan kelas sentimen selanjutnya akan dipartisi menjadi *training* dan *testing* dengan split data 80:20 dan *k-fold cross validation*. Dalam analisis ini dibagi banyaknya data latih adalah 80% dan data uji 20% untuk masing-masing data sedangkan untuk *k-fold cross validation* akan digunakan *10-fold cross validation*, yaitu membagi data keseluruhan menjadi 10 bagian dimana masing-masing bagian akan menjadi *training* dan *testing*. Namun, permasalahan yang sering muncul pada sebagian besar studi analisis sentimen, yaitu kebanyakan data *review* cenderung tidak seimbang (*imbalanced datasets*) dari segi jumlah tiap kelas individu, misalnya cenderung ke arah kelas positif atau sebaliknya. Oleh karena itu, metode *Synthetic Minority Oversampling Technique* (SMOTE) digunakan untuk menangani kasus data yang tidak seimbang. SMOTE adalah salah satu pendekatan teknik *oversampling* yang efektif dan baik digunakan untuk menangani *imbalance dataset* karena dapat menangani *overfitting* pada saat proses *oversampling* untuk kelas minoritas. Pendekatan SMOTE dilakukan dengan cara menambah banyak data kelas minoritas agar seimbang dengan kelas mayoritas, hal tersebut dilakukan dengan cara membangkitkan data buatan (*synthetic*) berdasarkan tetangga terdekat (*k-nearest neighbors*) sehingga diharapkan pendekatan teknik tersebut dapat berimbang pada hasil performa klasifikasi (Chawla et al. 2002).

Proses pembelajaran (*learning*) dilakukan dengan menggunakan beberapa metode klasifikasi *supervised learning*, seperti *Support Vector Machine* (SVM), *Naïve Bayes*, dan *Logistic Regression*. Saat melakukan pengukuran ketepatan klasifikasi, perlu diketahui jumlah pada setiap kelas prediksi dan kelas aktualnya karena hal tersebut dilakukan untuk melihat performa klasifikasi yang telah dilakukan. Tabel *confusion matrix* berikut ini dapat digunakan untuk menghitung ketepatan klasifikasi.

**Tabel 1.** *Confusion Matrix*

| Aktual  | Prediksi       |                |
|---------|----------------|----------------|
|         | Positif        | Negatif        |
| Positif | True Positive  | False Negative |
| Negatif | False Positive | True Negative  |

Namun, apabila terjadi kasus ketidakseimbangan data (*imbalance dataset*) maka dalam memilih model mana yang terbaik dapat dilakukan dengan menggunakan nilai AUC (*Area Under Curve*) sebagai dasar pengukurannya dan nilai akurasi dinilai kurang tepat di dalam *imbalance dataset* karena hanya

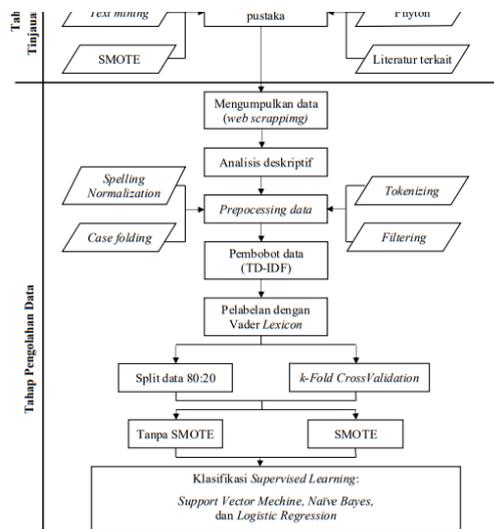
mempelajari data mayoritas saja sehingga hasil yang didapatkan tidak memiliki informasi penting karena mungkin saja terjadi bias atau *overfitting* (Zhou and Liu 2006). AUC adalah kriteria evaluasi yang menggunakan sensitivitas atau spesifisitas sebagai dasar pengukuran (He and Ma 2013). Nilai AUC pada umumnya berada pada interval 0,5 – 1,0. Berikut ini tabel keterangan untuk masing-masing interval nilai AUC dapat dilihat pada Tabel 2 di bawah ini (Bekkar, Djemaa, and Alitouche 2013).

**Tabel 2.** Kategori Nilai AUC

| Nilai AUC | Keterangan |
|-----------|------------|
| 0,9 – 1,0 | Excellent  |
| 0,8 – 0,9 | Good       |
| 0,7 – 0,8 | Fair       |
| 0,6 – 0,7 | Poor       |
| 0,5 – 0,6 | Failure    |

Penelitian ini akan menggunakan teknik SMOTE dan non-SMOTE yang nantinya akan digunakan untuk membandingkan performa klasifikasinya ketika diterapkan pada *imbalance dataset* sehingga dapat digunakan untuk menentukan metode klasifikasi terbaik.

