PARTICLE SWARM OPTIMIZATION UNTUK MENINGKATKAN AKURASI PREDIKSI PEMASARAN BANK

Ridwansyah[1]; Esty Purwaningsih[2]

[1]Program Studi Teknik Informatika
STMIK Nusa Mandiri Jakarta
http://www.nusamandiri.ac.id
rdwansyah@gmail.com

[2]Program Studi Manajemen Informatika
AMIK BSI Tangerang
http://www.bsi.ac.id
esty.epw@bsi.ac.id

Abstract— *Bank marketing can be through the promotion of a product by generating data and information that is quite complex and broad so that will find difficulties in analyzing it. Neural network can solve the problem especially the large data sample that can be in the data set but still lacks the selection of the appropriate and optimal feature on the weight of the attributes used so as to cause the prediction accuracy level becomes less high, PSO algorithm can increase the attribute weight and improve the accuracy a larger algorithm and data classification. The results of the study using Neural network and PSO Algorithm show better improvement and impressive in performance model of neural network which has weakness in feature selection..*

Keyword: *Bank Marketing, Data Mining, Neural Network, Particle Swarm Optimization.*

Intisari—Pemasaran bank dapat melalui promosi sebuah produk dengan menghasilkan data dan informasi yang cukup kompleks dan luas sehingga akan merasa kesulitan dalam menganalisisnya. Neural network dapat menyelesaikan masalah khususnya sampel data besar yang di dapat di uci data set namun masih memiliki kekurangan pada pemilihan fitur yang sesuai dan optimal pada bobot atribut yang digunakan sehingga menyebabkan tingkat akurasi prediksi menjadi kurang tinggi, Algoritma PSO dapat meningkatkan bobot atribut dan meningkatkan akurasi suatu algoritma dan klasifikasi data yang lebih besar. Hasil penelitian menggunakan Neural network dan Algoritma PSO menunjukkan perbaikan yang lebih baik dan mengesankan dalam kinerja model neural network yang memiliki kelemahan dalam pemilihan fitur.

Kata Kunci*:* Data Mining*, Neural Network, Particle Swarm Optimization,* Pemasaran Bank*.*

PENDAHULUAN

Perusahaan untuk meningkatkan bisnis dapat mempromosikan produk atau jasa dengan cara menggunakan pemasaran secara langsung. Ada dua pendekatan untuk sebuah promosi: pemasaran massal dan pemasaran langsung, namun sekarang ini dimana sebuah produk sangat banyak dan pasarnya sangat kompetitif maka pemasaran massal menjadi kurang efektif tingkat respon dari orang yang membeli produk setelah melihat promosinya berbeda dengan pemasaran langsung dengan memilih pelanggan tertentu sebagai target promosi(Ling & Li, 1998) dengan menghubungi mereka untuk tujuan tertentu. Pemasaran merupakan strategi khas untuk meningkatkan bisnis dengan menggunakan pemasaran langsung saat menargetkan segmen pelanggan(Moro & Laureano, 2011).

Data mining telah digunakan dalam pemasaran langsung untuk menargetkan pelanggan, data mining juga bisa digunakan untuk menemukan pola(Agrawal, Ghosh, Imielinski, Iyer, & Swami, 1992). *Neural Network* lebih flesksibel yaitu tidak ada batasan apriori yang dikenakan bila dibandingkan dengan pemodelan statistic klasik, sehingga Neural Network cenderung memberikan prediksi yang akurat(Moro & Laureano, 2011). Salah satu keuntungan Neural Network yang sering disebutkan bahwa mereka dapat bekerja secara efektif dengan data yang tidak terdistribusi secara normal(Satapathy, Chittineni, Mohan Krishna, Murthy, & Prasad Reddy, 2012). Neural network memiliki kelebihan pada prediksi non linear, memiliki performa yang lebih baik dan memiliki kemampuan dengan tingkat kesalahan rendah(Moro, Cortez, & Rita, 2014).

