

SENTIMEN ANALISIS TRANSPORTASI ONLINE MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE, NAIVE BAYES DAN KNN

Beni Rahmatullah¹; Pungkas Budiyo²; Suwanda Aditya Saputra³
Universitas BSI
beni.brh@bsi.ac.id¹, pungkas.pby@bsi.ac.id², suwanda.sdz@bsi.ac.id³

ABSTRAK : Transportasi merupakan kebutuhan yang paling digunakan dalam aktifitas sehari-hari dalam bekerja, berbelanja dan kegiatan yang lainnya. Transportasi online menjadi pilihan masyarakat. Aplikasi seperti Gojek dan Grab saat ini yang paling banyak diminati. Saran dan kritik dari pelanggan guna memperbaiki sistem dan pelayanan yang diberikan. Dalam hal ini penulis memperoleh data dari komentar di twitter dan menggunakan algoritma Naive Bayes, Support Vector Machine dan K-Nearest Neighbor sebagai perbandingan akurasi yang dihasilkan. Ketiga algoritma tersebut menggunakan salah satu fitur seleksi *Particle Swarm Optimization*. Hasil akurasi yang didapat membuktikan dengan fitur seleksi *Particle Swarm Optimization* lebih tinggi dan akurasi yang paling tinggi oleh Algoritma Support Vector Machine (PSO) dengan Akurasi 98 % dan AUC 0.988.

Kata Kunci: SVM, NB, KNN, PSO, Online

1. PENDAHULUAN

Sosial media telah menjadi informasi yang dibutuhkan seiring dengan berkembangnya teknologi dan internet. Media sosial seperti blog, situs jejaring sosial (Joshi, Prajapati, & Vala, 2017), komunitas konten, dan dunia virtual menjadi salah satu sumber paling kuat untuk berita, berbelanja, industri, dan banyak lagi (Duwairi, Marji, Sha'Ban, & Rushaidat, 2014). Analisis sentimen juga telah diterapkan pada berbagai rekayasa perangkat lunak (Lin, Zampetti, Bavota, Di Penta, Lanza, & Oliveto, 2018). Analisis sentimen adalah penelitian dalam bidang pemrosesan teks (Medhat, Hassan, & Korashy, 2014).

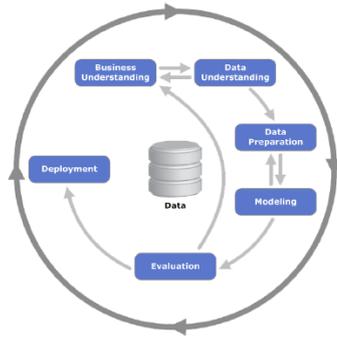
Hadirnya toko *online* seperti tokopedia, bukalapak dan shopee membuat peluang transportasi online tumbuh cepat di Indonesia. Dengan adanya transportasi online memudahkan masyarakat dalam pengiriman barang dan pergi bekerja. Meskipun transportasi online banyak diminati sebagian masyarakat tentu tidak semuanya memberikan komentar positif sebagian masyarakat mungkin ada yang memberikan komentar negative, maka dari itu dibutuhkan suatu teknik dan metode untuk mengetahui presentase klasifikasi komentar-komentar yang diberikan masyarakat pada transportasi online.

Algoritma *Naive Bayes* (Vadivukarassi, Puviarasan, & Aruna, 2017) dan *Support Vector Machine* (Al-Amrani, Lazaar, & El Kadiri, 2018) dioptimalkan dengan *Particle Swarm Optimization* (Wahyudi & Kristiyanti, 2016) adalah metode

penambahan data untuk klasifikasi yang digunakan dalam penelitian ini. Pemrosesan sentimen terhadap teks adalah salah satu bagian penting dalam proses komputasi (Ren & Wang, 2017) dan pelabelan teks menjadi positive dan negative (Stine, 2019). Data yang sudah didapat dilabeli positif dan negatif. Setelah itu dilakukan *preprocessing* baik itu mengubah kata tidak baku menjadi baku atau biasa disebut normalisasi menggunakan kamus dan mencari akar kata yaitu *stemming* dengan bantuan aplikasi Sastrawi Master. Selanjutnya dilakukan juga tokenisasi dan filter terhadap kalimat, kemudian menghilangkan kata-kata yang umum digunakan dan tidak mempunyai Informasi yang berharga pada suatu konteks atau biasa disebut *stopword removal*. Algoritma klasifikasi yang digunakan yaitu *Naive Bayes*, *Support Vector Machine* dan *KNN*.

