

SURAT TUGAS
1544/B.01/LPPM-UBSI/IV/2023

Tentang

PENELITIAN YANG DIPUBLIKASIKAN DALAM JURNAL ILMIAH
Periode Maret - Agustus 2023

Menulis pada IJCIT (Indonesian Journal on Computer and Information Technology)
Volume 8 No 1 Mei 2023 (P-ISSN : 2527-449X | E-ISSN : 2549-7421)

Judul :

Penerapan Metode Naive Bayes Dengan PSO Untuk Pemilihan Peminatan Jurusan Pada SMK

- Menimbang :
1. Bahwa perlu diadakan pelaksanaan Tridharma Perguruan Tinggi dalam bentuk Penelitian.
 2. untuk Keperluan pada butir 1 (satu) diatas, maka perlu dibentuk tugas yang berkaitan dengan penelitian yang dipublikasikan dalam Jurnal Ilmiah.

MEMUTUSKAN

- Pertama : Menugaskan kepada saudara
A. Gunawan M.Kom
Sebagai Penulis yang mempublikasikan Penelitiannya pada Jurnal Ilmiah.
- Kedua : Mempunyai tugas sbb:
Melaksanakan Tugas yang diberikan dengan penuh rasa tanggung jawab.
- Ketiga : Keputusan ini berlaku sejak tanggal ditetapkan, dengan ketentuan apabila dikemudian hari terdapat kekeliruan akan diubah dan diperbaiki sebagaimana mestinya

Jakarta, 3 April 2023

LPPM Universitas Bina Sarana Informatika

Ketua



Taufik Baidawi, M.Kom

Tembusan

- Rektor Universitas Bina Sarana Informatika
- Arsip
- Ybs

IJCIT (Indonesian Journal on Computer and Information Technology)

Journal Homepage: <http://ejournal.bsi.ac.id/ejournal/index.php/ijcit>

Penerapan Metode Naïve Bayes Dengan PSO Untuk Pemilihan Peminatan Jurusan Pada SMK

Erika Mutiara¹, Lis Saumi Ramdhani², Rusda Wajhillah³, Jamal Maulana Hudin⁴, A Gunawan⁵

^{1,4}Sistem Informasi Kampus Kota Sukabumi, Universitas BSI, Sukabumi, Indonesia

^{3,4}Sistem Informasi Akuntansi Kampus Kota Sukabumi, Universitas BSI, Sukabumi, Indonesia

⁵Ilmu Komputer Kampus Kota Sukabumi, Universitas BSI, Sukabumi, Indonesia

e-mail: erika.emb@bsi.ac.id¹, lis.lud@bsi.ac.id², rusha.rwh@bsi.ac.id³, jamal.jml@bsi.ac.id⁴, a.gunawan.agn@bsi.ac.id⁵

ABSTRAK

Siswa SMP yang berencana melanjutkan jenjang pendidikannya ke SMK harus memutuskan program studi apa yang akan dipilih untuk melanjutkan pendidikannya kelak. Hal ini merupakan sesuatu yang cukup sulit diputuskan oleh kebanyakan siswa terutama bagi mereka yang tidak memiliki banyak referensi. Kebanyakan keputusan siswa, dipengaruhi oleh pendapat orang tua, teman atau figur yang diidolakan. Penelitian ini bertujuan untuk membangun sistem pendukung keputusan dengan metode Naive Bayes sebagai algoritma terpilih berbasis Particle Swarm Optimization (PSO) dalam pemilihan jurusan yang sesuai dengan kemampuan akademik siswa. Pada penelitian ini dilakukan optimasi metode Naive Bayes dengan menggunakan metode Particle Swarm Optimization untuk seleksi atribut untuk meningkatkan akurasi prediksi yang diaplikasikan terhadap data nilai calon siswa SMK. Setelah dilakukan pengujian dengan dua model, yaitu metode naive bayes dan optimasi naive bayes menggunakan particle swarm optimization hasil yang diperoleh menunjukkan bahwa dengan menggunakan Naive Bayes nilai accucary yang dihasilkan sebesar 73,09% sedangkan pengujian dengan menggunakan optimasi Naive Bayes dengan Particle Swarm Optimization didapatkan nilai accuracy sebesar 80,43%. Sehingga kedua metode tersebut memiliki perbedaan tingkat akurasi sebesar 7,34%.

