

## Analisis Sentimen Aplikasi Duolingo Menggunakan Metode Naive Bayes dan Synthetic Minority Over Sampling Technique

Saifurrachman Chohan<sup>1</sup>, Arifin Nugroho<sup>2</sup>, Achmad Maezar Bayu Aji<sup>3</sup>, Windu Gata<sup>4</sup>

<sup>1</sup>STMIK Nusa Mandiri  
e-mail: 14002305@nusamandiri.ac.id

<sup>2</sup>STMIK Nusa Mandiri  
e-mail: 14002306@nusamandiri.ac.id

<sup>3</sup>STMIK Nusa Mandiri  
e-mail: achmad.azb@nusamandiri.ac.id

<sup>4</sup>STMIK Nusa Mandiri  
e-mail: Windu@nusamandiri.ac.id

**Abstract** - Learning foreign languages is one of the important concerns in order to compete with the international world. Time and cost limitations make the application of learning English on mobile devices more popular as a medium for learning foreign languages. One application that is often used to learn foreign languages on mobile devices is duolingo. This study aims to find and compare classifications in sentiment analysis from user reviews of Duolingo applications obtained from Google Playstore. Text mining is used to divide the reviews given by users into two groups, namely positive reviews and negative reviews. Rapid miner is used to find and compare two different classification methods between datasets that use Naive Bayes Classification only and data sets that use the naive bayes algorithm with the syntetic minority over-sampling technique (SMOTE). The results of the two methods in this study found that the highest results were obtained using the Naive Bayes algorithm with a syntetic minority over-sampling technique (SMOTE) where has the rate of accuracy 91.95% and 0,740 for AUC.

**Keywords:** text mining, naive bayes, smote method, playstore, duolingo.

**Intisari** - Belajar bahasa asing menjadi salah satu perhatian penting agar dapat bersaing dengan dunia internasional. Keterbatasan waktu dan biaya membuat aplikasi belajar bahasa inggris pada perangkat mobile lebih disukai sebagai media pembelajaran bahasa asing. Salah satu aplikasi yang sering di gunakan untuk belajar bahasa asing pada perangkat mobile adalah duolingo. Penelitian ini bertujuan untuk menemukan dan membandingkan klasifikasi dalam sentimen analisis dari ulasan pengguna aplikasi duolingo yang didapat dari google playstore. Text mining digunakan untuk membagi ulasan yang diberikan pengguna menjadi dua kelompok yaitu ulasan positif dan ulasan negatif. Rapid miner digunakan untuk mencari dan membandingkan dua metode klasifikasi yang berbeda antara dataset yang menggunakan Naive Bayes Classification saja dan data set yang menggunakan algoritma naive bayes dengan syntetic minority over-sampling technique (SMOTE). Hasil dari dua metode dalam penelitian ini menemukan bahwa hasil tertinggi didapatkan menggunakan algoritma naive bayes dengan syntetic minority over-sampling technique (SMOTE) dimana memiliki tingkat akuransi 91,95% dan AUC sebesar 0.740.

**Kata Kunci:** text mining, naive bayes, metode smote, playstore, duolingo.

### PENDAHULUAN

Teknologi Berkembang Dengan Cepat Tidak Terkecuali Dalam Dunia Pendidikan, Peran Teknologi Dalam Pendidikan Salah Satunya Yaitu Mobile Learning, Penggunaan Mobile Learning Menjadi Salah Satu Media Pendukung Pembelajaran. Dengan Mobile Learning Maka Proses Belajar Menjadi Lebih Mudah, Seperti Waktu

Yang Tidak Terikat, Tidak Terbatas Ruang Dan Waktu, Dan Dapat Diakses Dari Mana Saja (Irfani, 2020)

Mobile learning (m-learning) adalah pembelajaran yang memanfaatkan perangkat komunikasi bergerak seperti telepon seluler. Pengguna smartphone dengan platform android mendominasi pasar dunia dengan jumlah pengguna sebanyak 74,3% pada bulan januari 2020, selanjutnya IOS berada

dibawahnya dengan jumlah pengguna sebanyak 24,76% (<https://gs.statcounter.com/os-market-share/mobile>, diakses pada bulan Mei 2020)

