

Optimasi Metode Particle Swarm Optimization (PSO) Pada Prediksi Penilaian Apartemen

Lala Nilawati¹, Yuni Eka Achyani²

¹Universitas Bina Sarana Informatika
e-mail: lala.lni@bsi.ac.id

²Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri
e-mail: yuni.yea@nusamandiri.ac.id

Cara Sitasi: Nilawati, L., & Achyani, Y. E. (2019). Optimasi Metode Particle Swarm Optimization (PSO) Pada Prediksi Penilaian Apartemen. *Paradigma - Jurnal Komputer dan Informatika*, 21(2), 227-234. doi:10.31294/p.v21i2.6159

Abstract - One property that is currently being glimpsed by investors is an apartment. Property consulting companies as one of the service provider companies that become a link between apartment owners and apartment enthusiasts, have an important task in terms of providing information about the assessment of the offered institutions. This study will conduct a trial on the accuracy of apartment assessment predictions using the Support Vector Machine (SVM) method, then will be compared again with the results of the accuracy of the assessment method Support Vector Machine (SVM) combined using the optimization method Particle Swarm Optimization (PSO). The results of the combination of the application of SVM and PSO are used to optimize attribute selection in apartment valuation to improve the accuracy of using the SVM method. This study shows that the Particle Swarm Optimization (PSO) Support Vector Machine (SVM) method is a pretty good method of data classification, because it can be seen from the increase in accuracy of 2.84% and AUC of 0.003. Subjects (attributes) that affect apartment valuation are seen from rent prices (price range), city (apartment location), size (area), furnishing (equipment), bedroom (number of bedrooms), bathroom (number of bathrooms) and maid bedroom (number of maid rooms). The results of the attribute testing showed that city attributes (apartment locations), furnishing (equipment) and maid bedroom (number of maid rooms) greatly influenced the valuation of an apartment.

Keywords: Support Vector Machine (SVM), Particle Swarm Optimization (PSO), Apartment Rating

PENDAHULUAN

Apartemen merupakan salah satu jenis properti yang saat ini banyak dilirik oleh para investor, sehingga banyak para perusahaan konsultan properti berlomba-lomba dalam menawarkan apartemen yang terbaik yang dimiliki. Perusahaan konsultan properti selalu melakukan penilaian untuk semua jenis properti yang ditawarkan, tidak terkecuali jenis apartemen. Menurut Arikunto menyebutkan bahwa "Penilaian adalah mengambil suatu keputusan terhadap sesuatu dengan ukuran baik buruk dan penilaian bersifat kualitatif" (Arikunto, 2009). Salah satu faktor untuk mengetahui nilai pasar apartemen bisa dilihat melalui pendekatan *sales comparison approach* (pendekatan data pasar). Pendekatan *sales comparison approach* (pendekatan data pasar) adalah salah satu pendekatan yang banyak digunakan dalam penilaian properti karena mudah digunakan dan data yang diperoleh berasal langsung dari pasar (Janssen, 2003). Teknik regresi adalah salah satu teknik yang digunakan dalam mengurangi subjektivitas pada pendekatan *sales comparison approach*, teknik ini menentukan

faktor-faktor fisik yang mempengaruhi nilai pasar apartemen yaitu lokasi, fasilitas dan fisik bangunan apartemen (Betts & Elly, 2001). Irooth dan Anastasia dalam penelitian mengenai penilaian apartemen di Surabaya menyimpulkan bahwa variabel *view* taman, *view* golf, jumlah kamar, luas, akses ke pusat perbelanjaan, akses ke rumah sakit, akses ke sekolah, dan akses ke tempat kerja berpengaruh pada nilai apartemen (Irooth & Anastasia, 2017). Pada penelitian ini subjek-subjek (atribut) yang mempengaruhi penilaian apartemen dilihat dari *rent price* (range harga), *city* (lokasi apartemen), *size* (luas), *furnishing* (perlengkapan), *bedroom* (jumlah kamar tidur), *bathroom* (jumlah kamar mandi) dan *maid bedroom* (jumlah kamar pembantu). Prediksi penilaian apartemen ini akan diuji lewat perbandingan hasil akurasi penilaian, yang diperoleh menggunakan metode *Support Vector Machine* (SVM), kemudian akan dibandingkan lagi dengan hasil akurasi penilaian metode *Support Vector Machine* (SVM) yang dikombinasikan menggunakan optimasi metode *Particle Swarm Optimization* (PSO). Hasil kombinasi penerapan

SVM dan PSO digunakan untuk optimasi seleksi atribut dalam penilaian apartemen untuk meningkatkan akurasi penggunaan metode SVM.