Dalam penelitian ini, *word cloud* akan digunakan untuk memvisualisasikan hasil analisis klasifikasi. *Word cloud* adalah representasi dari data yang menunjukkan sekumpulan kata-kata penting dan sering muncul dalam data. Kata-kata yang sering muncul ditandai dengan banyaknya jumlah kata dan di cetak dengan ukuran besar dalam *wordcloud* (Castellà and Sutton 2014). Tahap analisis yang diterapkan dalam penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

### 3. Hasil dan Pembahasan

#### 3.1 Text preprocessing

Sebelum mengklasifikasikan dokumen teks, diperlukan *text preprocessing* untuk menyeragamkan ejaan kata maupun bentuk huruf, mengurangi kosakata, menghilangkan karakter selain huruf sehingga teks data lebih terstruktur dan informatif.

Pada tahap ini, metode *text preprocessing* digunakan untuk pembersihan data. Beberapa tahapan yang akan dilakukan antara lain *spell normalization*, *case folding*, *tokenization*, dan *filtering*. Hasil dari *text preprocessing* ditunjukkan pada Tabel 3 dibawah ini.

**Tabel 1.** Ilustrasi *Text Preprocessing*

| Review | Hasil Text Preprocessing |
|--------|--------------------------|
|--------|--------------------------|

|  |   |
|--|---|
| I'm a platinum user, been using the app for years. Nowadays, it is too much unnecessary update. Please refrain doing cosmetic interface updates, it makes the app getting slower and buggy. I am appreciating the functionalities update, but overall I don't see significant differences over the time while using this app but the cosmetics interface and numbers of useless promo banner (for me). | platinum user using year nowadays much unnecessary update please refrain cosmetic interface update make getting slower buggy appreciating functionality update overall see significant difference time using cosmetic interface number useless promo banner |
|--|---|

### 3.2 Analisis Sentimen

Penelitian ini menggunakan *libraryVADER lexicon* untuk mengotomatiskan proses pelabelan kategori sentimen pada data *review* bahasa Inggris. Data *review* hanya akan dilabeli sebagai dua jenis sentimen, yaitu sentimen positif dan sentimen negatif. Setelah proses pelabelan kategori menggunakan analisis sentimen, dapat dilihat jumlah perbandingan *review* berlabelkan positif dan negatif pada Tabel 4 di bawah ini.

**Tabel 2.** Perbandingan Jumlah Data *Review* pada Kelas Sentimen

| Sentimen | Tokopedia |
|----------|-----------|
| Positif  | 2598      |
| Negatif  | 527       |
| Jumlah   | 3125      |

Berdasarkan teks *review* yang akan dianalisis dan sudah melewati tahap *preprocessing* dapat dilakukan simulasi perhitungan skor sentimen. Rumus yang digunakan untuk menghitung skor sentimen adalah sebagai berikut (Bekkar et al. 2013) :

$$\text{Skor} = (\text{banyak kata positif}) - (\text{banyak kata negatif})$$

Jika, skor akhir simulasi perhitungan  $> 0$ , maka hasil klasifikasi *review* adalah sentimen positif. Adapun hasil simulasi perhitungan skor sentimen dapat dilihat sebagai berikut.

**Tabel 3.** Simulasi Perhitungan Skor Sentimen

| Teks <i>Review</i>   | Skor      | Label Sentimen |
|--|-----------|----------------|
| platinum user using year nowadays much unnecessary update <b>please</b> refrain cosmetic interface update make getting slower buggy <b>appreciating</b> functionality update overall see <b>significant</b> difference time using cosmetic interface <b>number</b> <b>useless</b> promo banner | 4 - 1 = 3 | Positif        |

Kata yang dicetak tebal berwarna hijau menunjukkan bahwa kata tersebut memiliki makna positif dalam *lexicon*, sedangkan kata yang berwarna merah memiliki makna negatif. Apabila skor akhir simulasi perhitungan, sehingga hasil klasifikasi *review* termasuk sentimen positif. Berdasarkan hasil pelabelan kategori kelas sentimen, dapat diketahui bahwa jumlah *review* positif lebih banyak daripada jumlah *review* negatif, dengan diperoleh rasio perbandingan 5:1. Ketidakseimbangan banyak data pada kelas positif dan negatif merupakan indikasi bahwa penelitian ini mengalami kasus ketidakseimbangan (*imbalance*). Oleh karena itu, dibuat data *synthetic (oversampling)* pada kelas minoritas (negatif) dengan menggunakan metode SMOTE pada metode klasifikasi yang digunakan.

