VOL. 5. NO. 2 FEBRUARI 2020
P-ISSN: 2685-8223 | E-ISSN: 2527-4864

# PENERAPAN PARTICLE SWARM OPTIMIZATION TERHADAP SUPPORT VECTOR MACHINE PADA REVIEW PENGGUNA TRANSPORTASI UDARA

**Retno Sari[1]; Ratih Yulia Hayuningtyas[2]**

[1,2]Program Studi Teknik Informatika
STMIK Nusa Mandiri
www.nusamandiri.ac.id
[1]retno.rnr@nusamandiri.ac.id, [2]ratih.ryl@nusamandiri.ac.id

***Abstract***—*Air transportation is currently one of the alternative choice in travel by general public. To choose air transportation in accordance with the choice, the buyer can see a review of air transportation. Reviews obtained of text information from varios sources. Sometimes reviews about air transportation make it difficult for buyers to draw conclusions about air transportation information, so we need a method to determine the accuracy of an information. In this study air transportation review uses data of 100 reviews which are used as a dataset and then classified into 50 positive reviews and 50 negative reviews. The dataset will be tested with preprocessing steps and methods. The purpose of this research is the application of Particle Swarm Optimization to Support Vector Machine can increase the value of accuracy. This test resulted in an accuracy rate of 61,03% with AUC 0,953 with Support Vector Machine algorithm, while an accuracy of 71,00% with AUC 0,976 with the Support Vector Machine algorithm optimized with Particle Swarm Optimization. From the test results above, it can be concluded that the research using Support vector Machine algorithm which is optimized with Particle Swarm Optimization has an accuracy level of 9.97%.*

***Keywords****: Particle Swarm Optimization, Review, Support Vector Machine.*

**Intisari**—Transportasi udara saat ini menjadi salah satu alternative pilihan dalam perjalanan oleh masyarakat umum. Untuk memilih transportasi udara yang sesuai dengan pilihan, pembeli dapat melihat sebuah review dari transportasi udara. Review yang didapat berupa informasi teks dari berbagai sumber. Terkadang review mengenai transportasi udara menyulitkan pembeli untuk menarik kesimpulan mengenai informasi transportasi udara, maka diperlukan suatu metode untuk menentukan tingkat akurasi dari sebuah informasi. Dalam penelitian ini review transportasi udara menggunakan data sebanyak 100 review yang dijadikan dataset kemudian diklasifikasikan kedalam 50 review positif dan 50 review negatif. Dataset tersebut akan diuji dengan tahapan preprocessing dan metode. Tujuan dari penelitian ini yaitu penerapan *Particle Swarm Optimization* terhadap *Support Vector Machine* dapat meningkatkan nilai akurasi. Pengujian ini menghasilkan tingkat akurasi sebesar 61,03% dengan AUC 0,953 dengan algoritma *Support Vector Machine*, sedangkan akurasi sebesar 71,00% dengan AUC 0,976 dengan algoritma *Support Vector Machine* yang di optimasi dengan *Particle Swarm Optimization*. Dari hasil pengujian diatas dapat disimpulkan penelitian dengan algoritma *Support Vector Machine* yang di optimasi dengan Particle Swarm Optimization tingkat akurasinya naik sebesar 9,97%.

***Kata Kunci****: Particle Swarm Optimization, Review, Support Vector Machine.*

## PENDAHULUAN

Transportasi umum saat ini memiliki banyak pilihan baik transportasi darat, laut maupun udara. Transportasi udara saat ini menjadi salah satu alternative pilihan yang banyak dipilih oleh masyarakat umum dikarenakan dapat menghemat waktu perjalanan yang tidak membutuhkan waktu lama untuk berpergian ke suatu tempat.

Untuk memilih maskapai penerbangan yang sesuai dengan keinginan, calon pembeli dapat melihat opini maskapai penerbangan untuk mengetahui apakah maskapai tersebut sesuai dengan keinginan. Informasi berupa teks merupakan informasi penting dan didapatkan dari berbagai sumber (Ling, N. Kencana, & Oka, 2014).

Situs ulasan online semakin meningkat popularitasnya karena banyak orang yang mencari masukan dari sesama pengguna mengenai layanan dan produk (Brody & Elhadah, 2010).