Pada penelitian yang dilakukan oleh Jun, dkk menggabungkan algoritma CART dan *Particle Swarm Optimization* (PSO) adaptif sebagai alat optimasi yang memungkinkan bisa menjadi alternatif yang baik, dalam mengoptimalkan aturan klasifikasi berbasis pohon keputusan (Jun, Cho, & Lee, 2013). Algoritma C4.5 yang berbasis PSO sudah diuji dan memperoleh akurasi yang paling tinggi, dengan nilai akurasi mencapai 89,09% dalam prediksi untuk pencapaian hafalan Al-Qur'an di Pesantren Tahfih Daarul Qur'an (Muchlis & Pahlevi, 2018). Penggunaan *Support Vector Machine* berbasis *Particle Swarm Optimization* untuk data yang terkait kampanye pemasaran langsung, dari lembaga perbankan Portugis yang diperoleh dari UCI Repository tahun 2012, didapat nilai akurasi sebesar 89,38% dan ini membuktikan ada kenaikan nilai akurasi dari yang sebelumnya sebesar 88,71% dibandingkan dengan metode *Support Vector Machine* tunggal (Achyani, 2018). Penelitian Noor dalam memprediksi siswa bermasalah, menggunakan perbandingan algoritma C4.5 dengan C4.5 berbasis PSO (*Particle Swarm Optimazion*) menunjukkan bahwa kombinasi C4.5 dan *Particle Swarm Optimization* menjadi lebih baik dalam memprediksi nilai akurasi daripada menggunakan hanya metode C4.5 saja, yang mampu meningkatkan nilai akurasi cukup tinggi yaitu sebesar 35,7% (Noor, 2018). *Particle Swarm Optimization* (PSO) dikombinasikan dengan algoritma *Naive Bayes* dalam memprediksi kredit macet pada koperasi memberikan nilai akurasi sebesar 89% dengan diagnosa klasifikasi baik (*good classification*). Performansi kinerja metode PSO-*Naive Bayes* dengan *Sample Stratified* menghasilkan peningkatan akurasi sebesar 21,03% dan AUC sebesar 0,069 dibandingkan dengan metode *Naive Bayes* tunggal dan menjadi solusi pemecahan permasalahan dalam memprediksi kredit macet pada koperasi (Mardiana, 2018).

Data Mining

Data mining atau lebih di kenal juga dengan sebutan *Knowledge Discovery in Databases* (KDD) merupakan metode untuk memperoleh pengetahuan dari *database* yang ada. Istilah *Knowledge Discovery in Database* (KDD) dan *data mining* seringkali digunakan secara bergantian untuk menjelaskan proses penggalan informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain, dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining* (Bramer, 2007). *Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan (Bramer, 2007), yaitu:

1. *Description* (Deskripsi)

Yaitu proses untuk menemukan suatu karakteristik penting dari data dalam suatu basis data. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan. Tujuan dari Deskripsi adalah mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan/kriteria yang dapat mudah dimengerti.

2. *Classification* (Klasifikasi)

Dalam klasifikasi, terdapat target variabel kategori. Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya.

3. *Estimation* (Estimasi)

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

4. *Prediction* (Prediksi)

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa mendatang.

5. *Clustering* (Pengklusteran)

Pengklusteran merupakan pengelompokan yang memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan *record* atau pengamatan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu *cluster* dan semakin besar perbedaan tiap *cluster* maka kualitas analisis *cluster* semakin baik.

6. *Association* (Asosiasi)

Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut.

K-Fold Cross Validation

Cross Validation (CV) adalah salah satu metode statistik yang digunakan dengan tujuan untuk mengevaluasi kinerja model atau algoritma, dimana data dipisahkan menjadi dua data *subset* yaitu data proses pembelajaran dan data validasi/evaluasi (Witten, Frank, & Hall, 2011). Jenis *Cross Validation* (CV)-*K Fold* dapat dipilih berdasarkan pada ukuran *dataset*, dan biasanya digunakan karena dapat mengurangi waktu komputasi dengan tetap menjaga keakuratan estimasi. *Cross Validation* (CV) hal

utamanya dapat digunakan untuk melakukan prediksi model dan memperkirakan seberapa akurat sebuah model prediktif ketika dijalankan dalam praktiknya. Pada permasalahan prediksi sebuah model biasanya diberikan *dataset* yang nantinya digunakan dalam menjalankan pelatihan (*dataset* pelatihan), serta kumpulan data yang tidak diketahui (atau data yang pertama kali dilihat) terhadap model yang diuji (pengujian *dataset*). Metode *3-Fold Cross Validation* (CV) membagi sebuah himpunan secara acak menjadi tiga *subset* yang saling bebas. Cara kerja metode ini akan melakukan pengulangan sebanyak tiga kali untuk pelatihan dan pengujian. Pada setiap ulangan, disisakan satu *subset* untuk pengujian dan *subset* lainnya untuk pelatihan. Tingkat akurasi dihitung dengan membagi jumlah keseluruhan klasifikasi yang benar dengan jumlah semua instance pada data awal.

