

JUTIM (Jurnal Teknik Informatika Musirawas) is published twice a year, in June and December.

JUTIM (Jurnal Teknik Informatika Musirawas) published by LPPM Universitas Bina Insan, Lubuklinggau, addressed on street Jendral Besar H.M Soeharto Kelurahan Lubuk Kupang Kecamatan Lubuklinggau Selatan I, Kota Lubuklinggau, Provinsi Sumatera Selatan, Indonesia, Tel / fax : 0733 / 3280300, email lppm@univbinainsan.ac.id

Profile Google Scholar JUTIM (Jurnal Teknik Informatika Musirawas) [in here](#)

JUTIM (Jurnal Teknik Informatika Musirawas) telah memiliki ISSN Cetak dan ISSN Online, [ISSN Cetak = 2541-1888](#) [ISSN Online = 2614-5782](#)



PKP|INDEX

JUTIM (Jurnal Teknik Informatika Musirawas)
Jalan Jendral Besar H.M Soeharto Kelurahan Lubuk Kupang Kecamatan Lubuklinggau
Barat I Kota Lubuklinggau Provinsi Sumatera Selatan
email : lppm@univbinainsan.ac.id
<http://jurnal.univbinainsan.ac.id/index.php/jutim>

Principal Contact

Susanto
Editor In Chief
Universitas Bina Insan Lubuklinggau
Phone 081996247882
santo.calem@gmail.com

Support Contact

Fido Rizki
Phone 082179654408
fidorizki@gmail.com

DAFTAR ISI

DOI: <https://doi.org/10.32767/jutim.v6i1>

Article

IMPLEMENTASI METODE FUZZY INFERENCE SYSTEM UNTUK REKOMENDASI SISWA KELAS UNGGUL PADA SMA NEGERI 2 KEPAHIANG

Juju Jumadi, Devi Sartika

1-9 | published: 2021-05-26

MENGOPTIMALKAN METODE USABILITY TESTING DENGAN PENAMBAHAN ATRIBUT COMFORTABLY

Eny Jumiati

10-20 | published: 2021-05-25

TEXT MINING MENGGUNAKAN NAIVE BAYES BERBASIS PARTICLE SWARM OPTIMIZATION UNTUK SENTIMENT RESTAURANT

Rizki Aulianita, Achmad Maezar Bayu Aji, Yuni Eka Achyani

21-29 | published: 2021-06-19

IMPLEMENTASI METODE HYBRID AHP-TOPSIS PADA PEMERINGKATAN NEGARA ASEAN DALAM PENANGANAN PANDEMI COVID-19

Aditya Wafda Nahari, Setyawan Wibisono

30-41 | published: 2021-05-26

PENERAPAN FITUR SELEKSI FORWARD SELECTION UNTUK MENENTUKAN KEMATIAN AKIBAT GAGAL JANTUNG MENGGUNAKAN ALGORITMA C4.5

Elin Nurlia, Ultach Enri

42-50 | published: 2021-05-26

PERANCANGAN PERANGKAT LUNAK SISTEM PEMESANAN
PADA PELANGI CAKE

Dedi Haryanto, Zulhipni Reno Saputra Elsi

51-60 | published: 2021-05-25

PERANCANGAN APLIKASI E-CATALOG PENJUALAN FIF GROUP
CABANG LUBUKLINGGAU

Zulfauzi Zulfauzi

61-71 | published: 2021-06-28

SISTEM AKUNTABILITAS KINERJA INSTANSI PEMERINTAH
(SAKIP) BERBASIS WEB MOBILE

Budi Santoso, Satrianansyah Satrianansyah

72-80 | published: 2021-05-26

TEXT MINING MENGGUNAKAN NAIVE BAYES BERBASIS PARTICLE SWARM OPTIMIZATION UNTUK SENTIMENT RESTAURANT

Rizki Aulianita^{1*}, Achmad Maezar Bayu Aji², Yuni Eka Achyani³

^{1,2,3} Fakultas Teknologi Informasi Program Studi Sistem Informasi, Universitas Nusa Mandiri,
e-mail: rizki.rzk@nusamandiri.ac.id¹ achmad.azb@nusamandiri.ac.id²,
yuni.yea@nusamandiri.ac.id³

