

PREDIKSI CALON NASABAH GADAI POTENSIAL PADA PT. PEGADAIAN (PERSERO) MENGGUNAKAN *SUPPORT VECTOR MACHINE* DENGAN ALGORITMA GENETIKA

Muhammad Abdul Aziz^{1*}, Rukun Santoso², Budi Warsito³

^{1,2,3} Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro

*e-mail : muhammadazizch09@gmail.com

DOI: 10.14710/J.GAUSS.XX.X.XX-XX

Article Info:

Received:

Accepted:

Available Online:

Keywords:

Prediction; Customers;

PT Pegadaian; Support Vector

Machine; Genetic Algorithm

Abstract: PT Pegadaian is a company that offers business capital loans to customers by providing collateral or a pawn system. Customers who cannot repay loans and are in arrears are considered non-potential. Non-potential customers are harm the company because it reduces capital and profits. Statistical methods are needed to predict potential. Predictions are made using the classification method, in this research is *Support Vector Machine* (SVM). However SVM has a weakness in determining optimal parameters, so optimization is carried out using genetic algorithms to help find optimal classification parameters. Genetic algorithms are used to find optimal solutions to problems by imitating observed processes. This study uses customer data and proof of pawn PT Pegadaian Kebumen on January 3rd 2022 – November 29th 2022. The support vector machine model is formed with a percentage of 80% for training data and 20% for test data. The conclusion in this study was that SVM-GA increased the accuracy obtained from 86,6% being 94,2%.

1. PENDAHULUAN

PT Pegadaian menyediakan pinjaman modal usaha dalam bentuk uang dengan penyerahan barang jaminan dari pelaku usaha atau biasa disebut pinjaman dengan sistem gadai. Setiap orang yang menggunakan atau menerima fasilitas baik dalam bentuk jasa maupun produk Pegadaian dinamakan Nasabah PT Pegadaian. Nasabah yang melakukan gadai diharuskan untuk mengembalikan uang pinjaman sesuai kesepakatan dengan ditambah sewa modal sesuai ketentuan. Nasabah gadai yang tidak melunasi atau menunggak sehingga barang yang dijamin berada dalam keadaan lelang dapat dikatakan sebagai nasabah tidak potensial gadai. Metode statistika diperlukan untuk mengetahui nasabah yang potensial dan tidak potensial. Prediksi adalah proses memperkirakan sesuatu pada masa yang akan datang secara sistematis (Herdianto, 2013). Prediksi dilakukan dengan menggunakan metode klasifikasi sehingga dapat mengolah data-data Pegadaian yang berkaitan tentang nasabah gadai. *Support Vector Machine* (SVM) merupakan metode klasifikasi dengan tingkat akurasi sangat baik karena mampu mengolah data berdimensi tinggi (Siringoringo & Jamaludin, 2019). Akurasi SVM yang baik sangat bergantung pada nilai parameter C (cost) dan γ (gamma). Permasalahan dalam pemilihan parameter yang optimal menjadi fokus pada penelitian ini. Algoritma genetika digunakan sebagai metode optimalisasi dalam membantu pencarian parameter optimal dalam SVM.

Algoritma genetika berjalan untuk menemukan solusi terbaik dari permasalahan dengan meniru proses yang telah diamati. Kriteria dalam mendapatkan solusi yang paling optimal apabila nilai *fitness* telah mencapai konvergensi. Penelitian terhadap algoritma genetika telah dilakukan oleh Harafani dan Maulana (2019) yang meneliti mengenai algoritma genetika sebagai pengoptimasi parameter pada SVM untuk memprediksi kesuburan. Hasil penelitian

yang dilakukan diketahui setelah menggunakan algoritma genetika tingkat akurasi meningkat menjadi 89% dari sebelumnya 88%.

2. TINJAUAN PUSTAKA

PT Pegadaian resmi didirikan sejak 1 April 1901. Pegadaian bertindak sebagai perusahaan keuangan non-bank satu-satunya yang menyediakan penyaluran dana untuk pembiayaan terhadap masyarakat dengan sistem gadai. Pinjaman gadai yang ditawarkan antara lain gadai emas, gadai non emas, gadai kendaraan, gadai tabungan emas, gadai angsuran emas baik konvensional maupun berbasis syariah, dan pembiayaan porsi haji serta pembiayaan wisata religi (Pegadaian, 2022). Nasabah gadai yang memberikan kontribusi dengan menggunakan produk gadai dan selalu memenuhi kewajibannya dalam mencicil kredit, memperpanjang, maupun melunasi pinjaman disebut nasabah potensial gadai. Transaksi gadai antara nasabah dengan perusahaan dibuktikan dengan Surat Bukti Gadai (SBG). SBG merupakan surat perjanjian utang piutang dan keterangan kredit yang berisi rincian pinjaman dan data diri nasabah. Prediksi dilakukan untuk mengetahui nasabah yang potensial sehingga memudahkan perusahaan dalam mengambil keputusan maupun kebijakan. Prediksi dimulai dengan mengambil data yang kemudian dilakukan proses *pre-processing*.

Pre-processing yang dilakukan dalam penelitian ini adalah normalisasi. Normalisasi merupakan proses merubah nilai atribut dalam sebuah dataset dengan memberikan rentang nilai sehingga data dari suatu atribut tidak mendominasi atribut lain. Rentang nilai yang biasanya digunakan yaitu $[-1,1]$ atau $[0,1]$. Penelitian ini menggunakan metode *min-max normalization* dengan rumus persamaan (1):

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

keterangan:

x' = Hasil normalisasi

x = Nilai atribut data

x_{min} = Nilai minimal data

x_{max} = Nilai maksimal data

Setelah melakukan *pre-processing* maka dilakukan proses prediksi yaitu menggunakan klasifikasi *Support Vector Machine* (SVM). Klasifikasi dilakukan dengan membentuk model yang dapat mengelompokkan objek sesuai dengan keadaannya (Elly et al., 2015). Klasifikasi bertujuan untuk membagi dan mengelompokkan objek berdasarkan kesamaan yang dimiliki. Analisis klasifikasi adalah metode untuk menganalisis keterkaitan antara beberapa variabel prediktor dan satu variabel respon yang merupakan variabel kualitatif.

Pengklasifikasian data dilakukan dengan dua tahapan yaitu:

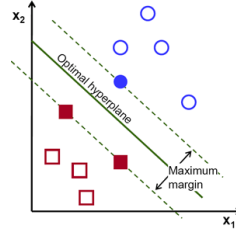
1. Proses *Training*

Data *training* merupakan tahapan pembentukan model yang akan diuji menggunakan data *testing* (Raharjo & Winarko, 2014). *Training* dilakukan untuk melatih algoritma dalam menentukan model yang sesuai.

2. Proses *Testing*

Proses testing diartikan sebagai proses pengujian menggunakan data testing terhadap model yang diperoleh pada proses *training*. *Testing* bertujuan untuk mengetahui dan menguji performa algoritma dari model yang sudah dilatih sebelumnya.

Support Vector Machine (SVM) merupakan algoritma yang dapat digunakan untuk klasifikasi dan regresi dengan cara melakukan prediksi (Santosa, 2007). SVM secara umum dilakukan dengan tujuan untuk memperoleh *hyperplane* fungsi pemisah (*decision boundary*) terbaik pada ruang input.



Gambar 1. Konsep SVM (Santosa, 2007)

Misalkan terdapat set data *training* (\mathbf{x}_i, y_i) , $i=1,2,\dots,l$ dengan data *input* $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\} \in R^n$ yang merupakan data *training* dan *output* yang bersangkutan $Y = \{y_1, \dots, y_l\} \in \{\pm 1\}$ menyatakan label kelas. Himpunan data *training* dari dua kelas memiliki rumus:

$$D = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_l, y_l)\} \quad (2)$$

keterangan:

\mathbf{x}_i = Himpunan dataset ke- i

y_i = Label kelas data ke- i

l = Banyaknya data

Linear SVM digunakan untuk data yang dapat dipisahkan dengan menggunakan satu garis lurus atau dikenal dengan sebutan linier. Rumus *hyperplane* klasifikasi *linear SVM* menurut Santosa (2007):

$$f(\mathbf{x}) = \mathbf{w}^T \cdot \mathbf{x}_i + b \quad (3)$$

dengan \mathbf{w}^T adalah besaran bobot (vektor bobot) sedangkan b adalah bias atau *error*.

Klasifikasi SVM memisahkan kelas menggunakan rumus pertidaksamaan (4)

$$g(\mathbf{x}) = \begin{cases} [(\mathbf{w}^T \cdot \mathbf{x}_i) + b] \geq +1, \text{ untuk } y_i = +1 \\ [(\mathbf{w}^T \cdot \mathbf{x}_i) + b] \leq -1, \text{ untuk } y_i = -1 \end{cases} \quad (4)$$

Dari persamaan (4) didapatkan pertidaksamaan baru untuk klasifikasi kelas data yaitu:

$$y_i[(\mathbf{w}^T \cdot \mathbf{x}_i) + b] \geq 1, \text{ dengan } i = 1, 2, 3, \dots, l \quad (5)$$

Nilai margin dapat dihitung menggunakan rumus panjang garis ke titik pusat:

$$\frac{1-b-(-1-b)}{\|\mathbf{w}\|^2} = \frac{2}{\|\mathbf{w}\|^2} \quad (6)$$

Hyperplane terbaik didapatkan dengan memaksimalkan margin menggunakan metode *Quadratic Programming (QP)* seperti persamaan (7) (Santosa, 2007)

$$\min \tau(\mathbf{w}) = \min \frac{1}{2} \|\mathbf{w}\|^2 \quad (7)$$

Persamaan (6) dapat diselesaikan dengan menggunakan *lagrange multiplier* (Vapnik, 1995). Sehingga didapatkan persamaan (8):

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^l \alpha_i [y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) - 1], \quad i=1, 2, \dots, l \quad (8)$$

dengan a_i merupakan pengganda fungsi *lagrange*. Persamaan (8) dioptimalkan dengan diturunkan terhadap w dan b , dan L memiliki gradien 0.

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^l a_i y_i = 0 \quad (9)$$

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^l a_i x_i y_i = 0 \quad (10)$$

Berdasarkan syarat optimal persamaan (10), maka diperoleh hasil tranformasi rumus dualitas *Lagrange Multiplier* menurut (Hastie et al., 2001) seperti pada persamaan (11):

$$L_d = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j (x_i^T \cdot x_j) \quad (11)$$

Kemudian diperoleh fungsi dual problem dari persamaan (12) yang dimaksimalkan:

$$\max_a L_d = \max \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j (x_i^T \cdot x_j) \quad (12)$$

dengan batas $a_i \geq 0$ untuk $i = 1, 2, \dots, l$ dan $\sum_{i=1}^l a_i y_i = 0$

Pemisahan *hyperplane* memiliki kemungkinan terdapat *outlier* yang menyebabkan kesalahan pelabelan pada kelas data sehingga dibutuhkan *soft margin* dengan menambahkan variabel kendur atau *slack variable* ($\xi_i > 0$). Diperoleh rumus persamaan (Gunn, 1998):

$$y_i[(w^T \cdot x_i) + b] \geq 1 - \xi_i \quad ; i = 1, 2, \dots, l \quad (13)$$

Persamaan memaksimalkan margin yang didapat setelah menggunakan *soft margin*:

$$\min \tau(w, \xi_i) = \min \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \right) \quad (14)$$

dengan nilai C adalah parameter besarnya pinalti yang ditentukan oleh pengguna terhadap kesalahan pada klasifikasi data.

Apabila tidak dapat dipisahkan menggunakan cara linier maka perlu proses pengubahan data menjadi dimensi yang lebih tinggi menggunakan metode *kernel*. Metode *kernel* atau *kernel trick* merupakan fungsi pemetaan data yang dirumuskan pada persamaan (15):

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \quad (15)$$

Penelitian yang dilakukan menggunakan fungsi *kernel Gaussian Radial Basis Function* (RBF) yang memiliki rumus:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (14)$$

Fungsi pemisah terbaik (*hyperplane*) untuk klasifikasi *non*-linier menggunakan fungsi kernel adalah:

$$f(x_j) = \text{sign}(\sum_{i=1}^m a_i y_i K(x_i, x_j) + b) \quad (15)$$

Kinerja klasifikasi diukur dengan membandingkan data asli dan data yang dihasilkan oleh model. Proses membandingkan dilakukan menggunakan bantuan tabulasi silang atau *confusion matrix* yang berisi informasi mengenai kelas data asli dan kelas data hasil prediksi model klasifikasi (Prasetyo, 2012).

Tabel 1. *Confusion matrix*

		Prediksi	
		Positif	Negatif
Aktual	Positif	TP	FN
	Negatif	FP	TN

Prasetyo (2012) mengatakan bahwa akurasi mencerminkan ketepatan pada klasifikasi yang dilakukan. Akurasi merupakan besaran tingkat ketepatan klasifikasi dari tiap kelas yang didapat dengan nilai sebenarnya, dapat dihitung menggunakan rumus $\frac{TP + TN}{TP + FP + TN + FN}$.

Penelitian dilakukan menggunakan optimasi Algoritma Genetika untuk membantu dalam menentukan parameter klasifikasi. Algoritma Genetika (GA) adalah suatu algoritma yang digunakan sebagai metode optimasi dengan pedoman bahwa individu dapat bertahan apabila memiliki nilai kelayakan atau *fitness* yang tinggi, nilai *fitness* yang rendah berdampak kepada kelangsungan hidup individu yang pendek dan akan mati (Goldberg, 1989).

Struktur dasar GA menurut Haupt (2004) terdiri atas beberapa langkah antara lain:

1. Inisialisasi populasi
Membangkitkan populasi untuk memulai proses algoritma genetika secara acak.
2. Evaluasi
Populasi dievaluasi dengan melihat nilai *fitness* yang merupakan ukuran kinerja kelayakan sebuah kromosom untuk dihilangkan atau dipelihara agar bertahan hidup. Nilai *fitness* dari suatu individu berbanding lurus dengan hasil yang diberikan. Individu yang baik memiliki nilai *fitness* yang cenderung tinggi.
3. Seleksi
Seleksi bertujuan mendapatkan kromosom induk yang baik untuk digunakan dalam proses crossover dan mutasi. Semakin baik kualitas kromosom induk maka akan menghasilkan keturunan yang baik. Ada beberapa metode seleksi antara lain:
 - 1) *Random Selection* (Seleksi Acak)
 - 2) *Elitism*
 - 3) *Roulette Wheel*
 - 4) *Steady-State*
 - 5) *Rank*
 - 6) *Tournament*
4. Crossover
Proses crossover merupakan penyilangan pasangan kromosom dengan membentuk dua kromosom anak baru. *Crossover* bergantung pada nilai P_c (Probabilitas *crossover*) yang telah ditentukan sebelumnya. Nilai P_c optimal terletak antara nilai 0,75 sampai 0,95 (Schaffer et al., 1989). Jenis-jenis *crossover* yang dapat digunakan antara lain:
 - 1) *Uniform*
 - 2) *One Point*
 - 3) *Two Point*
 - 4) *Multi Point*
5. Mutasi
Proses mutasi merupakan tahapan pergantian gen secara acak dengan nilai lain. Mutasi gen dilakukan dengan menukar nilai gen dengan nilai inversnya, apabila nilai gennya bernilai 0 maka diganti menjadi 1. Beberapa metode yang dapat digunakan pada tahapan mutasi antara lain:
 - 1) *Random*
 - 2) *Random Only Improving*
 - 3) *Sistematic Only Improving*
6. Evaluasi populasi baru
Populasi dievaluasi kembali untuk melihat kinerja populasi hasil algoritma genetika.
7. Ulangi langkah seleksi sampai evaluasi hingga syarat berhenti terpenuhi.

3. METODE PENELITIAN

Penelitian ini menggunakan data berupa dokumen surat bukti gadai dan database nasabah dari PT Pegadaian Cabang Kebumen. Data surat bukti gadai yang digunakan merupakan data kredit dalam kurun waktu 3 Januari 2022 – 29 November 2022. Data yang didapatkan sebanyak 3.548 data dengan perbandingan 80%: 20% pada data *training* dan data *testing* sebesar. Variabel penelitian yang digunakan terbagi menjadi variabel bebas dan variabel terikat. Variabel terikat berupa label kelas (potensial atau tidak potensial), sedangkan variabel bebas terdapat enam variabel terdiri dari uang pinjaman (x_1), kredit aktif (x_2), kredit tidak aktif (x_3), tingkat pendidikan (x_4).

Adapun tahapan analisis yang dilakukan pada penelitian ini:

1. Melakukan *pre-processing* data dengan melakukan transformasi berupa normalisasi data.
2. Melakukan penanganan *imbalanced class* dataset dengan *resampling*.
3. Membagi data menjadi data *training* dan data *testing*.
4. Melakukan inisiasi parameter awal C dan gamma.
5. Melakukan klasifikasi menggunakan SVM non-linier.
6. Menentukan nilai Pc, Pm, dan popsize.
7. Melakukan proses optimasi GA.
8. Melakukan klasifikasi menggunakan SVM-GA.
9. Melakukan evaluasi kinerja model klasifikasi dengan *confusion matrix*.
10. Mendapatkan kesimpulan.

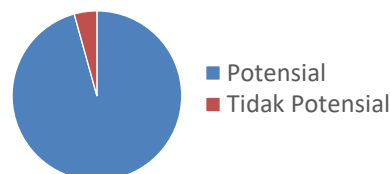
4. HASIL DAN PEMBAHASAN

Normalisasi data dilakukan sebagai *pre-processing* dengan mengubah dataset menjadi atribut yang sama. *Min-max normalization* digunakan untuk mengubah nilai menjadi rentang [0,1].

Tabel 2. Hasil Normalisasi Data

x_1	x_2	x_3	x_4
0,001685362	0	1	Tidak Sekolah
0,004586396	0,217391304	0,954545455	SMP
0,031054871	0,130434783	0,909090909	Tidak Sekolah
\vdots	\vdots	\vdots	\vdots
0,008316296	0,043478261	0	SMA

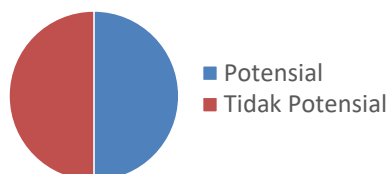
Setelah proses normalisasi, data dibagi menjadi dua yaitu data *training* dan data *testing*. Pembagian data pada penelitian ini dilakukan dengan metode *trial and error* dan didapatkan perbandingan terbaik yaitu data *training* sebesar 80% dan data *testing* sebesar 20%.



Gambar 2. Hasil pembagian data

Resampling dibutuhkan karena hasil pembagian data menunjukkan adanya *imbalanced class* dataset yang terjadi akibat data negatif hanya berjumlah sebesar 4,31% dari total data atau sejumlah 153 data.

Metode yang digunakan adalah *oversampling* dan *undersampling* menggunakan *software R Studio*. Proses penanganan *imbalanced class* data menghasilkan jumlah data positif dan negatif menjadi sama yaitu sebanyak 3395 data.



Gambar 3. Hasil setelah *resampling*

Selanjutnya menentukan nilai parameter yang digunakan pada penelitian. Parameter penelitian menggunakan nilai *default* pada parameter *C (cost)* dan γ (*gamma*) yaitu $C = 1$ dan $\gamma = 0,000147$. Hasil dari *confusion matrix* menggunakan nilai *default* dapat dilihat pada Tabel 3.

Tabel 3. Nilai *Confusion Matrix* pada Kernel RBF pada SVM

		Prediksi	
		Positif	Negatif
Aktual	Positif	548	51
	Negatif	131	628

Tabel 3. menunjukkan bahwa nasabah potensial yang diprediksi potensial berjumlah 548 orang, nasabah tidak potensial yang diprediksi tidak potensial berjumlah 628 orang, nasabah potensial yang diprediksi tidak potensial berjumlah 51 orang dan nasabah tidak potensial yang diprediksi potensial berjumlah 131 orang. Performa yang dihasilkan oleh klasifikasi SVM dapat dilihat pada Tabel 4.

Tabel 4. Performa Klasifikasi SVM dengan kernel RBF

Ukuran Kinerja Klasifikasi	Nilai (%)
<i>Accuracy</i>	86,60
<i>Precision</i>	80,70
<i>Recall</i>	91,48

Tahapan selanjutnya yaitu melakukan pencarian parameter optimal untuk klasifikasi menggunakan algoritma genetika (GA). Parameter GA untuk proses pencarian menggunakan rentang $C = 0 - 2^6$ dan $\gamma = 0 - 2^5$.

Hasil pencarian menggunakan GA pada kernel RBF menghasilkan parameter optimal pada kombinasi nilai $C = 40,537$ dan $\text{Gamma} = 19,775$. Hasil dari *confusion matrix* menggunakan nilai hasil GA dapat dilihat pada Tabel 5.