Support Vector Machine (SVM)

Support Vector Machine (SVM) dikenal juga dengan *support vector network* yang merupakan metode *supervised* terkait dengan *learning algorithm* untuk analisa pola data yang digunakan untuk klasifikasi dan *regresi* (Khan, Badruddin, & Bashier, 2007). SVM merupakan usaha mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dua buah *class* pada input *space* (Nugroho, Witarto, & Handoko, 2003). SVM saat pertama kali diperkenalkan oleh Vapnik pada tahun 1992, hanya dapat mengklasifikasikan data ke dalam dua kelas (klasifikasi biner). Secara prinsip SVM adalah *linear classifier* bidang dalam komputer sains, yang memetakan suatu data ke dalam konsep tertentu yang telah didefinisikan sebelumnya (*Pattern Recognition*) (Nugroho et al., 2003). Aplikasi *pattern recognition* sangat luas, di antaranya mengenali suara dalam sistem sekuriti, membaca huruf dalam OCR, mengklasifikasikan penyakit secara otomatis berdasarkan hasil diagnosa kondisi medis pasien dan sebagainya. *Pattern recognition* dilakukan dengan mentransformasikan data pada input *space* ke ruang yang berdimensi lebih tinggi, dan optimisasi dilakukan pada ruang *vector* yang baru tersebut. Hal ini membedakan SVM dari solusi *pattern recognition* pada umumnya, yang melakukan optimisasi parameter pada ruang hasil transformasi yang berdimensi lebih rendah daripada dimensi input *space*. SVM menerapkan strategi *Structural Risk Minimization* (SRM) yang bertujuan menemukan *hyperplane* terbaik yang memisahkan dua buah *class* pada input *space*. Beberapa kelebihan SVM dapat dilihat sebagai berikut:

1. Generalisasi

Yaitu kemampuan suatu metode (SVM, *neural network*, dsb.) untuk mengklasifikasikan suatu *pattern*, yang tidak termasuk data yang dipakai dalam fase pembelajaran metode itu. Strategi pembelajaran pada SVM meminimalkan *error* pada *training-set*, juga meminimalkan faktor kedua. Strategi ini disebut *Structural Risk Minimization* (SRM), dan dalam SVM diwujudkan dengan memilih *hyperplane* dengan margin terbesar. Berbagai studi empiris menunjukkan bahwa pendekatan SRM pada SVM

memberikan *error generalisasi* yang lebih kecil daripada yang diperoleh dari strategi ERM pada *neural network* maupun metode yang lain.

2. Curse of dimensionality

Didefinisikan sebagai masalah yang dihadapi suatu metode *pattern recognition* dalam mengestimasi parameter. *Curse of dimensionality* sering dialami dalam aplikasi di bidang *biomedical engineering*, karena biasanya data biologi yang tersedia sangat terbatas, dan penyediaannya memerlukan biaya tinggi. SVM merupakan salah satu metode yang tepat dipakai untuk memecahkan masalah berdimensi tinggi, dalam keterbatasan sampel data yang ada (Vapnik, 1999).

3. Landasan teori

SVM memiliki landasan teori yang dapat dianalisa dengan jelas, dan tidak bersifat *black box*.

4. Feasibility

SVM dapat diimplementasikan relatif mudah, karena proses penentuan *support vector* dapat dirumuskan dalam *QP problem*. Jika kita memiliki *library* untuk menyelesaikan *QP problem*, dengan sendirinya SVM dapat diimplementasikan dengan mudah.

Untuk kelemahan nya bisa dilihat jika SVM Sulit dipakai dalam problem berskala besar. Skala besar dalam hal ini dimaksudkan dengan jumlah *sample* yang diolah. Selain itu walaupun SVM telah dimodifikasi agar dapat menyelesaikan masalah dengan *class* lebih dari dua, antara lain strategi *One versus rest* dan *strategi Tree Structure*, tetapi masing-masing strategi ini memiliki kelemahan dan perlu pengembangan lebih lanjut.

