

Komparasi Algoritma Klasifikasi C4.5 dan Naïve Bayes Berbasis Particle Swarm Optimization Untuk Penentuan Resiko Kredit

Achmad Rifai¹⁾, Rizki Aulianita²⁾

STMIK Nusa Mandiri Jakarta¹⁾, STMIK Nusa Mandiri Jakarta²⁾
achmad.acf@nusamandiri.ac.id¹⁾, rizki.rzk@nusamandiri.ac.id²⁾

Abstrak – Perkreditan memiliki peran penting dalam pembiayaan pembangunan infrastruktur dan perekonomian nasional sekarang ini. Begitu juga dengan koperasi yang merupakan pembiayaan berskala kecil yang digemari masyarakat menengah kebawah terutama dalam fasilitas simpan pinjam. Namun, fasilitas peminjaman memiliki risiko yang besar dalam kerugian yang harus ditanggung oleh koperasi, jadi dalam mengurangi risiko dalam peminjaman diperlukan metode dalam menjalankannya. Dalam penelitian ini dilakukan dengan menggunakan komparasi algoritma klasifikasi data mining C4.5 dan Naïve Bayes kemudian dilakukan optimasi berbasis Particle Swarm Optimization. Berdasarkan hasil pengujian bahwa nilai akurasi algoritma C4.5 sebesar 85.40% dan nilai akurasi algoritma Naïve Bayes sebesar 85,09%. Dari kedua algoritma tersebut kemudian dilakukan kombinasi dengan optimasi Particle Swarm Optimization, dengan hasil algoritma C4.5+PSO memiliki nilai tertinggi berdasarkan nilai accuracy sebesar 87.61%, AUC sebesar 0.860 dan precision sebesar 88.96% sedangkan nilai recall tertinggi diperoleh oleh algoritma Naïve Bayes+PSO sebesar 96.75%.

Kata Kunci: Kredit, Data Mining, Algoritma C4.5, Naïve Bayes, Optimasi

Abstract – Credit has an important role in financing infrastructure development and the current national economy. Likewise with the cooperative which is a small-scale financing that favored the middle to lower society, especially in savings and loan facilities. However, lending facilities have a great risk in losses that must be borne by the cooperative, so in reducing the risk in lending method required in running it. In this research is done by using comparison of data mining classification algorithm C4.5 and Naïve Bayes then done by Particle Swarm Optimization. Based on the test results that the accuracy of C4.5 algorithm of 85.40% and the accuracy of Naïve Bayes algorithm of 85.09%. The two algorithms are then combined with Particle Swarm Optimization optimization, where the C4.5 + PSO algorithm has the highest value based on accuracy value of 87.61%, AUC is 0.860 and precision is 88.96% while the highest recall value is obtained by Naïve Bayes + PSO algorithm 96.75%.

Key Word: Credit, Data Mining, Algorithm C4.5, Naïve Bayes, Optimization

1. PENDAHULUAN

Koperasi merupakan salah satu ciri khas bangsa Indonesia. Banyak para debitur yang melakukan simpan pinjam melalui koperasi dengan pertimbangan bunga yang relatif ringan dengan asaz kekeluargaan. Perkembangan koperasi yang signifikan menyebabkan banyak permasalahan yang muncul diantaranya jumlah piutang lebih besar (Saraswati & Yadnyana, 2014), kredit macet dan hutang tak tertagih. Agar sistem perkreditan berjalan stabil dan lancar maka perlu analisis kredit yang dilakukan oleh pihak koperasi.

Penelitian yang berkaitan dengan metode klasifikasi data mining tentang analisa resiko telah banyak dibahas oleh peneliti sebelumnya seperti penelitian dengan melakukan berbagai metode seperti Logistic Regression, Neural Network, Radial Basis Function, Case-base Reasoning, Decision Trees dan Support Vector Machine untuk menguji tingkat akurasi klasifikasi dengan data historis oleh lembaga keuangan Jerman dengan

hasil Decision Tree secara signifikan dapat memperoleh nilai akurasi tertinggi dengan melebihi 0.05 dibanding model yang lainnya (Zurada, 2010). 85.20% untuk dataset yang bersumber dari Australia diperoleh akurasi paling tinggi dari Naive Bayes Clasification, sedangkan penelitian lain melakukan perbandingan dari beberapa algoritma atau metode diantaranya adalah Decision Tree, Boosting, Bagging dan Random Forest (Sawant & Chawan, 2013). Penelitian yang lain dengan perbandingan algoritma Decision Tree, Bayesian, Neural Network dan Support Vector Machine (SVM) diketahui bahwa algoritma klasifikasi C4.5 dan Support Vector Machine memiliki nilai accuracy tertinggi sebesar 85,4% (Rifai, 2016).