2. METODE PENELITIAN

Penelitian ini termasuk kedalam *machine learning*, sumber data yang menjadi dataset pada penelitian ini bersumber dari komentar terhadap transportasi online yang terdapat pada media sosial *twitter*. Adapun yang menjadi metode penelitiannya adalah dengan menggunakan *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Metode tersebut memiliki beberapa tahapan seperti yang terlihat pada gambar 1 (IBM, 2011).



Sumber: (IBM, 2011)

Gambar 1. Metode CRISP-DM

1. Business Understanding

Pada tahap *business understanding*, adalah bagian pemahaman tentang objek penelitian yang dilakukan. Pada penelitian ini pemahaman terhadap objek penelitian dilakukan dengan menggali informasi melalui beberapa situs media sosial dengan objek komentar terhadap transportasi online. Motivasi dalam fase ini adalah komentar yang disajikan biasanya berupa teks pada media *digital* yang dikelompokkan berdasarkan isi diskusi masing-masing kategori komentar. Media online tidak hanya terbatas sebagai media sosial saja tapi juga dapat digunakan untuk melihat masalah yang terjadi dan bahkan dapat digunakan untuk melihat penggunaan transportasi online. Analisis sentimen ini dilakukan untuk menemukan metode klasifikasi yang dapat membantu menentukan komentar positif dan negatif. Pada tahap ini juga dibuat pemahaman untuk menemukan metode klasifikasi terbaik sehingga dapat membantu selama proses pengolahan data yang akan dilakukan dengan membandingkan hasil algoritma yang digunakan dan untuk meningkatkan kinerja metode klasifikasi dapat dilakukan dengan menggunakan pemilihan fitur (*feature selection*).

2. Data Understanding

Pada tahap *data understanding*, proses pengambilan data mentah dilakukan sesuai dengan atribut yang diperlukan. Data diperoleh dari twitter sebanyak 50 komentar positive dan 50 komentar negative.

3. Data Preparation

Tahap data preparation adalah tahap dengan proses persiapan data yang bertujuan untuk mendapatkan data bersih dan siap digunakan dalam penelitian. Pada tahap awal penambangan teks, tahap *pre-processing* teks akan dilakukan, pada tahap ini, peneliti akan menggunakan alat bantu Gata Framework Textmining (Kurniawan, Gata, Puspitawati, Tabrani, & Novel, 2019). Pada tahap ini peneliti akan melakukan beberapa proses preprocessing teks pada dataset komentar, antara lain: *Case folding*, *Tokenizing*, *Tagging*, *Filtering*,

Stemming. Pembahasan tahapan-tahapan tersebut akan dijelaskan lebih mendetail pada bagian berikutnya.

4. Modeling

Merupakan tahap pemilihan teknik penambangan dengan menentukan algoritma yang akan digunakan. Penelitian ini menggunakan *tools* yang digunakan untuk melakukan pemodelan sesuai dengan teknik yang telah ditentukan, *tools* tersebut adalah RapidMiner versi 8.2. Penelitian ini menggunakan 2 algoritma klasifikasi sebagai modelnya. Algoritma klasifikasi yang digunakan yaitu, *Naïve Bayes* (NB), *Support Vector Machine* (SVM) dan K-Nearest Neighbor (KNN) yang dioptimasi dengan menggunakan *Particle Swarm Optimization* (PSO). Hasil pengujian setiap model adalah untuk mengklasifikasikan komentar positif dan komentar negatif untuk mendapatkan nilai akurasi terbaik pada setiap algoritma.

5. Evaluation

Tahap *evaluation* bertujuan untuk menentukan kegunaan model yang telah berhasil dibuat pada langkah *modeling* sebelumnya. Penelitian ini menggunakan tahap evaluasi dengan *10 fold-cross validation*. Proses validasi terdiri dari dua subproses yaitu, data pelatihan (*training set*) dan data pengujian (*testing set*). Subproses pelatihan digunakan untuk melatih model algoritma yang sudah ditentukan sebelumnya pada tahap *modeling* dengan data pelatihan yang ada. Setelah model algoritma dilatih pada tahap subproses pelatihan, selanjutnya akan model tersebut akan dilakukan pengujian pada subproses pengujian.