Particle Swarm Optimization (PSO)

Metode PSO diperkenalkan oleh Dokter Kennedy dan Elbert pada tahun 1995 berdasarkan penelitian yang dilakukan terhadap perilaku kawanan burung dan ikan dan merupakan metode optimasi *heuristic global* (Bai, 2010). PSO adalah algoritma iteratif berbasis populasi. Populasi terdiri dari banyak partikel, di mana diinisialisasi dengan populasi solusi acak dan digunakan untuk memecahkan masalah optimasi (Abraham, Grosan, & Ramos, 2006). Setiap partikel mewakili solusi kandidat dan bergerak menuju posisi optimal dengan mengubah posisinya sesuai dengan kecepatan partikel terbang melalui ruang pencarian dengan kecepatan yang dinamis disesuaikan untuk perilaku historis. Oleh karena itu, partikel memiliki kecenderungan untuk terbang menuju daerah pencarian yang lebih baik dan lebih baik selama proses pencarian (Abraham et al., 2006). Rumus untuk menghitung perpindahan posisi dan kecepatan partikel yaitu:

$$V_i(t) = V_i(t-1) + c_1 r_1 [X_{pbest} - X_i(t)] + c_2 r_2 [X_{gbest} - X_i(t)] \dots (1)$$

$$X_i(t) = X_i(t-1) + V_i(t) \dots \dots \dots (2)$$

Dimana:

$V_i(t)$ = kecepatan partikel i saat iterasi t

$X_i(t)$ = posisi partikel i saat iterasi t

c_1 dan $c_2 = learning\ rates$ untuk kemampuan individu (*cognitive*) dan pengaruh sosial (*group*)

r_1 dan $r_2 = bilangan\ random$ yang berdistribusi *uniformal* dalam *interval* 0 dan 1

X_{pbest} = posisi terbaik partikel i

X_{gbest} = posisi terbaik global

Confusion matrix

Confusion matrix digunakan untuk melakukan evaluasi terhadap model klasifikasi berdasarkan perhitungan objek testing yang ditabulasikan kedalam tabel dimana akan diprediksi benar dan tidak benar (Gorunescu, 2011). *Confusion matrix* merupakan data set hanya memiliki dua kelas, kelas

CLASSIFICATION	PREDICTED CLASS		
	Class = YES	Class = NO	
OBSERVED CLASS	Class = YES	a (<i>true positive-TP</i>)	b (<i>false negative-FN</i>)
	Class = NO	c (<i>false positive-FP</i>)	d (<i>true negative-TN</i>)

yang satu sebagai positif dan kelas yang lain sebagai negatif. Terdiri dari empat sel yaitu *True Positives* (TP), *False Positives* (FP), *True Negatives* (TN) dan *False Negatives* (FN) (Bramer, 2007).

Gambar 1 Confusion matrix untuk 2 model kelas
Sumber: (Gorunescu, 2011)

Kurva ROC

ROC mengekspresikan *confusion matrix* (Vercellis, 2009). Kurva ROC digunakan untuk menunjukkan visualisasi dari akurasi model dan perbandingan perbedaan antar model klasifikasi. ROC adalah grafik dua dimensi dengan *false positives* sebagai garis horizontal dan *true positives* untuk mengukur perbedaan performansi metode yang digunakan. Kurva ROC adalah teknik untuk memvisualisasi dan menguji kinerja pengklasifikasian berdasarkan performannya (Gorunescu, 2011). Model klasifikasi yang lebih baik adalah yang mempunyai kurva ROC lebih besar (Vercellis, 2009). Performa keakurasian AUC dapat diklasifikasikan menjadi lima kelompok (Gorunescu, 2011) yaitu:

- a. 0.90 – 1.00 = Unggul
- b. 0.80 – 0.90 = Baik
- c. 0.70 – 0.80 = Cukup
- d. 0.60 – 0.70 = Kurang
- e. 0.50 – 0.60 = Gagal

METODOLOGI PENELITIAN

Data yang digunakan pada penelitian ini berasal dari perusahaan konsultan properti yang berlokasi di daerah kebayoran baru pada tahun 2015. Data yang akan diolah berupa data penilaian apartemen berjumlah 1727 record dan terdiri dari 7 atribut, dengan rincian 5 atribut bertipe numerik (*Rent_Price*, *Size*, *Bedroom*, *Bathroom*, dan *Maid_Bathroom*) dan

2 bertipe kategorikal (*City* dan *Furnishing*). Terdapat dua metode yang akan digunakan yaitu metode *Support Vector Machine* (SVM) dan metode *Support Vector Machine* (SVM) yang dikombinasikan dengan metode *Particle Swarm Optimization* (PSO). Untuk membandingkan atau mengkomparasi dalam penelitian ini akan menggunakan *framework RapidMiner* versi 5.3.1 sehingga akan ditemukan algoritma mana yang paling akurat.