Analisis sentimen atau *opinion mining* merupakan proses mengolah data tekstual untuk mendapatkan informasi yang ada pada opini yang

diberikan (I. Rozi, Pramono, & Dahlan, 2012). Analisis sentimen berguna untuk menganalisis ulasan untuk diterjemahkan menjadi sesuatu yang bermakna (Monarizqa, Nugroho, & Hantono, 2014). Analisis sentimen memiliki tugas mengelompokan polaritas dari suatu dokumen, kalimat atau pendapat (Ling et al., 2014).

Sudah dilakukan penelitian sebelumnya mengenai analisis sentimen opini maskapai penerbangan. Penelitian sebelumnya yaitu analisis sentimen tentang opini maskapai penerbangan pada dokumen twitter menggunakan algoritma Support Vector Machine (SVM) (Pravina, Cholissodin, & Adikara, 2019), analisis sentimen pada opini pengguna maskapai penerbangan menggunakan Hybrid Cuckoo Search (N. F. Rozi, Arianto, & Hapsari, 2019).

Algoritma Support Vector Machine (SVM) memiliki tujuan merancang cara pembelajaran komputasi yang efisien dalam pemisahan hyperplane didalam ruang fitur berdimensi tinggi (Drajana, 2017). Algoritma Support Vector Machine merupakan sistem pembelajaran yang menggunakan fungsi linear didalam sebuah fitur yang memiliki dimensi tinggi dan dilatih dengan menggunakan algoritma pembelajaran. Pemilihan fungsi kernel yang tepat dan sesuai akan menentukan feature space dimana fungsi klasifier akan dicari (Parapat, Furqon, & Sutrisno, 2018).

Particle Swarm Optimization (PSO) memiliki kecepatan partikel dalam ruang pencarian dan bergerak dengan kecepatan dinamis, parikel memiliki kencenderungan untuk bergerak menuju daerah pencarian yang lebih baik selama proses pencarian (Muhamad et al., 2017). Particle Swarm Optimization (PSO) telah digunakan sebagai teknik yang efektif di berbagai bidang termasuk dalam pemilihan fitur (Xue, Zhang, Member, & Browne, 2012).

Dalam penelitian analisis sentimen tentang opini maskapai penerbangan pada dokumen twitter menggunakan algoritma *Support Vector Machine* (SVM) akurasi yang didapat masih tergolong kecil yaitu sebesar 40% (Pravina et al., 2019).

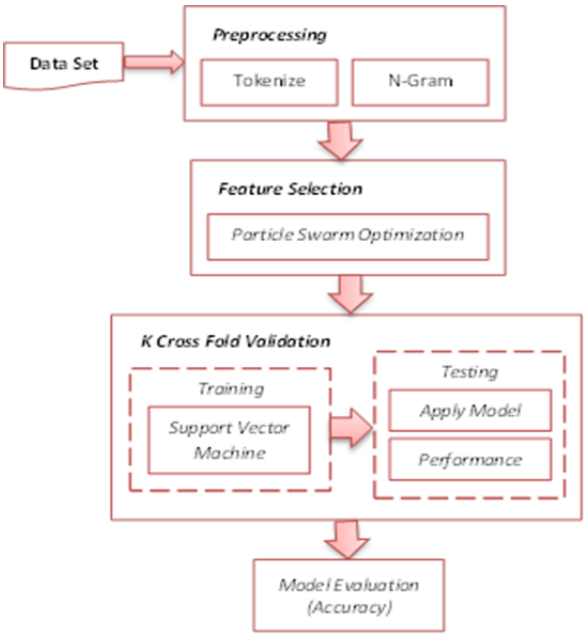
Pada penelitian sebelumnya yang diuraikan diatas, penelitian dengan menggunakan algoritma *Support Vector Machine* (SVM) tingkat akurasinya hanya 40%. Penelitian ini berjudul penerapan *Particle Swarm Optimization* terhadap *Support Vector Machine* pada review pengguna transportasi udara memiliki tujuan agar nilai akurasi dapat meningkat.