Abstrak

Berbagai website dengan hasil review dan ulasan memudahkan kita dalam menentukan sebuah keputusan. Namun keputusan tersebut belumlah maksimal dan akurat. Seperti review makanan di yelp.com. Pengguna cukup banyak melihat review pada web tersebut sebelum memutuskan untuk memesan makanan. Permasalahan disini adalah jika hasil review terbukti kurang objektif, maka hasil keputusan menjadi tidak akurat. Tujuan dari penelitian ini untuk membantu para pengguna review untuk menghasilkan sebuah keputusan yang optimal. Naïve Bayes terbukti sebagai salah satu metode klasifikasi text yang menghasilkan akurasi tinggi. Sedangkan Particle Swarm Optimization dikenal sebagai algoritma optimasi yang baik untuk penyelesaian masalah berdasarkan parameter proses yang ada. Pada penelitian ini akan digunakan metode Naïve Bayes yang dilakukan ujicoba menggunakan Particle Swarm Optimization untuk pembobotan sehingga hasil akurasinya lebih tinggi. Berdasarkan hasil pengolahan data maka dihasilkan nilai akurasi Naïve Bayes sebesar 81.00% dan 83.80% adalah hasil pengolahan akurasi untuk Naïve Bayes Berbasis Particle Swarm Optimization (PSO). Kesimpulan pada percobaan metode Naïve Bayes tersebut yaitu bahwa PSO dapat meningkatkan nilai optimasi dari sebuah algoritma Naïve Bayes sehingga mampu diterapkan sebagai solusi untuk pemecahan masalah di atas.

Kata kunci : Text mining; Naïve Bayes; Particle Swarm Optimization

Abstract

Various websites with the results of reviews and reviews make it easier for us to make decisions. However, this decision has not been optimal and accurate. Like the food reviews on yelp.com. Users see quite a number of reviews on the web before deciding to order food. The problem here is that if the results of the review are proven to be less objective, the results of the decisions will be inaccurate. The purpose of this study is to help user reviews to produce optimal decisions. This study will use the Naïve Bayes method which will be tested using Particle Swarm Optimization for weighting so that the accuracy results are higher. Based on the results of data processing, the Nave Bayes accuracy value is 81.00% and 83.80% is the result of the accuracy of Nave Bayes processing based on Particle Swarm Optimization (PSO). The conclusion of the Nave Bayes method experiment is that PSO can increase the optimization value of a Nave Bayes algorithm so that it can be applied as a solution to the above problem solving.

Keywords : Text mining; Naïve Bayes; Particle Swarm Optimization

I. PENDAHULUAN

Text mining merupakan penemuan pada scope rumpun ilmu computer science yang berfungsi untuk mengekstrak informasi berupa tulisan dari sumber yang berbeda[1].

Kegunaan text mining salah satunya yaitu dapat memberikan alternative keputusan bagi penggunaannya. Berbagai website penyedia ulasan atau review seperti amazon.com, sociolla.com,traveloka.com dan yelp.com. Masing-masing website tersebut dapat memberikan review sehingga para user dapat memanfaatkan review tersebut dalam mengambil keputusan.

Analisis sentiment banyak digunakan untuk meneliti sebuah produk tertentu, seperti twitter sentiment analisis [1]media social [2], mobile user [3]dan sentiment restaurant [4]. Hasil dari analisa tersebut berguna bagi para marketing atau pengguna lain dalam menjaring pelanggan potensialnya, retail, produsen produk dalam upaya memahami opini publik dari pelanggan serta membantu dalam pengambilan keputusan

Dengan adanya text mining, penambahan text tersebut dapat dilakukan dengan mudah, akurat dan efisien. Salah satu metode yang paling banyak digunakan untuk text mining yaitu Naïve Bayes.

Naïve Bayes biasa digunakan untuk text kategorisasi, dengan skala kategori yang tinggi. Naïve Bayes memberikan saran text klasifikasi terbaik dari berbagai metode yang ada untuk meningkatkan akurasi [5]

Pada penelitian ini dikategorikan sentiment text positive dan sentiment text negative, menggunakan supervised learning dengan metode Naïve Bayes dilakukan uji coba untuk dapat menghasilkan klasifikasi dengan akurasi tinggi.

Review restoran yang yang paling banyak dijadikan sumber rujukan yaitu

www.yelp.com. Website tersebut dikunjungi oleh 83 juta pengunjung per bulannya melalui perangkat mobile di tahun 2015. (about yelp.com, 2015). Frekuensi pengguna dan pengaruh yang dirasakan yelp.com terhadap review restoran memberikan dampak positif pada masyarakat [6]. Dengan adanya review restoran ini, juga akan memudahkan para *customer* sebelum datang ke restoran tujuan, sehingga dapat lebih mengetahui mengenai restoran yang ingin dituju. Untuk itu penulis ingin melakukan klasifikasi teks dari review yang diperoleh melalui situs online di atas.