Penelitian ini menggunakan metode penelitian eksperimen, yakni melibatkan penyelidikan perlakuan pada parameter atau variabel tergantung dari peneliti itu sendiri, dan menggunakan tes yang dikendalikan oleh si peneliti itu sendiri. Langkah-langkah yang dilakukan pada proses penelitian adalah sebagai berikut:

1. Pengumpulan Data

Ada dua tipe dalam pengumpulan data, yaitu pengumpulan data primer dan pengumpulan data sekunder. Data primer adalah data yang dikumpulkan pertama kali untuk melihat apa yang sesungguhnya terjadi. Sedangkan data sekunder adalah data yang sebelumnya pernah dibuat oleh seseorang baik diterbitkan atau tidak.

2. Pengolahan Data Awal

Pengolahan awal data meliputi proses input data ke format yang dibutuhkan, pengelompokan dan penentuan atribut data, serta pemecahan data (*split*) untuk digunakan dalam proses pembelajaran (*training*) dan pengujian (*testing*).

3. Model yang diusulkan

Pada tahap ini data dianalisis, kemudian dikelompokkan variabel mana yang berhubungan dengan satu sama lainnya. Setelah data dianalisis lalu diterapkan model-model yang sesuai dengan jenis data. Pembagian data kedalam data latihan (*training data*) dan data uji (*testing data*) juga diperlukan untuk pembuatan model.

4. Eksperimen dan Pengujian Model

Pada bagian ini dijelaskan tentang langkah-langkah eksperimen meliputi cara pemilihan arsitektur yang tepat dari model atau metode yang diusulkan sehingga didapatkan hasil yang dapat membuktikan bahwa metode yang digunakan adalah tepat.

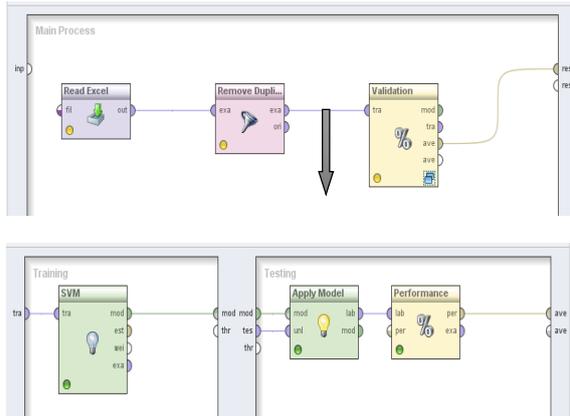
5. Evaluasi dan Validasi Hasil

Pada bagian ini dilakukan evaluasi dan validasi hasil penerapan terhadap model penelitian yang dilakukan untuk mengetahui tingkat keakurasian model.

HASIL DAN PEMBAHASAN

1. Hasil Eksperimen dan Pengujian Metode

Berikut adalah gambar pengujian Algoritma *Support Vector Machine* (SVM) menggunakan metode *K-Fold Cross Validation* dengan menggunakan *RapidMiner*:

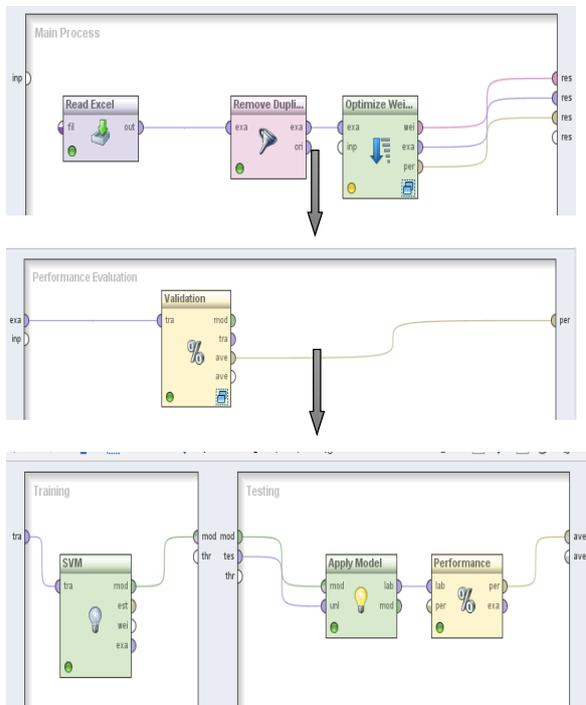


Sumber: Hasil Penelitian (2019)

Gambar 2. Pengujian K-Fold Cross Validation algoritma Support Vector Machine

Hasil terbaik pada eksperimen SVM diatas adalah C=0.0 dan Epsilon=0.0 dihasilkan *accuracy* 79.20% dan AUC 0.770. Lalu diikuti SVM dengan C=1.0 dan Epsilon=0.0 dihasilkan *accuracy* 78.16% dan AUC 0.770, kemudian SVM dengan C=1.0 dan Epsilon=1.0 dihasilkan *accuracy* 73.42% dan AUC 0.500.

