

ANALISIS SENTIMEN PENGGUNA *WEBSITE* SISKAMENGGUNAKAN ALGORITMA *NAIVE BAYES* DENGAN OPTIMASI *PARTICLE SWARM OPTIMIZATION*

Khalidah Rossa Amaliah^{1*}, Ultach Enri², Sofi Defiyanti³

^{1, 2, 3} Universitas Singaperbangsa Karawang; Jl. HS. Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361, Telp. (0267) 641177

Keywords:

Analisis Sentimen;
Naïve Bayes;
Particle Swarm Optimization.

Correspondent Email:

khalidahrossaam06@gmail.com



JITET is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Abstrak. *Website Siska* merupakan sistem informasi akademik Universitas Singaperbangsa Karawang yang digunakan untuk menunjang aktivitas akademik mahasiswa. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aspek performa *Website Siska* menggunakan algoritma *Naive Bayes*, yang dioptimalkan dengan metode *Particle Swarm Optimization* (PSO). Data diperoleh dari kuesioner yang diisi oleh 384 mahasiswa aktif, kemudian dianalisis dengan pendekatan *Knowledge Discovery in Databases* (KDD) yang mencakup tahapan *data selection*, *pre-processing*, *transformation* menggunakan TF-IDF, *data mining*, dan *evaluation*. Untuk mengatasi ketidakseimbangan data, digunakan teknik SMOTE. Pembagian data yang digunakan adalah 80% untuk pelatihan dan 20% untuk pengujian. Hasil menunjukkan bahwa akurasi algoritma *Naive Bayes* pada *Website Siska* mencapai 88%, dan meningkat menjadi 90% setelah dilakukan optimasi dengan PSO. Temuan ini menunjukkan bahwa penerapan PSO dapat meningkatkan performa model klasifikasi dalam analisis sentimen, khususnya pada aspek performa sistem informasi akademik. Penelitian ini memberikan kontribusi penting dalam upaya peningkatan kualitas layanan *Website Siska* melalui pemanfaatan teknik optimasi pada analisis data opini pengguna.

Abstract. *SISKA* is the academic information system of Universitas Singaperbangsa Karawang used to support student academic activities. This study aims to analyze user sentiment regarding the performance aspect of the *SISKA* website using the *Naive Bayes* algorithm, optimized with the *Particle Swarm Optimization* (PSO) method. Data were collected through questionnaires completed by 384 active students and analyzed using the *Knowledge Discovery in Databases* (KDD) approach, which includes the stages of *data selection*, *pre-processing*, *transformation* using TF-IDF, *data mining*, and *evaluation*. To address class imbalance, the SMOTE technique was applied. The dataset was divided into 80% for training and 20% for testing. Results show that the *Naive Bayes* algorithm achieved an accuracy of 88% on the *SISKA* website, which increased to 90% after optimization with PSO. These findings indicate that applying PSO can enhance the performance of classification models in sentiment analysis, particularly regarding the performance aspect of academic information systems. This research provides valuable contributions toward improving the quality of *SISKA* services through the use of optimization techniques in analyzing user opinion data.

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah membawa dampak signifikan

terhadap berbagai bidang, termasuk dunia pendidikan. Di lingkungan perguruan tinggi, sistem informasi akademik berbasis web menjadi salah satu sarana penting dalam menunjang aktivitas akademik mahasiswa, dosen, maupun tenaga kependidikan. Sistem ini dirancang untuk memudahkan pengelolaan data akademik seperti jadwal kuliah, KRS, KHS, nilai, hingga informasi akademik lainnya secara terpusat dan *real-time*.

Website SISKA (Sistem Informasi Akademik) merupakan sistem yang digunakan oleh Universitas Singaperbangsa Karawang (UNSIKA) dalam mendukung proses akademik. Seiring dengan peningkatan jumlah pengguna dan kompleksitas fitur yang disediakan, evaluasi terhadap kualitas sistem, terutama dari sisi pengguna, menjadi hal yang krusial. Salah satu pendekatan yang dapat digunakan adalah analisis sentimen terhadap opini pengguna. Melalui analisis ini, dapat diketahui bagaimana persepsi mahasiswa terhadap kinerja dan performa sistem.