Berdasarkan penelitian sebelumnya bahwa algoritma klasifikasi atau metode yang memiliki akurasi terbaik hasilnya berbeda-beda sehingga belum diketahui metode yang paling akurat dalam menentukan kelayakan pemberian kredit. Untuk

itu dalam penelitian ini akan dilakukan perbandingan metode Algoritma C4.5 dibandingkan dengan metode *Naïve Bayes* dengan masing-masing akan dioptimasi menggunakan *Particle Swarm Optimization* sehingga diperoleh metode dengan akurasi prediksi penentuan resiko kredit. Penelitian ini menggunakan data kredit yang bersumber dari koperasi pegawai Dirjen Pajak pada tahun 2010, yang bersumber dari pemberian kredit oleh koperasi dan pembayaran pinjaman oleh anggota atau nasabah.

2. LANDASAN TEORI

2.1. Data Mining

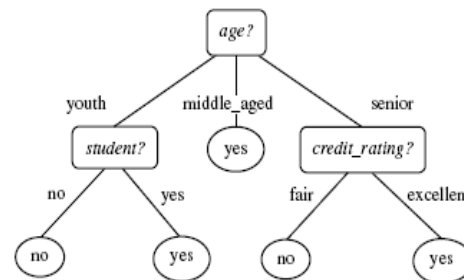
“*Data mining* adalah proses menelusuri pengetahuan baru, pola dan tren yang dipilah dari jumlah data yang besar yang disimpan dalam repositori atau tempat penyimpanan dengan menggunakan teknik pengenalan pola serta statistik dan tehnik matematika” (Larose, 2006).

2.2. Algoritma Klasifikasi Data Mining

“Klasifikasi adalah proses menempatkan obyek atau konsep tertentu, kedalam satu set kategori berdasarkan sifat obyek atau konsep yang bersangkutan” (Gorunescu, 2011). Klasifikasi *Data mining* adalah suatu metode pembelajaran, untuk memprediksi nilai dari sekelompok atribut dalam menggambarkan dan membedakan kelas data atau konsep yang bertujuan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui (Vercellis, 2009).

2.3. Algoritma C4.5

Pohon keputusan sendiri merupakan pendekatan *divide and conquer* dalam mempelajari masalah dari sekumpulan data independen yang digambarkan dalam bagan pohon (Witten, Frank, & Hall, 2011). Pohon keputusan dianggap sebagai salah satu pendekatan yang paling populer, dalam klasifikasi pohon keputusan terdiri dari sebuah node yang membentuk akar, node akar tidak memiliki inputan. Node lain yang bukan sebagai akar tetapi memiliki tepat satu inputan disebut node internal atau test node, sedangkan node lainnya dinamakan daun. Daun mewakili nilai target yang paling tepat dari salah satu *class* (Maimon & Rokach, 2010).



Sumber : (Witten et al., 2011)

Gambar 1. Contoh konsep pohon keputusan

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5, yaitu (Kusri & Luthfi, 2009):

1. Menyiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas-kelas tertentu.
2. Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai gain dari masing-masing atribut, nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung gain dari atribut, hitung dahulu nilai entropy yaitu:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i$$

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi S

pi : proporsi dari Si terhadap S

3. Kemudian hitung nilai gain dengan metode *informasi gain*:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

|Si| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

4. Ulangi langkah ke-2 hingga semua tupelo terpartisi.
5. Proses partisi pohon keputusan akan berhenti saat:
 - a. Semua tupel dalam node N mendapat kelas yang sama.
 - b. Tidak ada atribut di dalam tupel yang dipartisi lagi.
 - c. Tidak ada tupel di dalam cabang yang kosong.

