

OPINION MINING FUNGSI KPI (KEY PERFORMANCE INDIKATOR) DENGAN ALGORITMA NAÏVE BAYES CLASIFIER DAN SUPPORT VECTOR MACHINE (SVM)

Elwin Dewi Nurhazizah¹, Ira Puspitasari²

Program Studi Pengembangan Sumber Daya Manusia Sekolah, Pascasarjana Universitas Airlangga
Email: elwindewinurhazizah@gmail.com; ira-p@fst.unair.ac.id

Kata kunci:

Key Performance Indicators, Opinion Mining, Naïve Bayes, SVM

ABSTRAK

Kinerja diukur untuk mengetahui sejauh mana tujuan direalisasikan sehingga manajemen bisa bertindak cepat dalam mengambil keputusan. Ukuran kinerja pada setiap instansi adalah key performance Indikator (KPI). Penelitian dilakukan dengan menggunakan metodologi CRISP-DM (Cross Industry Standard Process Model for Data Mining). Pengambilan data dilakukan pada Twitter dengan mengambil Tweet Opini penggunaan KPI (Key Performance Indicator) dan akan dianalisis menggunakan RapidMiner dengan algoritma Naive Bayes dan Support Vector Machine (SVM). Hasil penelitian didapatkan akurasi dengan menggunakan algoritma SVM dengan SMOTE menunjukkan hasil 72.32% lebih tinggi dibandingkan algoritma Naïve Bayes yang memperoleh hasil 60.95%. AUC algoritma Naïve Bayes dengan SMOTE sebesar 0.875 dan AUC algoritma Support Vector Machine (SVM) dengan SMOTE sebesar 0.772. Dari hasil ini dapat diketahui bahwa algoritma SVM akan memprediksi lebih baik daripada algoritma Naïve Bayes pada penelitian ini untuk menganalisis sentiment pada fungsi penggunaan KPI (*Key Performance Indicator*).

Keywords:

Key Performance Indicators, Opinion Mining, Naïve Bayes, SVM

ABSTRACT

*Performance is measured to determine the extent to which goals are realized so that management can act quickly in making decisions. The performance measure for each agency is a key performance indicator (KPI). The research was conducted using the CRISP-DM (Cross Industry Standard Process Model for Data Mining) methodology. Data is collected on Twitter by taking Opinion Tweets using KPI (Key Performance Indicator) and will be analyzed using RapidMiner with the Naive Bayes algorithm and Support Vector Machine (SVM). The research results obtained accuracy using the SVM algorithm with SMOTE showing a result of 72.32% higher than the Naïve Bayes algorithm which obtained a result of 60.95%. The AUC of the Naïve Bayes algorithm with SMOTE is 0.875 and the AUC of the Support Vector Machine (SVM) algorithm with SMOTE is 0.772. From these results it can be seen that the SVM algorithm will predict better than the Naïve Bayes algorithm in this study to analyze sentiment in the use of KPI (*Key Performance Indicator*) functions.*

PENDAHULUAN

Karyawan, guru, siswa atau berbagai posisi lainnya akan memiliki standard ketercapaian dari kinerja setiap karyawannya agar mampu mencapai visi dan misi dari tempat bekerja. Kinerja diukur untuk mengetahui sejauh mana tujuan direalisasikan sehingga manajemen bisa bertindak

Opinion Mining Fungsi Kpi (Key Performance Indikator) Dengan Algoritma Naïve Bayes Clasifier Dan Support Vector Machine (Svm)

cepat dalam mengambil keputusan. Manfaat pengukuran kinerja dapat dirasakan secara jangka panjang karena lingkungan bisnis berubah secara dinamis. Dengan mengetahui hasil pengukuran kinerja penyebab rendahnya kinerja dapat segera diketahui sehingga dengan cepat diperbaiki, baik karena kelambatan maupun penyimpangan yang terjadi. Ukuran kinerja pada setiap instansi adalah *key performance Indikator* (KPI). *Key performance indicator* adalah metrik finansial ataupun nonfinansial yang digunakan untuk membantu suatu organisasi menentukan dan mengukur kemajuan terhadap sasaran organisasi (Parmenter, 2007).