### 3.3 Klasifikasi

Penelitian ini akan menggunakan teknik SMOTE dan non-SMOTE yang nantinya akan digunakan untuk membandingkan perform klasifikasinya ketika diterapkan pada *imbalance dataset*. Teknik SMOTE dilakukan dengan cara membangkitkan data buatan (*synthetic*) pada data kelas minoritas berdasarkan tetangga terdekat (*k-nearest neighbors*) agar seimbang dengan kelas mayoritas, teknik SMOTE hanya bekerja atau dilakukan pada data *training* yang digunakan.

Sebelum dilakukan prediksi klasifikasi pada penelitian ini data terlebih dahulu dipartisi menjadi dua jenis, yaitu dengan split data dan *k-fold cross validation*. Data yang telah melewati tahapan pre-

processing, selanjutnya akan dilanjutkan dengan proses pembelajaran (*learning*) menggunakan metode klasifikasi *supervised learning* diantaranya SVM, *Naïve Bayes*, dan *Logistic Regression*.

Selanjutnya dilakukan prediksi klasifikasi dengan algoritma *supervised learning*. Dari setiap metode klasifikasi yang digunakan didapat nilai AUC, *recall*, *precision*, dan *accuracy*. Pada penelitian ini ditemui terjadi kasus ketidakseimbangan data (*imbalance dataset*) maka dalam memilih model mana yang terbaik dapat dilakukan dengan menggunakan nilai AUC (*Area Under Curve*) sebagai dasar pengukurannya (Liu 2010) karena nilai akurasi dinilai kurang tepat di dalam *imbalance dataset* karena hanya mempelajari data mayoritas saja sehingga hasil yang didapatkan tidak memiliki informasi penting karena mungkin saja terjadi bias atau *overfitting*.

### 3.3.1. Split Data 80:20

Analisis ini akan mempartisi dataset dengan rasio 80% sebagai data latih (*training*) dan 20% sebagai data uji (*testing*). Berikut ini merupakan tabel perbandingan hasil split data jika tidak menggunakan teknik SMOTE dan dengan menggunakan teknik SMOTE.

**Tabel 4.** Perbandingan Hasil Algoritma tanpa SMOTE dan dengan SMOTE pada Tokopedia

|           | Metode              | AUC    | Recall | Precision | Accuracy |
|-----------|---------------------|--------|--------|-----------|----------|
| Non SMOTE | SVM                 | 0,9558 | 0,9846 | 0,9127    | 0,9088   |
|           | Naïve Bayes         | 0,9206 | 0,9962 | 0,8368    | 0,8352   |
|           | Logistic Regression | 0,9589 | 1,0000 | 0,8595    | 0,8640   |
| SMOTE     | SVM                 | 0,9572 | 0,9596 | 0,9433    | 0,9184   |
|           | Naïve Bayes         | 0,9424 | 0,8500 | 0,9757    | 0,8576   |
|           | Logistic Regression | 0,9613 | 0,9365 | 0,9568    | 0,9120   |

Berdasarkan Tabel 6, hasil yang diperoleh akan dibandingkan kinerja perform klasifikasinya dengan menggunakan nilai AUC (*Area Under Curve*). Hasil ini menunjukkan bahwa SMOTE dapat meningkatkan akurasi model pada data yang tidak seimbang (*imbalance*), terbukti dengan nilai AUC pada model dengan SMOTE lebih tinggi daripada nilai AUC yang dihasilkan oleh model non-SMOTE.

Secara keseluruhan, metode SVM merupakan metode yang tepat untuk mengklasifikasikan data review pengguna Tokopedia. Metode tersebut dipilih karena memiliki nilai AUC sebesar 0,9545 yang menunjukkan bahwa metode ini memiliki kinerja klasifikasi yang sangat baik dan penggunaan SMOTE dapat meningkatkan akurasi model pada data *imbalance* hasil review pengguna marketplace Tokopedia.

### 3.3.2. K-Fold Cross Validation

Partisi data *k-fold cross validation* akan digunakan pada pengujian kedua karena metode ini sering digunakan untuk mengurangi bias yang terkait dengan data random. Dalam penelitian ini, akan digunakan *10-fold cross validation* dengan membagi dataset menjadi 10 bagian (*fold*) yang berbeda dengan ukurannya sama dan melakukan pelatihan dan pengujian pada model sebanyak 10 kali.