2.4. Structural Equation Modeling (SEM)/Model Persamaan Struktura

Teknik *Naïve Bayes* (NB) adalah salah satu bentuk sederhana dari *Bayesian* yang jaringan untuk klasifikasi. Sebuah jaringan *Bayes* dapat dilihat sebagai diarahkan sebagai tabel dengan distribusi probabilitas gabungan lebih dari satu set diskrit dan variabel stokastik (Pearl 1988) (Liao, 2007).

$$P(x|y) = \frac{P(y|x) P(x)}{P(y)}$$

Keterangan

- y = data dengan kelas yang belum diketahui
 x = hipotesis data y merupakan suatu kelas spesifik
 P(x|y) = probabilitas hipotesis x berdasarkan kondisi y (*posteriori probability*)
 P(x) = probabilitas hipotesis x (*prior probability*)
 P(y|x) = probabilitas y berdasarkan kondisi pada hipotesis x
 P(y) = probabilitas dari y

3. METODE PENELITIAN

Terdapat empat metode penelitian yang umum digunakan yaitu *Reserch of Action*, *Experiment*, *Case of Study*, dan *Survey* (Dawson, 2019). Adapun metode penelitian yang digunakan adalah bentuk *Experiment* penelitian yaitu peneliti mengendalikan penyelidikan hubungan kausal dengan menggunakan tes (Dawson, 2019). Penelitian ini menggunakan Model *Cross-Standard Industry for Data Mining* (CRISP-DM) yang terdiri dari 5 fase, yaitu (Larose, 2006) :

A. Business Understanding

Dalam menjaga kedisiplinan dan kepatuhan bagi nasabah yang melakukan pinjaman harus mengikuti prosedur pinjaman meliputi:

- 1) Permohonan Pinjaman
Nasabah mengisi formulir yang berisi pinjaman yang diajukan secara tertulis yang disediakan oleh koperasi.
- 2) Legalitas Pinjaman
Pinjaman hanya berlaku untuk satu anggota dan calon anggota yang bersifat pribadi.
- 3) Analisis Pinjaman
Setelah data terkumpul maka analisis menilai apakah nasabah dapat memenuhi kriteria pinjaman yang baik atau tidak yang meliputi:
 - a) Kemauan / niat bayar
 - b) *Character* (ahlak)
 - c) Integritas
 - d) Analisis jaminan
- 4) Pencairan Dana Pinjaman

Setelah analisis melakukan pengujian terhadap kelayakan kredit maka dana pinjaman dapat dicairkan dan dapat diambil di koperasi.

B. Data Understanding

Data yang diperoleh dari Koperasi Pegawai Kantor Pusat Dirjen Pajak adalah data kredit nasabah tahun 2010. Setelah dilakukan proses *preparation data*, kemudian 15 atribut yang digunakan dalam data status kredit nasabah. Variabel prediktor atau pemrediksi (*predictor variabel*) merupakan variabel yang digunakan dalam penelitian ini. Variabel prediktor yaitu umur, jenis kelamin, status pernikahan, jumlah anak, pendidikan, gaji, plafon, jangka waktu, pokok, bunga, jumlah angsuran, jaminan, tujuan dan *prosentase*. Sedangkan variabel tujuan adalah status kredit.

C. Data Preparation

Jumlah data awal yang diperoleh dari pengumpulan data yaitu sebanyak 366 *record*, namun tidak semua data dapat digunakan dan tidak semua atribut digunakan karena harus melalui beberapa tahap pengolahan awal data (*preparation data*). Data yang berkualitas diperoleh dengan menggunakan beberapa teknik (Vercellis, 2009) yaitu:

1) Data Validation

Mengidentifikasi dan menghapus data yang *noise* dan data yang tidak lengkap (*missing value*)

2) Data integration and transformation

Meningkatkan akurasi dan efisiensi algoritma. Peneliti menggunakan data yang bernilai kategorikal.