Penelitian yang dilakukan oleh (Chandra Babu & Shantharajah, 2019) menggunakan teknik klasifikasi algoritma berupa Decision Tree, Naive Bayes dan Support Vector Machine (SVM) pada KPI (*key performance indicator*) menunjukkan bahwa penggunaan algoritma dapat memprediksi performa dari para siswa dan meningkatkan cara belajarnya. Hal tersebut selaras dengan penelitian yang dilakukan oleh Sultan (2017) dengan judul *data mining approach for detecting key performance indicators* dengan menggunakan algoritma FP-Growth mendapatkan hasil bahwa penggunaan data science membuat proses dalam penilaian KPI lebih efisien dan memberikan strategi untuk dapat meningkatkan kinerja. Penelitian lain yang dilakukan Gelbard (2017) dengan judul *sentiment analysis in organizational work: towards an ontology of people analytics* dengan menggunakan naïve bayes dengan mengambil data pada email para karyawan dari hasil penelitian tersebut didapatkan bahwa sentiment negatif pada KPI yang ada menunjukkan kepuasan kerja yang rendah. Adanya sentiment analisis pada KPI (*key performance indicator*) digunakan untuk memahami faktor manusia dalam suatu organisasi dan keterikatan karyawan pada KPI (*key performance indicator*).

Penelitian terkait dengan sentiment analisis pada twitter dilakukan (Zemnickis et al., 2020) dengan judul *A Little Bird Told Me: Discovering KPIs from Twitter Data* dalam penelitian tersebut menggunakan data twitter sebagai data utama, hasil penelitian menunjukkan penggunaan post context akan menghasilkan warna emosional dari para customer restoran. Penelitian lain dilakukan oleh (Valera & Patel, 2017) dengan judul *A Peculiar Sentiment Analysis Advancement in Big Data* didapatkan hasil bahwa adanya sentiment analisis sebagai solusi yang digunakan untuk mengukur sentiment negatif dan positif dari product Amazon dari competitor lain berdasarkan KPI (*key performance indicator*) bisnis (Wu & Chen, 2014).

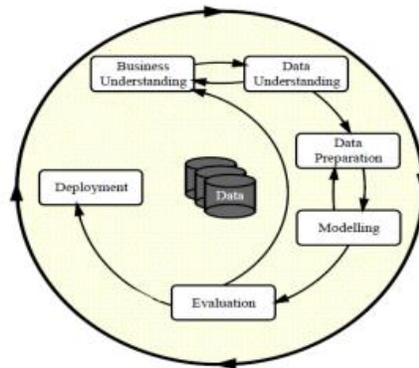
Pendapat mengenai keefektifan atau kritik dalam penggunaan *key performance indicator* (KPI) memiliki pengetahuan tersembunyi yang dapat diekstraksi dengan *text mining* menggunakan *tools* berupa RapidMiner (Badawy et al., 2018). Pendekatan yang dilakukan dalam pengolahan data *text mining* ini sering disebut dengan *opinion mining* atau sentiment analisis (Bouazizi & Ohtsuki, 2015). Analisis sentimen adalah bidang studi yang menganalisis pendapat, sentimen, penilaian, sikap dan emosi seseorang terhadap suatu entitas dan atributnya yang diungkapkan dalam sebuah teks tertulis (Liu, 2015). Salah satu algoritma yang dapat digunakan untuk melakukan sentimen analisis adalah Naïve Bayes dimana pada algoritma ini sistem akan melakukan klasifikasi melalui probabilitas kemungkinan dari data yang diperoleh, algoritma lain yang dapat digunakan dalam teknik analisis sentimen adalah Support Vector Machine (SVM),

dimana pada algoritma tersebut proses pembobotan dilakukan melalui pembentukan patern garis untuk dilakukan pemobobotan dan klasifikasi.

Dari beberapa penelitian diatas disimpulkan bahwa penggunaan dari data mining dapat memberikan keefektifan dalam KPI. Sentimen analisis dalam KPI (*key performance indicator*) memberikan sentiment negatif dan positif pada penerapan KPI namun terbatas pada data dan analisis. Kontribusi yang dapat diberikan dari penelitian ini dengan membandingkan dua metode analisis data mining menggunakan algoritman naïve bayes dan SVM pada data twitter dapat menjadi lebih efektif mengenai kegunaan dari KPI (*key performance indicator*). Dengan demikian penelitian ini bertujuan untuk meninjau keefektifan dari KPI (*key performance indicator*) berdasarkan sentiment analisis publik melalui data twitter.

METODE

Penelitian ini dilakukan dengan menggunakan metodologi CRISP-DM (*CrossIndustry Standard Process Model for Data Mining*). CRISP-DM adalah proses pemodelan yang menyediakan data kerangka pertambangan yang dapat digunakan dalam teknologi dan sektor industri untuk meningkatkan efektivitas biaya, keandalan, pengulangan, dan kecepatan untuk proyek data mining yang besar (Wirth, 2000). Secara umum dijelaskan tentang data mining framework dalam enam tahap yakni business understanding, data understanding, data preparation, modelling, evaluation, deployment bisa terlihat pada Gambar 1.