Naïve Bayes dikenal sebagai teknik yang baik dalam menghasilkan probabilitas klasifikasi dan telah digunakan pada banyak aplikasi. Naïve Bayes juga telah terbukti sebagai metode dengan algoritma yang efisien dan mudah dalam klasifikasi text [7].

Keuntungan Naïve Bayes yaitu teknik yang sangat intuisi, sederhana untuk diaplikasikan dengan hasil dari variasi yang lebar sehingga menghasilkan kecepatan dan akurasi tinggi [8]

Particle Swarm Optimization diinspirasi dari perilaku kolektif hewan social, yang dapat digunakan untuk menyelesaikan masalah optimasi dengan partikel yang bergerombol. [9]

II. TINJAUAN PUSTAKA

Dalam penelitian ini digunakan beberapa teori ilmiah dan kajian Pustaka terdahulu yang relevan sesuai dengan pembahasan peneliti. Di bawah ini merupakan teori terkait, yang digunakan dalam menunjang penelitian ini yaitu:

2.1 Text Mining

Penambahan text yang dieksploitasi untuk mendapatkan pengetahuan dalam jumlah besar [10]. Informasi dapat diekstraksi secara otomatis dari berbagai

literatur yang tujuannya adalah mendapatkan sebuah hipotesa baru untuk dilakukan eksperimen secara konvensional. [11] Dalam text mining terdapat proses spesifik yaitu pre-processing dan algoritma yang spesifik [12]

2.2 Naïve Bayes

Metode memiliki Teknik yang baik dalam penyelesaian masalah. Berasal dari teorema bayes yang berarti semua instance dan label dianggap dapat berdiri sendiri sehingga bisa menghasilkan sebuah nilai dari class yang ada [13]. Naïve bayes sudah banyak diujicoba untuk memecahkan masalah big data klasifikasi text dengan menghasilkan nilai skala akurasi tinggi [14]

2.3 Particle Swarm Optimization

Dalam PSO, setiap individu diperlakukan sebagai partikel, yang merupakan solusi dalam masalah ini. Partikel didefinisikan sebagai titik point $X_i(X_{i1}, X_{i2}, \dots, X_{id}, \dots, X_{iD})$ dalam ruang dimensi. Disini ruang dimensi hanya sebagai lingkungan pencarian. Sementara itu setiap partikel memiliki kecepatan sendiri seperti $V_i(V_{i1}, V_{i2}, \dots, V_{id}, \dots, V_{iD})$. Dengan demikian memungkinkan partikel untuk bergerak dalam pencarian ruang. Dari penjelasan di atas, *array solution* diinisialisasi secara acak dan kemudian memperbaharui posisi kecepatan mereka masing-masing [15].

Penelitian terbaru menunjukkan bahwa PSO memiliki 2 teknik untuk seleksi fitur sehingga akan menyeleksi fitur yang informative kemudian sisanya dipangkas menggunakan PSO untuk menghasilkan final subset, penelitian ini menggunakan Reuter-21578 dataset. Secara substansi hasilnya PSO ini dapat mengurangi dimensi fitur dan meningkatkan akurasi untuk klasifikasi teks [16].

2.4 Penelitian Terdahulu

Pada penelitian terdahulu, digunakan objek portal berita, tentang tentang sentimen terhadap sikap dan tindakan pejabat negara dalam menjalankan tugasnya menggunakan metode Naive Bayes dan PSO. Hasil akurasi 76% sehingga metode tersebut mampu memberikan sebuah solusi dalam mengklasifikasikan opini publik menjadi berita penyelenggara Negara dapat lebih akurat dan optimal. [17]

Penelitian terdahulu berikutnya yaitu menggunakan sumber text berupa dokumen yang disinyalir memiliki hoax sehingga diperlukan klasifikasi hoax. Penelitian tersebut difokuskan untuk meningkatkan akurasi menggunakan Naïve Bayes dan dioptimasi oleh PSO dalam klasifikasi text. Hasil penelitian tersebut bahwa penggunaan PSO mampu meningkatkan nilai akurasi pada Naïve Bayes untuk klasifikasi text hoax [18]

Penelitian yang digunakan sebagai bahan rujukan berikutnya sentiment analisis review hotel, pengujian algoritmanya menggunakan Naïve Bayes dan PSO. Hasil akurasi menunjukkan bahwa PSO dapat meningkatkan hasil akurasi Naïve Bayes untuk klasifikasi teks.

III. METODOLOGI PENELITIAN

Sumber text yang digunakan penulis sebagai rujukan berasal dari yelp.com. Data yang digunakan untuk penelitian ini berjumlah 500 data yang terdiri dari dua bagian yaitu 250 sentiment positif dan 250 sentiment negatif.