Selanjutnya, setelah data pengujian dilakukan, proses silang dilakukan yaitu, data pengujian digunakan sebagai data pelatihan dan sebaliknya data pelatihan sebagai data pengujian, proses tersebut dilakukan sebanyak 10 kali untuk setiap bagian, sehingga setiap bagian dari 10 bagian telah diuji untuk model algoritma yang digunakan.

6. Deployment

Tahap *deployment* adalah tahap yang digunakan untuk membuat sebuah model implementasi yang dibuat dalam sebuah *tools* yang dapat dibangun dengan berbagai jenis pemrograman. Pembuatan model implementasi ini menggunakan hasil dari proses eksperimen dan evaluasi sebagai sumber data acuan.

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data yang telah diambil dari komentar di twitter seperti yang telah disebutkan pada bagian *data understanding* diatas. Adapun data yang diambil secara keseluruhan adalah

berjumlah 100 data komentar. Kemudian data tersebut pada tahap awal akan dilakukan *data cleansing* dan dilakukan pemberian label dengan label positif atau negatif pada setiap komentar. Dari tahap awal tersebut didapatkan data sebanyak 100 komentar yang telah berlabel maka, data tersebutlah yang akan menjadi dataset pada penelitian ini.

1. Pre-processing

Pembahasan pada tahapan ini adalah proses awal pengolahan dataset sebelum dapat diproses untuk klasifikasi dengan algoritma *Naïve Bayes* (NB), *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) yang dioptimalkan dengan *Particle Swarm Optimization* (PSO). Penelitian ini menggunakan beberapa tahapan *pre-processing* untuk dataset teks komentar, berikut ini adalah tahapan tersebut:

a. Case Folding

Tahap ini adalah proses untuk mengkonversi teks yang tidak beraturan dalam penggunaan huruf di penulisan teks komentar, sehingga teks komentar yang ditulis tidak konsisten. Proses *case folding* ini berfungsi untuk mengubah huruf pada teks komentar kedalam bentuk standar yaitu, menjadi huruf kecil seluruhnya. Hasil dari proses *case folding* tersebut dapat dilihat dibawah ini.

Tabel 1. Perbandingan Teks Komentar Sebelum dan Sesudah Proses Tokenizing

Sebelum	Sesudah
Gojek sudah menonaktifkan akun driver dengan inisial AY sebagai respon aduan dari Agnes.https://t.co/ZYbdHbWj5F	Gojek sudah menonaktifkan akun driver dengan inisial AY sebagai respon aduan dari Agnes https t co ZYbdHbWj F
tq grab driver	tq grab driver
Oknum Driver Gojek Diduga Tipu Pemesan Go Food hingga Rp 9 Juta https://t.co/K04NYfuyO0#LINETODAY - tiati gaes kal...https://t.co/13y7n6aFy5	Oknum Driver Gojek Diduga Tipu Pemesan Go Food hingga Rp Juta https t co K NYfuyO LINETODAY tiati gaes kal https t co y n aFy
bikin emosi mana ada tambahan biaya tiba2 trus driver gaada konfirmasi, bener2 yaaa driver grab td tuh, mana gue ga... https://t.co/dlqkpmRRP6	bikin emosi mana ada tambahan biaya tiba trus driver gaada konfirmasi bener yaaa driver grab td tuh mana gue ga https t co dlqkpmRRP

Sumber : (Rahmatullah & Rosiyadi, 2020)

b. Tokenizing

Tahap *Tokenizing* adalah tahap pemotongan *string input* berdasarkan tiap kata yang menyusunnya. Hasil dari proses tersebut dapat dilihat dibawah ini.