Duolingo adalah aplikasi gratis yang dibuat oleh Luis Von Ahn dan Severin Hacker pada bulan November 2011. Slogannya adalah "Pendidikan bahasa gratis untuk dunia." Menurut situs webnya, ia memiliki lebih dari 30 juta pengguna terdaftar. Menawarkan beberapa bahasa untuk bahasa Inggris penutur serta orang lain untuk penutur non-Inggris (Munday, 2015). Belajar menggunakan aplikasi Duolingo selama 34 jam setara dengan belajar di universitas selama satu semester atau 11 minggu (Vesselinov & Grego, 2012).

Diantara penelitian sebelumnya yang sudah ada adalah dengan judul "Optimization Sentiments of Analysis from Tweets in myXLCare using Naive Bayes Algorithm and Synthetic Minority Over Sampling Technique Method", penelitian ini membahas tentang analisis komentar pelanggan kartu seluler XL pada sosial media Twitter dan hasil yang di dapatkan yaitu dengan tingkat akurasi sebesar 86.33%, precision sebesar 82,85% dan recall ratio sebesar 92,38 (Saputra et al., 2020).

Dan ada juga penelitian yang menggunakan Naive Bayes dengan judul "Sentimen Analisis Operasi Tangkap Tangan KPK Menurut Masyarakat Menggunakan Algoritma Support Vector Machine, Naive Bayes Berbasis Particle Swarm Optimization", penelitian ini membahas tentang pesan masyarakat yang disampaikan melalui kolom komentar pada media sosial Twitter terhadap operasi tangkap tangan Komisi Pemberantasan Korupsi. Pada penelitian ini hasil yang didapatkan sebagai berikut, dari data 78 tweet positif dan 78 data tweet negatif, menghasilkan akurasi sebesar 80.75% dan AUC 0.867 untuk algoritma SVM (Support Vector Machine) dan untuk algoritma Naive Bayes menghasilkan 76.92 % dan AUC 0.729. memiliki selisih akurasi sebesar 3.3 % dan setelah dioptimalisasi dengan operator Weight Partial Swarm Optimization untuk SVM menghasilkan akurasi 83.79% dan AUC 0.910, sedangkan Naive Bayes menghasilkan akurasi sebesar 80.13% dan 0.771 (Hernawati & Windu, 2019).

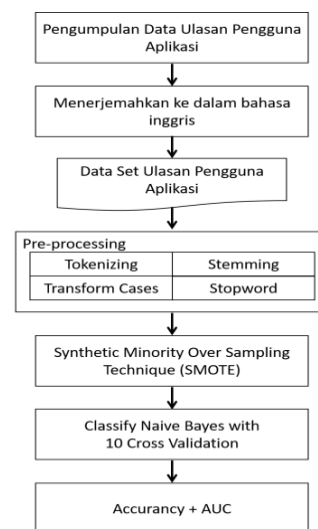
Pada penelitian ini studi kasus yang di analisa adalah sentimen pengguna aplikasi m-learning duolingo dari ulasan pada google play store dengan pemodelannya akan menggunakan Naive Bayes dan Synthetic Minority Over Sampling Technique. Naive Bayes merupakan pendekatan statistik yang fundamental dalam mengenal pola (Pattern Recognition). Pendekatan ini di dasarkan pada kualifikasi trade-off antara berbagai keputusan

klasifikasi dengan menggunakan probabilitas dan ongkos yang di timbulkan dalam keputusan-keputusan tersebut (Santosa & Umam, 2018).

Pada penelitian ini akan dibahas tahapan – tahapan yang akan dilalui untuk melakukan pemrosesan analisis sentimen terhadap ulasan – ulasan yang ada pada aplikasi duolingo di googleplay. Tahapan tersebut akan diawali dengan tahapan preprocessing sampai tahap analisis menggunakan metode Naive Bayes Classifier dengan Synthetic Minority Over Sampling Technique (SMOTE) dan dalam penelitian ini juga akan menampilkan hasil perhitungan jika tanpa menggunakan SMOTE dan akan dibandingkan dengan hasil yang menggunakan SMOTE.