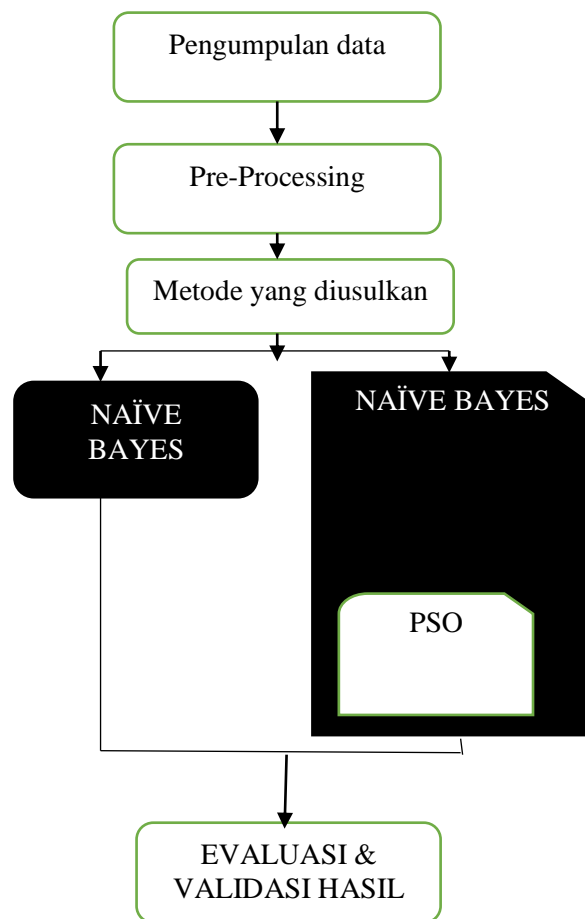
Data yang sudah diambil kemudian dilakukan pengujian dengan metode yang akan digunakan. Berikut adalah rinciannya:

1. Pengumpulan data

Data dikumpulkan dalam bentuk .txt kemudian text yang ada disimpan dalam dua kategori klasifikasi yaitu sentimen positif dan sentiment negatif.

2. Pre processing
Proses ini dilakukan untuk menghilangkan noise yang ada. dengan cara tokenize, stopword removal dan stemming.
3. Pengujian data
Setelah data selesai dilakukan pre-processing, Langkah berikutnya yaitu pengujian data.
Data diuji dengan menggunakan metode yang diusulkan yaitu Naïve Bayes
4. Pembobotan
Pada bagian ini dilakukan pemberian Particle Swarm Optimization sebagai pembobotan untuk meninggikan nilai akurasi yang dihasilkan oleh metode Naïve Bayes. Hasil akurasi dari Naïve Bayes dan Naïve bayes berbasis PSO kemudian dibandingkan. Eksperimen dan pengujian data dilakukan dengan RapidMiner
5. Evaluasi dan Validasi Hasil Pengujian Data
Terakhir, validasi digunakan untuk melihat perbandingan dari model yang awal Naïve bayes kemudian dibandingkan dengan model yang diusulkan yaitu Naïve Bayes berbasis PSO. *Cross Validation* digunakan dalam Teknik pada validasi ini. *Confusion Matrix* digunakan saat menghitung nilai akurasi pada pengujian NB dan NB+PSO. Curve ROC merupakan gambaran visual dari hasil tampilan akurasi yang dihitung menggunakan metode yang ada.

Berikut adalah diagram dari metode yang digunakan pada penelitian ini:



Gambar 1. Kerangka Penelitian

IV. HASIL DAN PEMBAHASAN

Data training yang digunakan pada saat pengujian data diambil dari *restaurant.com* dan *yelp.com*. Pengujian data, dilakukan dengan menggunakan *review restaurant* (500 *data training*, yang terdiri dari 250 review negatif dan 250 review positif) kemudian dilakukan *testing* dan *training dataset* sehingga didapatkan *accuracy* dan AUC. Berikut akan dijelaskan lebih rinci mengenai hasil penelitian yang diperoleh.