Penelitian tentang data mining untuk pemasaran langsung dengan menggunakan decision tree dan naïve bayes, dari Ridwansyah algoritma decision tree lebih baik dari pada naïve

bayes(Ling & Li, 1998). Namun pada tahun 2014 dengan data yang sedikit berbeda, penelitian tentang prediksi keberhasilan telemarketing bank menggunakan regresi logistic, decision tree, neural network, dan svm dengan ke empat model yang diuji menghasilkan bahwa algoritma neural network memberikan hasil akurasi terbaik dengan AUC 0,8 atau 80%(Moro et al., 2014).

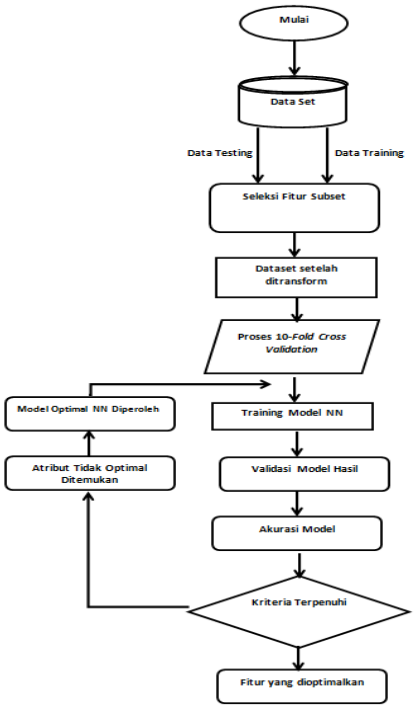
Optimasi metaheuristic yang paling banyak digunakan untuk seleksi fitur meliputi algoritma genetika, optimasi partikel swarm dan optimasi koloni semut. Algoritma neural network dapat menyelesaikan masalah khususnya sampel data besar yang di dapat di uci data set, tetapi neural network memiliki kekurangan pada sulitnya pemilihan fitur yang sesuai dan optimal pada bobot atribut yang digunakan sehingga menyebabkan tingkat akurasi prediksi menjadi kurang tinggi. Dalam rumusan masalah tersebut maka peneliti menggunakan algoritma particle swarm optimization (PSO)(Saputra, 2017). Algoritma PSO dapat meningkatkan bobot atribut dan meningkatkan akurasi suatu algoritma dan klasifikasi data yang lebih besar(Cao, Cui, Shi, & Jiao, 2016). Karena konsep sederhana, implementasi mudah, dan konvergensi cepat, particle swarm optimization (PSO) dapat diterapkan untuk berbagai aplikasi di berbagai bidang untuk memecahkan masalah optimasi(Vieira, Mendonça, Farinha, & Sousa, 2013). PSO digunakan untuk optimasi dalam menentukan parameter yang optimal, akan tetapi parameter yang digunakan memiliki ketergantungan yang sensitif(Yusup, Zain, & Hashim, 2012).

Persaingan dan krisis keuangan yang ada saat ini mempunyai tekanan besar bagi suatu bank untuk meningkatkan asset keuangan. Pemasaran bank dapat melalui promosi sebuah produk membantu dalam meningkatkan pendapatan, mengurangi biaya dan meningkatkan kepuasan pelanggan sehingga dapat menghasilkan data dan informasi yang cukup kompleks dan luas. Dalam rumusan masalah tersebut maka peneliti menggunakan Neural network dapat menyelesaikan masalah khususnya sampel data besar yang di dapat di uci data set dan Algoritma PSO dapat meningkatkan bobot atribut dan meningkatkan akurasi suatu algoritma dan klasifikasi data yang lebih besar.

BAHAN DAN METODE

Penelitian yang dilakukan menggunakan penelitian eksperimen. Penelitian eksperimen melibatkan penyelidikan perlakuan pada parameter atau variabel tergantung dari penelitian dan menggunakan tes yang dikendalikan oleh si peneliti. Salah satu masalah yang paling penting untuk penelitian prediksi pemasaran bank adalah penggunaan kumpulan data yang bersifat pribadi suatu bank, beberapa bank mengembangkan suatu model untuk memprediksi pemasaran dengan menggunakan kepemilikan dan mempresentasikan model ini dalam konferensi, namun tidak mudah membandingkan Ridwansyah seperti itu dengan hasil model kita sendiri karena untuk mendapatkan suatu data mereka sulit didapatkan, maka dari itu kami mencoba beberapa metode penelitian sebagai berikut:

1. Pengumpulan data yang digunakan adalah data sekunder yang didapat dari data UCI dengan jumlah data sebanyak 16384 *record,* dengan variabel sebagai penentu terdiri dari 20 variabel pemasaran bank seperti, *age, job, marital, education, default, housing, loan, contact, month, day of week, duration, campaign, pdays, previous, poutcome, employment variation rate, consumer price index, confidence price index, euribor 3 month rate, dan nr. Employed*. Dan 1 variabel tujuan yaitu *yes* atau *no*.
2. Pengolan data awal dengan data yang didapat di validasi, dintegrasi dan transformasi, setelah itu pengurangan ukuran dan diskritisasi.
3. Model yang kami diusulkan dalam penelitian ini dengan menerapkan neural network berbasis particle swarm optimization, pada gambar 1.



Sumber: (Ridwansyah & Purwaningsih, 2018)

Gambar1. Model Penelitian Neural Network berbasis Particle Swarm Optimization

Menggambarkan metode yang disulkan dalam penelitian ini yaitu metode particle swarm optimization pada neural network yang menunjukan diagram aktifitas. Pada pengolahan data set awal, data set berupa data pengujian (*testing*) dan pembelajaran (*training*), setiap partikel mewakili subset fitur yang merupakan solusi dari kandidat data set ditransformasikan kedalam range 0 dan 1, kemudian data set dibagi dengan metode 10-fold cross validation, setelah itu data set diuji dengan model neural network dan memvalidasi model yang dihasilkan untuk mendapatkan akurasi model, jika kriteria akurasi terpenuhi maka menghasilkan fitur yang di optimalkan, jika kriteria akurasi tidak terpenuhi maka mencari atribut yang tidak di optimal, sehingga menemukan atribut yang tidak optimal dan menghilangkan atribut yang tidak optimal tersebut sehingga model optimasi neural network didapatkan, sehingga dilakukan pengujian ulang dengan model neural network.

Model validasi yang kami gunakan adalah 10-cross-validasi untuk mempelajari dan menguji data. Kami membagi data pelatihan menjadi 10 bagian yang sama dan kemudian melakukan proses pembelajaran sebanyak 10 kali, setiap kali kami memilih bagian lain dari data set untuk pengujian (*testing*) dan menggunakan sembilan bagian yang tersisa untuk pembelajaran (*training*). Kami menggunakan 10-cross-validation, lihat tabel 1:

Tabel 1. Ilustrasi 10 *Fold Cross Validation*

Validasi Ke-n	Data Set
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Sumber: (Ridwansyah & Purwaningsih, 2018)

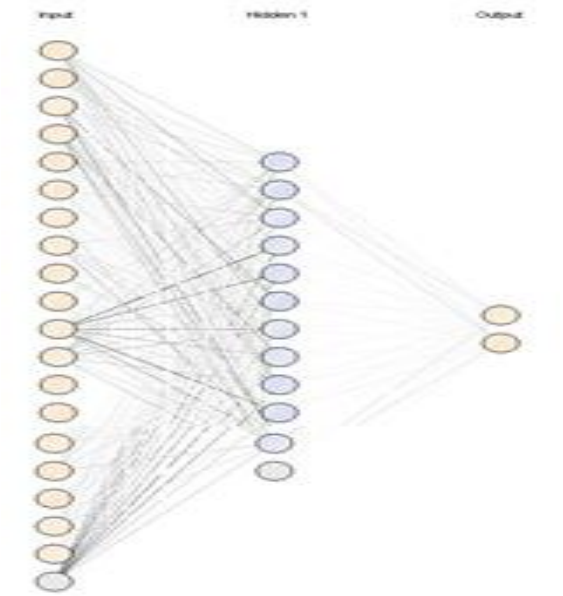
Metode 10 *Fold Cross Validation* telah menjadi metode standar untuk pembelajaran dan pengujian data. Penggunaan data set bersifat publik membuat penelitian yang dapat di teliti ulang, tidak dapat dipungkiri, dan dapat diverifikasi(Catal & Diri, 2009).