Berikut adalah gambar pengujian algoritma *support vector machine* (SVM) berbasis *Particle Swarm Optimization* (PSO) menggunakan metode *K-Fold Cross Validation* dengan menggunakan *RapidMiner*.



Sumber: Hasil Penelitian (2019)

Gambar 3. Pengujian K-Fold Cross Validation algoritma Support Vector Machine berbasis PSO

Nilai *training cycles* dalam penelitian ini ditentukan dengan cara melakukan uji coba memasukkan C dan epsilon. Selanjutnya dilakukan

observasi terhadap variabel C dan ϵ dari *Support Vector Machine* dan *Particle Swarm Optimization*. Berikut ini adalah hasil dari percobaan yang telah dilakukan untuk penentuan nilai *training cycles*.

Hasil terbaik pada eksperimen SVM berbasis PSO diatas adalah dengan C=0.0 dan Epsilon=0.0 dihasilkan *Accuracy* sebesar 82.04% dengan AUC sebesar 0.773. Maka nilai-nilai tersebutlah yang akan digunakan pada penelitian ini.

2. Evaluasi dan Validasi Model

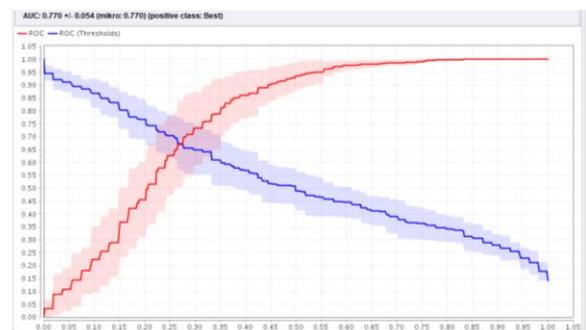
a. Hasil Pengujian Metode Support Vector Machine (SVM)

Gambar 4 menunjukkan hasil dari *confusion matrix* algoritma *Support Vector Machine*. Berdasarkan Gambar 3, dapat diketahui bahwa dari 1.787 data, 193 data diprediksi *Good* sesuai dengan prediksi yang dilakukan dengan metode SVM, kemudian 72 data di prediksi *Good*, tetapi ternyata hasil prediksinya adalah *Best*. Kemudian 1.467 data diprediksi *Best*, sesuai dengan prediksi yang dilakukan dengan metode SVM, dan 364 data diprediksi *Best* tetapi ternyata prediksinya *Good*.

accuracy: 79.20% +/- 10.88% (mikro: 79.20%)			
	true Good	true Best	class precision
pred Good	193	72	72.83%
pred Best	364	1467	80.12%
class recall	34.66%	95.32%	

Sumber: Hasil Penelitian (2019)

Gambar 4. Hasil Pengujian dengan Metode Support Vector Machine (SVM)



Sumber: Hasil Penelitian (2019)

Gambar 5. Kurva ROC

b. Hasil Pengujian Metode Support Vector Machine (SVM) berbasis Particle Swarm Optimization

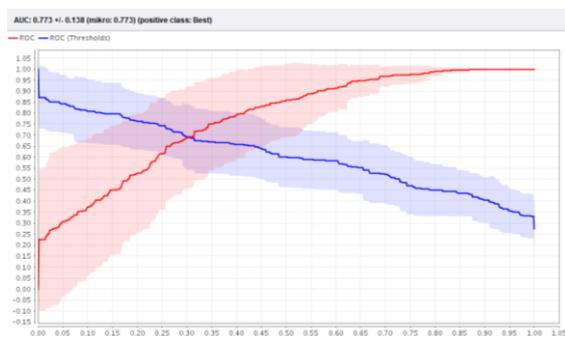
Gambar 6 menunjukkan hasil dari *confusion matrix* algoritma *Support Vector Machine* berbasis *Particle Swarm Optimization* (PSO). Berdasarkan Gambar 5, dapat diketahui bahwa dari

1.787 data, 272 data diprediksi *Good* sesuai dengan prediksi yang dilakukan dengan metode SVM, kemudian 56 data di prediksi *Good*, tetapi ternyata hasil prediksinya adalah *Best*. Kemudian 1194 data diprediksi *Best*, sesuai dengan prediksi yang dilakukan dengan metode SVM, dan 265 data diprediksi *Best* tetapi ternyata prediksinya *Good*.

accuracy: 82.04% +/- 3.33% (mikro: 82.04%)			
	true Good	true Best	class precision
pred. Good	272	56	82.93%
pred. Best	265	1194	81.84%
class recall	50.65%	95.52%	