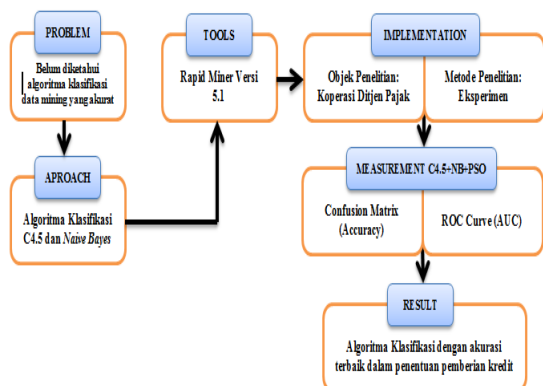
3) Data size reduction

Memperoleh data set dengan jumlah atribut dan *record* yang lebih sedikit tetapi bersifat informatif. Pada penelitian ini menggunakan 15 atribut untuk data *training*. Setelah dilakukan *preprocessing* maka data sebanyak 366 *record* direduksi dengan menghilangkan duplikasi menjadi 348 *record* untuk data *training*.

D. Modeling

Metode yang diusulkan yaitu dengan penerapan metode atau algoritma C4.5 dan *Naïve Bayes* kemudian dilakukan optimasi berbasis *Particle Swarm Optimization* dalam penilaian kelayakan pemberian kredit. Dimulai dari pembagian dataset dengan metode 10 *cross validation* yaitu data testing dan data training, kemudian data training diproses dengan membandingkan dari kedua metode tersebut ditambahkan optimasi PSO, sehingga menghasilkan model evaluation yang diukur

dengan nilai *Confusion Matrix* dan *ROC curve* dapat dilihat pada Gambar 2.



Sumber: Hasil Penelitian (2018)

Gambar 2. Model yang Diusulkan

E. Evaluation

Evaluasi terhadap model yang terbentuk akan dilakukan dengan pengukuran akurasi dan area di bawah kurva. Akurasi diukur dengan menggunakan *confusion matrix*, dan nilai *AUC* akan diukur dengan menggunakan *ROC Curve*. *Confusion matrix* akan menggambarkan hasil akurasi mulai dari prediksi positif yang benar (tp), prediksi positif yang salah (tn), prediksi negatif yang benar (fp), dan prediksi negatif yang salah (fn). Sehingga model yang terbentuk dapat langsung diuji dengan data yang secara acak dipisahkan dengan 10 *folds cross validation*. Dengan *confusion matrix*, nilai akurasi dari model akan dibandingkan antara algoritma klasifikasi *C4.5*, *Naïve Bayes*, *C4.5 Optimasi PSO* dan *Naïve Bayes Optimasi PSO*. Hasil pengukuran dengan akurasi tertinggi akan digunakan sebagai penentuan resiko kredit.

3. HASIL DAN PEMBAHASAN

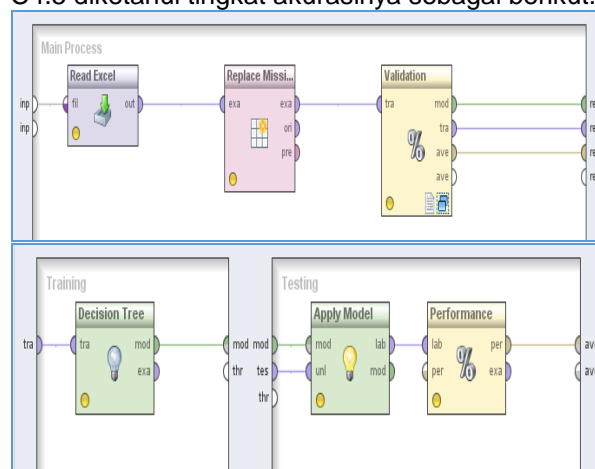
Pada penelitian ini, metode yang digunakan adalah penelitian eksperimen dimana penelitian percobaan melibatkan penyelidikan hubungan kausal menggunakan tes dikendalikan oleh peneliti itu sendiri. Pada penelitian tentang penentuan pemberian kredit ini menggunakan data berdasarkan koperasi dengan memasukan data uji yang berasal dari data *training*. Dimana data akhir tersebut didapat setelah melakukan proses *preprocessing*. Hasil dengan dilakukannya *preprocessing* ternyata terdapat sebanyak 366 *record* direduksi dengan menghilangkan duplikasi menjadi 315 *record* untuk data *training* data maka digunakan metode *cross validation* untuk menguji tingkat akurasi. Setelah didapat data yang akan