Tabel 5. Nilai *Confusion Matrix* pada Kernel RBF pada SVM-GA

		Prediksi	
		Positif	Negatif
Aktual	Positif	624	24
	Negatif	55	655

Performa yang dihasilkan oleh klasifikasi SVM-GA dapat dilihat pada Tabel 6.

Tabel 6. Hasil Perhitungan *Confusion Matrix* SVM-GA dengan kernel RBF

Ukuran Kinerja Klasifikasi	Nilai (%)
<i>Accuracy</i>	94,20
<i>Precision</i>	96,29
<i>Recall</i>	91,89

Selanjutnya akan dibandingkan nilai akurasi kedua metode yang telah digunakan. Hasil dari perbandingan kedua metode terdapat pada Tabel 7.

Tabel 7. Perbandingan Hasil Klasifikasi SVM dan SVM-GA

Ukuran Kinerja Klasifikasi	Metode	
	SVM	SVM-GA
<i>Accuracy</i>	86,60	94,20
<i>Precision</i>	80,70	96,29
<i>Recall</i>	91,48	91,89

Berdasarkan Tabel 7. memperlihatkan bahwa akurasi terbaik dihasilkan oleh *kernel* RBF pada SVM-GA dengan besaran akurasi mencapai 94,20% sehingga model SVM-GA merupakan model yang paling baik dalam memprediksi nasabah gadai potensial pada PT Pegadaian.

5. KESIMPULAN

Berdasarkan analisis dan pembahasan pada nasabah potensial PT Pegadaian diperoleh kesimpulan bahwa nilai akurasi dari model yang dihasilkan sangat baik. Hasil prediksi menggunakan algoritma SVM tanpa optimasi menghasilkan nilai akurasi sebesar 86,6%, sedangkan prediksi menggunakan algoritma SVM-GA menghasilkan nilai akurasi sebesar 94,2%. Sehingga dapat dinyatakan bahwa algoritma SVM-GA lebih baik dalam memprediksi calon nasabah gadai potensial pada PT Pegadaian.

DAFTAR PUSTAKA

- Cortes, C., & Vapnik, V. (1995), Support-Vector Networks, *Machine Learning*, 20(3).
- Elly. S., Sabariah. M. K., & Gozali. A. A. 2015. *Impelentasi Metode Support Vector Machine Untuk Melakukan Klasifikasi Kemacetan Lalu Lintas Pada Twitter*. eProceedings of Engineering 2.1.
- Goldberg, D, E, (1989), Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, MA, 1989, *NN Schraudolph and J.*, 3.
- Hastie. T., Tibshirani. R., Friedman. J. H., & MyiLibrary. 2001. *The elements of statistical learning data mining. inference. and prediction : with 200 full-color illustrations*. Springer Series in Statistics.
- Haupt. R. L. H. and S. E. 2004. *Practical Genetic Algorithms MATLAB Code*. John Wiley & Sons.
- Herdianto. 2013. *Prediksi Kerusakan Motor Induksi Menggunakan Metode Jaringan Saraf Tiruan Backpropagation*. Fakultas Teknik. Universitas Sumatera Utara. Medan.
- Pegadaian. 2022. *Sejarah Perusahaan Pegadaian*. www.pegadaian.co.id/
- Raharjo. S., & Winarko. E. 2014. *Klasterisasi. klasifikasi dan peringkasan teks berbahasa indonesia*. Kommit 2014. 8(Kommit).

- Santosa, B, (2007), Feature Selection with Support Vector Machines Applied on Tornado Detection, *IPTEK The Journal for Technology and Science*, 18(1).
- Schaffer. J. D., Caruana. R. A., Eshelman. L. J., & Das. R. 1989. *A study of control parameters affecting online performance of genetic algorithms for function optimization*. Proceedings of the Third International Conference on Genetic Algorithms.
- Siringoringo. R., & Jamaludin. J. 2019. *Text Mining dan Klasterisasi Sentimen Pada Ulasan Produk Toko Online*. Jurnal Teknologi Dan Ilmu Komputer Prima (JUTIKOMP). 2(1).
- Yanaeng, S., Saelee, S. dan Samai, W., 2014. Automatic Medical Case Study Essay Scoring by Support Vector Machine and Genetic Algorithms. *International Journal of Information and Education Technology*, IV(2), pp. 132-137.