Selanjutnya dilakukan prediksi klasifikasi dengan algoritma *Support Vector Machine*, *Naïve Bayes*, dan *Logistic Regression*. Dari setiap metode klasifikasi yang digunakan didapat nilai AUC, *recall*, *precision*, dan *accuracy*.

**Tabel 5.** Perbandingan Hasil Algoritma tanpa SMOTE dan dengan SMOTE Pada Tokopedia

|           | Metode              | AUC    | Recall | Precision | Accuracy |
|-----------|---------------------|--------|--------|-----------|----------|
| Non SMOTE | SVM                 | 0,9802 | 0,9895 | 0,9827    | 0,9744   |
|           | Naïve Bayes         | 0,9599 | 0,9965 | 0,9286    | 0,9263   |
|           | Logistic Regression | 0,9827 | 0,9964 | 0,9384    | 0,9391   |
| SMOTE     | SVM                 | 1,0000 | 0,9605 | 0,9781    | 0,9455   |
|           | Naïve Bayes         | 0,9649 | 0,9233 | 0,9888    | 0,9199   |
|           | Logistic Regression | 0,9982 | 0,9527 | 0,9850    | 0,9455   |

Berdasarkan Tabel 7, diperoleh bahwa metode yang tepat untuk mengklasifikasikan data *review* pengguna Tokopedia adalah metode *Support Vector Machine*. Metode ini dipilih karena memiliki nilai AUC sebesar 1,000 yang menunjukkan bahwa metode ini memiliki kinerja klasifikasi yang sangat baik, dan penggunaan SMOTE dapat meningkatkan akurasi model pada data *imbalance* dari *review* pengguna terhadap *marketplace* Tokopedia.

### 3.4 Visualisasi

Dalam penelitian ini, *word cloud* akan digunakan untuk memvisualisasikan hasil analisis klasifikasi. Tujuan dari visualisasi tersebut adalah untuk mengekstrak informasi berupa topik yang sering dibahas oleh pengguna *marketplace* Tokopedia sehingga dapat diambil informasi yang dianggap penting dari sekian banyak teks *review* yang ada.



**Gambar 2.** (a) *Word Cloud* Sentimen Positif Tokopedia; (b) *Word Cloud* Sentimen Negatif Tokopedia

Berdasarkan Gambar 2(a) informasi mengenai kata-kata yang paling sering muncul, yaitu *good*, *easy*, *use*, *shopping*, *best*, *thank*, *great*, dan *helpful*. Artinya, kata-kata tersebut menjadi kata yang sering dibicarakan oleh pengguna karena dicetak lebih besar pada *word cloud* daripada kata-kata lainnya. Hal tersebut didapatkan informasi bahwa penilaian positif dari pengguna Tokopedia berkaitan dengan aplikasi mudah digunakan, bagus, membantu pengguna dalam berbelanja, aplikasi selalu *update*, transaksi pada aplikasi ini mudah, dan banyak menyediakan produk. Sedangkan gambar *word cloud* sentimen negatif dapat diketahui bahwa kata *seller*, *update*, *bad*, *service*, *error*, *problem*, *time*, dan *notification* yang terdapat pada sentimen negatif merupakan kata yang sering dibicarakan karena pada *word cloud*, kata-kata tersebut tercetak paling besar dibanding kata-kata lain sehingga *review* negatif pengguna sering mengeluhkan tentang hal-hal tersebut.

## 4. Simpulan

Kesimpulan dari penelitian diperoleh bahwa hasil tersebut menunjukkan bahwa partisi data menggunakan *10-fold cross validation* menghasilkan performa klasifikasi lebih baik dibandingkan split data 80:20 karena konsep *k-fold cross validation* membagi dataset menjadi beberapa bagian (*k-subset*) dengan ukuran yang sama dan melakukan pelatihan dan pengujian pada model sebanyak *k* yang ditentukan.

Penggunaan SMOTE terbukti dapat meningkatkan akurasi model pada data yang tidak seimbang (*imbalance*) dilihat dari nilai AUC pada model dengan SMOTE lebih tinggi daripada nilai AUC yang dihasilkan oleh model non-SMOTE sehingga penerapan SMOTE lebih efektif dalam meningkatkan ketepatan akurasi klasifikasi. Metode klasifikasi yang tepat digunakan untuk klasifikasi *review* pengguna Tokopedia adalah *k-fold cross validation* dengan SMOTE, yaitu metode klasifikasi *Support Vector Machine* dengan nilai AUC sebesar 1,000, sehingga metode tersebut tergolong memiliki kinerja klasifikasi yang sangat baik. Secara umum *review* pengguna lebih banyak bersentimen positif yang berkaitan dengan kemudahan penggunaan aplikasi tersebut, sedangkan *review* negatif berkaitan dengan sistem aplikasi *error* dan sering terjadi *lagging/bug*.