## METODOLOGI PENELITIAN

Di dalam penelitian ini kami mencoba menggunakan metode Naive Bayes dan Synthetic Minority Over Sampling Technique (SMOTE) untuk mendapatkan akurasi terbaik dalam menganalisa ulasan – ulasan pengguna aplikasi m-learning duolingo pada googleplay. Adapun metode penelitian yang dilakukan dapat dilihat pada gambar berikut:



Gambar 1. Metode Penelitian

### 1. Pengumpulan Data

Metode pengumpulan data dapat diartikan sebagai serangkaian cara yang digunakan peneliti untuk mengumpulkan data (Nanda & Sari, 2020). Pengumpulan data pada penelitian ini didapatkan dari ulasan - ulasan pengguna aplikasi duolingo pada web googleplay yang di dapatkan dengan menggunakan teknik web scraping, suatu teknik yang digunakan untuk mengekstrak sejumlah besar data dari situs web dimana data yang sudah diekstraksi disimpan ke sebuah file lokal di komputer atau ke database dalam format tabel (spreadsheet) (Perdana & Irwansyah, 2019), kemudian diproses menggunakan aplikasi rapidminer untuk kemudian dilanjutkan ke tahap berikutnya yaitu preprocessing.

## 2. Menterjemahkan ke dalam bahasa Inggris.

Pada tahap ini, kami mencoba menyamakan semua ulasan menjadi berbahasa Inggris dengan melakukan proses penerjemahan ulasan-ulasan yang masih berbahasa Indonesia ke dalam bahasa Inggris.

## 3. Preprocessing Data

Preprocessing Data adalah tahapan untuk melakukan persiapan data yang akan dilakukan pemodelan[6]. Pada tahapan ini mencakup beberapa kegiatan dalam membentuk data dan juga melakukan penyisiran untuk membersihkan data hingga nantinya data dapat diproses ke tahap berikutnya. Berikut ini adalah tahapan – tahapan yang dilakukan dalam preprocessing data:

### a. Tokenizing

Langkah pertama dari preprocessing text pada penelitian ini adalah Tokenization yaitu suatu metode untuk memecah teks menjadi komponen yang lebih kecil (kata, kalimat, bigrams) (Rizaldi & Putranto, 2017).

### b. Stemming

Setelah tahap tokenisasi, tahapan berikutnya adalah stemming, Teknik stemming dilakukan untuk alasan mereduksi term menjadi bentuk dasarnya. Term yang ada pada dokumen dan query memiliki banyak varian morfologik maka akan sulit term-term tersebut dianggap ekuivalen. Namun dalam beberapa kasus tertentu varian morfologik term-term memiliki interpretasi semantik yang sama dan dapat dikategorikan ekuivalen. Terdapat perbedaan algoritma stemming untuk bahasa yang satu dengan yang lainnya. Proses stemming pada teks berbahasa Indonesia lebih rumit/kompleks karena terdapat variasi imbuhan yang harus dibuang untuk mendapatkan root word dari sebuah kata (Anggraini & Suroyo, 2019).

### c. Transform Cases

Dalam proses ini kata – kata yang tidak relevan akan diubah, seperti kata-kata yang mengandung huruf besar dikonversi menjadi huruf kecil mereka dapat berhubungan dengan sentimen (Wardhani et al., 2018)

### d. Stopwords

Stopwords removal adalah sebuah proses untuk menghilangkan kata yang 'tidak relevan' pada hasil parsing sebuah dokumen teks dengan cara membandingkannya dengan Stoplist yang ada. Kata-kata yang terlalu sering muncul dalam belum tentu berguna dalam proses retrieval. Kata-kata yang tidak berguna nantinya akan dibuang dan tidak dijadikan index term (Anna & Hendini, 2018)

## 4. Synthetic Minority Over Sampling Technique (SMOTE)

Setelah proses Preprocessing selesai, pada tahap selanjutnya penulis menggunakan optimasi dengan menggunakan Synthetic Minority Over Sampling Technique (SMOTE) yang merupakan suatu

pendekatan untuk pengklasifikasian dari dataset yang tidak seimbang dimana kategori klasifikasi tidak dipresentasikan secara merata. Kombinasi dengan metode ini telah dibuktikan dapat menghasilkan kinerja pengklasifikasian yang lebih baik (Chawla et al., 2002).

## 5. Classify Naive Bayes With 10 Fold Cross Validation

Pada tahapan pemodelan penelitian ini kami menggunakan teorema Naive Bayes dengan proses Cross Validation sebanyak 10 kali, klasifikasi Naive Bayes sendiri adalah metode klasifikasi yang berdasarkan probabilitas dan Teorema Bayes dengan asumsi bahwa setiap variable X bersifat bebas atau berdiri sendiri dan tidak ada kaitannya dengan variable lainnya[20]. Dapat di lihat dalam persamaan 1 berikut :

$$P(R|S) = \frac{P(R)P(S|R)}{P(S)} \dots\dots\dots (1)$$

Keterangan :

R : Data yang belum diketahui kelasnya

S : Hipotesis pada data R yang merupakan class khusus

P(R|S) : Nilai probabilitas pada hipotesis R yang berdasarkan kondisi S

P(R) : Nilai probabilitas pada hipotesis R

P(S|R) : Nilai probabilitas S yang berdasarkan dengan kondisi hipotesis R

P(S) : Nilai probabilitas S

## 6. Evaluasi Accuracy dan AUC

Tahapan terakhir yakni perhitungan Accuracy. Dimana Accuracy adalah perhitungan terhadap proporsi dari jumlah total prediksi yang benar dan dirumuskan dalam persamaan 2 berikut : [21]

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \dots\dots\dots (2)$$

Keterangan :

TP: True Positive

TN: True Negative

FP : False Positive

FN: False Negative

Dan area under the curve (AUC) adalah suatu daerah di bawah receiver operating characteristic (ROC). Receiver operating characteristic (ROC) merupakan kurva yang dihasilkan dari tarik ulur antara sensitivitas dan spesifisitas pada berbagai titik potong. Nilai AUC secara teoritis berada di antara 0 dan 1. Nilai AUC memberikan gambaran tentang keseluruhan pengukuran atas kesesuaian dari model yang digunakan. Semakin besar area under curve maka semakin baik variabel yang diteliti dalam memprediksi kejadian (Maskoen & Purnama, 2018).

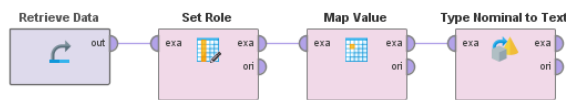
## HASIL DAN PEMBAHASAN

Pada tahapan ini peneliti akan menjelaskan hasil mengenai proses yang dilakukan pada penelitian yang sudah di paparkan pada metode penelitian sebelumnya. Berikut hasil dan pembahasan dari metode penelitian pada gambar

1. Pengumpulan Data

Pengumpulan data yang telah dilakukan oleh peneliti bertujuan membentuk dataset untuk digunakan dalam penelitian ini dan data diambil dari ulasan-ulasan web google play pada halaman aplikasi duolingo.

Data-data tersebut disimpan dalam sebuah file excel agar mudah dalam pemrosesan data di rapidminer. Data yang sudah disimpan dalam rapid miner kemudian dapat dipanggil kembali menggunakan fungsi retrieve untuk kemudian dilanjutkan ke proses berikutnya seperti pada gambar berikut:

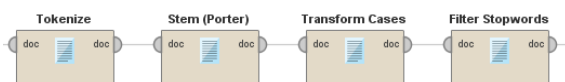


Gambar 2. Proses retrieve dataset pada Rapidminer

Setelah tahap retrieve data excel yang sudah di simpan pada rapidminer, maka selanjutnya dilakukan penentuan attribute name dan target role untuk menentukan label nya pada proses Set Role, kemudian masuk ke dalam proses MAP Value untuk merubah label yang bernilai true menjadi Positive dan False menjadi Negative dan berikutnya masuk pada tahapan selanjutnya yaitu Type Nominal to Text, pada tahapan ini bertujuan merubah nilai yang bersifat nominal menjadi sebuah teks.

2. Proses Preprocessing Data

Setelah tahapan explorasi data atau Load data telah selesai dilakukan, maka setelah itu kita bisa masuk ke tahap selanjutnya yaitu melakukan Preprocessing Data. Desain tahapan preprocessing data ini dapat dilihat pada gambar berikut:



Gambar 3. Tahapan Preprocessing Data

Pada gambar diatas terdapat beberapa proses dalam Preprocessing, antara lain adalah Tokenize, Stem (Porter), Transform Cases, Stopwords.

a. Tokenize

Pada proses ini ulasan akan di pecah menjadi per kata, menghilangkan tanda baca dan angka, berikut contoh dari proses tokenize:

Tabel 1. Proses Tokenize

Ulasan	Tokenize
I Really like this application, It really started the lessons	I Really like this application It really started the lessons

from scratch, Helped me learn a lot...	from scratch Helped me learn a lot
--	------------------------------------

b. Stem (Porter)

Setelah itu dilanjutkan dengan proses stemming (porter), pada proses ini kata yang masih memiliki imbuhan akan dikembalikan ke kata dasar nya dalam bahasa inggris. Proses stemming setelah tokenize dapat dilihat pada tabel berikut:

Tabel 2. Proses Stemming

Tokenize	Stemming
I Really like this application It really started the lessons from scratch Helped me learn a lot	I Realli like thi applic realli start the lesson from scratch Help me learn a lot

c. Transform Cases.

Setelah itu dilakukan transform cases untuk merubah huruf besar menjadi huruf kecil, tabel perubahan dari ulasan yang sudah di stemming untuk diproses dengan transform cases dapat dilihat pada tabel berikut:

Tabel 3. Transform Cases

Stemming	Transform Cases
I Realli like thi applic It realli start the lesson from scratch Help me learn a lot	i realli like thi applic it realli start the lesson from scratch help me learn a lot

d. Filter Stopwords

Setelah transform cases, penulis melakukan tahap terakhir pada preprocessing yaitu menghilangkan stopwords, perubahan setelah menghilangkan stopwords dapat dilihat pada tabel berikut:

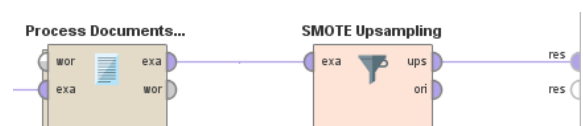
Tabel 4. Filter Stopwords

Transform Cases	Filter Stopwords
i realli like thi applic it realli start the lesson from scratch help me learn a lot	i realli thi applic realli start lesson scratch help learn lot

3. Dengan SMOTE dan tanpa SMOTE

Setelah tahapan preprocessing, maka tahapan di lanjutkan dengan proses pengoptimasian yaitu dengan cara mengkombinasikan metode SMOTE sebagai penyeimbang kelas dengan mereplikasi kelas minoritas sebanyak kelas mayoritas.

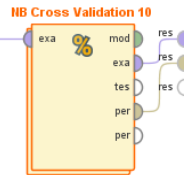
Seperti yang sudah peneliti sampaikan di akhir pendahuluan bahwa peneliti akan melakukan perbandingan pemodelan menggunakan optimisasi SMOTE dengan yang tidak menggunakan optimisasi SMOTE. Desain dari pemodelan yang diterapkan penulis pada rapid miner dapat dilihat pada gambar berikut:



Gambar 4. Tampilan desain dengan SMOTE pada RapidMiner

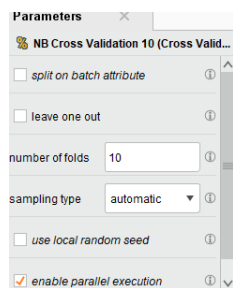
4. Proses Validasi

Pada tahap ini validasi yang digunakan adalah operator k-10 Fold cross validation, tampilan desain pada rapidminer dapat dilihat pada gambar berikut:



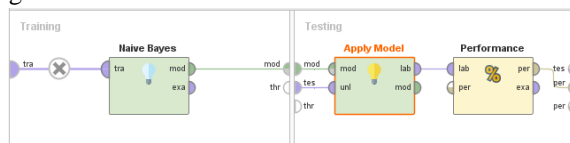
Gambar 5. Tampilan desain proses validasi pada rapidminer

Konfigurasi parameter validasi ini dapat dilihat pada gambar berikut:



Gambar 6. Tampilan Parameters Cross Validation

Di dalam operator validasi tersebut terdapat beberapa proses yang digunakan untuk menghasilkan nilai dari evaluasi. Proses – proses yang ada di dalam operator validasi tersebut pada rapid miner dapat dilihat pada gambar berikut:



Gambar 7. Proses - proses di dalam operator Validation

5. Evaluasi

Seperti yang sudah penulis sampaikan sebelumnya di dalam pendahuluan, bahwa hasil dari penelitian ini adalah penerapan algoritma Naive Bayes dalam analisa sentimen berdasarkan evaluasi nilai Accuracy dan AUC (Area Under Curve) dengan membandingkan hasil dari yang menggunakan optimasi SMOTE dan yang tidak menggunakan optimasi SMOTE. Hasil akurasi tanpa menggunakan SMOTE dapat dilihat pada tabel berikut:

Tabel 5. Hasil Accuracy tanpa menggunakan SMOTE

	NB
Accuracy	77.14%
AUC	0.526

Sedangkan untuk hasil akurasi dengan menggunakan SMOTE dapat dilihat pada tabel berikut:

Tabel 6. Hasil Accuracy dengan penggunaan SMOTE

	NB + SMOTE
Accuracy	91.95%
AUC	0.740

Grafik hasil AUC tanpa menggunakan SMOTE dapat dilihat pada gambar berikut:



Gambar 8. Grafik AUC tanpa menggunakan SMOTE

Sedangkan grafik hasil AUC dengan menggunakan SMOTE dapat dilihat pada gambar berikut:



Gambar 9. Grafik AUC dengan menggunakan SMOTE

Pada tabel berikut dapat dilihat hasil dari pemodelan dataset NB tanpa SMOTE dan dengan SMOTE:

Tabel 7. Hasil Pemodelan Dataset

	NB	NB + SMOTE
Accuracy	77.14%	91.95%
AUC	0.526	0.740

Berdasarkan tabel diatas dapat diketahui terdapat peningkatan yang cukup signifikan bila menggunakan SMOTE yaitu peningkatan akurasi sebesar 14.81% dan peningkatan AUC sebesar 0.214.

**KESIMPULAN**

Setelah melakukan penelitian untuk data set ulasan pengguna aplikasi m-learning duolingo pada googleplay dengan menggunakan pemodelan Naive Bayes Classifier tanpa Synthetic Minority Over Sampling Technique dan dibandingkan dengan model Naive Bayes Classifier dengan Synthetic Minority Over Sampling Technique, dapat disimpulkan bahwa untuk penggunaan Synthetic Minority Over Sampling Technique dapat meningkatkan nilai dari Accuracy dan AUC. Dan peningkatan yang dihasilkan cukup signifikan dari 77.14% untuk Accuracy dan 0.526 untuk AUC pada pemodelan Naive bayes tanpa SMOTE, menjadi 91.95% untuk Accuracy dan 0.740 untuk AUC pada pemodelan Naive bayes dengan SMOTE. Penggunaan Synthetic Minority Over Sampling Technique memiliki peran penting dalam penelitian

ini terhadap hasil dari kedua pemodelan tersebut, dan memberikan dampak kenaikan sebesar 14.81% untuk Accuracy dan 0.214 AUC terhadap hasil akhir.