## BAHAN DAN METODE

Pada penelitian ini dilakukan metode penelitian eksperimen, berikut tahapannya:

a. Pengumpulan Data
   Pada penelitian ini penulis mengambil data dari situs Review www.tripadvisor.com yang terdiri dari 50 review positif dan 50 review negative dengan teks berbahasa Indonesia.
b. Pengolahan Awal Data
   Preprocessing yang dilakukan pada dataset, yaitu:
   1) *Tokenize*
      Pada tahap ini menghilangkan tanda baca maupun simbol apapun yang bukan huruf. Sehingga membentuk kumpulan kata yang memiliki arti.
   2) *N-Gram*
      Pada tahap *N-Gram* merupakan tahap pemotongan karakter dalam suatu string atau potongan n kata pada suatu kalimat. Pada proses ini dilakukan n-gram dengan bi-gram yaitu sebuah *string* dipotong menjadi 2 bagian.
c. Metode yang digunakan
   Metode yang digunakan yaitu *Support Vector Machine* (SVM) yang dioptimasi dengan *Particle Swarm Optimization*
d. Eksperimen dan Pengujian Metode
   Eksperimen pada penelitian ini menggunakan Rapid Miner 5 dalam pengolahan data.
e. Evaluasi dan Validasi Hasil
   Validasi dilakukan menggunakan *fold cross validation*, untuk melakukan pengukuran akurasi menggunakan *confusion matrix* fan kurva ROC untuk mengukur nilai AUC.

Berikut adalah gambar model dari penelitian ini:




Sumber: (Sari & Hayuningtyas, 2019)
Gambar 1. Model Penelitian


Terakreditasi Sinta Peringkat 3 (Sinta 3) berdasarkan SK Dirjen Penguatan RisBang Kemenristekdikti Nomor 28/E/KPT/2019, Tanggal 26 September 2019

## HASIL DAN PEMBAHASAN

Penelitian ini membahas mulai dari pengumpulan data, pengolahan data, metode yang digunakan, eksperiman dan pengujian metode, evaluasi dan validasi hasil.

Data pada penelitian ini terdiri dari 50 data review positif dan 50 data review negative dengan teks berbahasa Indonesia. Masing-masing data review dikelompokkan ke dalam folder positif untuk review positif dan ke dalam folder negative untuk review negative. Pada penelitian ini penulis mengambil data dari situs review www.tripadvisor.com

Proses pengolahan awal data terdiri dari proses *tokenize* dan n-gram. Pada proses tokenize semua review dihilangkan tanda baca maupun simbol apapun yang bukan huruf. Sehingga membentuk kumpulan kata yang memiliki arti. Pada table 1 dapat dilihat hasil proses *tokenize*.

Tabel 1. Hasil Proses *Tokenize*

| Ulasan | *Tokenize* |
|---|---|
| Kualitas pelayanan kru, khususnya pramugaranya sangat buruk karena sangat tidak ramah, jutek, kurang beretika dan bersikap kurang santun terhadap penumpang. Jadi, saat itu saya cuma bertanya apakah toiletnya kosong. Malah pramugara tersebut menjawab dengan kasar, ketus dan kurang sopan. Kebetulan juga, pintu toilet agak keras dibuka dan pramugara tersebut diam saja, tidak mau membantu saya untuk membukakan pintu tersebut. Malah saya melihat ekspresi mukanya yang terlihat kesal dan marah, seperti orang sress. Selain itu, juga ada delay yang cukup lama. Untuk fasilitas pesawat terlalu standar dan tidak ada jatah makan. | Kualitas pelayanan kru khususnya pramugaranya sangat buruk karena sangat tidak ramah jutek kurang beretika dan bersikap kurang santun terhadap penumpang Jadi saat itu saya cuma bertanya apakah toiletnya kosong Malah pramugara tersebut menjawab dengan kasar ketus dan kurang sopan Kebetulan juga pintu toilet agak keras dibuka dan pramugara tersebut diam saja tidak mau membantu saya untuk membukakan pintu tersebut Malah saya melihat ekspresi mukanya yang terlihat kesal dan marah seperti orang sress Selain itu juga ada delay yang cukup lama Untuk fasilitas pesawat terlalu standar dan tidak ada jatah makan |

Sumber: (Sari & Hayuningtyas, 2019)

Proses pengolahan data awal selanjutnya yaitu gram. Pada proses ini memotong kataker dalam suatu string atau potongan n kata pada suatu kalimat. Pada proses ini string dipotong menjadi 2 bagian. Pada table 2 dapat dilihat hasil proses N-Gram.