Model evaluasi yang kami terapkan yaitu area under curve (AUC) sebagai indicator akurasi dalam eksperimen kami untuk mengevaluasi kinerja klasifikasi *neural network.* AUC memiliki potensi untuk secara signifikan memperbaiki pemusatan di setiap eksperimen empiris dalam pemasaran bank dan memiliki interpretasi statistik yang jelas. Panduan untuk mengklasifikasikan keakuratan uji diagnosa menggunakan AUC adalah sistem tradisional yang disajikan dibawah ini(Gorunescu, 2011).

0.90 – 1.00 = *Excellent Classification*
0.80 – 0.90 = *Good Classification*
0.70 – 0.80 = *Fair Classification*
0.60 – 0.70 = *Poor Classification*
0.50 – 0.60 = *Failure*

HASIL DAN PEMBAHASAN

Setelah melakukan hasil eksperimen yang kami lakukan pada data pemasaran bank yang didapat dari UCI *machine learning repository* dengan penerapan model *neural network* berdasarkan nilai akurasi dalam penelitian kami maka terbentuk arsitektur *neural network* seperti pada Gambar2.

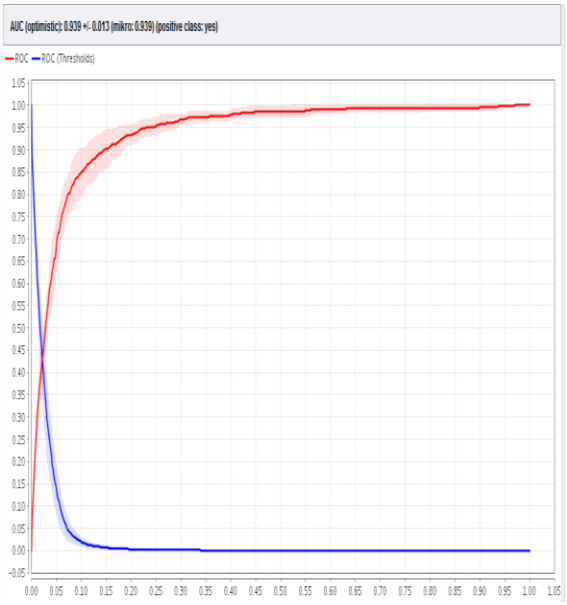


Sumber: (Ridwansyah & Purwaningsih, 2018)
Gambar2. Arsitektur *Neural Network*

Pada desain model algoritma *neural network* diatas dilakukan proses training model dengan memberikan:

1. *Hidden Layer Size* : 9
2. *Training Cycles* : 500
3. *Learning Rate* : 0.3
4. *Momentum* : 0.2

Proses training seperti itu mendapatkan nilai *accuracy* sebesar 95.67% dan grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0,939 dengan diagnosa hasilnya *Excellent classification* lihat Gambar3.



Sumber: (Ridwansyah & Purwaningsih, 2018)
Gambar3. nilai AUC dalam Grafik ROC algoritma *neural network*

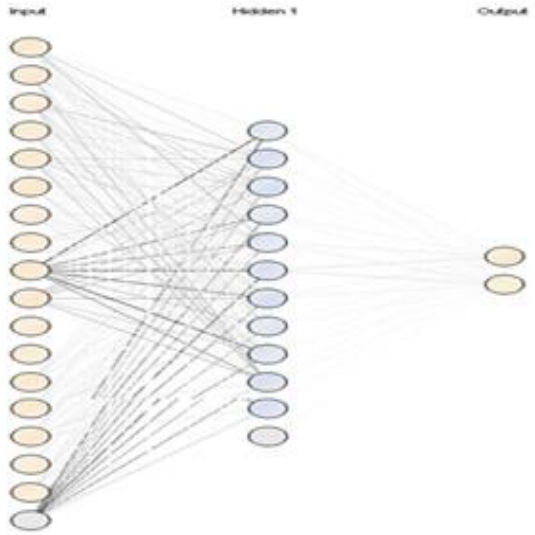
Berdasarkan hasil dari uji coba dengan menggunakan algoritma *neural network* maka data *training* tersebut akan diseleksi dengan attribut yang digunakan yaitu *age, job, default, housing, loan, contact, month, day of week, duration, campaign, previous, poutcome, emp var rate, cons price, cons conf,* marital, education, pdays, euribor3m *dan nr employee* dan 1 atribut sebagai label yaitu hasil.