## REFERENSI

- Anggraini, N., & Suroyo, H. (2019). Comparison of Sentiment Analysis against Digital Payment “T-cash and Go-pay” in Social Media Using Orange Data Mining. *Journal of Information Systems and Informatics*. <https://doi.org/10.33557/journalisi.v1i2.21>
- Anna, A. (AMIK B. P., & Hendini, A. (AMIK B. P. (2018). Implementasi Vector Space Model Pada Sistem Pencarian Mesin Karaoke. *Evolusi : Jurnal Sains Dan Manajemen*, 6(1), 1–6. <https://doi.org/10.31294/evolusi.v6i1.3535>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*. <https://doi.org/10.1613/jair.953>
- Hernawati, & Windu. (2019). *Sentimen Analisis Operasi Tangkap Tangan KPK Menurut Masyarakat Menggunakan Algoritma Support Vector Machine, Naive Bayes Berbasis Particle Swarm Optimization*. 14. <https://doi.org/10.30998/faktorexacta.v12i3.4992>
- Irfani, F. F. (2020). ANALISIS SENTIMEN REVIEW APLIKASI RUANGGURU MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE. *JBMI (Jurnal Bisnis, Manajemen, Dan Informatika)*. <https://doi.org/10.26487/jbmi.v16i3.8607>
- Maskoen, T. T., & Purnama, D. (2018). Area Under the Curve dan Akurasi Cystatin C untuk Diagnosis Acute Kidney Injury pada Pasien Politrauma. *Majalah Kedokteran Bandung*. <https://doi.org/10.15395/mkb.v50n4.1342>
- Munday, P. (2015). THE CASE FOR USING DUOLINGO AS PART OF THE LANGUAGE CLASSROOM EXPERIENCE. *RIED. Revista Iberoamericana de Educación a Distancia*. <https://doi.org/10.5944/ried.19.1.14581>
- Nanda, Y. A. I., & Sari, B. W. (2020). NAIVE BAYES ALGORITHM IMPLEMENTATION TO DETECT HUMAN PERSONALITY DISORDERS. *Jurnal Techno Nusa Mandiri*. <https://doi.org/10.33480/techno.v17i1.1239>
- Perdana, R. P., & Irwansyah, I. (2019). Implementasi Asisten Virtual Dalam Komunikasi Pelayanan Pelanggan (Studi Kasus Pada Layanan Pelanggan Telkomsel). *Jurnal Komunikasi*. <https://doi.org/10.24912/jk.v11i2.5491>
- Rizaldi, T., & Putranto, H. A. (2017). Perbandingan Metode Web Scraping Menggunakan CSS Selector dan Xpath Selector. *Teknika*. <https://doi.org/10.34148/teknika.v6i1.56>
- Santosa, B., & Umam, A. (2018). *Data Mining dan Big Data Analytics Edisi 2* (Isa (Ed.); Edisi 2). Penebar Media Pustaka.
- Saputra, D. D., Gata, W., Wardhani, N. K., Parthama, K. S., Setiawan, H., Budilaksono, S., Yogatama, D., Hadiyatna, A., Purnamasari, E. P., Pratama, B., & Novianti, D. (2020). Optimization Sentiments of Analysis from Tweets in myXLCare using Naïve Bayes Algorithm and Synthetic Minority over Sampling Technique Method. *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/1471/1/012014>
- Vesselinov, R., & Grego, J. (2012). Duolingo Effectiveness Study. *City University of New York, USA*.
- Wardhani, N. K., Rezkiani, Kurniawan, S., Setiawan, H., Gata, G., Tohari, S., Gata, W., & Wahyudi, M. (2018). Sentiment analysis article news coordinator minister of maritime affairs using algorithm naive bayes and support vector machine with particle swarm optimization. *Journal of Theoretical and Applied Information Technology*.

## PROFIL PENULIS

Saifurrachman Chohan

Mahasiswa Program Studi Magister Ilmu Komputer, Fakultas Ilmu Komputer, STMIK Nusa Mandiri. Jl. Kramat Raya No.18, RT.5/RW.7, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450.

Arifin Nugroho

Mahasiswa Program Studi Magister Ilmu Komputer, Fakultas Ilmu Komputer, STMIK Nusa Mandiri. Jl. Kramat Raya No.18, RT.5/RW.7, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450.

Sulaeman Hadi Sukmana

Dosen STMIK Nusa Mandiri. Jl. Kramat Raya No.18, RT.5/RW.7, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450.

Windu Gata

Dosen STMIK Nusa Mandiri. Jl. Kramat Raya No.18, RT.5/RW.7, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450.