diproses kemudian dilakukan pengujian terhadap empat algoritma yang berbeda dengan menggunakan *software data mining Rapid Miner 5.3* untuk menguji yaitu algoritma *C4.5*, *Naïve Bayes*, *C4.5+PSO* dan *Naïve Bayes+PSO*. Hasil algoritma dengan akurasi yang terbaik nantinya akan digunakan dalam penentuan kelayakan kredit pada koperasi.

3.1. Algoritma C4.5

Berdasarkan teori diatas, dalam penelitian ini terdapat dua macam variabel, yaitu variabel dependen (bergantung) dan variabel independen (bebas), antara lain:

Berdasarkan uji hasil dari data koperasi yang sudah dilakukan *preprocessing* maka perhitungan akurasi data *training* menggunakan Algoritma *C4.5* diketahui tingkat akurasinya sebagai berikut.



Gambar 3. Pengujian Algoritma C4.5

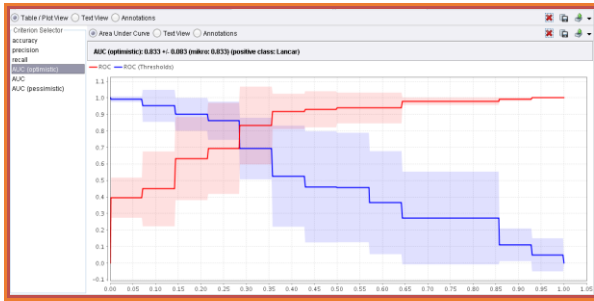
Dari hasil uji dengan menggunakan *Rapid Miner 5.3* seperti gambar 3 diatas maka didapat *Confusion Matrix* untuk Algoritma *C4.5* seperti pada tabel 1.

Tabel 1. Confusion Matrix Algoritma C4.5

	True Bermasalah	True Lancar
Pred Bermasalah	38	14
Pred Lancar	32	231
Class Recall	54.29%	94.29%

$$Accuracy = \frac{(TN+TP)}{(TN+FN+TP+FP)} = \frac{231+38}{(231+14+38+32)} = 85.40\%$$

Setelah didapat akurasi sebesar 85.40% algoritma *C4.5* juga diketahui *ROC Curve* sebesar 0.833 dapat dilihat pada gambar 4.



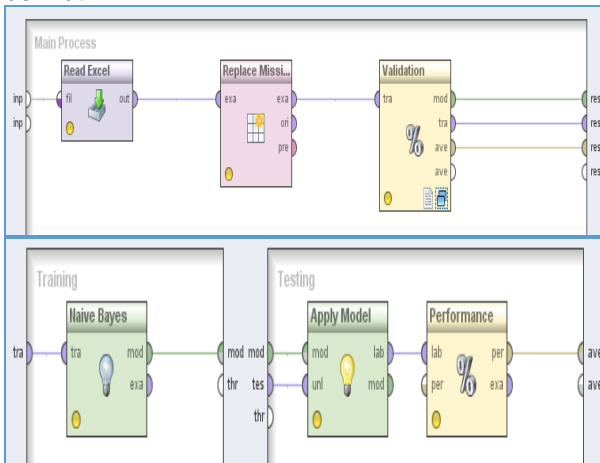
Gambar 4. Grafik AUC Algoritma C4.5



Gambar 6. Grafik AUC Algoritma Naive Bayes

3.2. Algoritma Naive Bayes

Berdasarkan uji hasil dari data koperasi yang sudah dilakukan preprocessing maka perhitungan akurasi data training menggunakan Algoritma Naive Bayes diketahui tingkat akurasinya sebagai berikut.



Gambar 5. Pengujian Algoritma Naive Bayes

Dari hasil uji dengan menggunakan Rapid Miner 5.3 seperti gambar 5 diatas maka didapat Confusion Matrix untuk Algoritma Naive Bayes seperti pada tabel 2.