Tabel 2. Perbandingan Teks Komentar Sebelum dan Sesudah Proses Tokenizing

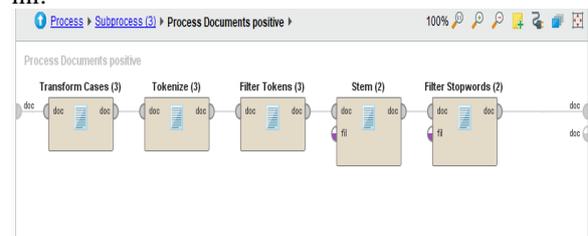
Sebelum	Sesudah
---------	---------

Gojek sudah menonaktifkan akun driver dengan inisial AY sebagai respon aduan dari Agnes.https://t.co/ZYbdHbWj5F	Gojek sudah menonaktifkan akun driver dengan inisial AY sebagai respon aduan dari Agnes https t co ZYbdHbWj F
---	---

tq grab driver	tq grab driver
Oknum Driver Gojek Diduga Tipu Pemesan Go Food hingga Rp 9 Juta https://t.co/K04NYfuyO0#LINETODAY - tiati gaes kal...https://t.co/13y7n6aFy5	Oknum Driver Gojek Diduga Tipu Pemesan Go Food hingga Rp Juta https t co K NYfuyOLINETODAY tiati gaes kal https t co y n aFy
bikin emosi mana ada tambahan biaya tiba2 trus driver gaada konfirmasi, bener2 yaaa driver grab td tuh, mana gue ga... https://t.co/dlqkpmRRP6	bikin emosi mana ada tambahan biaya tiba trus driver gaada konfirmasi bener yaaa driver grab td tuh mana gue ga https t co dlqkpmRRP

Sumber : (Rahmatullah & Rosiyadi, 2020)

Tabel 1 dan Tabel 2 diolah dengan menggunakan tools yang terdapat pada rapidminer seperti gambar dibawah ini:



Sumber : (Rahmatullah & Rosiyadi, 2020)

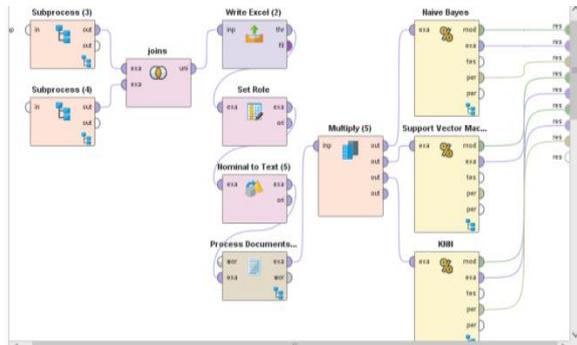
Gambar 2. Tahap Preprocessing.

2. Pembuatan Model Klasifikasi

Tahap ini pada penelitian adalah untuk membuat model dengan menggunakan algoritma klasifikasi untuk dataset teks komentar yang sudah melalui tahap *pre-processing*. Tahap ini menggunakan tiga algoritma klasifikasi yaitu, *Naive Bayes* (NB), *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) pada gambar 3 dan yang dioptimalkan dengan *Particle Swarm Optimization* (PSO) pada gambar 4 menggunakan *tools* Rapidminer versi 8.0 untuk mengolah dataset teks komentar yang sudah melalui tahap *data preparation* dengan *pre-processing* teks.

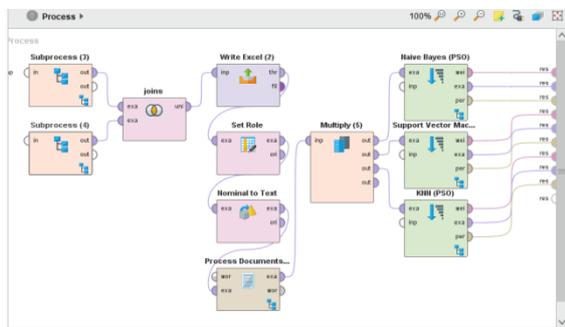
Tahap pertama dari proses ini adalah data teks komentar akan diunggah ke dalam tools dengan menggunakan file excel yang kemudian akan di proses dengan algoritma *Naive Bayes* (NB), *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) untuk mendapatkan hasil awal dari masing-masing algoritma tersebut. Setelah tahap pertama dilakukan maka, penelitian ini melanjutkan dengan melakukan perbandingan kedua algoritma dengan menambahkan algoritma optimalisasi dengan menggunakan *Particle Swarm Optimization* (PSO). Langkah penggunaan PSO dalam proses modelling ini bertujuan untuk

meningkatkan nilai akurasi hasil dari klasifikasi algoritma NB, SVM dan KNN.