Kata Kunci: jurusan, seleksi atribut, naive bayes, particle swarm optimization, program studi

ABSTRACTS

The junior high school students who plan to continue their education to the vocational high school must decide what study program they will choose to continue their education later. This is something that is quite difficult for most students to decide, especially for those who don't have many references. Most students' decisions, sometimes influenced by the opinions of parents, friends or idolized figures. This study aims to build a decision support system using the Naive Bayes method as the selected algorithm based on Particle Swarm Optimization (PSO) in selecting majors according to students' academic abilities. In this study, the optimization of the Naive Bayes method was carried out by using the Particle Swarm Optimization method for attribute selection to improve the accuracy of predictions applied to prospective Vocational School student value data. After testing with two models, namely the Naive Bayes method and the Naive Bayes optimization using particle swarm optimization results obtained showed that using Naive Bayes the accucary value generated was 73.09% while testing using Naive Bayes optimization with Particle Swarm Optimization obtained values accuracy is 80.43%. So that both methods have different accuracy rates of 7.34%.

Keywords: attribute selection, major, naive bayes, particle swarm optimization, study program



1. PENDAHULUAN

Sekolah Menengah Kejuruan (SMK) merupakan sekolah menengah setingkat dengan SMA yang bertujuan untuk membekali siswa supaya siap untuk memasuki lapangan kerja atau siap terjun di dunia usaha, oleh karena itu maka sangatlah penting kesesuaian jurusan bagi siswa-siswi yang ada di SMK (Juansen, 2021). Dalam UUSPN Nomor 20 Pasal 3 menjelaskan dalam Pasal 15 tahun 2003 menyatakan bahwa sekolah menengah kejuruan (SMK) adalah satuan tingkat pendidikan menengah untuk mempersiapkan peserta didik untuk bekerja pada bidang tertentu (Irwanto, 2021).

Hambatan dalam menentukan jurusan sering dialami oleh siswa yang akan meneruskan studi ke jenjang Sekolah Menengah Kejuruan (SMK), hambatan dapat berasal dari dalam maupun luar diri siswa (Prabowo & Subiyanto, 2017). Padahal penjurusan pada sekolah menengah kejuruan memiliki tujuan antara lain mengelompokkan siswa sesuai kecakapan, kemampuan, bakat, dan minat yang relatif sama. Dan tentunya agar siswa bisa terarah dalam menerima pelajaran yang sesuai dengan kemampuan dan bakat yang dimiliki oleh siswa. Penjurusan ini diselenggarakan untuk menyeleksi dan mengumpulkan kemampuan peserta didik yang sama untuk menempuh satu program pendidikan yang sama juga (Rahmayu & Serli, 2018). Membantu mempersiapkan siswa melanjutkan studi dan memilih dunia kerja juga membantu memperkokoh keberhasilan dan kecocokan atas prestasi yang akan dicapai di waktu mendatang.

Antusias siswa lulusan SMP untuk masuk ke SMK cukup besar, para siswa SMP yang melanjutkan jenjang pendidikannya ke SMK ini harus memutuskan program studi apa yang akan mereka pilih. Tetapi banyak siswa kurang matang untuk memilih jurusan yang sesuai dengan kemampuannya, akibatnya banyak siswa yang gagal ditengah jalan ketika mereka sudah diterima di sekolah SMK tersebut, serta banyak juga kasus siswa yang tidak cocok dengan jurusan yang dipilihnya ketika telah memperoleh pelajaran disekolah (Rusdiansyah, 2017).

Jumlah jurusan yang ada di SMK terdiri dari tiga jurusan yaitu jurusan Farmasi, Keperawatan dan TKJ (Teknik Komputer dan Jaringan). Pemilihan jurusan masing-masing

jurusan didasarkan pada ketentuan nilai yang sudah disyaratkan (Hasanah et al., 2018). Proses pemilihan jurusan saat ini memiliki kelemahan diantaranya membutuhkan waktu yang sangat lama, juga hasil yang didapat kurang akurat dikarenakan masih menggunakan sistem manual sehingga bisa saja terjadi banyak kekeliruan karena belum tersedianya aplikasi khusus untuk mendukung perhitungan tersebut (Sucipto, 2016).