Gambar 1. CRISP-DM

1. Business Understanding

Pada penelitian ini business understanding menggambarkan mengenai pemahaman mengenai subjek penelitian yang diambil melalui data twitter. Data yang diambil dari twitter berupa opini yang berupa tweet mengenai KPI (*key performance indicator*) pada twitter. Opini mining atau sentiment analisis digunakan untuk mengetahui pendapat seseorang tersebut termasuk pendapat yang positif atau negatif. Pada titik ini, dipahami bahwa menemukan metode klasifikasi terbaik dapat membantu memproses informasi yang akan dibuat untuk membandingkan hasil algoritma yang digunakan.

2. Data Understanding

Tahap selanjutnya adalah mengambil data mengenai komentar twitter mengenai KPI (*key performance indicator*) menggunakan tools berupa RapidMiner. Untuk penelitian ini, komentar yang didapatkan melalui twitter sebanyak 794 komentar. Data tersebut diambil pada range waktu selama satu minggu. Setelah mengumpulkan informasi berupa komentar dilakukan cleaning data karena data yang ada masih acak dan banyak isi komentar yang tidak sesuai dengan isi penelitian. Setelah melakukan pembersihan data tersisa 494 data.

a. Data Preparation

Tahap selanjutnya adalah melakukan persiapan data yang telah didapat agar dapat terolah pada saat melakukan pemodelan. Tahap preprocessing atau pemrosesan data mencakup kegiatan membangun data dan juga membersihkan data agar siap untuk dikelola ke tahap modelling atau pemodelan data. Berikut ini tahap pemrosesan data antara lain:

- 1) Regex Removal: menghilangkan regular expression yang ada dalam teks.
- 2) Remove URL: menghapus URL yang terdapat dalam teks.
- 3) Annotation Removal: menghilangkan tanda @ (annotation) yang ada dalam teks.
- 4) Tokenization: proses tokenisasi pada data teks adalah melakukan pemecahan sekumpulan kalimat menjadi potongan karakter atau kata-kata sesuai kebutuhan yang sering disebut token, sehingga menjadi kata yang memiliki arti tertentu.
- 5) Stemming: menghilangkan imbuhan yang terdapat pada masing-masing kata sehingga menjadi kata dasar, juga bertujuan untuk membersihkan suatu kata dengan pengejaan yang kurang tepat.
- 6) Transform Case: mengubah semua huruf besar atau kapital dalam data menjadi huruf kecil ataupun sebaliknya. Hal ini dilakukan agar ketika masuk ke tahap pemodelan klasifikasi terdapat keseragaman huruf dan tidak terjadi kesalahan dalam proses tokenize, yang biasa digunakan adalah mengubah semua huruf menjadi huruf kecil (lower case).
- 7) Filter Stopwords: sebuah proses untuk menghilangkan kata-kata yang tidak mempunyai arti yang biasanya merupakan kata sambung, kata keterangan dan sebagainya pada hasil parsing sebuah dokumen teks dengan cara membandingkannya dengan stoplist berisi kata-kata yang terlalu sering muncul dalam dokumen-dokumen, belum tentu berguna dalam proses retrieval, kemungkinan besar tidak akan memberikan pengaruh prediksi.
- 8) Filter Tokens (By Length): menghilangkan kata dengan panjang huruf tertentu. Misalnya minimal 2 karakter dan maksimal 25 karakter. Artinya kata yang panjangnya hanya 1 karakter dan lebih dari 25 karakter akan dihilangkan.
- 9) Labelling: merupakan tahap dimana hasil dari tahapan sebelumnya akan dilakukan perhitungan terhadap polarity dari ulasan yang terambil, sehingga dapat menghasilkan dua kategori yaitu label positif dan negatif, untuk label netral (nilai = 0) tidak diproses.

b. Modeling

Pada tahap modelling ini akan dilakukan analisis data berdasarkan algoritma yang telah ditentukan yakni Naïve Bayes dan Support Vector Machine (SVM). Hasil pengujian model yang dilakukan adalah mengklasifikasikan label positif dan negatif pada komentar opini mengenai kegunaan KPI (*key performance indicator*) di Twitter menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM) untuk mendapatkan nilai akurasi terbaik dengan menggunakan tools RapidMiner.

c. Evaluation

Tahap evaluasi bertujuan untuk menentukan kegunaan model yang berhasil dibuat pada langkah pemodelan sebelumnya. Dalam penelitian ini, tahap evaluasi digunakan secara bersamaan dengan Over-Sampling Minoritas Sintetis Teknik (SMOTE). RapidMiner digunakan untuk membantu membandingkan SMOTE, vector drive algoritma dan algoritma minoritas buatan dan dukungan mesin untuk menemukan dua pengelompokan yang berbeda metode antara algoritma Naïve Bayes dan kumpulan data dan membandingkannya dengan Naïve Bayes algoritma. Tujuan penggunaan SMOTE adalah untuk meningkatkan akurasi hasil yang diperoleh dari algoritma dan menemukan algoritma terbaik metode untuk penelitian ini.