Tabel2. Bobot Atribut Metode *Neural Network* Berbasis PSO

Age	0,671
Job	0,984
Marital	0
Education	0
Default	0,136
Housing	0,533
Loan	0,617
Contact	0,988
Month	0,918
Day_of_week	1
Duration	0,321
Campaign	1
Pdays	0
Previous	0,613
Poutcome	1
Emp_var_ra	0,001
Cons_price	0,195
Cons_conf	0,617
Euribor3m	0
Nr_employee	0,058

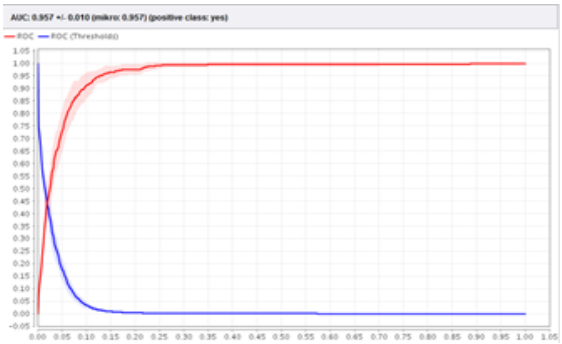
Sumber: (Ridwansyah & Purwaningsih, 2018)

Dari pengujian *neural network* berbasis *particle swarm optimization* yang diperoleh ada beberapa atribut yang berpengaruh terhadap bobot atribut yaitu: *age, job, default, housing, loan, contact, month, day of week, duration, campaign, previous, poutcome, emp var rate, cons price, cons conf dan nr employee*. beberapa atribut yang tidak berpengaruh terhadap bobot atribut yaitu: *marital, education, pdays dan euribor3m.* maka terbentuk arsitektur *neural network* seperti pada Gambar5.



Sumber: (Ridwansyah & Purwaningsih, 2018)
Gambar5. Arsitektur *Neural Network* Berbasis *Particle Swarm Optimization*

Setelah melakukan hasil eksperimen yang di lakukan pada data pemasaran bank yang didapat dari UCI *machine learning repository* dengan penerapan model *neural network* berbasis *particle swarma optimization* maka atribut yang tidak berpengaruh dihilangkan dan menghasilkan peningkatan akurasi dengan nilai akurasi 95,69% dan grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0,957 dengan diagnosa hasilnya tetap *Excellent classification* lihat Gambar3.



Sumber: (Ridwansyah & Purwaningsih, 2018)
Gambar6. nilai AUC dalam Grafik ROC algoritma *neural network* berbasis PSO

Pengujian model setelah melakukan eksperimen dengan penerapan *neural network* dan *neural network* berbasis *particle swarm optimization* dengan membandingkan nilai *accuracy* dan AUC, dalam menentukan tingkat keakurasian dari perbandingan tersebut maka confusion matrix yang didapat baik model *neural network* itu sendiri ataupun *neural network* berbasis *particle swarm optimization* berdasarkan 16384 data yang diolah, dengan hasil perbandigan pada tabel 3 dan tabel 4.

Tabel2. Nilai Akurasi dengan NN

Accuracy: 95,67%			
	true no	*true*_yes	*class precision*
pred.	15397	425	97.31%
*pred.*yes	285	276	49.20%
class recall	98.18%	39.37%	

Sumber: (Ridwansyah & Purwaningsih, 2018)

$$Akurasi = \frac{(\mathrm{TN} + \mathrm{TP})}{(\mathrm{TN} + \mathrm{FN} + \mathrm{TP} + \mathrm{FP})}$$

$$= \frac{(15397 + 276)}{(15397 + 285 + 276 + 425)}$$

$$= \mathbf{0{,}9567}$$

Berdasarkan Tabel2 dapat diklasifikasikan sebagai berikut:
T*rue positive* (tp) = 15397 *record*
False Negative (fn) = 425 *record*
True negative (tn) = 276 *record*
False Negative (fp) = 285 *record*
dari hasil klasifikasi menunjukkan bahwa, tingkat akurasi dengan menggunakan algoritma *neural network* adalah sebesar 95,67%, dan dapat dihitung untuk mencari nilai *accuracy, sensitivity, specificity, ppv* dan *npv* hasilnya pada persamaan berikut:

Tabel3. Nilai Akurasi dengan NN+PSO

Accuracy: 95,69%			
	true no	*true*_yes	*class precision*
pred.	15444	468	97.06%
*pred.*yes	238	233	49.47%
class recall	98.48%	33.24%	

Sumber: (Ridwansyah & Purwaningsih, 2018)

$$Akurasi = \frac{(\mathrm{TN} + \mathrm{TP})}{(\mathrm{TN} + \mathrm{FN} + \mathrm{TP} + \mathrm{FP})}$$

$$= \frac{(15444 + 233)}{(15444 + 238 + 233 + 468)}$$

$$= \mathbf{0{,}9569}$$

Berdasarkan Tabel3 dapat diklasifikasikan sebagai berikut:
T*rue positive* (tp) = 15444 *record*
False Negative (fn) = 468 *record*
True negative (tn) = 233 *record*
False Negative (fp) = 238 *record*
dari hasil klasifikasi menunjukkan bahwa, tingkat akurasi dengan menggunakan algoritma *neural network* adalah sebesar 95,69%, dan dapat dihitung untuk mencari nilai *accuracy, sensitivity, specificity, ppv* dan *npv* hasilnya pada persamaan berikut:

KESIMPULAN

Metode yang digunakan dalam makalah ini untuk mengintegrasikan teknik optimasi partikel dengan meningkatkan akurasi pemasaran bank. Partikel swarm optimasi diterapkan untuk mengatasi masalah yang ada pada *neural network* dengan melakukan analisa data UCI dengan jumlah data sebanyak 16384 *record*. Hasil percobaan menunjukan bahwa metode yang diusulkan *neural network* dengan *particle swarm optimization* membuat peningkatan kinerja prediksi yang mengesankan untuk pengklasifikasian data dibandingkan dengan metode *neural network* yang sebelumnya dan dengan hasil akurasi tersebut akan membantu pihak bank dalam menentukan parameter mana saja yang dibutuhkan untuk kepentingan dalam pengambilan keputusan sehingga dapat memprediksi sebuah bank akan mengalami kerugian yang mengakibatkan bangkrut atau tidak.

REFERENSI

Agrawal, R., Ghosh, S., Imielinski, T., Iyer, B., & Swami, A. (1992). An Interval Classi er for Database Mining Applications. *Proceedings of the 18th VLDB Conference Vancouver*.

Cao, J., Cui, H., Shi, H., & Jiao, L. (2016). Big data: A parallel particle swarm optimization-back-propagation neural network algorithm based on MapReduce. *PLoS ONE*, *11*(6), 1–17.

Catal, C., & Diri, B. (2009). No Title. *Information Sciences*, *179*(8), 1040–1058.

Gorunescu, F. (2011). *Data Mining: Concepts,*

Models and Techniques. *Data mining - Concepts, Models and Technique*.

Ling, C., & Li, C. (1998). Data Mining for Direct Marketing: Problems and Solutions. *Proceedings of KDD 98'*, 73–79.

Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, *62*, 22–31. Elsevier B.V.

Moro, S., & Laureano, R. M. S. (2011). Using Data Mining for Bank Direct Marketing: An application of the CRISP-DM methodology. *European Simulation and Modelling Conference*, (Figure 1), 117–121.

Ridwansyah, R., & Purwaningsih, E. (2018). *Laporan Akhir Penelitian PDY*. Jakarta.

Saputra, E. P. (2017). Prediksi Keberhasilan Telemarketing Bank Untuk. *Jurnal Ilmu Pengetahuan dan Teknologi Komputer*, *2*(2), 66–72.

Satapathy, S. C., Chittineni, S., Mohan Krishna, S., Murthy, J. V. R., & Prasad Reddy, P. V. G. D. (2012). Kalman particle swarm optimized polynomials for data classification. *Applied Mathematical Modelling*, *36*(1), 115–126. Elsevier Inc.

Vieira, S. M., Mendonça, L. F., Farinha, G. J., & Sousa, J. M. C. (2013). Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients. *Applied Soft Computing*, *13*(8), 3494–3504.

Yusup, N., Zain, A. M., & Hashim, S. Z. M. (2012). Overview of PSO for optimizing process parameters of machining. *Procedia Engineering*, *29*, 914–923.