Untuk itu, diperlukan sebuah sistem pendukung keputusan yang dapat membantu para siswa dalam memilih program studi yang sesuai dengan kemampuan akademik serta minat dan bakatnya. Penelitian ini merekomendasikan model pemilihan jurusan pada Sekolah Menengah Kejuruan dengan mengimplementasikan algoritma optimasi Particle Swarm Optimization (PSO) dan metode algoritma klasifikasi Naïve Bayes dikarenakan Naïve Bayes merupakan salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk machine learning dan data mining (Syarli & Muin, 2016). Sistem pendukung keputusan ini diharapkan mampu membantu para calon siswa SMK dalam memilih program studi yang sesuai dengan kemampuan akademik serta minat dan bakatnya.

Tujuan dari penelitian ini adalah untuk membangun suatu sistem pendukung keputusan dengan metode Naive Bayes sebagai algoritma terpilih berbasis Particle Swarm Optimization (PSO) dalam pemilihan jurusan yang sesuai dengan kemampuan akademik siswa. Penelitian ini menggunakan data primer yang diperoleh dari hasil riset pada salah satu sekolah menengah kejuruan swasta di Sukabumi, sebagai salah satu acuan.

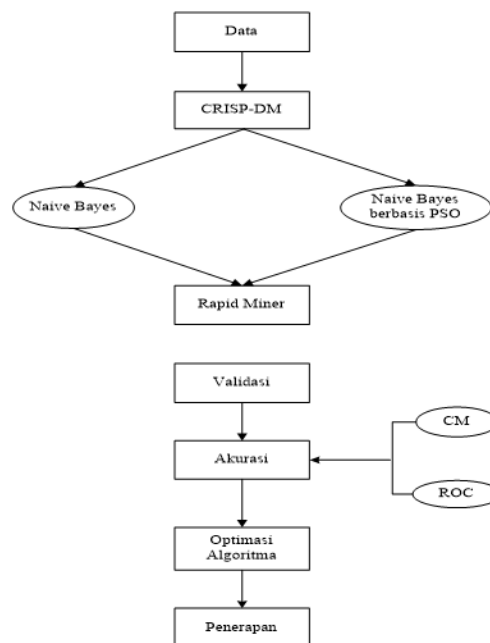
2. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini yaitu metode experiment. Metode experiment merupakan suatu penelitian yang menilai suatu kondisi tertentu dikendalikan sehingga satu atau beberapa variable dapat dikontrol (Nurhafida & Sembiring, 2021). Sedangkan untuk pengujian data menggunakan metode eksperimen dengan model Cross Industry Standard Process for Data Mining (CRISP-DM) yang terdiri dari enam tahap (Wiratama & Pradnya, 2022), yaitu: (1) Tahap Business Understanding, data set yang digunakan pada penelitian ini ialah data primer