Sumber : (Rahmatullah & Rosiyadi, 2020)

Gambar 3. Design Model Menggunakan Algoritma Naive Bayes, Support Vector Machine dan KNN.

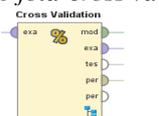


Sumber : (Rahmatullah & Rosiyadi, 2020)

Gambar 4. Design Model Menggunakan Algoritma Naive Bayes, Support Vector Machine dan KNN dengan PSO.

3. Evaluasi Model Klasifikasi

Tahapan evaluasi bertujuan untuk menentukan nilai kegunaan dari model yang telah berhasil dibuat pada langkah sebelumnya. Untuk evaluasi digunakan *10 fold-cross validation*.



Sumber : (Rahmatullah & Rosiyadi, 2020)

Gambar 5. Operator *Cross Validation*.

Pada pengujian ini, data digunakan adalah data bersih yang telah melalui *pre-processing*. Data tersebut diambil dari operator *Read Excel*, hal ini dilakukan karena dataset disimpan dalam bentuk Excel.

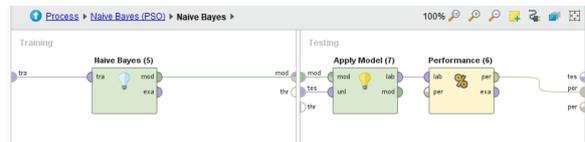


Sumber : (Rahmatullah & Rosiyadi, 2020)

Gambar 6. Operator *Read Excel*.

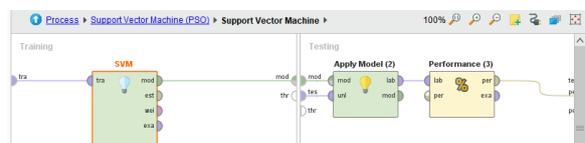
Process documents from files untuk mengkonversi file menjadi dokumen. Proses validasi terdiri dari data pelatihan dan data pengujian. Pada

tahap ini juga menggunakan *Set Role* yang berfungsi untuk menentukan *field* pada kelas kemudian menggunakan *Particle Swarm Optimization (PSO)* agar akurasi yang dihasilkan lebih tinggi dari hasil pemodelan yang telah dilakukan sebelumnya. Desain model klasifikasi yang di evaluasi menggunakan PSO dapat dilihat pada gambar 7, 8 dan 9.



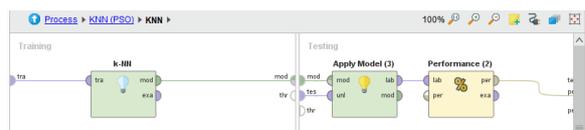
Sumber : (Rahmatullah & Rosiyadi, 2020)

Gambar 7. Proses Evaluasi *10 Fold Cross-Validation* untuk algoritma NB dengan PSO.



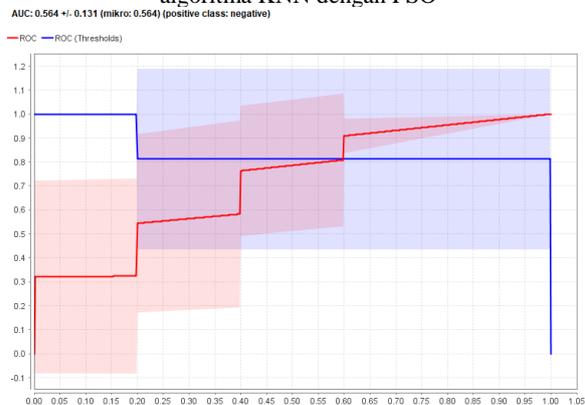
Sumber : (Rahmatullah & Rosiyadi, 2020)

Gambar 8. Proses Evaluasi *10 Fold Cross-Validation* untuk algoritma SVM dengan PSO.