Penelitian analisis sentimen telah banyak dilakukan menggunakan algoritma *Naïve Bayes* karena keefektifannya dalam menangani data teks dan kemudahan implementasinya [1][3]. Namun, tantangan utama dalam penggunaan algoritma ini adalah bagaimana mengoptimalkan performanya terutama dalam konteks klasifikasi yang melibatkan opini beragam. Oleh karena itu, penelitian ini menggunakan teknik *Particle Swarm Optimization* (PSO) untuk meningkatkan akurasi model *Naïve Bayes* sebagaimana disarankan dalam penelitian sebelumnya [4][6].

Dalam penelitian ini, data opini mahasiswa difokuskan pada aspek performa *Website* SISKA, karena performa sistem merupakan indikator utama dalam keberhasilan implementasi sistem informasi [2]. Selain itu, penggunaan teknik SMOTE untuk mengatasi ketidakseimbangan data kelas diharapkan dapat memberikan hasil klasifikasi yang lebih akurat [9][11].

Dengan demikian, tujuan utama penelitian ini adalah untuk mengetahui bagaimana penerapan *Naïve Bayes* yang dioptimalkan dengan PSO dalam menganalisis sentimen mahasiswa terhadap performa *Website* SISKA, menggunakan skenario pembagian data latih dan uji sebesar 80:20.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen merupakan proses otomatis dalam mengidentifikasi dan mengelompokkan opini pengguna terhadap suatu topik, layanan, atau produk menjadi sentimen positif, negatif, atau netral. Teknik ini banyak digunakan untuk mengevaluasi tanggapan pengguna terhadap suatu sistem berbasis teks, seperti ulasan, komentar, atau kuesioner [5].

2.2 Naïve Bayes

Naïve Bayes adalah algoritma klasifikasi probabilistik yang berbasis pada Teorema Bayes dengan asumsi independensi antar fitur. Algoritma ini sangat populer dalam tugas-tugas klasifikasi teks seperti analisis sentimen karena kemampuannya dalam memberikan hasil yang cukup baik dengan waktu komputasi yang efisien [7].

2.3 Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) merupakan metode optimasi berbasis populasi yang terinspirasi dari perilaku sosial organisme seperti burung dan ikan. PSO digunakan dalam penelitian ini untuk mengoptimalkan parameter model *Naïve Bayes*, dengan tujuan meningkatkan akurasi klasifikasi [4].

2.4 Synthetic Minority Over-sampling Technique (SMOTE)

SMOTE (*Synthetic Minority Over-sampling Technique*) adalah teknik yang digunakan untuk mengatasi ketidakseimbangan kelas pada data. Teknik ini bekerja dengan membuat sampel sintetis dari kelas minoritas sehingga distribusi data menjadi lebih seimbang dan model klasifikasi dapat belajar lebih baik [11].

2.5 Website SISKA

Website SISKA sebagai objek penelitian merupakan sistem informasi akademik berbasis web yang digunakan di UNSIKA. Evaluasi terhadap performa sistem informasi akademik sangat penting karena berkaitan langsung dengan efektivitas dan efisiensi kegiatan akademik mahasiswa [2][8].

3. METODE PENELITIAN

Penelitian ini mengumpulkan data pengguna *Website Siska* melalui kuesioner yang dibagikan kepada mahasiswa aktif unsika. Data yang dikumpulkan merupakan opini pengguna mengenai *Website Siska*. Proses ini menerapkan metode KDD, meliputi *data selection*, *pre-processing*, *transformation*, *data mining*, dan *evaluation*, untuk mengolah data mentah menjadi informasi yang bermanfaat.

3.1. Data Selection

Data dikumpulkan melalui penyebaran kuesioner kepada mahasiswa aktif Universitas Singaperbangsa Karawang. Kuesioner difokuskan untuk menggali opini mahasiswa mengenai aspek performa *Website Siska*. Total responden yang berhasil dikumpulkan sebanyak 384 orang.

3.2. Pre-Processing

Tahap *pre-processing* merupakan langkah awal dalam analisis sentimen yang bertujuan untuk menyaring dan menyeleksi informasi relevan dari dokumen. Proses ini dirancang untuk menghapus elemen yang tidak diperlukan agar data siap untuk dianalisis lebih lanjut. Pada penelitian ini, proses *pre-processing* mencakup *data cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*.

3.3. Transformation

Pada tahap ini, data opini yang telah melalui *pre-processing* diubah menjadi representasi numerik menggunakan metode TF-IDF. Setiap dokumen teks dikonversi menjadi vektor berdasarkan frekuensi kata, kemudian diberi label sentimen positif atau negatif sesuai kategori yang telah ditentukan sebelumnya.

3.4. Data Mining

Pada tahap data mining, analisis menggunakan algoritma *Naïve Bayes* yang dioptimalkan menggunakan *Particle Swarm Optimization* (PSO). Data yang telah melalui tahap *pre-processing* dan transformasi menggunakan metode TF-IDF, serta diseimbangkan dengan teknik SMOTE, kemudian dibagi menjadi 80% *data training* dan 20% *data testing*. Proses klasifikasi dilakukan dalam dua tahap, yaitu klasifikasi awal dengan algoritma *Naïve Bayes*, kemudian

dilanjutkan dengan optimasi menggunakan PSO guna meningkatkan akurasi model.

3.5. Evaluation

Pada tahap evaluasi, kinerja model dianalisis dengan membandingkan hasil klasifikasi sentimen terhadap aspek performa *Website Siska* sebelum dan sesudah dilakukan optimasi menggunakan *Particle Swarm Optimization* (PSO). Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, dan *recall* yang diperoleh dari *confusion matrix*. Metrik-metrik ini digunakan untuk mengukur seberapa baik model dalam mengklasifikasikan data sentimen secara tepat. Hasil evaluasi pada skenario pembagian data 80% *training* dan 20% *testing* menunjukkan perbedaan performa antara model *Naïve Bayes* standar dan model yang telah dioptimasi, di mana model hasil optimasi diharapkan memiliki akurasi yang lebih tinggi. Evaluasi ini bertujuan untuk memastikan bahwa penggunaan PSO memberikan peningkatan signifikan terhadap kualitas klasifikasi sentimen yang dilakukan.

4. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengklasifikasikan sentimen pengguna *Website Siska* pada aspek performa, menggunakan algoritma *Naïve Bayes* yang dioptimalkan dengan *Particle Swarm Optimization* (PSO). Tahapan penelitian mengacu pada metode KDD yang meliputi:

4.1 Data Selection

Proses seleksi data dilakukan dengan mengumpulkan tanggapan dari 384 mahasiswa aktif Universitas Singaperbangsa Karawang melalui *google formulir*. Kuesioner tersebut dirancang untuk mengidentifikasi opini mahasiswa mengenai *Website Siska* terhadap performa *Website*. Opini yang diseleksi berkaitan dengan persepsi pengguna terhadap kecepatan akses, kestabilan sistem, serta kehandalan fungsi layanan pada *website*. Hanya tanggapan yang relevan terhadap kategori ini yang diproses lebih lanjut untuk dianalisis.

4.2 Pre-processing

Data mentah yang telah dipilih selanjutnya melalui tahapan *pre-processing* untuk menghilangkan elemen yang tidak relevan dan menyiapkan data dalam bentuk siap analisis.

Proses ini meliputi pembersihan data (*cleaning*), normalisasi huruf (*case folding*), pemisahan kata (*tokenizing*), penghapusan kata tidak penting (*stopword removal*), serta reduksi kata ke bentuk dasarnya (*stemming*). Data yang sudah bersih dapat dilanjutkan untuk tahap selanjutnya.

Performa
[optimal]
[sesuai, penuh, butuh]
[oke, trouble, bagi, login, pakai, hp]
[fungsi, performa, sistem, informasi, akademik, siska, penuh, butuh, guna]
[kalo, keliru, mudah, paham]
[fungsi, dasar, guna, fiturnya, absen]
[sesuai, fitur, baik, admin]
[fungsi, jalan, error, periode, khusus]
[bagus]
[fungsi]

Gambar 4. 1 Hasil Pre-Processing

4.3 Transformation

Setelah tahap *pre-processing*, dilakukan pelabelan data menjadi positif dan negatif, serta data teks diubah ke dalam bentuk numerik menggunakan pendekatan representasi vektor. Dalam hal ini digunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk menghasilkan bobot dari setiap kata berdasarkan tingkat kepentingannya dalam dokumen. Hasil transformasi ini berupa vektor fitur yang dapat digunakan dalam proses klasifikasi oleh algoritma Naïve Bayes. Karena saat setelah melakukan pelabelan data didapatkan bahwa data mengalami ketidakseimbangan antara positif dan negatif, maka dilakukan teknik smote untuk mengatasinya.

TF-IDF Performa:													
	absen	absensi	access	admin	administrasi	agam	ai	ajar	aju	\			
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
akademik ... virtual wadah wajah waktu war web webnya website \													
0	0.000000	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
1	0.000000	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
2	0.000000	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
3	0.430773	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
4	0.000000	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
websitenya wisuda													
0	0.0	0.0											
1	0.0	0.0											
2	0.0	0.0											
3	0.0	0.0											
4	0.0	0.0											

Gambar 4. 2 Hasil Transformation

4.4 Data Mining

Pada tahap ini, proses klasifikasi dilakukan menggunakan algoritma *Naïve Bayes* dan *Naïve Bayes* yang dioptimasi menggunakan PSO. Data yang telah melalui tahap *pre-processing* dan *transformation* dimasukkan ke dalam model untuk dipelajari. Tujuan dari tahap ini adalah untuk mengelompokkan opini mahasiswa mengenai performa Website Siska ke dalam dua kategori sentimen, yaitu positif dan negatif. Dalam penelitian ini, menggunakan pembagian data dilakukan dengan rasio 80% sebagai data latih dan 20% sebagai data uji, sesuai dengan skenario yang telah ditentukan. Hasil dari tahap ini digunakan sebagai dasar untuk membandingkan performa sebelum dan sesudah dilakukan optimasi menggunakan PSO.

Sentimen_Prediksi_Performa	
268	Negatif
250	Positif
355	Positif
332	Positif
56	Positif

Gambar 4. 3 Hasil Data Mining

4.5 Evaluation

Evaluasi dilakukan untuk mengukur seberapa baik model klasifikasi dalam mengenali sentimen pengguna terhadap aspek performa Website Siska. Metode evaluasi yang digunakan adalah *confusion matrix*, yang menghasilkan metrik-metrik penting seperti akurasi, presisi, dan recall.

Pada tahap awal, model *Naïve Bayes* dijalankan tanpa optimasi dan menghasilkan akurasi sebesar 81%. Angka ini menunjukkan bahwa model sudah cukup baik dalam mengklasifikasikan opini positif dan negatif. Namun, setelah dilakukan optimasi menggunakan algoritma *Particle Swarm Optimization* (PSO), akurasi meningkat menjadi 82%. Peningkatan ini menunjukkan bahwa PSO mampu membantu model *Naïve Bayes* dalam menghasilkan klasifikasi yang lebih akurat.

Selain akurasi, peningkatan juga terlihat pada nilai presisi dan recall, yang berarti model menjadi lebih tepat dalam memprediksi kelas sentimen dan lebih sensitif terhadap opini yang sebenarnya. Evaluasi ini menunjukkan bahwa kombinasi *Naïve Bayes* dan PSO merupakan

pendekatan yang efektif untuk analisis sentimen dalam konteks sistem informasi akademik.

```

=== Evaluasi Sentimen Performa (Rasio 80:20) ===
Akurasi: 0.81
Presisi: 0.89
Recall: 0.81
    
```

Gambar 4. 4 Evaluasi Model *Naïve Bayes*

```

=== Evaluasi Model Performa 80:20 ===
Accuracy : 0.82
Precision : 0.83
Recall   : 0.82
    
```