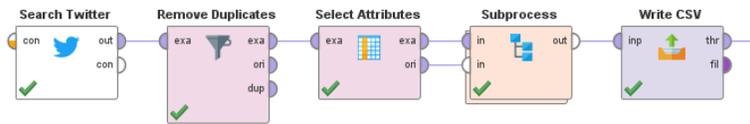
d. Deployment

Deployment adalah fase terakhir dari pendekatan CRISP-DM. Pada tahap ini digunakan untuk membuat model implementasi yang dibuat dalam sebuah alat yang dapat dibangun dengan berbagai jenis pemrograman. Deployment yang digunakan pada penelitian ini adalah tahap implementasi dari hasil perbandingan 2 algoritma yang kemudian dari 2 algoritma yang memiliki nilai akurasi tinggi dapat digunakan sebagai pengembangan prediksi kegunaan KPI (*key performance indicator*) pada seluruh kalangan.

HASIL DAN PEMBAHASAN

Pada penelitian ini mengambil data dari tweet para pengguna twitter mengenai opini mereka mengenai penggunaan KPI (*key performance indicator*). Data yang didapatkan dari komentar youtube sebanyak 794 data. Pada data tersebut telah dilakukan cleansing data dan menghasilkan 494 data komentar dengan menggunakan aplikasi RapidMiner dengan *process document* yang digunakan untuk menghilangkan karakter yang tidak dibutuhkan dalam data seperti hastag, mention, special karakter, username twitter, dan link. Selain hal tersebut pada tahap cleansing data dilakukan dengan Tokenize untuk menghapus kata lain selain Bahasa Indonesia, menghapus kata yang memiliki kurang dari 3 kata dengan *filter token by length*, dan melakukan *transform case* untuk mengubah kata menjadi huruf kecil. Tahap data *preparation* yang telah dilakukan terdapat 494 data yang telah memiliki label dan menjadi dataset pada penelitian ini. Proses pengambilan data pada Twitter digambarkan sebagai berikut:

Opinion Mining Fungsi Kpi (Key Performance Indikator) Dengan Algoritma Naïve Bayes Classifier Dan Support Vector Machine (Svm)

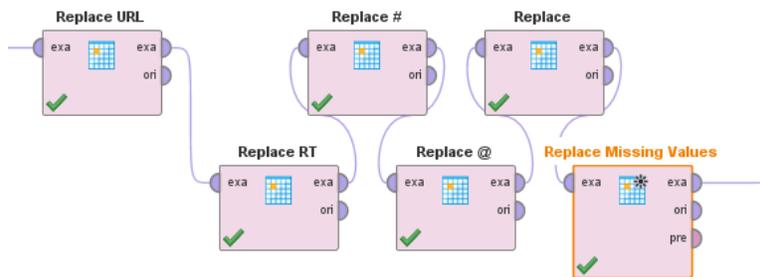


Gambar 2. Crawling data twitter

Pengolahan data untuk penelitian ini menggunakan sentiment analisis. Sentimen analisis pada penelitian ini digunakan untuk mengetahui opini mengenai penggunaan KPI (*key performance indicator*) dengan menggunakan dua algoritma yakni Naïve Bayes dan Support Vector Machine. Setelah melakukan scraping data melalui twitter terdapat beberapa tahap yang akan dilakukan yang akan dipaparkan lebih lanjut pada bagian ini.

a. Pre-processing

Pada tahap preprosesing dilakukan penghapusan pada data-data kotor dan data yang tidak memiliki arti. Dataset yang telah diambil melalui twitter harus dilakukan pembersihan data dengan melakukan replace URL, hastag, RT, @ dan symbol lain yang tidak memiliki arti dengan memasukan data pada subprocess ketika proses crawling data. Berikut merupakan tahap preprosesing yang dilakukan pada RapidMiner.