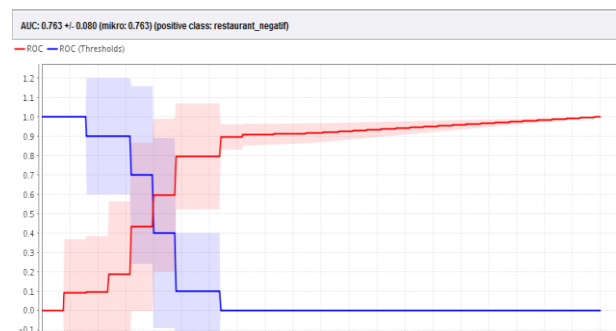
Berikut merupakan tahapan-tahapan dalam melakukan pengolahan data:

1. Pengumpulan Data
Review restaurant, masing-masing dikelompokkan dengan cara disimpan

- ke dalam satu folder dengan format .txt dengan label negatif dan positif.
2. Pengolahan Data Awal (Preprocessing)
Berikut merupakan tahapan yang dilakukan dalam preprocessing:
 - a. Tokenization
Dalam proses tokenization ini, semua kata yang ada di dalam tiap dokumen dikumpulkan dan dihilangkan tanda baca, serta dihilangkan jika terdapat simbol, karakter khusus atau apapun yang bukan huruf.
 - b. Stopwords Removal
Dalam proses ini, kata-kata yang tidak relevan akan dihapus, seperti kata the, on, of, to, for, was, will, with, yang merupakan kata-kata yang tidak mempunyai makna tersendiri jika dipisahkan dengan kata yang lain dan tidak terkait dengan kata sifat yang berhubungan dengan sentimen.
 - c. Stemming
Dalam proses ini kata-kata akan dikelompokkan ke dalam beberapa kelompok yang memiliki kata dasar yang sama, seperti awesome menjadi awesom, unique menjadi uniku.
 3. Metode yang diusulkan NB+PSO
Indikator yang diujikan pada metode Naive Bayes adalah Laplace Correction. Pengujian data dengan menggunakan Metode Naive Bayes menghasilkan Accuracy sebesar 81.00% dan AUC sebesar 0.763. Hasil pengujian data dengan Metode Naive Bayes berbasis PSO menghasilkan Accuracy=83.80% dan AUC sebesar 0.784. Hal ini menunjukkan bahwa dengan menggunakan optimasi Particle Swarm Optimization dapat meningkatkan akurasi yang lebih baik.

4. Analisis Pengujian Naïve Bayes

Pada pengujian ini, data yang digunakan adalah data bersih yang telah melalui preprocessing. Process document from files yaitu mengkonversi files menjadi dokumen, kemudian proses selanjutnya set role dengan menentukan field yang akan dijadikan sebagai kelas dan yang terakhir adalah proses validasi. Proses Validasi terdiri dari data training dan data testing. Berikut akan dijabarkan perhitungan pada tingkat accuracy, class precision dan class recall. Hasil dari pengujian model yang dilakukan adalah mengklasifikasikan review restaurant negatif dan review restaurant positif menggunakan Naive Bayes (NB) dan Naive Bayes (NB) Particle Swarm Optimization (PSO)



Gambar 2. ROC Naïve Bayes Algoritm
Sumber: Hasil Pengolahan Data (2020)

Akurasi yang diperoleh yaitu 81.00% dari 250 data review positif restoran dan 250 data review negatif restoran. Data review yang dihasilkan RapidMiner dengan model Naive Bayes menunjukkan bahwa klasifikasi untuk review positif yang sesuai prediksi yaitu **203**. Data review negatif yang termasuk ke dalam prediksi positif yaitu **48**. Dara review positif yang termasuk ke dalam prediksi negatif yaitu **47** dan data review negatif yang termasuk ke dalam prediksi negatif yaitu **202**.

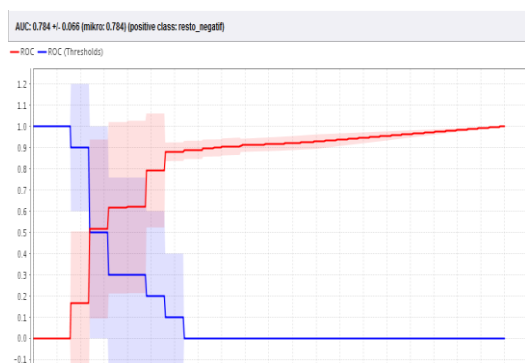
Tabel 1. Confusion Matrix Model Naïve Bayes

Accuracy : 81.00% +/- 6.15% (mikro: 81.00%)			
	True Positif	True Negatif	Class Precision
Prediksi Positif	203	48	80.88%
Prediksi Negatif	47	202	81.12%
Class Recall	81.20%	80.80%	

Sumber: Hasil pengolahan data (2020)