Tabel 2. Confusion Matrix Algoritma Naive Bayes

	True Bermasalah	True Lancar
Pred Bermasalah	38	15
Pred Lancar	32	230
Class Recall	54.29%	94.88%

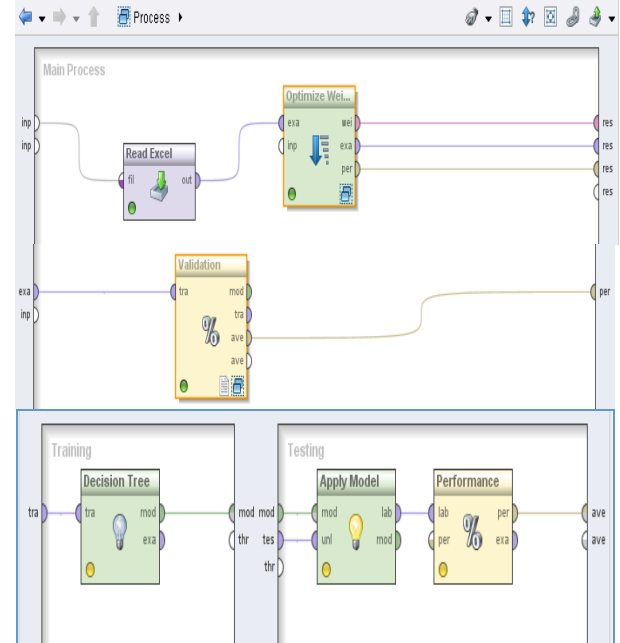
$$Accuracy = \frac{(TN+TP)}{(TN+FN+TP+FP)} = \frac{230+38}{(230+15+38+32)} = 85.09\%$$

Setelah didapat akurasi sebesar 85.09% algoritma Naive Bayes juga diketahui ROC Curve sebesar 0.835 dapat dilihat pada gambar 6.

3.3. Particle Swarm Optimization

Setelah dilakukan uji dari kedua metode atau algoritma klasifikasi C4.5 dan Naive Bayes maka pada penelitian ini akan diterapkan optimasi menggunakan Particle Swarm Optimization untuk meningkatkan nilai akurasi dari masing-masing metode yaitu:

A. Algoritma C4.5 Dengan Particle Swarm Optimization



Gambar 7. Pengujian 10-Fold Validation Model Algoritma C4.5+PSO

Dari hasil uji dengan menggunakan Rapid Miner 5.3 seperti gambar 7 diatas maka didapat Confusion Matrix untuk Algoritma Naive Bayes seperti pada tabel 3.

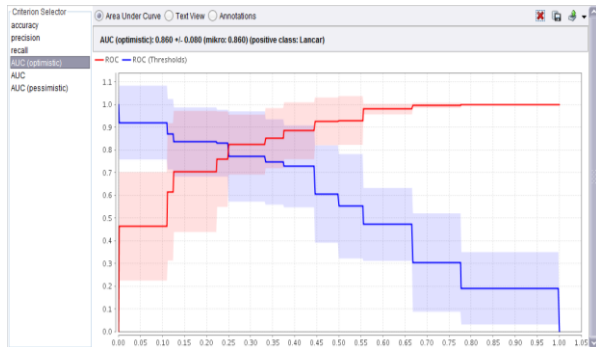
Tabel 3. Confusion Matrix Algoritma C4.5+PSO

	True Bermasalah	True Lancar
Pred Bermasalah	40	9
Pred Lancar	30	236
Class Recall	57.14%	96.33%