dengan jumlah data sebanyak 311 record, terdiri dari 16 variabel atau atribut dan 1 class yang bernilai Farmasi, Keperawatan dan TKJ. (2) Tahap Data Understanding, variabel tujuan yaitu variabel yang dijadikan sebagai Farmasi, Keperawatan dan TKJ. Sedangkan Variabel prediktor atau pemrediksi (predictor variable) yaitu variabel yang dijadikan dasar sebagai penentu jurusan apa yang sesuai untuk calon siswa, dan variabel yang digunakan sebagai prediktor oleh peneliti yaitu nilai rata-rata raport pelajaran Matematika, nilai rata-rata raport pelajaran Bahasa Inggris, nilai rata-rata raport pelajaran Bahasa Indonesia, nilai rata-rata raport pelajaran IPA, nilai rata-rata raport pelajaran PAI, nilai rata-rata raport pelajaran PPKn, nilai rata-rata raport pelajaran IPS, nilai rata-rata raport pelajaran Seni Budaya, nilai rata-rata raport pelajaran TIK, nilai rata-rata raport pelajaran PJOK, nilai UN Bahasa Inggris, nilai UN Bahasa Indonesia, nilai UN Matematika, nilai UN IPA dan nilai tes tertulis (Placement Test). (3) Tahap Data Preparation, jumlah data yang diperoleh pada penelitian ini sebanyak 311 record, baik siswa yang diterima di jurusan Farmasi, Keperawatan dan TKJ. Setelah dilakukan preprocessing data yang didapat dari data internal sekolah sebanyak 311 record direduksi dengan menghilangkan duplikasi menjadi 308 record, 98 record masuk pada jurusan Farmasi, 103 record masuk pada jurusan Keperawatan dan 107 record masuk pada jurusan TKJ. Setelah itu pada penelitian ini data training dan data testing akan dipisah dengan menggunakan 10-fold cross validation dataset tersebut akan dibagi menjadi 10 bagian dan akan dilakukan pengulangan sebanyak 10 pengulangan. (4) Tahap Modelling, Tahap modelling dilakukan untuk menerapkan teknik yang tepat guna mendapatkan hasil yang optimal dalam pemilihan jurusan yang tepat. Pada penelitian ini model yang digunakan yaitu algoritma terpilih Naïve Bayes dan algoritma Naïve Bayes berbasis Particle Swarm Optimization (PSO) yang kemudian menghasilkan sejumlah aturan. Tahap ini dilakukan dengan menggunakan software rapid miner. (5) Tahap Evaluasi (Evaluation), Pada tahap ini dilakukan pengujian terhadap model-model untuk mendapatkan informasi model yang akurat. Evaluasi dan validasi menggunakan Confusion Matrix dan Kurva ROC. (6) Tahap Deployment, Pada tahap ini diterapkan model yang memiliki akurasi tinggi atau yang paling baik untuk membantu pemilihan peminatan

jurusan yang sesuai dengan kemampuan akademik dengan menggunakan data baru.

Langkah-langkah penelitian yang dilakukan pada penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Bagan Langkah Penelitian

Berdasarkan gambar 1, langkah-langkah pemikiran pada penelitian ini dapat dijelaskan sebagai berikut: data mentah direduksi terlebih dahulu, kemudian dipisahkan menjadi data training dan data testing. Dengan menggunakan metode CRISP-DM kemudian data training diuji dengan metode naive bayes dan naive bayes berbasis PSO. Tools yang digunakan untuk menguji kedua metode tersebut menggunakan rapid miner. Hasil dari masing-masing pengujian tersebut kemudian divalidasi menggunakan tools yang sama. Berdasarkan evaluasi pengujian akan diperoleh nilai akurasi, dimana evaluasi yang digunakan menggunakan dua measurement yaitu berdasarkan evaluasi Confusion Matrix (CM) dan Evaluasi ROC Curve. Setelah diperoleh evaluasi dari optimasi algoritma yang diharapkan, kemudian akan diterapkan pada pembuatan aplikasi.

3. HASIL DAN PEMBAHASAN

Tujuan dari penelitian ini adalah menerapkan Particle Swarm Optimization untuk mengeliminasi atribut input pada metode Naive Bayes, untuk meningkatkan akurasi prediksi jurusan yang sesuai dengan kemampuan akademik calon siswa SMK. Hasil dari penelitian

ini berupa hasil proses pengolahan kualitatif dan kuantitatif yang telah dikumpulkan dengan perhitungan berdasarkan model yang diusulkan. Penelitian ini akan dilakukan terhadap semua dataset yang tersedia. Eksperimen dan pengujian dalam penelitian dilakukan prediksi terhadap dataset dengan Naive Bayes tanpa PSO, dan metode seleksi atribut dengan menggunakan PSO.

3.1. Eksperimen dan Pengujian Metode Naive Bayes

Pembuatan model naive bayes dilakukan pada dataset yang terdiri dari 16 atribut yang merupakan atribut dari pemilihan jurusan dan class yang merupakan hasil akhir prediksi. Dalam membuat model naive bayes terlebih dahulu kita mencari nilai probabilitas hipotesis untuk masing-masing kelas $P(H)$. Hipotesis yang ada yaitu siswa yang diterima di jurusan Farmasi, Keperawatan dan TKJ. Data training yang digunakan dengan total data yaitu 308 dengan 98 record masuk pada jurusan Farmasi, 103 record masuk pada jurusan Keperawatan dan 107 record masuk pada jurusan TKJ., perhitungan probabilitas:

$$P(\text{Farmasi}) = 98 : 308 = 0,318318$$

$$P(\text{Keperawatan}) = 103 : 308 = 0,334415584$$

$$P(\text{TKJ}) = 107 : 308 = 0,347402597$$

Setelah probabilitas untuk tiap hipotesis diketahui langkah selanjutnya adalah menghitung probabilitas kondisi tertentu (probabilitas X) berdasarkan probabilitas tiap hipotesis (probabilitas H) atau dinamakan probabilitas prior dengan menggunakan naive bayes. Perhitungan diatas dapat dibuatkan model Naive Bayes dengan Framework Rapidminer versi 8.2. Sehingga didapatkan hasil seperti gambar 2.