5. Analisis Pengujian NB Optimasi PSO

Hasil pengujian data training metode Naive Bayes berbasis Particle Swarm Optimization menggunakan *Set Role* yang berfungsi untuk menentukan field pada kelas kemudian diberikan optimasi menggunakan *Particle Swarm Optimization* agar akurasi yang dihasilkan lebih tinggi. Pengukuran akurasi tersebut, akan dijabarkan melalui Kurva ROC dan Confusion Matrix di bawah ini:



Gambar 3. ROC NB+PSO

Sumber: Hasil pengolahan data (2020)

Kurva ROC yang dihasilkan berdasarkan pengujian data pada gambar di atas, menunjukkan bahwa ada peningkatan pada **akurasi** menggunakan Naive Bayes berbasis Particle Swarm Optimization sebesar **83.80%** dan AUC sebesar **0.784**

Data training yang digunakan terdiri dari **250** data review positif

mengenai review restoran dan **250** data review negatif pada review restoran. Data review negatif, setelah melalui beberapa tahap pengolahan pada RapidMainer dengan model Naive Bayes berbasis Particle Swarm Optimization, diklasifikasikan untuk review positif yang sesuai prediksi sebanyak **211** data, kemudian **42** data yang diprediksi positif namun masuk kedalam kategori review negatif. Sedangkan untuk data review positif, yang diprediksi bahwa data tersebut negatif adalah **39**, dan untuk prediksi review negatif yang masuk dalam prediksi review negatif adalah **208** data, hasil akurasi yang muncul adalah **83.80%**.

Tabel 2. Confusion Matrix Model Naïve Bayes + PSO

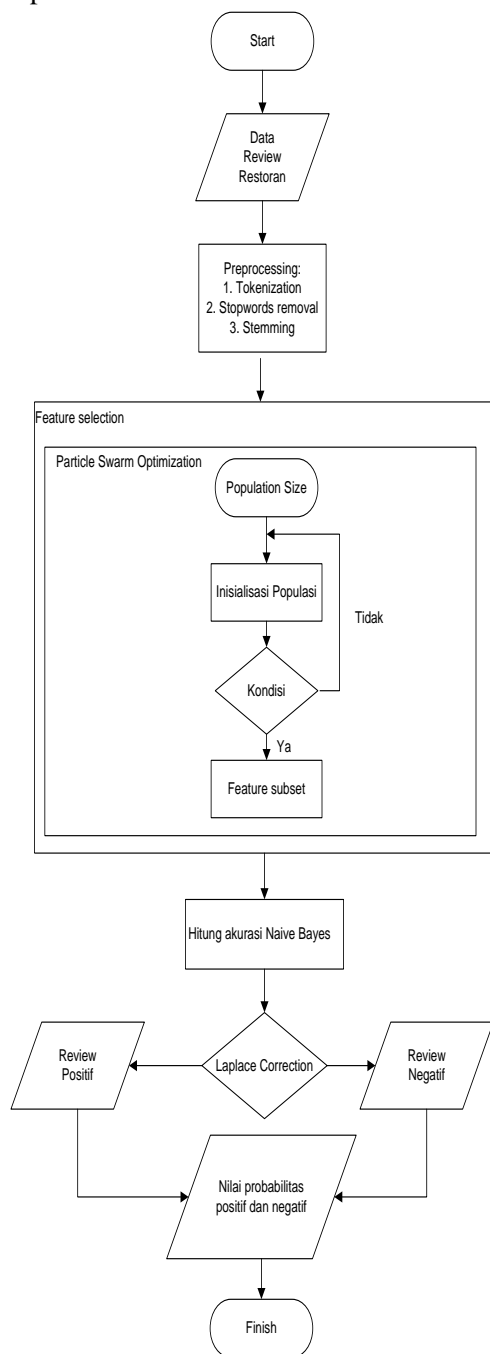
Accuracy : 83.80% +/- 3.84% (mikro: 83.80%)			
	True Positif	True Negatif	Class Precision
Prediksi Positif	211	42	83.40%
Prediksi Negatif	39	208	84.21%
Class Recall	84.40%	83.20%	

Sumber: Hasil pengolahan data (2020)

Dalam penelitian ini, hasil perhitungan metode Naive Bayes (NB) memiliki Accuracy sebesar 81.00% dan AUC sebesar 0.763 sedangkan Metode Naive Bayes berbasis Particle Swarm Optimization (NB PSO) menghasilkan Accuracy sebesar 83.80%.

Hal ini menunjukkan bahwa penggunaan optimasi Particle Swarm Optimization dapat meningkatkan nilai akurasi pada Naive Bayes untuk text mining sentiment analisis restaurant.