Gambar 3. Tahap pembersihan data

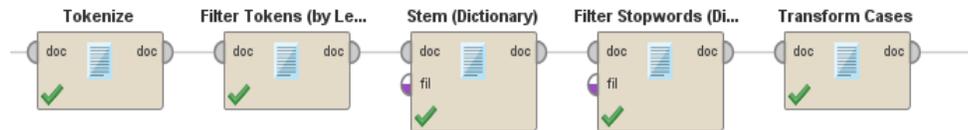
Hasil dari tahap pembersihan data yang telah dilakukan dengan menggunakan tools Rapid Miner dapat terlihat seperti tabel berikut:

Tabel 1. Hasil pembersihan data

Sebelum	Sesudah
@khaims80@AfiqMahadhir	Nk achieve kpi laaa nnti masa
@fahmi_fadzil@Injang_Nation	performancereview dia terpaling
@kkd_gov Nk achieve kpi laaa..	perform nk bonus X nnti melopong
nnti masa performance review, dia	
terpaling perform.. nk bonus... X	
nnti, melopong.. ??????????????	

b. Tokenizing

Pada tahap ini dilakukan dengan menggunakan memecah data menjadi kata yang bermakna. Penggunaan operator *Process Document from Data*, dilakukan beberapa proses yang digunakan untuk membersihkan data agar menjadi vektor yang dapat digunakan sebagai perhitungan algoritma Naïve bayes dan Support vector machine. Berikut merupakan struktur dari tahap *pre processing* kedua;



Gambar 4. Tahap Preprocessing

Hasil dari tahap preporcessig yang telah dilakukan dengan menggunakan tools RapidMiner dapat terlihat seperti tabel berikut:

Tabel 2. Hasil Tokenizing

Sebelum	Sesudah
Nk achieve kpi laaa nnti masa performance review dia terpaling perform nk bonus X nnti melopong	Mau mencapai kpi ketika waktu evaluasi kinerja ia paling kinerja mau bonusnya tidak kosong

c. Filter Token by Leght dan Stopword

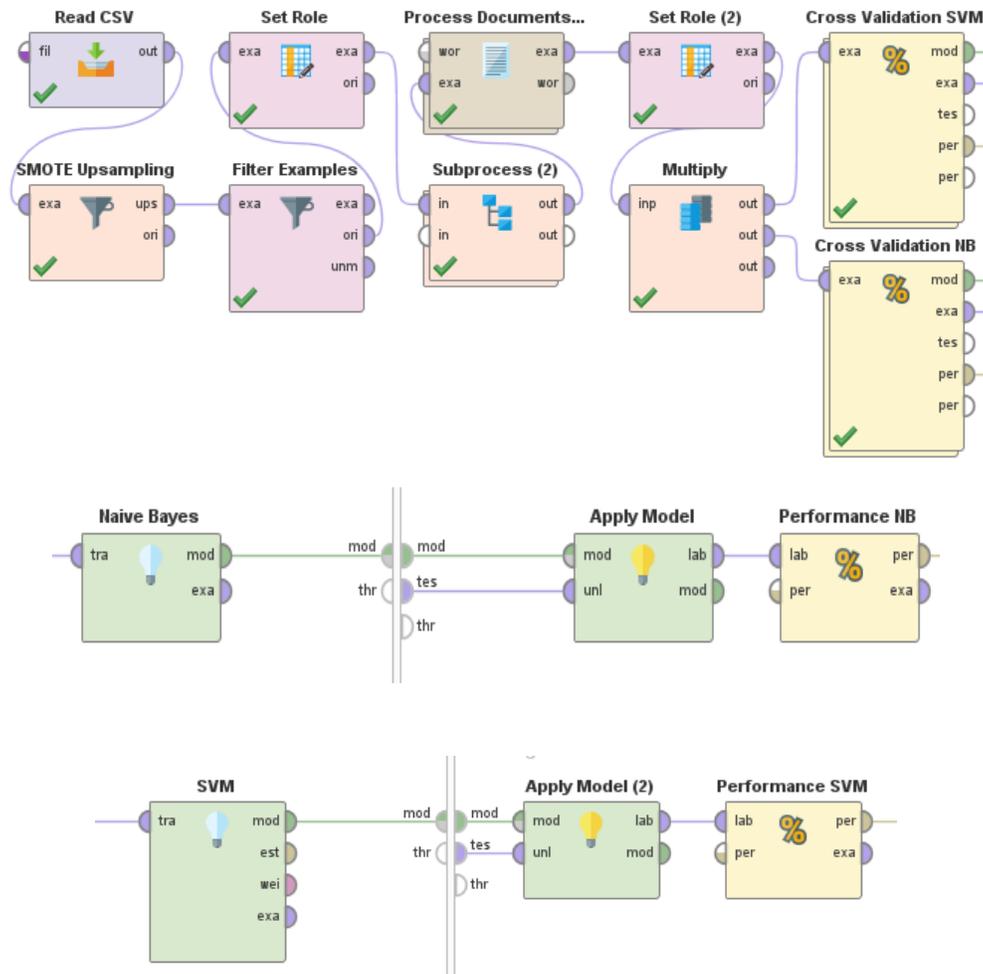
Pada tahap ini peneliti memiih untuk menghilangkan karakter yang kurang dari 2 dan lebih dari 25, seperti kata di, ada, oleh, yang merupakan katakata yang tidak mempunyai makna tersendiri jika dipisahkan dengan kata yang lain dan tidak terkait dengan kata sifat yang berhubungan dengan sentiment dan melakukan stopwords untuk menghapus kalimat yang dianggap tidak memiliki arti bermakan. Hasil dari *filter token by leght dan stopword* dituliskan pada tabel berikut:

Tabel 3. Hasil Filtering

Sebelum	Sesudah
Mau mencapai kpi ketika waktu evaluasi kinerja ia paling kinerja mau bonusnya tidak kosong	Mau mencapai kpi ketika waktu evaluasi kinerja paling kinerja mau bonusnya tidak kosong

1) Model Klasifikasi

Langkah selanjutnya pada penelitian ini dilakukan dengan membuat model dengan menggunakan algoritma klasifikasi untuk menganalisis komentar dataset yang telah melalui tahap preprocessing. Algoritma klasifikasi yang digunakan dalam penelitian ini yakni naïve bayes dan support vector machine (SVM). Pada penelitian ini menggunakan RapidMiner untuk memproses dataset yang telah melalui proses preprocessing. Untuk lebih jelasnya terkait dengan model algoritma pada penelitian ini digambarkan pada gambar 4.



Gambar 4. Model Algoritma Naïve Bayes dan SVM

Tahap pertama dari proses ini adalah mengolah data text komentar twiiter ke dalam tools RapidMiner dengan menggunakan file CSV yang nantinya akan di proses untuk dianalisis lebih lanjut menggunakan algoritma Naïve Bayes (NB) dan Support Vector Machine (SVM) untuk mendapatkan hasil awal dari setiap algoritma, seperti dapat terlihat pada Gambar 2. Setelah tahap pertama dilakukan penelitian ini dilanjutkan dengan membandingkan kedua algoritma dengan menambahkan Minoritas Sinteti Algoritma Over-Sampling Technique (SMOTE). Langkah

penggunaan SMOTE dalam pemodelan proses bertujuan untuk meningkatkan nilai akurasi dari hasil klasifikasi NB dan SVM algoritma, prosesnya dapat dilihat pada Gambar 4.

2) Evaluasi Model

Pada tahap evaluasi model klasifikasi digunakan cross validation. Cross validation digunakan untuk validasi atau evaluasi kinerja model, dan berfungsi sebagai pemilihan rasio untuk memisahkan data pengujian dan pelatihan. Cross validation dapat menghasilkan estimasi akurasi dari proses dataset. Cara kerja dari k-fold cross validation adalah dengan model dilatih oleh subset data training dan divalidasi oleh subset data testing.

Dari tabel 4 kita dapat mengetahui akurasi dari algoritma Naïve Bayes sebesar 90.07%. Hal tersebut menandakan bahwa akurasi yang dimiliki algoritma tersebut sangat baik. Setelah melakukan analisis dengan menggunakan SMOTE pada akurasi didapatkan hasil yang dapat dilihat pada tabel 5 bahwa algoritma Naïve Bayes tingkat akurasinya menurun menjadi 60.95%. Hal tersebut menandakan bahwa algoritma Naïve Bayes dengan menggunakan SMOTE kurang baik digunakan dalam penelitian ini.

Tabel 4. Accuracy Naïve Bayes

Accuracy 90.07%			
	Positive (Actual)	Negative (Actual)	Class Precision
Positive (Prediction)	234	0	100%
Negative (Prediction)	44	165	78.95%
Class Recall	84.17%	100%	

Tabel 5. Accuracy Naïve Bayes dengan SMOTE

Accuracy 60.95%+/- 7.79% (micro average: 60.96%)			
	Positive (Actual)	Negative (Actual)	Class Precision
Positive (Prediction)	156	64	70.91%
Negative (Prediction)	173	214	55.30%
Class Recall	47.42%	76.98%	

Dari tabel 6 kita dapat mengetahui akurasi dari algoritma Support Vector Machine (SVM) sebesar 67.00%. Hal tersebut menandakan bahwa akurasi yang dimiliki algoritma tersebut baik. Setelah melakukan analisis dengan menggunakan SMOTE pada akurasi didapatkan hasil yang dapat dilihat pada tabel 7 bahwa algoritma SVM tingkat akurasinya mengalami peningkatan

Opinion Mining Fungsi Kpi (Key Performance Indikator) Dengan Algoritma Naïve Bayes Clasifier Dan Support Vector Machine (Svm)

menjadi 72.32%. Hal tersebut menandakan bahwa algoritma SVM menggunakan SMOTE sangat baik digunakan dalam penelitian ini.