SimpleDistribution	
Distribution model for label attribute jurusan	
Class KEP (0.334)	15 distributions
Class TKJ (0.347)	15 distributions
Class FAR (0.318)	15 distributions

Gambar 2. Hasil Perhitungan Probabilitas dengan Rapid Miner

3.2. Hasil Pengujian dengan Metode Naive Bayes

Hasil dari uji coba yang dilakukan yaitu untuk menghasilkan nilai accuracy dan nilai AUC (*Area Under Curve*).

a. Evaluasi Model dengan Confusion Matrix

Model confusion matrix akan membentuk matrix yang terdiri dari true positif atau tupel positif dan true negatif atau tupel negatif, kemudian masukkan data testing yang sudah disiapkan kedalam confusion matrix sehingga didapatkan hasil pada gambar 3.

accuracy: 73.05% +/- 7.08% (micro average: 73.05%)				
	true KEP	true TKJ	true FAR	class precision
pred. KEP	74	6	10	82.22%
pred. TKJ	10	84	21	73.04%
pred. FAR	19	17	67	65.05%
class recall	71.84%	78.50%	68.37%	

Sumber : Data Hasil Olahan dengan Rapid Miner

Gambar 3. Confusion Matrix Metode Naive Bayes

Berdasarkan gambar 3, data testing terdapat rincian jumlah True Keperawatan (TP) 74, False Keperawatan 16, True TKJ 84, False TKJ 31, True Farmasi 67, False Farmasi 36. Dari data tersebut maka dapat dihitung nilai accuracy, sensitivity dan spesificity. Data hasil olahan dapat dilihat pada tabel 1.

Tabel 1. Nilai Accuracy , Sensitivity, Spesificity, PPV dan NPV Naive Bayes

	Nilai
Accuracy	0,7305
Sensitivity	0,7343
Specificity	0,8655

Secara manual data tersebut dihitung dengan menggunakan persamaan (2.9), (2.10) dan (2.11) sebagai berikut:

$$Accuracy = \frac{TK + TP + FN}{TK + TP + FN + FP + TP + FP} = \frac{74 + 84 + 67}{74 + 84 + 67 + 16 + 31 + 36} = 0,7305$$

Dikarenakan dalam penelitian ini kasusnya adalah klasifikasi multiclass maka nilai Sensitivity dan Spesificity harus dihitung untuk setiap class. Sehingga perlu dilakukan perhitungan penjumlahan Sensitivity dan Spesificity untuk setiap class, seperti berikut:

$$\text{Sensitivity Keperawatan} = \frac{TP}{TP+FN} = \frac{74}{74+16} = 0,8222$$

$$\text{Sensitivity Farmasi} = \frac{TP}{TP+FN} = \frac{67}{67+36} = 0,6504$$

$$\text{Sensitivity TKJ} = \frac{TP}{TP+FN} = \frac{84}{84+31} = 0,7304$$

Dari Sensitivity ketiga class tersebut, maka diperoleh nilai rata-rata Sensitivity dalam kasus ini yaitu $(0,8222 + 0,6504 + 0,7304)/3 = 0,7343$.

$$\text{Spesificity Keperawatan} = \frac{TN}{TN+FP} = \frac{189}{189+29} = 0,8669$$

$$\text{Spesificity Farmasi} = \frac{TN}{TN+FP} = \frac{174}{174+31} = 0,8487$$

$$\text{Spesificity TKJ} = \frac{TN}{TN+FP} = \frac{170}{170+23} = 0,8808$$

Dari Spesificity ketiga class tersebut, maka diperoleh nilai rata-rata Spesificity dalam kasus ini yaitu $(0,8669+0,8487+0,8808)/3 = 0,8655$. Untuk kasus klasifikasi multiclass perhitungan kinerja (Performance) hanya menggunakan fungsi Confusion Matrix saja. Sehingga kinerja rata-rata yang dapat dihitung adalah accuracy saja.