Tabel 2. Hasil Proses Bigram

| Ulasan | Bigram |
|---|---|
| Gila Lion nih ya, udah tau Adek saya masih SMA sendirian lagi, lagi antre buat check-in malah di close karena waktu nya udh mepet. Eh Lion gk usah sok2an tutup, saya tau pesawat lu PASTI delay, padahal masih ada waktu 30 menit, GK MUNGKIN pesawat lu tuh tepat waktu, yakin saya pasti delay dan penumpang masih bisa masuk. Fix terakhir kali keluarga saya pakai Lion Air. Udh itu kecelakaan mulu lagi, bukannya gara2 itu customer servicenya dibenerin, eh malah nambah2 alasan buat customer tuh kecewa. Saya doain aja kalian tutup | Gila Gila_Lion Lion Lion_nih nih nih_ya ya ya_udah udah udah_tau tau tau_Adek Adek Adek_saya saya saya_masih masih masih_SMA SMA SMA_sendirian sendirian sendirian_lagi lagi lagi_lagi lagi lagi_antre antre antre_buat buat buat_check check check_in in in_malah malah malah_di di di_close close close_karena karena karena_waktu waktu waktu_nya nya nya_udh udh udh_mepet mepet mepet_Eh Eh Eh_Lion Lion Lion_gk gk gk_usah usah usah_sok sok sok_an an an_tutup tutup tutup_saya saya saya_tau tau tau_pesawat pesawat pesawat_lu lu lu_PASTI PASTI PASTI_delay delay delay_padahal padahal padahal_masih masih masih_ada ada ada_waktu waktu waktu_menit menit menit_GK GK GK_MUNGKIN MUNGKIN MUNGKIN_pesawat pesawat pesawat_lu lu lu_tuh tuh tuh_tepat tepat tepat_waktu waktu waktu_yakin yakin yakin_saya saya saya_pasti pasti pasti_delay delay delay_dan dan dan_penumpang penumpang penumpang_masih masih masih_bisa bisa bisa_masuk masuk |

Terakreditasi Sinta Peringkat 3 (Sinta 3) berdasarkan SK Dirjen Penguatan RisBang Kemenristekdikti Nomor 28/E/KPT/2019, Tanggal 26 September 2019

| | |
|---|---|
| | masuk_Fix Fix Fix_terakhir terakhir terakhir_kali kali kali_keluarga keluarga keluarga_saya saya saya_pakai pakai pakai_Lion Lion Lion_Air Air Air_Udh Udh Udh_itu itu itu_kecelakaan kecelakaan kecelakaan_mulu mulu mulu_lagi lagi lagi_bukannya bukannya bukannya_gara gara gara_itu itu itu_customer customer customer_servicenya servicenya servicenya_dibenerin dibenerin dibenerin_eh eh eh_malah malah malah_nambah nambah nambah_alasan alasan alasan_buat buat buat_customer customer customer_tuh tuh tuh_kecewa kecewa kecewa_Saya Saya Saya_doain doain doain_aja aja aja_kalian kalian kalian_tutup tutup |

Sumber: (Sari & Hayuningtyas, 2019)

Pada penelitian ini pengujian dilakukan dengan model *k fold cross validation* dimana proses ini membagi data secara acak kedalam k bagian. Proses pengujian dimulai dengan pembentukan model dengan data pada bagian pertama. 1 bagian menjadi data testing, data lainnya menjadi data training. Setelah dilakukan training dan testing maka dapat di ukur akurasinya.

Pengujian metode pada algoritma *Support Vector Machine* menggunakan *fold cross validation* 1-10. Berikut adalah hasil dari eksperimen pengujian metode *Support Vector Machine* dengan *fold cross validation*.

Tabel 3. Akurasi dengan *Support Vector Machine*

| K Fold Cross Validation | Akurasi | AUC |
|---|---|---|
| 10 | 60.00% | 0.964 |
| 9 | 58.08% | 0.972 |
| 8 | 58.09% | 0.979 |
| 7 | 61.02% | 0.983 |
| 6 | 61.03% | 0.953 |
| 5 | 60.00% | 0.970 |
| 4 | 58.00% | 0.966 |
| 3 | 58.05% | 0.954 |
| 2 | 54.00% | 0.940 |

Sumber: (Sari & Hayuningtyas, 2019)

Berdasarkan table akurasi dengan *Support Vector Machine* dapat dilihat *6 Fold Cross Validation* memiliki nilai akurasi terbesar yaitu sebesar 61.03%. Untuk meningkatkan akurasi maka agloritma *Support Vector Machine* dioptimasi dengan *Particle Swarm Optimization*.