Sumber: Hasil Penelitian (2019)

Gambar 6. Hasil Pengujian dengan Metode Support Vector Machine (SVM) berbasis Particle Swarm Optimization (PSO)



Sumber: Hasil Penelitian (2019)

Gambar 7. Kurva ROC

Hasil dari *Attribute Weight* yang didapat dari penelitian ini adalah sebagai berikut:

attribute	weight
nil_rent_p	0.705
nil_lota	1
nil_size	0.425
nil_furnis	1
nil_bedroom	0.684
nil_bathr	0.942
nil_maid	1

Sumber: Hasil Penelitian (2019)

Gambar 8. Attribute Weight dengan Algoritma SVM berbasis PSO

3. Analisa Evaluasi dan Validasi Model

Hasil Analisis dari metode Algoritma *Support Vector Machine* dan *Support Vector Machine* berbasis *Particle Swarm Optimization* dirangkum dalam tabel berikut:

Tabel 1. Komparasi Nilai Accuracy dan AUC

	Support Vector Machine	SVM Berbasis PSO	Peningkatan
Accuracy	79.20%	82.04 %	2.84%
AUC	0.770	0.773	0.003

Pada Tabel 1 dapat dilihat bahwa nilai *Accuracy* dan *AUC* Algoritma *Support Vector Machine* berbasis *Particle Swarm Optimization* lebih tinggi dibandingkan dengan penggunaan SVM tunggal. Penerapan Algoritma *Particle Swarm Optimization* itu sendiri menghasilkan peningkatan akurasi sebesar 2.84 % dan *AUC* sebesar 0.003.

Mengingat pentingnya seleksi atribut dalam *Support Vector Machine* (SVM) maka diterapkan *Particle swarm optimization* (PSO) untuk melakukan tugas tersebut. *Particle Swarm Optimization* (PSO) diketahui dapat digunakan sebagai teknik optimasi untuk mengoptimalkan subset fitur. Algoritma PSO sederhana dan memiliki kompleksitas yang lebih rendah, sehingga dapat memastikan solusi optimal dengan menyesuaikan pencarian global dan lokal, sehingga kinerja klasifikasi *Support Vector Machine* (SVM) dapat ditingkatkan. *Eksperiment* dilakukan kembali dengan menerapkan *Particle swarm optimization* (PSO) untuk seleksi atribut dalam *Support Vector Machine* (SVM) dan dilakukan penyesuaian pada parameter C dan ϵ . Dari 7 variabel *predictor* yaitu *rent price* (range harga), *city* (lokasi apartemen), *size* (luas), *furnisihing* (perlengkapan), *bedroom* (jumlah kamar tidur), *bathroom* (jumlah kamar mandi) dan *maid badroom* (jumlah kamar pembantu) dilakukan seleksi atribut sehingga menghasilkan terpilihnya 3 atribut yang digunakan dan sangat mempengaruhi penilaian sebuah apartemen yaitu *city* (lokasi apartemen), *furnisihing* (perlengkapan) dan *maid badroom* (jumlah kamar pembantu).

KESIMPULAN

Hasil dari penelitian ini mempunyai implikasi terhadap penilaian apartemen yakni meliputi implikasi terhadap aspek sistem penunjang keputusan penilaian apartemen, aspek manajerial, dan aspek terhadap penelitian-penelitian selanjutnya sebagai berikut:

1. Implikasi terhadap aspek sistem pendukung keputusan penilaian apartemen penerapan *Particle Swarm Optimization* (PSO) untuk seleksi atribut dalam *Support Vector Machine* (SVM) yang dilakukan dengan penyesuaian pada parameter C dan ϵ mampu menunjukkan aspek berupa atribut yang berpengaruh terhadap hasil dari prediksi penilaian

apartemen. Dari hasil ini menunjukkan bahwa metode *Support Vector Machine* berbasis *Particle Swarm Optimization* merupakan metode yang cukup baik dalam pengklasifikasian data. Dengan demikian metode *Support Vector Machine* berbasis *Particle swarm optimization* dapat memberikan pemecahan permasalahan analisa penilaian prediksi penilaian apartemen dan dapat mendukung pengambilan keputusan serta pengembangan sistem informasi manajemen pada Perusahaan Konsultasi Properti.