Berikut adalah diagram alir Tahapan Proses Klasifikasi Algoritma Naive Bayes berbasis Particle Swarm Optimization



Gambar 4. Diagram Alir Tahapan NB + PSO
Sumber: Hasil pengolahan data (2020)

Gambar di atas menjelaskan diagram alir tahapan pengujian metode Naive Bayes + Particle Swarm Optimization. Dimulai dari pengumpulan data review restoring berupa text review sebanyak 500 data, yang terdiri dari 250 data positif dan 250 data negatif, kemudian data tersebut diproses dengan menggunakan pre-processing yaitu tokenization, stopwords removal dan stemming. Kemudian hasil dari pre-processing tersebut dimasukan ke dalam data training menggunakan metode Naive Bayes, dengan menggunakan Laplace Corection. Laplace correction disini berfungsi untuk menghilangkan nilai probabilitas 0. Sehingga hasilnya non negative. Hasil dari pengolahan data menggunakan Naive Bayes tersebut berupa Probabilitas review positif dan probabilitas review negative. Pada saat data dimasukan ke dalam data training Naive Bayes, maka dilakukan pembobotan atau pemberian nilai optimasi dengan teknik Particle Swarm Optimization (PSO). Pemberian PSO ini dimaksudkan agar nilai akurasi yang dihasilkan oleh metode Naive Bayes menjadi lebih tinggi. PSO melakukan Feature selection, sesuai dengan sifat yang dimiliki oleh metode tersebut sehingga berguna untuk meningkatkan kinerja performa atau akurasi atau mengurangi proses yang dihasilkan sehingga lebih cepat dan hasil akurasi lebih tinggi.

V. KESIMPULAN

Berdasarkan penjelasan pada pengolahan data di atas. Memiliki hasil akurasi pada algoritma Naive Bayes sebesar 81.00% dan AUC sebesar 0.763, kemudian dibandingkan dengan Naive Bayes berbasis Particle Swarm Optimmmization dengan akurasi 83.80% dan AUC sebesar 0.784. Model di Naive Bayes berbasis Particle Swarm Optimization dapat memberikan

solusi terhadap permasalahan klasifikasi review restoran agar lebih akurat dan optimal. Particle Swarm optimization mampu menaikkan tingkat akurasi pada method Naïve bayes untuk klasifikasi teks.

VI. SARAN

Agar penelitian ini bisa ditingkatkan, berikut adalah saran-saran yang diusulkan:

- 1 Menggunakan metode pemilihan fitur yang lain, seperti Chi Square, Gini Index, Mutual Information, dan lain-lain agar hasilnya bisa dibandingkan
- 2 Menggunakan pengklasifikasi lain yang mungkin di luar Supervised learning. Sehingga bisa dilakukan penelitian yang berbeda dari umumnya yang sudah ada.
- 3 Menggunakan data review dari domain yang berbeda, misalnya review kamera, review film, review saham dan lain sebagainya

VII. DAFTAR PUSTAKA

- [1] T. Kwartler, "What is Text Mining?," *Text Min. Pract. with R*, pp. 1–15, 2017, doi: 10.1002/9781119282105.ch1.
- [2] R. Feldman and J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. United States Of America: Cambridge University Press, 2007.
- [3] H. Hashimi, A. Hafez, and H. Mathkour, "Selection criteria for text mining approaches," *Comput. Human Behav.*, vol. 51, pp. 729–733, Oct. 2015, doi: 10.1016/j.chb.2014.10.062.
- [4] M. Ghiassi, J. Skinner, and D. Zimbra, "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network," *Expert Syst. Appl.*, vol. 40, no. 16, pp. 6266–6282, Nov. 2013, doi: 10.1016/j.eswa.2013.05.057.
- [5] I. Habernal, T. Ptáček, and J. Steinberger, "Reprint of 'Supervised sentiment analysis in Czech social media,'" *Inf. Process. Manag.*, vol. 51, no. 4, pp. 532–546, Jul. 2015, doi: 10.1016/j.ipm.2015.05.006.
- [6] L. Zhang, K. Hua, H. Wang, G. Qian, and L. Zhang, "Sentiment Analysis on Reviews of Mobile Users," *Procedia Comput. Sci.*, vol. 34, pp. 458–465, 2014, doi: 10.1016/j.procs.2014.07.013.
- [7] H. Kang, S. J. Yoo, and D. Han, "Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews," *Expert Syst. Appl.*, vol. 39, no. 5, pp. 6000–6010, Apr. 2012, doi: 10.1016/j.eswa.2011.11.107.
- [8] A. Bagheri, M. Saraee, and F. de Jong, "Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews," *Knowledge-Based Syst.*, vol. 52, pp. 201–213, Nov. 2013, doi: 10.1016/j.knosys.2013.08.011.
- [9] A. Hicks, S. Comp, J. Horovitz, M. Hovarter, M. Miki, and J. L. Bevan, "Why people use Yelp.com: An exploration of uses and gratifications," *Comput. Human Behav.*, vol. 28, no. 6, pp. 2274–2279, Nov. 2012, doi: 10.1016/j.chb.2012.06.034.
- [10] S. Tan, "Neighbor-weighted K-nearest neighbor for unbalanced text corpus," *Expert Syst. Appl.*, vol. 28, no. 4, pp. 667–671, May 2005, doi: 10.1016/j.eswa.2004.12.023.
- [11] Z. Yao and C. Zhi-Min, "An Optimized NBC Approach in Text Classification," *Phys. Procedia*, vol. 24, pp. 1910–1914, 2012, doi: 10.1016/j.phpro.2012.02.281.
- [12] D. Farid, L. Zhang, C. Mofizur, M. A. Hossain, and R. Strachan, "Expert Systems with Applications Hybrid decision tree and naïve Bayes classifiers for multi-class

-
- classification tasks,” *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1937–1946, 2014, doi: 10.1016/j.eswa.2013.08.089.
- [13] H. Alshalabi, S. Tiun, N. Omar, and M. Albared, “Experiments on the Use of Feature Selection and Machine Learning Methods in Automatic Malay Text Categorization,” *Procedia Technol.*, vol. 11, no. Iccci, pp. 748–754, 2013, doi: 10.1016/j.protcy.2013.12.254.
- [14] Y. Zhang, S. Wang, P. Phillips, and G. Ji, “Binary PSO with mutation operator for feature selection using decision tree applied to spam detection,” *Knowledge-Based Syst.*, vol. 64, pp. 22–31, Jul. 2014, doi: 10.1016/j.knosys.2014.03.015.
- [15] J. Xiao, C. He, X. Jiang, and D. Liu, “A dynamic classifier ensemble selection approach for noise data,” *Inf. Sci. (Ny).*, vol. 180, no. 18, pp. 3402–3421, Sep. 2010, doi: 10.1016/j.ins.2010.05.021.
- [16] C. Kim, H. Li, S.-Y. Shin, and K.-B. Hwang, “An Efficient and Effective Wrapper based on Paired t-test for Learning Naive Bayes Classifiers from Large-scale Domains,” *Procedia Comput. Sci.*, vol. 23, pp. 102–112, 2013, doi: 10.1016/j.procs.2013.10.014.
- [17] G. Suresh kumar and G. Zayaraz, “Concept relation extraction using Naive Bayes classifier for ontology-based question answering systems,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 27, no. 1, pp. 13–24, Jan. 2015, doi: 10.1016/j.jksuci.2014.03.001.
- [18] P. Bermejo, J. a. Gámez, and J. M. Puerta, “Speeding up incremental wrapper feature subset selection with Naive Bayes classifier,” *Knowledge-Based Syst.*, vol. 55, pp. 140–147, Jan. 2014, doi: 10.1016/j.knosys.2013.10.016.
- [19] R. Liu, Y. Chen, L. Jiao, and Y. Li, “A particle swarm optimization based simultaneous learning framework for clustering and classification,” *Pattern Recognit.*, vol. 47, no. 6, pp. 2143–2152, 2014, doi: 10.1016/j.patcog.2013.12.010.
- [20] X. Bai, X. Gao, and B. Xue, “Particle Swarm Optimization Based Two-Stage Feature Selection in Text Mining,” *2018 IEEE Congr. Evol. Comput. CEC 2018 - Proc.*, pp. 1–8, 2018, doi: 10.1109/CEC.2018.8477773.
- [21] A. Idrus, H. Brawijaya, and Maruloh, “Sentiment Analysis of State Officials News on Online Media Based on Public Opinion Using Naive Bayes Classifier Algorithm and Particle Swarm Optimization,” *2018 6th Int. Conf. Cyber IT Serv. Manag. CITSM 2018*, no. Citsm, pp. 1–7, 2019, doi: 10.1109/CITSM.2018.8674331.
- [22] A. Pandhu Wijaya and H. Agus Santoso, “Improving the Accuracy of Naive Bayes Algorithm for Hoax Classification Using Particle Swarm Optimization,” *Proc. - 2018 Int. Semin. Appl. Technol. Inf. Commun. Creat. Technol. Hum. Life, iSemantic 2018*, pp. 482–487, 2018, doi: 10.1109/ISEMANTIC.2018.8549700.
-