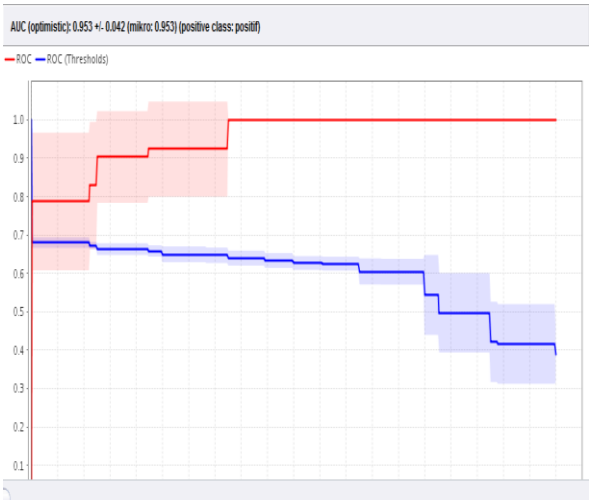
Dapat dilihat eksperimen menggunakan algoritma *Support Vector Machine* dengan menggunakan 6 *fold cross validation* didapat akurasi 61.03%. *confusion matrix* dengan menggunakan algoritma *Support Vector Machine* dapat dilihat dari table 4.

Tabel 4. *Confusion Matrix Support Vector Machine*

| **Accuracy : 61.03%+/-12.81%(mikro:74.26%)** | | |
|---|---|---|
| | **True negative** | **True positif** |
| **Pred negative** | **11** | **0** |
| **Pred positif** | **39** | **50** |

Sumber: (Sari & Hayuningtyas, 2019)

Grafik ROC dari eksperimen menggunakan algoritma *Support Vector Machine* dapat dilihat pada gambar 2.




Sumber: (Sari & Hayuningtyas, 2019)
Gambar 2. Grafik ROC *Support Vector Machine*

Pengujian metode pada algoritma *Support Vector Machine* yang dioptimasi dengan *Particle Swarm Optimization* menggunakan *fold cross validation* 1-10. Berikut adalah hasil dari eksperimen dari pengujian metode Support Vector


Terakreditasi Sinta Peringkat 3 (Sinta 3) berdasarkan SK Dirjen Penguatan RisBang Kemenristekdikti Nomor 28/E/KPT/2019, Tanggal 26 September 2019

Machine yang dioptimasi dengan *Particle Swarm Optimization* dengan populasi 5.

Tabel 5. Akurasi dengan *Support Vector Machine* + PSO

| K Fold Cross Validation | Akurasi | AUC |
|---|---|---|
| 10 | 71,00% | 0,976 |
| 9 | 69,02% | 0,972 |
| 8 | 65,87% | 0,982 |
| 7 | 70,07% | 0,983 |
| 6 | 69,91% | 0,972 |
| 5 | 64,00% | 0,984 |
| 4 | 71,00% | 0,973 |
| 3 | 64,05% | 0,969 |
| 2 | 61,00% | 0,952 |

Sumber: (Sari & Hayuningtyas, 2019)

Berdasarkan table akurasi dengan *Support Vector Machine* yang dioptimasi dengan *Particle Swarm Optimization* dapat dilihat 4 Fold Cross Validation memiliki nilai akurasi terbesar yaitu sebesar 71.00%,

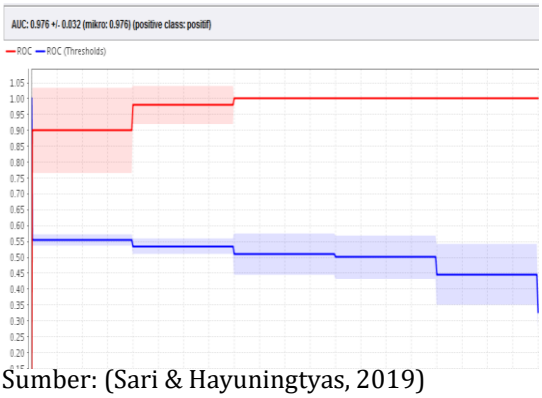
Dapat dilihat eksperimen menggunakan algoritma *Support Vector Machine* yang dioptimasi dengan *Particle Swarm Optimization* dengan menggunakan 4 fold cross validation didapat akurasi 71.00%. *confusion matrix* dengan menggunakan algoritma *Support Vector Machine* yang dioptimasi dengan *Particle Swarm Optimization* dapat dilihat dari table 6

Tabel 6. *Confusion Matrix Support Vector Machine* + PSO

| Accuracy : 71.00%+/-13.00%(mikro:71.00%) | | |
|---|---|---|
| | True negative | True positif |
| Pred negative | 21 | 0 |
| Pred positif | 29 | 50 |

Sumber: (Sari & Hayuningtyas, 2019)

Grafik ROC dari eksperimen menggunakan algoritma *Support Vector Machine* yang dioptimasi dengan *Particle Swarm Optimization* dapat dilihat pada gambar 3.