## Daftar Pustaka

- Berliana, G., Shaufiah, & Sa'adah, S. (2018). Klasifikasi Posting Tweet mengenai Kebijakan Pemerintah Menggunakan Naive Bayesian Classification. *E-Proceeding of Engineering*, 5(1), 1562–1569.
- Castellà, Q., & Sutton, C. (2014). Word storms: Multiples of Word Clouds for Visual Comparison of Documents. *WWW 2014 - Proceedings of the 23rd International Conference on World Wide Web*,

- 665–675. <https://doi.org/10.1145/2566486.2567977>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Feldman, R., & Sanger, J. (2006). The Text Mining Handbook. In *The Text Mining Handbook*. <https://doi.org/10.1017/cbo9780511546914>
- Hermanto, Kuntoro, A. Y., Asra, T., Pratama, E. B., Effendi, L., & Ocanitra, R. (2020). Gojek and Grab User Sentiment Analysis on Google Play Using Naive Bayes Algorithm and Support Vector Machine Based Smote Technique. *Journal of Physics: Conference Series*, 1641(1). <https://doi.org/10.1088/1742-6596/1641/1/012102>
- Jadhav, S. D., & Channe, H. P. (2016). Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques. *International Journal of Science and Research (IJSR)*, 5(1), 1842–1845. <https://doi.org/10.21275/v5i1.nov153131>
- Liu, L., & Ozsu, M. T. (2009). Encyclopedia of Database Systems. In *Encyclopedia of Database Systems*. Springer. <https://doi.org/10.1007/978-0-387-39940-9>
- Mohammad, S. M., Kiritchenko, S., & Zhu, X. (2013). NRC-Canada: Building the State-of-the-Art in Sentiment Analysis of Tweets. *\*SEM 2013 - 2nd Joint Conference on Lexical and Computational Semantics*, 2, 321–327.
- Moraes, R., Valiati, J. F., & Gavião Neto, W. P. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications*, 40(2), 621–633. <https://doi.org/10.1016/j.eswa.2012.07.059>
- Ramli, Yuniarti, D., & Goejantoro, R. (2013). Perbandingan metode klasifikasi Regresi Logistik dengan Jaringan Saraf Tiruan (studi kasus: pemilihan jurusan bahasa dan IPS pada SMAN 2 Samarinda tahun ajaran 2011/2012). *Jurnal EKSPONENSIAL Universitas Mulawarman*, 4(1), 17.
- Saputri, R. P., Winahju, W. S., & Fithriasari, K. (2019). Klasifikasi Sentimen Wisatawan Candi Borobudur pada Situs TripAdvisor Menggunakan Support Vector Machine dan K-Nearest Neighbor. *Jurnal Sains Dan Seni Its*, 8(2), 1. [http://ejournal.its.ac.id/index.php/sains\\_seni/article/view/44391](http://ejournal.its.ac.id/index.php/sains_seni/article/view/44391)
- Sari, E. D. N., & Irhamah. (2019). Analisis Sentimen Nasabah Pada Layanan Perbankan Menggunakan Metode Regresi Logistik Biner, Naïve Bayes Classifier (NBC), dan Support Vector Machine (SVM). *Jurnal Sains Dan Seni Its*, 8(2), 177.
- Sumpeno, S. (2009). Klasifikasi Emosi Untuk Teks Bahasa Indonesia. *Seminar*, c.
- Suyanto. (2019). *Data Mining untuk klasifikasi dan klusterisasi data*. Informatika.
- Tan, P.-P., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. Addison-Wesley Longman Publishing Co, 2005.
- Turland, M. (2010). *php|architect's Guide to Web Scraping with PHP*.
- Wardani, N. S., Prahutama, A., & Kartikasari, P. (2020). Analisis Sentimen Pemindahan Ibu Kota Negara Dengan Klasifikasi Naïve Bayes Untuk Model Bernoulli dan Multinomial. 9, 237–246.
- Wyatt, J. C., & Taylor, P. (2008). Decision Support Systems and Clinical Innovation. *Getting Research Findings into Practice: Second Edition*, 123–137. <https://doi.org/10.1002/9780470755891.ch11>