Sumber : (Rahmatullah & Rosiyadi, 2020)

Gambar 9. Proses Evaluasi *10 Fold Cross-Validation* untuk algoritma KNN dengan PSO



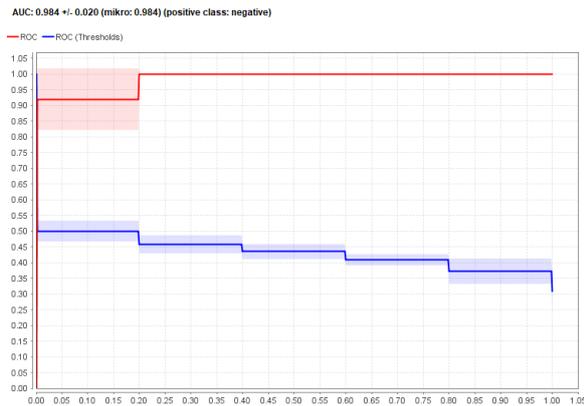
Sumber : (Rahmatullah & Rosiyadi, 2020)

Gambar 10. Kurva AUC *Naive Bayes*

Kategori Klasifikasi AUC:

1. 0.90 - 1.00 = *excellent classification*
2. 0.80 - 0.90 = *good classification*
3. 0.70 - 0.80 = *fair classification*
4. 0.60 - 0.70 = *poor classification*
5. 0.50 - 0.60 = *failure*

Kurva ROC *Naive Bayes* dengan nilai AUC (*Area Under Curve*) yang dihasilkan dari gambar 10 diatas sebesar 0.564 dimana diagnosa hasilnya adalah *failure*

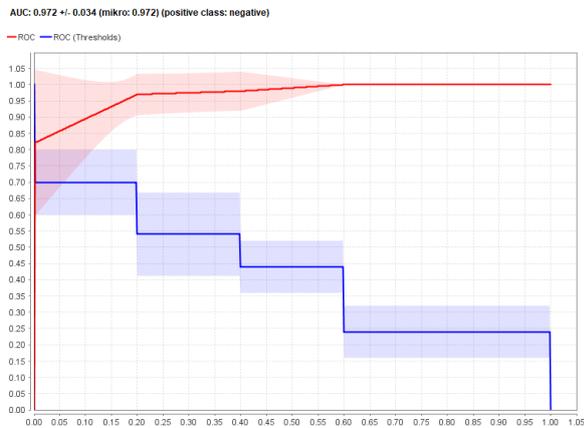


Sumber : (Rahmatullah & Rosiyadi, 2020)
Gambar 11. Kurva AUC Support Vector Machine

Kategori Klasifikasi AUC:

1. 0.90 - 1.00 = *excellent classification*
2. 0.80 - 0.90 = *good classification*
3. 0.70 - 0.80 = *fair classification*
4. 0.60 - 0.70 = *poor classification*
5. 0.50 - 0.60 = *failure*

Kurva ROC Support Vector Machine dengan nilai AUC (Area Under Curve) yang dihasilkan dari gambar 11 diatas sebesar 0.984 dimana diagnosa hasilnya adalah *excellent classification*.

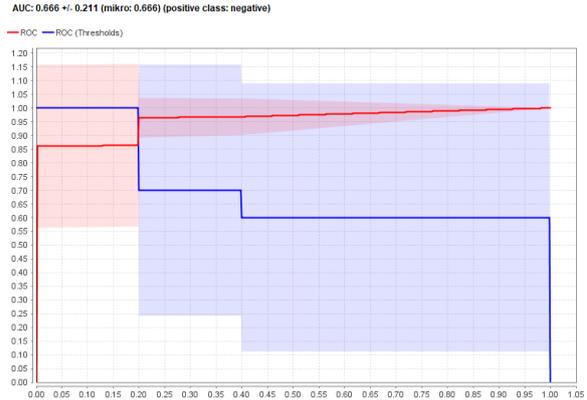


Sumber : (Rahmatullah & Rosiyadi, 2020)
Gambar 12. Kurva AUC KNN

Kategori Klasifikasi AUC:

1. 0.90 - 1.00 = *excellent classification*
2. 0.80 - 0.90 = *good classification*
3. 0.70 - 0.80 = *fair classification*
4. 0.60 - 0.70 = *poor classification*
5. 0.50 - 0.60 = *failure*

Kurva ROC KNN dengan nilai AUC (Area Under Curve) yang dihasilkan dari gambar 12 diatas sebesar 0.972 dimana diagnosa hasilnya adalah *excellent classification*

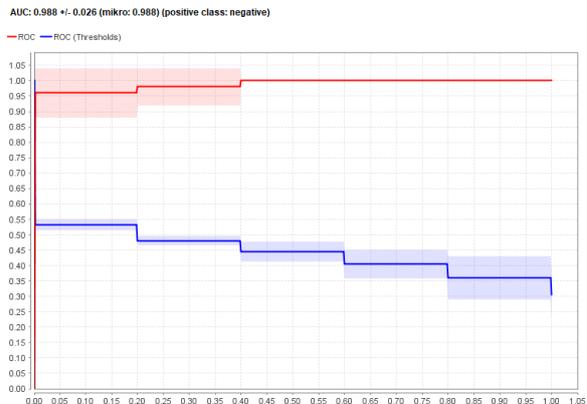


Sumber : (Rahmatullah & Rosiyadi, 2020)
Gambar 13. Kurva AUC Naive Bayes dan PSO

Kategori Klasifikasi AUC:

1. 0.90 - 1.00 = *excellent classification*
2. 0.80 - 0.90 = *good classification*
3. 0.70 - 0.80 = *fair classification*
4. 0.60 - 0.70 = *poor classification*
5. 0.50 - 0.60 = *failure*

Kurva ROC Naive Bayes dan PSO dengan nilai AUC (Area Under Curve) yang dihasilkan dari gambar 13 diatas sebesar 0.666 dimana diagnosa hasilnya adalah *poor classification*

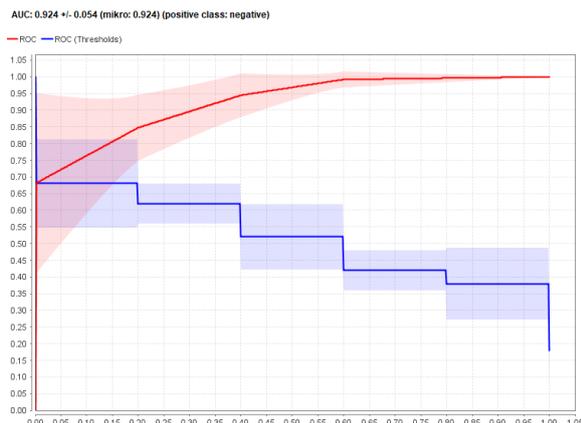


Sumber : (Rahmatullah & Rosiyadi, 2020)
Gambar 14. Kurva AUC Support Vector Machine dan PSO

Kategori Klasifikasi AUC:

1. 0.90 - 1.00 = *excellent classification*
2. 0.80 - 0.90 = *good classification*
3. 0.70 - 0.80 = *fair classification*
4. 0.60 - 0.70 = *poor classification*
5. 0.50 - 0.60 = *failure*

Kurva ROC Support Vector Machine dan PSO dengan nilai AUC (Area Under Curve) yang dihasilkan dari gambar 14 diatas sebesar 0.988 dimana diagnosa hasilnya adalah *excellent classification*.



Sumber : (Rahmatullah & Rosiyadi, 2020)
Gambar 15. Kurva AUC KNN

Kategori Klasifikasi AUC:

1. 0.90 - 1.00 = *excellent classification*
2. 0.80 - 0.90 = *good classification*
3. 0.70 - 0.80 = *fair classification*
4. 0.60 - 0.70 = *poor classification*
5. 0.50 - 0.60 = *failure*

Kurva ROC KNN dengan nilai AUC (*Area Under Curve*) yang dihasilkan dari gambar 15 di atas sebesar 0.924 dimana diagnosa hasilnya adalah *excellent classification*

Pada tahap ini menggunakan Role Set yang berfungsi untuk menentukan bidang di kelas kemudian menggunakan *Particle Swarm Optimization* (PSO) sehingga akurasi yang dihasilkan lebih tinggi dari hasil pemodelan sebelumnya. Kategori Klasifikasi memiliki nilai: 1). 0,90 - 1,00 = klasifikasi sangat baik. 2). 0,80 - 0,90 = klasifikasi yang baik. 3). 0,70 - 0,80 = klasifikasi yang adil. 4). 0,60 - 0,70 = klasifikasi buruk. 5). 0,50 - 0,60 = kegagalan.

Tabel 3. Nilai Akurasi dan AUC

Algoritma	Akurasi	AUC
Support Vector Machine	87%	0.984
Support Vector Machine + PSO	98%	0.988
KNN	94%	0.972
KNN + PSO	87%	0.924
Naive Bayes	81%	0.564
Naive Bayes + PSO	97%	0.666

Sumber : (Rahmatullah & Rosiyadi, 2020)

4. KESIMPULAN DAN SARAN

Dari hasil penelitian yang telah dilakukan terkait dengan sentimen analisis transportasi online pada komentar twitter, penggunaan Algoritma *Support Vector Machine* yang dioptimalisasikan dengan

Particle Swarm Optimization (PSO) memiliki tingkat akurasi yang lebih baik daripada Naive Bayes yang dioptimalisasikan dengan *Particle Swarm Optimization* (PSO) dan K-Nearest Neighbor *Particle Swarm Optimization* (PSO). Sementara *Particle Swarm Optimization* (PSO) memainkan peran penting dalam meningkatkan akurasi algoritma *Support Vector Machine*. Sehingga algoritma *Support Vector Machine* berbasis *Particle Swarm Optimization* (PSO) dapat menjadi solusi untuk klasifikasi dalam sentimen analisis transportasi online.

Selain kesimpulan di atas, penelitian ini juga berupaya memberikan saran untuk penelitian lebih lanjut, seperti melakukan pengujian dengan menggunakan algoritma yang berbeda dari penelitian ini dan tidak hanya melakukan analisis sentimen pada komentar yang berkaitan dengan transportasi online atau media sosial online lainnya namun juga bisa dengan topik yang berbeda misal seperti topik pelayanan pendidikan, pelayanan penjualan, pelayanan publik, dan lainnya.

5. DAFTAR PUSTAKA

- [1]. Joshi, M., Prajapati, P., Shaikh, A., & Vala, V. (2017). A Survey on Sentiment Analysis. *International Journal of Computer Applications*, 163(6), 34–38. doi:10.5120/ijca2017913552.
- [2]. Duwairi, R. M., Marji, R., Sha'Ban, N., & Rushaidat, S. (2014). Sentiment analysis in arabic tweets. In *2014 5th International Conference on Information and Communication Systems, ICICS 2014*. doi:10.1109/IACS.2014.6841964.
- [3]. Lin, B., Zampetti, F., Bavota, G., Di Penta, M., Lanza, M., & Oliveto, R. (2018). Sentiment analysis for software engineering (pp. 94–104). doi:10.1145/3180155.3180195.
- [4]. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. doi:10.1016/j.asej.2014.04.011.
- [5]. Vadivukarassi, M., Puviarasan, N., & Aruna, P. 2017. Sentimental Analysis of Tweets Using Naive Bayes Algorithm. *World Applied Sciences Journal* 35(1), pp.54-59.
- [6]. Al-Amrani, Y., Lazaar, M., & El Kadiri, E.K. 2018. Sentiment Analysis Using Hybrid Method of Support Vector Machine And

- Decision Tree. *Journal of Theoretical and Applied Information Technology* 96(7), pp.1886-1895.
- [7]. Wahyudi, M., & Kristiyanti, D. A. (2016). Sentiment analysis of smartphone product review using support vector machine algorithm-based particle swarm optimization. *Journal of Theoretical and Applied Information Technology*, 91(1), 189–201.
- [8]. Ren, F., & Wang, L. (2017). Sentiment analysis of text based on three-way decisions. *Journal of Intelligent and Fuzzy Systems*, 33(1), 245–254. doi:10.3233/JIFS-161522.
- [9]. Stine, R. A. (2019). Sentiment Analysis. *Annual Review of Statistics and Its Application*, 6(1), 287–308. doi:10.1146/annurev-statistics-030718-105242
- [10]. IBM. (2011). IBM SPSS Modeler CRISP-DM Guide. *Career: Data and Analytics*, 53.
- [11]. Kurniawan, S., Gata, W., Puspitawati, D. A., -, N., Tabrani, M., & Novel, K. (2019). Perbandingan Metode Klasifikasi Analisis Sentimen Tokoh Politik Pada Komentar Media Berita Online. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(2), 176–183. <https://doi.org/10.29207/resti.v3i2.935>