Tabel 6. Accuracy Support Vector Machine (SVM)
Accuracy 67.00%+/- 4.11% (micro average: 67.00%)

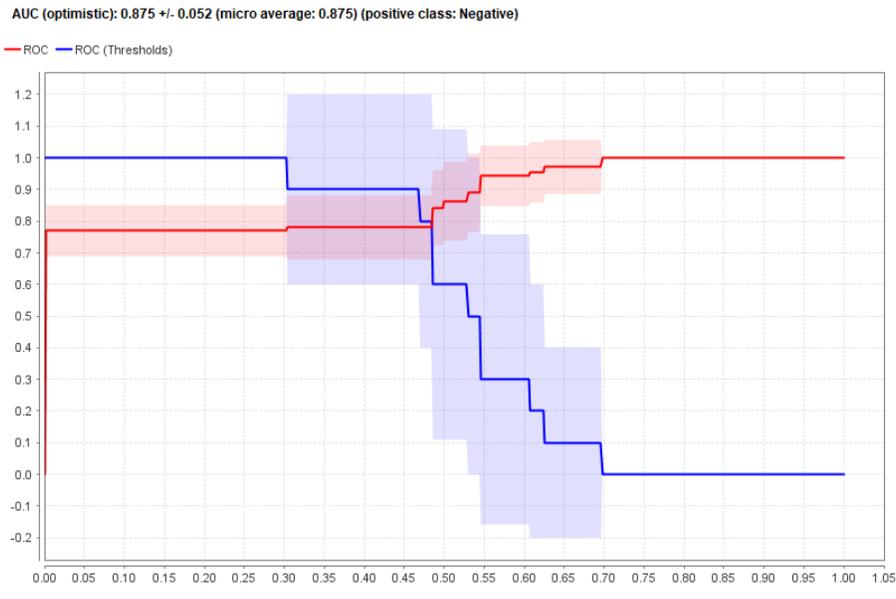
	Positive (Actual)	Negative (Actual)	Class Precision
Positive (Prediction)	321	155	67.44%
Negative (Prediction)	8	10	55.56%
Class Recall	97.57%	6.06%	

Tabel 7. Accuracy Support Vector Machine (SVM) dengan SMOTE
Accuracy 72.32%+/- 4.03% (micro average: 72.32%)

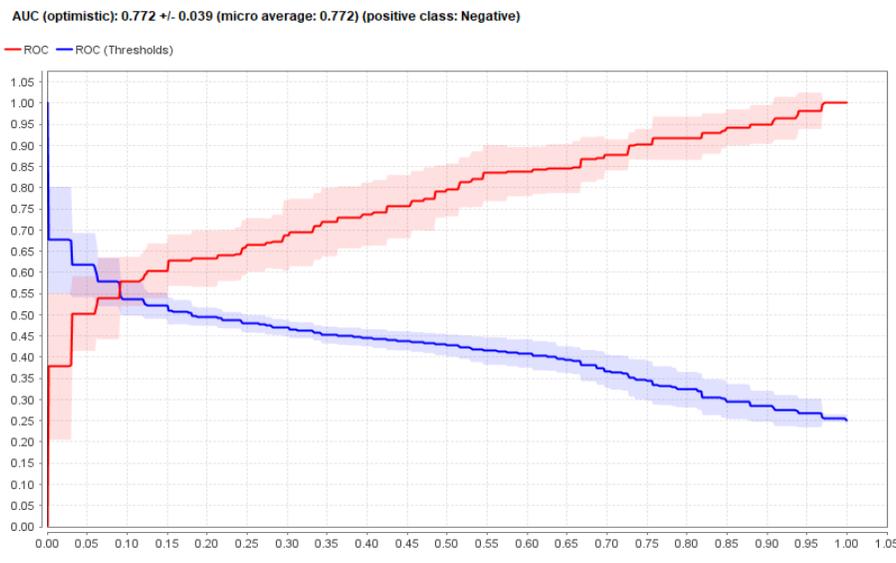
	Positive (Actual)	Negative (Actual)	Class Precision
Positive (Prediction)	275	144	70.69%
Negative (Prediction)	54	164	75.23%
Class Recall	83.59%	58.99%	