3.3. Eksperimen dan Pengujian Metode Naive Bayes berbasis PSO

Seleksi fitur yang digunakan pada penelitian ini yaitu particle swarm optimization. Dengan PSO data yang akan diolah akan diberikan bobot untuk membantu meningkatkan hasil perhitungan, pemberian bobot ini diberikan secara acak dengan menentukan nilai minimum dan maksimum bobot. Setelah itu setiap partikel akan memiliki bobot sendiri dalam dataset, dan metode naive bayes akan diterapkan dan dihitung tingkat akurasi. Setelah semua partikel dihitung akan dicari partikel dengan nilai akurasi terbaik. Perulangan selanjutnya partikel lainnya akan secara acak bergerak kearah partikel terbaik agar dapat menemukan bobot yang lebih baik lagi. Proses ini terus berulang sampai pada batas perulangan yang diijinkan.

Indikator untuk PSO yaitu population, inertia, dan global best weight ikut diuji untuk mendapatkan akurasi yang maksimal. Populasi yang digunakan adalah 5 populasi. Dalam penyesuaian indikator PSO nilai inertia 1,0 dan nilai global best dan local best masing-masing bernilai 1,0. Jumlah partikel yang diambil 5. Nilai C1 dan C2 masing-masing 0 karena partikel berada dalam putaran pertama.

Perhitungan pada keseluruhan data dilakukan dengan rapid miner, model yang terbentuk dengan metode naive bayes sendiri akan dibandingkan dengan model yang terbentuk setelah proses pembobotan atribut. Nilai bobot atribut akan diberikan kepada setiap partikel PSO. Model yang terbentuk akan diuji sehingga dapat diketahui akurasi model yang terbentuk, setelah itu partikel PSO akan bergerak kearah partikel terbaik agar akurasi model yang dihasilkan dapat lebih tinggi dari sebelumnya. Hasil dari proses optimasi fitur oleh PSO menggunakan rapid miner dapat dilihat pada gambar 4.

AttributeWeights (Optimize Selection)	
attribute	weight
rr_mtk	0
rr_bing	0
rr_bind	1
rr_ipa	1
rr_pai	0
rr_pkn	0
rr_ips	0
rr_sbk	0
rr_tik	0
rr_pjo	0
un_bind	0
un_bing	0
un_mtk	1
un_ipa	0
pt	0

Gambar 4. Hasil Proses Optimize Selection Menggunakan Rapid Miner

Hasil proses optimasi oleh PSO dapat diketahui atribut yang memiliki pengaruh terkuat dalam penelitian ini yaitu atribut yang memiliki nilai weight sama dengan 1. Dari 15 atribut yang ada hasil pembobotan atribut yaitu 12 atribut mempunyai bobot 0 dan 3 atribut mempunyai bobot 1 Rata-rata Raport Bahasa Indonesia, Rata-rata Raport IPA dan Nilai Ujian

Nasional Matematika. Hal ini dapat mempengaruhi peningkatan hasil accuracy pada proses klasifikasi sedangkan atribut yang bernilai 0 dapat dihilangkan karena tidak terlalu mempunyai pengaruh pada akurasi pemilihan jurusan.

3.4. Hasil Pengujian Menggunakan Metode Naive Bayes berbasis PSO

Hasil dari uji coba yang dilakukan yaitu untuk menghasilkan nilai accuracy menggunakan metode naive bayes berbasis PSO sebagai berikut:

a. Evaluasi Model dengan Confusion Matrix

Model confusion matrix akan membentuk matrix yang terdiri dari true positif atau tupel positif dan true negatif atau tupel negatif, kemudian masukan data testing yang sudah disiapkan kedalam confusion matrix sehingga didapatkan hasil seperti terlihat pada gambar 5.

accuracy: 80,43%				
	true KEP	true TKJ	true FAR	class precision
pred. KEP	21	1	0	95,45%
pred. TKJ	6	28	4	73,68%
pred. FAR	4	3	25	78,12%
class recall	67,74%	87,50%	86,21%	

Sumber: Data Hasil Olahan dengan Rapid Miner

Gambar 5. Confusion Matrix Metode Naive Bayes berbasis PSO

Berdasarkan gambar 5, data testing terdapat rincian jumlah True Positif (TP) Keperawatan 21, True Positif (TP) Farmasi 25, True Positif (TP) TKJ 28, False Negatif (FN) Keperawatan 1, False Negatif (FN) Farmasi 7, False Negatif (FN) TKJ 10, False Positif (FP) Keperawatan 10, False Positif (FP) Farmasi 4, False Positif (FP) TKJ 4 dan True Negatif (TN) Keperawatan 60, True Negatif (TN) Farmasi 56 dan True Negatif (TN) TKJ 50. Dari data tersebut maka akan dihitung nilai accuracy, sensitivity dan specificity. Data hasil olahan dapat dilihat pada tabel 2.

Tabel 2. Nilai Accuracy, Sensivity, Specificity Naive Bayes berbasis PSO

	Nilai
Accuracy	0,8043
Sensitivity	0,8241
Specificity	0,9054

Secara manual data tersebut dihitung menggunakan persamaan (2.9), (2.10), (2.11), (2.12) dan (2.13) sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{21 + 25 + 28}{(21 + 25 + 28) + (10 + 4 + 4)} = 0,8043$$

Dikarenakan dalam penelitian ini kasusnya adalah klasifikasi multiclass maka nilai Sensitivity dan Spesificity harus dihitung untuk setiap class. Sehingga perlu dilakukan perhitungan penjumlahan Sensitivity dan Spesificity untuk setiap class, seperti berikut:

$$Sensitivity\ Keperawatan = \frac{TP}{TP + FN} = \frac{21}{21 + 1} = 0,9545$$

$$Sensitivity\ Farmasi = \frac{TP}{TP + FN} = \frac{25}{25 + 7} = 0,7812$$

$$Sensitivity\ TKJ = \frac{TP}{TP + FN} = \frac{28}{28 + 10} = 0,7368$$

Dari Sensitivity ketiga class tersebut, maka diperoleh nilai rata-rata Sensitivity dalam kasus ini yaitu $(0,9545 + 0,7812 + 0,7368)/3 = 0,8241$.

$$Spesificity\ Keperawatan = \frac{TN}{TN + FP} = \frac{60}{60 + 10} = 0,8571$$

$$Spesificity\ Farmasi = \frac{TN}{TN + FP} = \frac{56}{56 + 4} = 0,9333$$

$$Spesificity\ TKJ = \frac{TN}{TN + FP} = \frac{50}{50 + 4} = 0,9259$$

Dari Spesificity ketiga class tersebut, maka diperoleh nilai rata-rata Spesificity dalam kasus ini yaitu $(0,8571 + 0,9333 + 0,9259)/3 = 0,9054$. Untuk kasus klasifikasi multiclass perhitungan kinerja (Performance) hanya menggunakan fungsi Confusion Matrix saja. Sehingga kinerja rata-rata yang dapat dihitung adalah accuracy saja.

3.5. Komparasi Model Naive Bayes dengan Metode Naive Bayes berbasis Particle Swarm Optimization

Hasil pengujian menyimpulkan bahwa atribut yang mempengaruhi pengujian adalah 4 atribut yang terdiri dari 3 atribut predictor dan 1 atribut class. Hasil pengujian model Naive Bayes tanpa seleksi atribut dibandingkan dengan model Naive Bayes dengan seleksi atribut menggunakan PSO dapat dilihat pada tabel 3.

Tabel 3. Pengujian Metode Naive Bayes dan Naive Bayes berbasis PSO

	Accuracy
Naive Bayes	73,09 %
Naive Bayes berbasis PSO	80,43%

Dari hasil pengujian yang terlihat pada tabel 3, dengan dilakukan evaluasi secara confusion matrix terbukti bahwa pengujian yang dilakukan oleh metode Naive Bayes berbasis Particle Swarm Optimization memiliki nilai akurasi yang lebih tinggi dibandingkan dengan menggunakan metode Naive Bayes saja. Nilai akurasi untuk metode naive bayes sebesar 73,09 % dan nilai akurasi model naive bayes dengan seleksi atribut menggunakan PSO sebesar 80,43%. Berdasarkan nilai tersebut diperoleh selisih akurasi sebesar 7,34%.

4. KESIMPULAN

Berdasarkan pembahasan yang telah diuraikan, maka dapat ditarik kesimpulan bahwa hasil penelitian untuk nilai akurasi metode naive bayes sebesar 73,09 %, sedangkan untuk nilai akurasi naive bayes berbasis PSO sebesar 80,43%. Penggunaan metode Particle Swarm Optimization pada metode Naïve Bayes meningkatkan akurasi klasifikasi sebesar 7,34%. Sehingga dapat disimpulkan bahwa penerapan teknik optimasi particle swarm optimization dapat meningkatkan nilai akurasi pada algoritma Naïve Bayes. Dari 15 atribut yang terdapat pada dataset nilai siswa yang diperoleh dari hasil riset, kemudian selanjutnya diseleksi menjadi hanya empat atribut yang digunakan dalam menentukan prediksi peminatan jurusan, atribut-atribut tersebut yaitu: nilai rata-rata raport untuk mata pelajaran IPA, nilai rata-rata raport untuk pelajaran Bahasa Indonesia, nilai ujian nasional matematika serta satu atribut class yaitu jurusan.

5. REFERENSI

- Hasanah, U., Nurcahyo, G. W., Santony, J., & Komputer, M. (2018). Indikator Pemilihan Jurusan Pada SMK Nusantara menggunakan Metode SAW. *Jurnal Mantik Penusa*, 22(1), 39–77. <https://e-jurnal.pelitanusantara.ac.id/index.php/mantik/article/view/354>
- Irwanto. (2021, Juli). Link And Match Pendidikan Kejuruan Dengan Dunia Usaha Dan Industri Di Indonesia. *Jurnal Inovasi Penelitian*, Vol.2 No.2(ISSN 2722-9475 (Cetak); ISSN 2722-9467 (Online)), 549-562.
- Juansen, M. (2021). Penentuan Jurusan Siswa SMK Dengan Menerapkan Metode Simple Additive Weighting. *JURSIMA (Jurnal Sistem Informasi Dan Manajemen)*, 9(1), 65–71. <https://doi.org/10.47024/JS.V9I1.244>
- Nurhafida, S. I., & Sembiring, F. (2021). Analisis Text Clustering Masyarakat Di Twitter Mengenai Mcdonald'sxpbs Menggunakan Orange Data Mining. *Prosiding Seminar Nasional Sistem Informasi Dan Manajemen Informatika Universitas Nusa Putra*, 1(01), 28–35. <https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/4>
- Prabowo, I. M., & Subiyanto. (2017, Juni). Sistem Rekomendasi Penjurusan Sekolah Menengah Kejuruan Dengan Algoritma C4.5. *Jurnal Kependidikan*, Volume 1, Nomor 1, 139-149.
- Rahmayu, M., & Serli, R. K. (2018, April). Sistem Pendukung Keputusan Pemilihan Jurusan Pada SMK Putra Nusantara Jakarta Menggunakan Metode Analytical Hierarchy Process (AHP). *Jurnal SIMETRIS*, Vol. 9 No. 1(ISSN: 2252-4983), 551-564.
- Rusdiansyah, R. (2017). Analisis Keputusan Menentukan Jurusan Pada Sekolah Menengah Kejuruan Dengan Metode Simple Additive Weighting. *Jurnal Techno Nusa Mandiri*, 14(1), 49–56. <https://doi.org/10.33480/TECHNO.V14I1.184>
- Sucipto, H. (2016, Juni). Sistem Pendukung Keputusan Penentuan Jurusan Sekolah Menengah Atas Dengan Metode SAW. *Jurnal Ilmiah SISFOTENIKAJ*, Vol. 6, No. 2,, 147-157.
- Syarli, & Muin, A. A. (2016, April). Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). *Jurnal Ilmiah Ilmu Komputer*, Vol. 2, No. 1((P) ISSN 2442-4512; (O) ISSN 2503-3832), 22-26.
- Wiratama, M. A., & Pradnya, W. M. (2022). Optimasi Algoritma Data Mining Menggunakan Backward Elimination untuk Klasifikasi Penyakit Diabetes. *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI*, 11(1), 1–12. <https://doi.org/10.23887/JANAPATI.V11I1.45282>