$$Accuracy = \frac{(TN+TP)}{(TN+FN+TP+FP)} = \frac{236+40}{(236+9+40+30)}$$

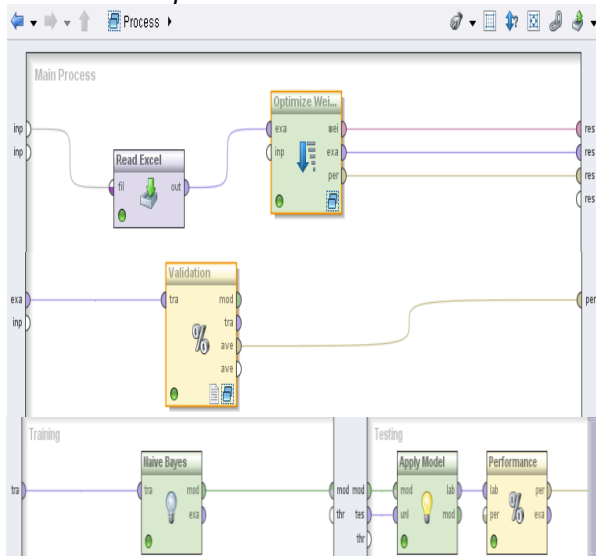
= 87.61%

Setelah didapat akurasi sebesar 87.61% algoritma C4.5+PSO juga diketahui ROC Curve sebesar 0.860 dapat dilihat pada gambar 8.



Gambar 8. Grafik AUC Algoritma C4.5+ PSO

B. Algoritma Naïve Bayes Dengan Particle Swarm Optimization



Gambar 9. Grafik AUC Algoritma Naïve Bayes + PSO

Dari hasil uji dengan menggunakan Rapid Miner 5.3 seperti gambar 9 diatas maka didapat Confusion Matrix untuk Algoritma Naïve Bayes + PSO seperti pada tabel 4.

Tabel 4. Confusion Matrix Algoritma Naïve Bayes + PSO

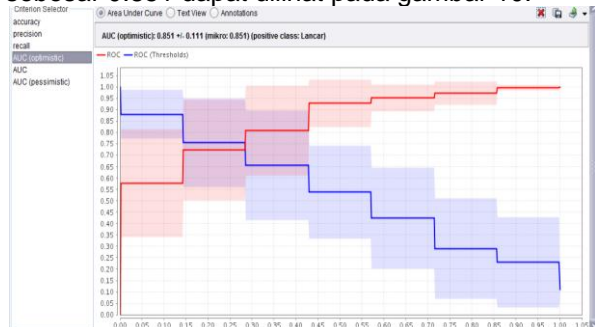
	True Bermasalah	True Lancar
--	-----------------	-------------

Pred Bermasalah	38	8
Pred Lancar	32	237
Class Recall	54.29%	96.73%

$$Accuracy = \frac{(TN+TP)}{(TN+FN+TP+FP)} = \frac{237+38}{(237+8+38+32)}$$

= 87.31%

Setelah didapat akurasi sebesar 87.31% algoritma Naïve Bayes + PSO juga diketahui ROC Curve sebesar 0.851 dapat dilihat pada gambar 10.



Gambar 10. Grafik AUC Algoritma Naïve Bayes + PSO

3.4. Analisis Hasil Komparasi

Berdasarkan uji hasil dari data koperasi yang sudah dilakukan preprocessing maka perhitungan akurasi data training menggunakan pengujian dengan mengukur kinerja algoritma C4.5 dan Naïve Bayes menggunakan 10-fold cross validation dapat dilihat hasilnya pada tabel 5, dimana dalam tabel tersebut dapat diketahui bahwa akurasi tertinggi diperoleh oleh algoritma C4.5 dengan optimasi Particle Swarm Optimization.

Tabel 5. Nilai Accuracy 10-fold cross validation

N-Fold	C4.5	NB	C4.5+PSO	NB+PSO
1	80.32%	81.60%	85.72%	86.35%
2	80.32%	81.60%	85.72%	86.35%
3	82.22%	82.54%	85.40%	86.35%
4	80.64%	81.88%	86.02%	87.00%
5	85.40%	83.49%	86.03%	86.98%
6	83.51%	84.48%	87.60%	86.65%
7	83.49%	84.13%	86.67%	87.00%
8	81.59%	85.09%	87.61%	87.00%
9	82.54%	83.49%	87.30%	87.30%
10	84.76%	83.84%	86.99%	87.31%

Sedangkan perbandingan nilai accuracy, auc, precision, dan recall keempat algoritma dapat dilihat pada tabel 6 dimana algoritma C4.5+PSO memiliki nilai tertinggi berdasarkan nilai accuracy,

AUC dan *precision* sedangkan nilai *recall* tertinggi diperoleh oleh algoritma *Naïve Bayes+PSO*.

Tabel 6. Perbandingan Nilai Semua Algoritma

	C4.5	NB	C4.5+PSO	NB+PSO
<i>Accuracy</i>	85.40%	85.09%	87.61%	87.31%
<i>AUC</i>	0.8350	0.8333	0.860	0.851
<i>Recall</i>	94.29%	93.88%	96.32%	96.75%
<i>Precision</i>	88.11%	87.83%	88.96%	88.56%

5. KESIMPULAN

Pada penelitian ini dilakukan pemodelan menggunakan algoritma *C4.5*, *Naïve Bayes*, *C4.5+PSO* dan *Naïve Bayes+PSO*, data yang digunakan adalah data nasabah koperasi dimana penelitian ini bertujuan untuk mengetahui nasabah mana yang nantinya berpotensi melakukan kredit macet atau tidak. Penelitian ini ditambahkan dengan penerapan algoritma *PSO* sebagai pembobotan atribut teknik klasifikasi data mining *C4.5* dan *Naïve Bayes*. Validasi model menggunakan *10fold cross-validation* dan evaluasi model menggunakan *confusion matrix* dan *kurva ROC*. Dari penelitian ini didapat algoritma *C4.5* yang dikombinasi dengan algoritma *PSO* memiliki nilai tertinggi berdasarkan nilai *accuracy* sebesar 87.61%, *AUC* sebesar 0.860 dan *precision* sebesar 88.96% sedangkan nilai *recall* tertinggi diperoleh oleh algoritma *Naïve Bayes+PSO* sebesar 96.75%.

Berdasarkan penelitian yang sudah dilakukan bahwa penggunaan optimasi *Particle Swarm Optimization* dapat meningkatkan akurasi dari algoritma *C4.5* dan algoritma *Naïve Bayes*. Dengan demikian metode algoritma *C4.5* berbasis *Particle Swarm Optimization* adalah metode yang terbaik dari keempat algoritma yang telah diujikan dan dapat digunakan untuk pemecahan masalah kelayakan penentuan kredit pada koperasi.

DAFTAR PUSTAKA

- [1] Dawson, C. W. (2019). *Projects in Computing and Information System A Student's Guide* (2nd ed.). England: Addison-Wesley.
- [2] Gorunescu, F. (2011). *Data Mining Concept, Models and Techniques* (12th ed.). Berlin Heidelberg: Springer Verlag.
- [3] Kusriani, & Luthfi, E. T. (2009). *Algoritma Data Mining*. Yogyakarta: Andi Offset.
- [4] Larose, D. T. (2006). *Data Mining Methods and Models*. Canada: John Wiley & Sons, Inc., Hoboken, New Jersey.
- [5] Liao, T. W. (2007). *Recent Advances in Data Mining of Enterprise Data* (6th ed.). Singapore: World Scientific.
- [6] Maimon, O., & Rokach, L. (2010). *Data Mining and knowledge Discovery Handbook* (Second). New York: Springer.
- [7] Rifai, A. (2016). Kajian Algoritma *C4.5 Naive Bayes*, *Neural Network* Dan *Svm* Dalam Penentuan Kelayakan Pemberian Kredit. *Jurnal Sistem Informasi Strmik Antar Bangsa*, V NO 2(2), 176–182.
- [8] Saraswati, L., & Yadnyana, I. K. (2014). Pengaruh Struktur Pengendalian Intern Terhadap Kelancaran Pengembalian Kredit Pada Koperasi Simpan Pinjam Di Kota Denpasar. *Jurnal Akuntansi Universitas Udayana*, 7(1), 122–134.
- [9] Sawant, A. A., & Chawan, P. M. (2013). Study of Data Mining Techniques used for Financial Data Analysis, 2(3), 503–509.
- [10] Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making* (First). Italy: John Wiley & Sons Ltd.
- [11] Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning and Tools* (Third). Burlington: Morgan Kaufmann.
- [12] Zurada, J. (2010). Could decision trees improve the classification accuracy and interpretability of loan granting decisions? *Proceedings of the Annual Hawaii International Conference on System Sciences*, 1–9. <https://doi.org/10.1109/HICSS.2010.124>