Gambar 4. 5 Evaluasi Model *Naïve Bayes* Setelah Dioptimasi Dengan PSO

5. KESIMPULAN

- a. Algoritma *Naïve Bayes* terbukti mampu melakukan klasifikasi sentimen terhadap opini mahasiswa mengenai aspek performa Website SISKAN dengan hasil akurasi awal sebesar 81%. Setelah dilakukan optimasi menggunakan algoritma Particle Swarm Optimization (PSO), akurasi meningkat menjadi 82%. Hal ini menunjukkan bahwa PSO efektif dalam meningkatkan kemampuan model *Naïve Bayes* dalam mengenali pola sentimen dari data teks.
- b. Pendekatan analisis sentimen berbasis teks yang digunakan dalam penelitian ini, dengan mengombinasikan *Naïve Bayes* dan PSO, memberikan hasil yang cukup baik dan dapat dijadikan sebagai metode evaluasi untuk menilai kualitas sistem informasi akademik. Hasil klasifikasi sentimen ini dapat menjadi acuan bagi pengelola Website SISKAN dalam memahami persepsi pengguna serta menentukan langkah perbaikan yang lebih tepat pada aspek performa sistem.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] Jatmiko, H. B., Kurniadi, N. T., & Maulana, D. (2022). Optimasi *Naïve Bayes* Dengan Particle Swarm Optimization Untuk Analisis Sentimen Formula E-Jakarta. *Jurnal*

Informatika dan Teknik Elektro Terapan, 10(1).

- [2] Shahputri, I., & Yamasari, A. (2024). Analisis Sentimen pada Sistem Akademik. *Jurnal Teknologi Informasi*, 11(2).
- [3] Ramadhani, A., & Suryono, R. R. (2024). Klasifikasi Teks Menggunakan *Naïve Bayes*. *Jurnal Sistem Informasi*, 13(1).
- [4] Hasibuan, M., & Fransisca, R. (2022). Penerapan PSO pada Model *Naïve Bayes* untuk Deteksi Stroke. *Jurnal Kesehatan dan Informatika*, 8(2).
- [5] Firdaus, A., & Istalama, F. (2021). Text Mining dalam Analisis Sentimen. *Jurnal Komputasi dan Informatika*, 7(3).
- [6] Risawati, D., Ernawati, R., & Maryani, T. (2020). Particle Swarm Optimization: Teori dan Aplikasi. *Jurnal Algoritma dan Pemrograman*, 9(1).
- [7] Rahayu, T., Fauzi, A., & Rahmat, H. (2022). Analisis Klasifikasi dengan *Naïve Bayes*. *Jurnal Teknologi dan Sistem Informasi*, 10(2).
- [8] Amrullah, A., Anas, S., & Hidayat, R. (2020). Proses Pre-processing dalam Analisis Teks. *Jurnal Informatika dan Data Mining*, 6(4).
- [9] Pramayasa, D., Maysanjaya, R., & Indradewi, A. (2023). Penerapan SMOTE untuk Ketidakseimbangan Kelas. *Jurnal Sains Komputer*, 5(2).
- [10] Yang, L., & Sihotang, H. (2022). Perhitungan Jumlah Sampel dengan Rumus Lemeshow. *Jurnal Statistika dan Riset Operasi*, 8(1).
- [11] Sajiwo, R., Rahmat, T., & Junaidi, F. (2024). SMOTE dalam Data Mining. *Jurnal Ilmu Komputer Terapan*, 7(1).
- [12] Putri, D.D., Nama, G.F., & Sulistiono, W.E. (2022). Analisis Sentimen Kinerja DPR pada Twitter Menggunakan *Naïve Bayes Classifier*. *Jurnal Informatika dan Teknik Elektro Terapan*, 10(1).
- [13] Nugroho, D., & Mulyanto, R. (2021). Implementasi *Naïve Bayes* untuk Klasifikasi Sentimen Ulasan Produk Online. *Jurnal Teknologi Informasi dan Komunikasi*, 10(2).
- [14] Lestari, I., & Pratama, H. (2023). Penerapan PSO untuk Optimasi Parameter dalam Model Klasifikasi Teks. *Jurnal Sains Data dan Algoritma*, 5(1).
- [15] Suryani, D., & Handoko, B. (2022). Evaluasi Kinerja Algoritma Klasifikasi pada Data Ketidakseimbangan Menggunakan SMOTE. *Jurnal Ilmu Komputer dan Rekayasa Sistem*, 6(2).