Gambar 5 menggambarkan mengenai grafik area under curve (AUC) hasil optimis metode algoritma Naïve Bayes dengan SMOTE sebesar 0.875. Dari hasil tersebut dapat diketahui bahwa hasil AUC menunjukkan perolehan nilai yang termasuk dalam kategori baik dengan rentangan diantara 0.80-0.90. Sedangkan pada gambar 6 menggambarkan mengenai grafik area under curve (AUC) hasil optimis metode algoritma Support Vector Machine (SVM) dengan SMOTE sebesar 0.772. Dari hasil tersebut dapat diketahui bahwa hasil AUC menunjukkan perolehan nilai yang termasuk dalam kategori rata-rata dengan rentangan diantara 0.70-0.80.

Opinion Mining Fungsi Kpi (Key Performance Indikator) Dengan Algoritma Naïve Bayes Classifier Dan Support Vector Machine (Svm)



Gambar 5. AUC Naïve Bayes



Gambar 6. AUC SVM

Berdasarkan komparasi kedua jenis Algoritma Naïve Bayes dan Support Vector Machine (SVM) dapat diketahui bahwa algoritma SVM seperti gambar 5 dan gambar 6 lebih baik dibandingkan dengan Naïve bayes karena mendapatkan nilai yang tinggi pada akurasi walaupun nilai AUC sedikit lebih rendah dibandingkan SVM.

KESIMPULAN

Berdasarkan hasil analisis menggunakan dua algoritma yakni Naïve Bayes dan Support Vector Machine (SVM) dapat diketahui bahwa penggunaan algoritma SVM dengan SMOTE lebih baik dibandingkan dengan algoritma Naïve Bayes. Hasil akurasi dengan menggunakan algoritma SVM dengan SMOTE menunjukkan hasil 72.32% lebih tinggi dibandingkan algoritma Naïve Bayes yang memperoleh hasil 60.95%. Dari hasil ini dapat diketahui bahwa algoritma SVM akan memprediksi lebih baik daripada algoritma Naïve Bayes pada penelitian ini untuk menganalisis sentiment pada fungsi penggunaan KPI (Key Performance Indicator).

Masukan yang dapat diberikan untuk penelitian selanjutnya dengan melakukan scraping data ke internal dari suatu organisasi akan lebih baik untuk memperoleh data penelitian. Pengembangan dalam tahap deployment pun akan dapat dilakukan dengan mengembangkan sebuah framework atau menggunakan salah satu framework tertentu.

DAFTAR PUSTAKA

- Ching Shan Wu, R. J.-C. (2014). KPIs (Key Performance Indicators) in Taiwan Basic Education . *Journal of Modern Education*, 565-578.
- G. Chandra Praba, M. K. (2021). Prediction and Analysis of Key Performance Indicators (Kpi) For Students using Data Science. *International Journal of Engineering Research & Technology (IJERT)*.
- Janis Zemnickis, L. N. (2020). A Little Bird Told Me: Discovering KPIs from Twitter Data. *Springer Nature Switzerland AG*, 161-175.
- Manisha Valera, Y. P. (2017). A Peculiar Sentiment Analysis Advancement in Big Data. *IOP Conf. Series: Journal of Physics*.
- Mohammed Badawy, A. A.-A. (2018). Exploring and Measuring the Key Performance Indicators in Higher Education Institutions. *IJICIS*.
- Mondher Bouazizi, T. O. (2015). Opinion Mining in Twitter How to Make Use of Sarcasm to Enhance Sentiment Analysis. *IEEE*, 1594-1597.
- Nehaya Sultan, A. K. (2017). Data Mining Approach for Detecting Key Performance Indicators. *Journal of Artificial Intelligence*.
- Rian Ardianto, T. R. (2020). Sentiment Analysis On E-Sports For Education Curriculum Using Naive Bayes And Support Vector Machine. *Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)*, 109-122.
- Roy Gelbard, R. R.-G. (2017). Sentiment analysis in organizational work: Towards an ontology of people analytics. *Expert Systems*, e12289. doi:10.1111.
- Senthil Murugan Nagarajan, U. D. (2018). Classifying streaming of Twitter data based on sentiment analysis using hybridization. *Springer*.

Opinion Mining Fungsi Kpi (Key Performance Indikator) Dengan Algoritma Naïve Bayes Clasifier Dan Support Vector Machine (Svm)

Yuzi Mahmud, N. S. (2021). Comparison of Machine Learning Algorithms for Sentiment Classification on Fake News Detection. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, 658-665.



work is licensed under a
Creative Commons Attribution-ShareAlike 4.0 International License