Sumber: (Sari & Hayuningtyas, 2019)

Gambar 3. Grafik ROC *Support Vector Machine* + PSO

## KESIMPULAN

Dari eksperimen pada penelitian ini review penerbangan dengan teks berbahasa menggunakan *Support Vector Machine* Indonesia menghasilkan akurasi 61.03% dengan menggunakan 6 fold cross validation. Sedangkan eksperimen dalam review Penerbangan dengan teks bahasa Indonesia menggunakan Support Vector Machine yang dioptimasi dengan Particle Swarm Optimization menghasilkan akurasi sebesar 71.00% dengan menggunakan 4 fold cross validation dan populasi 5. Dari kedua eksperimen tersebut diketahui dengan menggunakan *Particle Swarm Optimization* (PSO) pada *Support Vector Machine* meningkatkan akurasi sebesar 9,97%.

## REFERENSI

Brody, S., & Elhadah, N. (2010). An Unsupervised Aspect-Sentiment Model for Online Reviews_Slide. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL*, (June), 804–812. Retrieved from http://www.aclweb.org/anthology/N10-1122

Drajana, I. C. R. (2017). METODE SUPPORT VECTOR MACHINE DAN FORWARD SELECTION PREDIKSI PEMBAYARAN PEMBELIAN BAHAN BAKU. *ILKOM Journal Ilmiah*, *9*, 116–123.

Ling, J., N. Kencana, I. P. E., & Oka, T. B. (2014). Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier Dengan Seleksi Fitur Chi Square. *E-Jurnal Matematika*, *3*(3), 92. https://doi.org/10.24843/mtk.2014.v03.i03.p070

Monarizqa, N., Nugroho, L. E., & Hantono, B. S. (2014). Penerapan Analisis Sentimen Pada Twitter Berbahasa Indonesia Sebagai Pemberi Rating. *Jurnal Penelitian Teknik Elektro Dan Teknologi Informasi*, *1*, 151–155.

Muhamad, H., Prasojo, C. A., Sugianto, N. A., Surtiningsih, L., Cholissodin, I., Ilmu, F., … Optimization, P. S. (2017). OPTIMASI NAÏVE BAYES CLASSIFIER DENGAN MENGGUNAKAN PARTICLE. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *4*(3), 180–184.

Parapat, I. M., Furqon, M. T., & Sutrisno. (2018). Penerapan Metode Support Vector Machine ( SVM ) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2*(10), 3163–3169.

Pravina, A. M., Cholissodin, I., & Adikara, P. P. (2019). Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine ( SVM ). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, *3*(3), 2789–2797.

Rozi, I., Pramono, S., & Dahlan, E. (2012). Implementasi Opinion Mining (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik Pada Perguruan Tinggi. *Jurnal EECCIS*, *6*(1), 37–43.

Rozi, N. F., Arianto, F., & Hapsari, D. P. (2019). Analisis Sentimen Pada Opini Pengguna Maskapai Penerbangan Sentiment Analysis on Passenger Opinions At Airlines Company. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *6*(3), 321–326. https://doi.org/10.25126/jtiik.201961337

Sari, R., & Hayuningtyas, R. Y. (2019). *Penerapan Particle Swarm Optimization Terhadap Support Vector Machine Pada Review Pengguna Transportasi Udara. Laporan Akhir Penelitian Mandiri STMIK Nusa Mandiri Jakarta*. Jakarta.

Xue, B., Zhang, M., Member, S., & Browne, W. N. (2012). Particle Swarm Optimization for Feature Selection in Classification : A Multi-Objective Approach. *Ieee Transactions on Cybernetics*, *43*(6), 1–16.



Terakreditasi Sinta Peringkat 3 (Sinta 3) berdasarkan SK Dirjen Penguatan RisBang Kemenristekdikti
Nomor 28/E/KPT/2019, Tanggal 26 September 2019