2. Implikasi terhadap aspek manajerial
Berdasarkan hasil penelitian dapat diketahui bahwa metode *Support Vector Machine* berbasis *Particle Swarm Optimization* dapat mendukung pengambilan keputusan dan pengembangan sistem informasi manajemen pada Perusahaan Konsultasi Properti, untuk itu diperlukan peningkatan kemampuan *manajerial* dalam hal penilaian apartemen agar mampu membuat keputusan yang baik.
3. Implikasi terhadap penelitian-penelitian sebelumnya
Untuk mengetahui kehandalan metode, maka pada penelitian selanjutnya dapat dilakukan penggunaan data set lebih dari satu. Penelitian seperti ini dapat dikembangkan pada unit bisnis serupa atau yang lain. Pengembangan dapat dilakukan dengan pemilihan parameter menggunakan metode *Genetic Algorithm* untuk meningkatkan hasil optimasi. Penggunaan algoritma yang lain juga dapat digunakan misalkan dengan metode *Neural Network*, C4.5, *K-Nearest Neighbor* dan sebagainya.

REFERENSI

- Abraham, A., Grosan, C., & Ramos, V. (2006). *Swarm Intelligence in Data Mining*. London: Verlag Berlin Heidelberg, Springer.
- Achyani, Y. E. (2018). Penerapan Metode Particle Swarm Optimization Pada Optimasi Prediksi Pemasaran Langsung. *Jurnal Informatika*, 5(1), 1–11. <https://doi.org/10.31311/ji.v5i1.2736>
- Arikunto, S. (2009). *Dasar-dasar Evaluasi Pendidikan (edisi revisi)*. Jakarta: Bumi Aksara.
- Bai, Q. (2010). Analysis of Particle Swarm Optimization Algorithm. *Computer Dan Informasi Science*, 3(1).
- Betts, R. ., & Elly, J. . (2001). *Basic Real Estate Appraisal Fifth Edition*. New Jersey: Prentice-Hall, Inc Saddle River.
- Bramer, M. (2007). *Principles of Data Mining*. London: Springer.
- Gorunescu, F. (2011). *Data Mining: Concepts and Techniques*. Verlag berlin Heidelberg: Springer.
- Irooth, A. M., & Anastasia, N. (2017). Model Nilai Pasar Apartemen dan Kesediaan Membayar View Apartemen di Surabaya. *MAPPI Insight*, 1(1), 35–40.
- Janssen, C. T. . (2003). A Market Comparison Approach for Apartment Buildings. *The Canadian Appraiser*, 47(2), 32–37.
- Jun, C. H., Cho, Y. J., & Lee, H. (2013). Improving Tree-Based Classification Rules Using a Particle Swarm Optimization. *IFIP Advances in Information and Communication Technology*, 398(PART 2), 9–16. https://doi.org/10.1007/978-3-642-40361-3_2
- Khan, M. M., Badruddin, M., & Bashier. (2007). *Machine Learning: Algorithms and Applications*. New York: CRC Press.
- Mardiana, T. (2018). Optimasi Naïve Bayes Dengan Particle Swarm Optimization. *Jurnal Riset Informatika*, 1(1), 43–50.
- Muchlis, & Pahlevi, S. M. (2018). Prediksi Pencapaian Hafalan Al-Qur'an Menggunakan Metode C4.5 Berbasis PSO. *Seminar Nasional Sains Dan Teknologi 2018*, 1–4.
- Noor, H. (2018). Optimasi Model Klasifikasi C4.5 Dan Particle Swarm Optimization Untuk Prediksi Siswa Bermasalah. *Technologia : Jurnal Ilmiah*, 9(4), 228–237.
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). *Support Vector Machines : Teori Aplikasinya dalam Bioinformatika*. ilmukomputer.com.
- Vapnik, V. N. (1999). *The Nature of Statistical Learning Theory* (2 nd editi). New York Berlin Heidelberg: Springer-Verlag.
- Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. Southern Gate, Chichester, West Sussex: John Wiley & Sons, Ltd.
- Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning and Tools*. Burlington: Morgan Kaufmann Publisher.

PROFIL PENULIS

Lala Nilawati, M.Kom. Menempuh pendidikan Sarjana (S1) di Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri, Program Studi Sistem Informasi lulus tahun 2010, dan Program Magister (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri, Program Studi Ilmu Komputer lulus tahun 2015. Pernah bekerja menjadi Asisten Laboratorium Komputer Bina Sarana Informatika tahun 2006 sampai tahun 2007, tahun 2008 sampai sekarang aktif bekerja menjadi staff pengajar (dosen) di Universitas Bina Sarana Informatika program studi Sistem Informasi.

Yuni Eka Achyani, S.Kom. Tahun 2014 lulus dari Program Strata Satu (S1) Program Studi Sistem Informasi Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri Nusa Mandiri dan Tahun 2016 lulus dari Program Magister (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri. Saat ini bekerja sebagai Staff Pengajar (Dosen) di Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri.