

## Penentuan Kelayakan dan Besaran Pinjaman Pada Koperasi Di Banjarmasin Memanfaatkan Support Vector Machine (SVM) Dan Regresi Linier Berganda

Samsuri, S.Kom, M.Kom  
Jl Pangeran Hidayatullah, Banua Anyar, Banjarmasin  
Email : samsuri1479@gmail.com

### ABSTRAK

Resiko kredit menjadi hal yang penting untuk dikaji dalam sebuah lembaga keuangan, tak terkecuali Koperasi di Banjarmasin. Dalam penelitian ini akan digunakan Support Vector Machine (SVM) untuk melakukan klasifikasi pinjaman dan Regresi Linier Berganda (RLB) digunakan untuk mencari besaran pinjaman. Data pengajuan pinjaman yang digunakan mulai tahun 2006-2015 sebanyak 10.219 record data. Praprocessing data terdiri dari tahap field selection, data representation dan normalisasi data, menghasilkan dataset dengan 10.135 record yang terdiri dari 6.573 record untuk kelas "Layak" menerima pinjaman dan 3.562 records untuk kelas "Tidak Layak". Training dengan SVM diawali dengan proses estimasi parameter kemudian dilanjutkan dengan membentuk matriks kernel. Dalam penelitian ini digunakan 3 jenis kernel yaitu kernel linier, kernel polynomial dan kernel Radial Basis Function (RBF). Langkah selanjutnya adalah menghitung error dari matriks kernel dan menghitung perubahan nilai alpha. Nilai alpha selanjutnya digunakan untuk melakukan pengecekan kondisi pemberhentian iterasi. Dalam studi kasus menunjukkan bahwa perhitungan dengan kernel linier dan polynomial telah memenuhi kondisi pemberhentian sedangkan perhitungan dengan kernel RBF masih berlanjut. Pada kernel linier dan polynomial, semua nilai delta alpha menunjukkan angka negatif ( $<0$ ), sehingga nilainya lebih kecil dari nilai epsilon. Sedangkan pada kernel RBF menunjukkan nilai delta alpha positif ( $>0$ ), sehingga nilainya masih lebih besar dari nilai epsilon. Langkah terakhir adalah menghitung nilai alpha baru dan nilai bias. Hasil klasifikasi yang masuk ke dalam kelas "diterima", selanjutnya akan ditentukan besaran pinjaman dengan metode RLB. Dalam uji coba yang telah dilakukan, dihasilkan nilai akurasi yang baik. Dari uji coba klasifikasi dengan SVM didapatkan nilai akurasi = 72.79%. Sedangkan dari uji coba penentuan besaran pinjaman dengan Regresi Linier Berganda didapatkan Mean Square Error (MSE) = 0,011.

Kata Kunci : Klasifikasi, Machine Learning, SVM, RLB

### ABSTRACT

Credit risk is an important thing to study in a financial institution, including cooperatives. Cooperatives in Banjarmasin are still the people's choice to apply for loans. But the risk of credit or bad credit is still quite high in the last 5 years. The initial selection and screening process on the loan application data is absolutely necessary. In this study, Support Vector Machine (SVM) will be used to classify loan submission data and Multiple Linear Regression (MLR) is used to find the amount that can be accepted by borrowers of cooperatives in Banjarmasin. The loan application data used for 10 years starting from 2006 to 2015 is 10,219 record data. Data processing consists of field selection, data representation and data normalization stages, resulting in a dataset with 10,135 records consisting of 6,573 records for the "Eligible" class receiving loans and 3,562 records for the "Eligible" class. Training with SVM begins with the parameter estimation process and then proceeds to form a kernel matrix. The kernel matrix will measure the number of records trained. In this study, three types of kernels are used, namely linear kernels, polynomial kernels and Radial Basis Function (RBF) kernels. The next step is to calculate the error from the kernel matrix and calculate the change in alpha (delta alpha) value. The alpha delta value is then used to check the iteration termination condition. The case study shows that calculations with linear and polynomial kernels meet the stopping conditions while calculations with RBF kernels continue. In linear and polynomial kernels, all alpha delta values show a negative number ( $<0$ ), so that the value is smaller than the epsilon value. While the RBF kernel shows a positive alpha delta value ( $>0$ ), so the value is still greater than the epsilon value. The final step is to calculate the new alpha value and the bias value. The classification results that fall into the "accepted" class will then be determined the amount of the loan with the RLB method.

In the trials that have been carried out, a good accuracy value is produced. From the classification test with Support Vector Machine, the accuracy value is 72.79%. While the trial determination of the amount of the loan with Multiple Linear Regression obtained an Mean Square Error (MSE) value of 0,01100.

Keywords: Classification, Machine Learning, SVM, RLB

### I. PENDAHULUAN

Pemberian pinjaman atau biasa disebut dengan pemberian kredit merupakan proses yang penting disetiap lembaga keuangan. Resiko kredit merupakan salah satu yang dikaji dan diteliti oleh setiap lembaga keuangan [2]. Sitem keputusan pemberian pinjaman yang akurat adalah alasan penting untuk profitabilitas lembaga keuangan, kredibilitas nasabah untuk pemberian pinjaman tergantung pada beberapa parameter, parameter tersebut diantaranya

histori pinjaman, pekerjaan dan lain-lain[3].

Untuk mengurangi risiko kegagalan ini, kami telah mengajukan beberapa data model untuk membantu dalam mengidentifikasi kemampuan pelanggan untuk melunasi pinjaman kredit tepat waktu dengan menggunakan penilaian kredit dan mengklasifikasikan mereka sebagai kreditur yang baik, nasabah memiliki nilai bagus dan tidak memiliki kesalahan atau catatan kredit masa lalu atau kredit buruk jika nasabah adalah memiliki skor buruk dan

mungkin memiliki catatan masa lalu yang tidak baik dalam proses pinjaman maka lembaga keuangan dalam hal ini koperasi akan mampu memberikan kredit yang baik yang pada akhirnya akan menjadi menguntungkan dalam pendapatan tahunan mereka [4].

Berdasarkan kondisi yang telah dijelaskan diatas, maka dibutuhkan suatu sistem pengidentifikasian kelayakan pinjaman pada koperasi di Banjarmasin menggunakan Support Vector Machine (SVM) serta berapa besaran yang harus diberikan kepada peminjam atau anggota dengan metode Regresi Linier Berganda. Sistem ini digunakan untuk menentukan apakah anggota koperasi yang akan melakukan pinjaman dikategorikan layak atau tidak dan jika layak berapa besaran yang diterima.

## II. METODE PENELITIAN

### 2.1 Support Vector Machine (SVM)

SVM dalam machine learning dikenal juga dengan support vector network yang merupakan metode supervised terkait dengan learning algorithm untuk analisa pola data yang digunakan untuk klasifikasi dan regresi. SVM secara konseptual adalah mesin linier yang dilengkapi dengan fitur special dan didasarkan pada metode minimalisasi resiko struktural, serta teori pembelajaran statistic [10]. Dua sifat

khusus dari SVM yaitu [11]:

- Mencapai generalisasi yang tinggi dengan memaksimalkan margin
- Mendukung pembelajaran yang efisien dari fungsi nonlinier pada trik kernel sehingga membuat kinerja generalisasinya baik dalam menyelesaikan masalah pengenalan pola .

Untuk permasalahan klasifikasi, SVM mencoba untuk mencari garis pemisah yang optimal yang diekspresikan sebagai kombinasi linier dari subset data pelatihan dengan menyelesaikan masalah keterbatasan linier pemrograman kuadrat (QP) dengan margin maksimum antara dua kelas. Sementara untuk permasalahan regresi, diperkenalkan fungsi  $\epsilon$ -insensitive loss yang disebut sebagai SVM untuk regresi [11].

Klasifikasi data biasa digunakan pada sejumlah data yang telah di ketahui data induknya, untuk kemudian dijadikan data training/data model yang hasilnya akan menjadi keputusan prediksi dari sejumlah data yang serupa namun belum lengkap pada salah satu atributnya. SVM adalah sistem learning yang menggunakan sebuah ruang hipotesis fungsi linier dalam ruang fitur berdimensi tinggi, dilatih dengan menggunakan sebuah algoritma pembelajaran dari teori optimasi yang mengimplementasikan sebuah bias learning yang diturunkan dari teori learning statistika. Strategi learning yang diperkenalkan merupakan sebuah metode yang powerful dalam beberapa tahun sejak diperkenalkan dan telah melebihi sistem yang lain dalam berbagai aplikasi. Konsep dasar SVM adalah [10]:

- Class Separation, pada dasarnya, SVM mencari bidang hyperplane yang memisahkan secara optimal antara dua kelas dengan memaksimalkan margin antara titik terdekat kelas tersebut.
- Overlapping Classes, titik-titik data pada sisi

“salah” dari diskriminan margin diturunkan untuk mengurangi pengaruhnya (soft margin).

- Non Linearity, ketika tidak dapat ditemukan pemisah berbentuk linier, titik-titik data biasanya diproyeksikan ke dalam ruang dimensi yang lebih tinggi dimana titik-titik data secara efektif akan menjadi pemisah linier (proyeksi ini direalisasikan melalui teknik kernel).
- Problem Solution, semua tugas tersebut dapat diformulasikan sebagai permasalahan optimasi kuadrat yang dapat diselesaikan dengan teknik yang diketahui.

Algoritma yang dipakai pada tahap training adalah sebagai berikut [10]:

- Menginisialisasi  $\alpha_i=0,5$  dan parameter lain, misalnya  $\lambda = 0,5$ ,  $\gamma = 0,5$ ,  $c = 2$ , IterasiMax = 100, dan  $\epsilon = 0,001$ . Kemudian menghitung matriks Hessian dapat dihitung dengan rumus persamaan:  $D_{ij} = y_i y_j ( K( x_i , x_j ) + \lambda^2 )$
- Mulai dari data ke i sampai j , hitung menggunakan persamaan:
  - $E_i = \sum_{j=i}^n \alpha_j D_{ij}$
  - $\delta \alpha_i = \min \{ \max [ \gamma(1 - E_i), -\alpha_i ], C - \alpha_i$
  - $\alpha_i = \alpha_i + \delta \alpha_i$
- Langkah 2 dilakukan terus-menerus hingga kondisi iterasi maksimum tercapai atau  $\max ( | \delta \alpha_i | ) < \epsilon$

Selanjutnya didapatkan nilai support vector (SV), SV = (Threshold SV). Nilai Threshold SV didapatkan dari beberapa percobaan, biasanya digunakan threshold > 0. Kemudian dilakukan proses testing untuk mendapatkan keputusan di mana fungsi keputusan dapat di hitung dengan persamaan berikut:

$$f(x) = \sum_{i=1}^m \alpha_i y_i K(x_i, x) + b$$

- Dimana nilai b :  $-\frac{1}{2} [\sum_{i=1}^m \alpha_i y_i K(x_i, x^+) + \sum_{i=1}^m \alpha_i y_i K(x_i, x^-)]$

### 2.2 Regresi Linier Berganda

Regresi Linier yang dipakai adalah regresi Linier berganda dengan persamaan sebagai berikut [13]:

$$Y' = a + b_1 X_1 + b_2 X_2 + \dots + b_n X_n$$

Keterangan:

- Y' = Variabel dependen (nilai yang diprediksikan)
- X<sub>1</sub>, X<sub>2</sub>...X<sub>n</sub> = Variabel independen
- a = Konstanta (nilai Y' apabila X<sub>1</sub>, X<sub>2</sub>...X<sub>n</sub> = 0)
- b<sub>1</sub>, b<sub>2</sub>...b<sub>n</sub> = Koefisien regresi (nilai peningkatan ataupun penurunan)

Langkah-langkah dalam menyelesaikan Regresi Linier Berganda [13]:

1. Menentukan Variabel

TABEL I  
CONTOH TABEL KRITERIA

No	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	Dana
A <sub>1</sub>	1	3	2	1	1	1	2	2	8500000
A <sub>2</sub>	2	1	3	2	3	3	3	3	1000000
A <sub>3</sub>	3	2	1	3	1	1	1	1	9896250
A <sub>4</sub>	1	3	2	1	2	3	3	2	4500000
A <sub>5</sub>	2	1	3	3	3	1	1	3	3000000

2. Uji Asumsi Klasik
  - Uji Multikolinieritas
  - Uji autokorelasi
3. Uji Koefisien Regresi Linier Berganda
  - Uji Simultan (Uji F)

Adapun langkah-langkah pengujiannya sebagai berikut:

  - a. Menentukan Formulasi Hipotesis
    - H<sub>0</sub> : b<sub>i</sub>, artinya variabel independen tidak mempunyai pengaruh yang signifikan secara bersama-sama terhadap variabel dependen.
    - H<sub>0</sub> : b<sub>i</sub> ≠ 0, artinya variabel independen mempunyai pengaruh yang signifikan secara bersama-sama terhadap variabel dependen.
  - b. Menentukan derajat kepercayaan 95% (α = 0,05)
  - c. Menentukan signifikansi
    - Nilai signifikansi (P Value) < 0,05 maka H<sub>0</sub> ditolak dan H<sub>1</sub> diterima.
    - Nilai signifikansi (P Value) > 0,05 maka H<sub>0</sub> diterima dan H<sub>1</sub> ditolak.
  - d. Membuat kesimpulan
    - Bila (P Value) < 0,05 maka H<sub>0</sub> ditolak dan H<sub>1</sub> diterima. Artinya variabel independen secara simultan (bersama-sama) mempengaruhi variabel dependen.
    - Bila (P Value) > 0,05 maka H<sub>0</sub> diterima dan H<sub>1</sub> ditolak. Artinya variabel independen secara simultan (bersama-sama) tidak mempengaruhi variabel dependen.
  - Uji t
4. Analisis Regresi Linier Berganda
5. Pengujian Model

Metode yang digunakan dalam pengujian model adalah Root Mean Square Error (RMSE). Root Mean Square Error (RMSE) adalah ukuran yang digunakan sebagai pembeda antara nilai-nilai yang diprediksi dengan nilai-nilai yang sebenarnya [13].

$$RMSE = \frac{\sum_{i=1}^n \sqrt{(x_i - f_i)^2}}{n}$$

Dimana semakin besar nilai RMSE yang dihasilkan maka keakuratan suatu model semakin sedikit atau tidak akurat,

sedangkan semakin kecil nilai RMSE maka semakin baik akurasi suatu model regresi linier tersebut.

III. Pembentukan Dataset

Preprocessing data yang dilakukan meliputi proses selection, cleaning, conversion dan normalisasi/scaling data. Pemilihan (selection) data dari sekumpulan data operasional, pemilihan data sesuai tujuan yang diharapkan. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk pelatihan, seperti data atau informasi eksternal. Konversi adalah merubah nilai atribut disesuaikan dengan kebutuhan dalam algoritma SVM. Sedangkan normalisasi adalah proses merubah rentang nilai atribut. Hal ini dilakukan dengan merubah value data menjadi bilangan real agar proses perhitungan lebih mudah dan atribut yang rentang nilainya besar tidak mendominasi atribut yang rentang nilainya lebih kecil. Dengan proses normalisasi ini umumnya akan memberikan hasil yang lebih baik, dibandingkan dengan data yang tidak mengalami proses normalisasi.

TABEL 2  
TIPE DATA INPUT

No	Var	Atribut	Tipe data	Keterangan
1	X <sub>1</sub>	Jumlah Tanggungan	Numeric [0..9]	Normalisasi [0,1]
2	X <sub>2</sub>	Pekerjaan	Biner [PNS(0), Swasta(1)]	
3	X <sub>3</sub>	Pendapatan/bln	Numeric [1jt..∞]	Normalisasi [0,1]
4	X <sub>4</sub>	Level Pinjaman	numeric	Normalisasi [0,1]
5	X <sub>5</sub>	jangka waktu	Numeric [1..60]	Normalisasi [0,1]

TABEL 2  
 TIPE DATA INPUT (LANJUTAN)

No	Var	Atribut	Tipe data	Keterangan
6	X <sub>6</sub>	status pernikahan	Diskrit [menikah, belum menikah, bercerai]	Konversi [real]
7	X <sub>7</sub>	riwayat pinjaman	Biner [macet(0), lancar(1)]	
8	X <sub>8</sub>	status simpanan	Biner [tidak baik(0), baik(1)]	
9	X <sub>9</sub>	Rumah	Biner [milik sendiri(1), bukan(0)]	
10	X <sub>10</sub>	jenis jaminan	Biner [tanpa jaminan(0), jaminan(1)]	
11	X <sub>11</sub>	penilaian masyarakat	Diskrit [kurang, cukup, baik]	Konversi [real]
12	X <sub>12</sub>	tarif listrik	Biner [subsidi(1), non subsidi(0)]	
13	X <sub>13</sub>	daya listrik	Numeric	Normalisasi [0,1]
14	X <sub>14</sub>	pemakaian air	Numeric	Normalisasi [0,1]
15	X <sub>15</sub>	hutang ditempat lain	Biner [ada(1), tidak(0)]	
16	X <sub>16</sub>	Besar Pinjaman	Numeric	Normalisasi [0,1]

Normalisasi menggunakan metode mean normalization. Normalisasi dengan rumus z-score ini bertujuan untuk menskalakan data agar jatuh didalam suatu range kecil yang tertentu.

$$x' = \frac{x - \text{mean}}{\text{std}_{dev}}$$

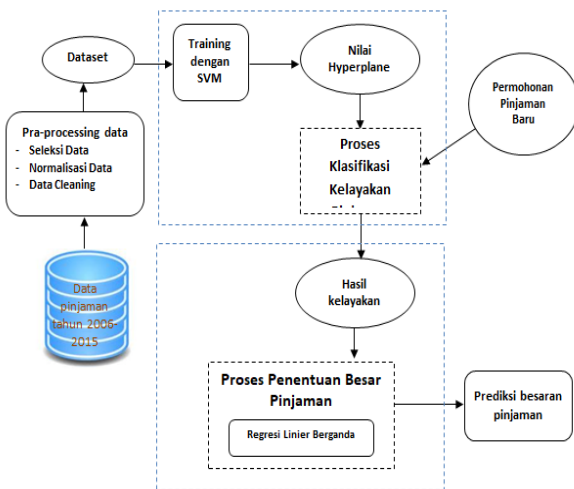
Value atribut dengan rentang yang besar akan diubah menjadi lebih kecil dan lebih bervariasi, dimana terdapat angka positif dan juga angka negatif. Dengan adanya angka negatif akan mencegah nilai menjadi besar saat harus di summing.

TABEL 3  
 SAMPEL HASIL PRAPROCESSING DATA

Thn	2015	2015	2014	2014	2013
X1	-0,12	-0,12	-0,72	1,673	-0,12
X2	0,71	0,71	-1,33	0,71	-1,32
X3	-1,18	0,248	-0,654	1,301	-1,105
X4	-0,431	-0,431	-0,431	0,799	-0,431
X5	-0,6	-0,6	-0,6	1,49	-0,6
X6	0,223	0,223	0,223	0,223	0,223
X7	0,409	0,409	0,409	-2,32	0,409
X8	0,487	0,487	0,487	0,487	0,487
X9	-0,881	1,077	-0,881	1,077	-0,881
X10	0,223	0,223	0,223	0,223	0,223
X11	0,382	0,382	0,382	0,382	0,382
X12	1,488	-0,638	-0,638	-0,638	-0,638
X13	-0,2012	1,14	-0,201	1,14	-0,201
X14	1,045	1,045	-1,414	-0,184	1,045
X15	0	0	0	0	0
X16	-0,697	-0,697	-0,455	1,118	-0,769
X17	0,562	0,562	0,562	-1,688	-1,688

#### IV. HASIL DAN PEMBAHASAN

Sistem yang diusulkan akan melakukan proses klasifikasi terhadap data permohonan pinjaman baru berdasarkan data permohonan yang telah ada sebelumnya. Dibutuhkan data histori yang cukup untuk dapat melakukan klasifikasi secara baik. Sistem akan melakukan pelatihan terhadap dataset permohonan pinjaman 2006 sampai 2015 yang telah tersimpan di dalam database. Hasil pelatihan dengan metode SVM adalah fungsi hyperplane yang memisahkan antar class yang ada.



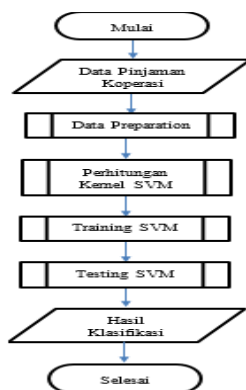
Gambar 1. Blok Diagram Sistem yang Diusulkan

Proses klasifikasi akan menghasilkan dua output, yaitu:

- a. Layak untuk mendapatkan pinjaman
- b. Tidak layak.

Apabila output yang dihasilkan dari proses klasifikasi adalah Layak untuk mendapatkan pinjaman, maka selanjutnya akan dilakukan proses penentuan besaran pinjaman menggunakan Regresi Linier Berganda.

Penggunaan SVM dapat digunakan dalam bentuk supervised maupun unsupervised learning. Pada prinsipnya semua memiliki tujuan untuk menyelesaikan permasalahan quadratic programming, karenanya proses pembelajarannya juga memiliki tahapan-tahapan yang hampir sama. Tahapan dalam pemodelan permohonan pinjaman koperasi di Banjarmasin ke dalam algoritma SVM ditunjukkan pada gambar 2.



Gambar 2. Alur Pemodelan Sistem dengan SVM

### A. Parameter dalam Algoritma Pembelajaran dengan SVM

Dalam tahap pelatihan (*training*) dataset permohonan sebagai berikut:

$$x = \{x_0, x_1, x_2, \dots, x_m\}; \text{ sampel training}$$

$y = (y_1, \dots, y_m) \in \{\pm 1\}$ ; label data training

$m = \text{jumlah data} = \text{sample}$

Kernel = jenis fungsi kernel

- Kernel Linier  $\Rightarrow K(x,y) = x \cdot y$
- Kernel Polynomial  $\Rightarrow K(x,y) = (x \cdot y + c)^d$
- Kernel Radial Basis Function (RBF)  
 $\Rightarrow K(x_i - x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right)$

$\alpha_i$  = lagrange multiplier, berguna untuk mencari support vector = 0.5

$\gamma$  = gamma, berguna dalam melakukan kontrol kecepatan training = 0.5

$\lambda$  = lamda, batas teoritis yang diturunkan = 0.5  $C = \text{complexity}$ , digunakan dalam proses training,

memberikan batas nilai alfa = 2

$\epsilon$  = epsilon, digunakan dalam mengukur error pada klasifikasi = 0.001

$\sigma$  = sigma, berpengaruh pada pemetaan data yang terbentuk dalam kernel RBF = 2

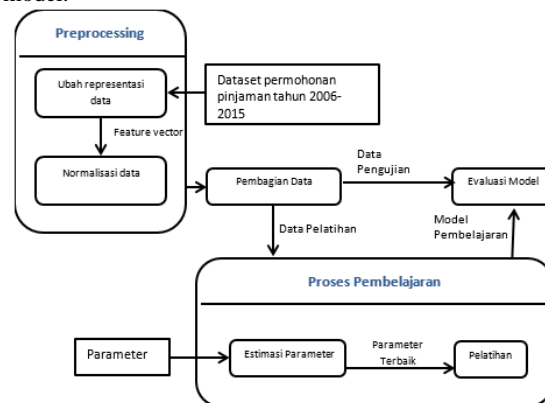
$d$  = derajat polynomial = 2

$b$  = bias

Itermax = 100

### B. Perancangan Proses Pembelajaran

Dataset hasil preprocessing data, selanjutnya akan memasuki proses klasifikasi data. Dalam proses klasifikasi dibagi menjadi 2 tahap yaitu tahap pelatihan dan tahap evaluasi model (testing). Sistem akan melakukan pelatihan data train dengan menggunakan metode support vector machine. Setelah melakukan training data, sistem akan menghasilkan model SVM yang kemudian akan dievaluasi dengan menggunakan data testing pada tahap pengujian model.



Gambar 3. Blok Diagram proses Pembelajaran dengan SVM

Pada tabel 4 ditunjukkan 3 records data sampling dataset permohonan pinjaman koperasi di Banjarmasin. Dengan menggunakan data pada tabel 4 bisa dihitung matriks kernel Hessian dengan dimensi 3 x 3. Digunakan 3 kernel yaitu linier, polynomial dan RBF.

#### 1. Membentuk Matriks kernel

- a. Kernel Linier

$$K = \begin{bmatrix} 32,41 & 3,73 & 1,78 \\ 3,727 & 3,06 & 0,2 \\ 1,785 & 0,2 & 2,12 \end{bmatrix}$$

b. Kernel Polynomial

$$K = \begin{bmatrix} 503,14 & -2,37 & 0,11 \\ -2,3681 & 37,66 & 1,316 \\ 0,1096 & 1,316 & 20,04 \end{bmatrix}$$

c. Kernel RBF

$$K = \begin{bmatrix} 3,1 & -0,35 & -0,697 \\ 0,2412 & 0,567 & 0,282 \\ -0,3706 & 0,467 & 0,567 \end{bmatrix}$$

TABEL 4  
SAMPel DATASET UNTUK TRAINING

Att	Record-1	Record-2	Record-3
x <sub>1</sub>	-1,31	1,67	-0,12
x <sub>2</sub>	0,72	0,72	0,72
x <sub>3</sub>	1,08	1	-1,18
x <sub>4</sub>	0,8	-0,43	-0,43
x <sub>5</sub>	-0,64	-0,64	-0,64
x <sub>6</sub>	0,22	0,22	0,22
x <sub>7</sub>	0,41	0,41	0,41
x <sub>8</sub>	0,49	0,49	0,49
x <sub>9</sub>	-0,88	1,08	-0,88
x <sub>10</sub>	0,22	0,22	0,22
x <sub>11</sub>	0,38	0,38	0,38
x <sub>12</sub>	-0,64	1,49	1,49
x <sub>13</sub>	-1,54	1,14	-0,2
x <sub>14</sub>	-1,41	-0,18	1,05
x <sub>15</sub>	0	0	0
x <sub>16</sub>	0,32	-0,7	-0,7
x <sub>17</sub>	-1,69	0,56	0,56

2. Menghitung nilai error

TABEL 5  
NILAI ERROR MATRIKS KERNEL

Kernel Linier	
A1	(32,41*0,5)+(3,73*0,5)+(1,785*0,5) = <b>18,96004</b>
A2	(3,723*0,5)+(3,059*0,5)+(0,198*0,5) = <b>3,491728333</b>
A3	(1,78*0,5)+(0,198*0,5)+(2,12*0,5) = <b>2,051306667</b>
Kernel Polynomial	
A1	(503,144*0,5)+(-2,368*0,5)+(0,11*0,5) = <b>250,4426994</b>
A2	(-2,368*0,5)+(37,66*0,5)+(1,316*0,5) = <b>18,30598704</b>
A3	(0,1095*0,5)+(1,316*0,5)+(20,044*0,5) = <b>10,73453682</b>
Kernel RBF	
A1	(3,1*0,5)+(-0,354*0,5)+(-0,697*0,5) = <b>1,024377091</b>
A2	(0,241*0,5)+(0,567*0,5)+(0,282*0,5) = <b>0,545103314</b>
A3	(-0,37*0,5)+(0,467*0,5)+(0,567*0,5) = <b>0,331497864</b>

Nilai error untuk setiap baris data pelatihan (3 baris data) dihitung dengan menjumlahkan hasil perkalian nilai elemen matriks Hessian dengan nilai  $\alpha$ . Terlihat pada tabel 4.3 bahwa nilai error terkecil terletak pada kernel RBF dengan nilai 0,33, 0,55 dan 1,024. Sedangkan error tertinggi terletak pada kernel Polynomial dengan nilai 250,44, 18,31 dan 10,73.

3. Menghitung nilai alpha

Langkah selanjutnya adalah menghitung nilai alpha/perubahan nilai alpha ( $\delta\alpha$ ). Dibutuhkan beberapa nilai parameter untuk menghitung alpha, yaitu  $\alpha = 0,5$ ,  $\gamma = 0,5$  dan  $C = 2$ . Nilai alpha dihitung dengan memilih nilai minimum dari nilai maksimum perkalian gamma dan 1 dikurangi error ( $\gamma*(1-E_i)$ ) dan negatif alpha ( $-\alpha$ ) dengan nilai  $C$  dikurangi  $\alpha$ . Nilai alpha selanjutnya digunakan untuk melakukan pengecekan kondisi pemberhentian iterasi. Kondisi pemberhentian iterasi apabila nilai maksimum  $\delta\alpha_i < \epsilon$  ( $\epsilon = 0.001$ ). Apabila nilai maksimum  $\delta\alpha_i > \epsilon$  ( $\epsilon = 0.001$ ) maka iterasi akan dilanjutkan kembali sampai nilai  $\alpha$  konvergen (tidak ada perubahan signifikan).

TABEL 6  
PENGECEKAN ITERASI 3 KERNEL

Kernel Linier	Cek Iterasi	
A1	MIN(MAX(0,5*(1-18,96);-0,5);2-0,5) = <b>-0,5</b>	Max(A1,A2,A3) < $\epsilon$
A2	MIN(MAX(0,5*(1-3,49);-0,5);2-0,5) = <b>-0,5</b>	Max(-0,5;-0,5;-0,5) = -0,5
A3	MIN(MAX(0,5*(1-2,051);-0,5);2-0,5) = <b>-0,5</b>	-0,5 < 0,001 (iterasi berhenti)
Kernel Polynomial		
A1	MIN(MAX(0,5*(1-250,44);-0,5);2-0,5) = <b>-0,5</b>	Max(A1,A2,A3) < $\epsilon$
A2	MIN(MAX(0,5*(1-18,31);-0,5);2-0,5) = <b>-0,5</b>	Max(-0,5;-0,5;-0,5) = -0,5
A3	MIN(MAX(0,5*(1-10,734);-0,5);2-0,5) = <b>-0,5</b>	-0,5 < 0,001 (iterasi berhenti)
Kernel RBF		
A1	MIN(MAX(0,5*(1-1,024);-0,5);2-0,5) = -0,012188546	Max(A1,A2,A3) < $\epsilon$
A2	MIN(MAX(0,5*(1-0,545);-0,5);2-0,5) = 0,227448343	Max(0,012;0,227;0,334) = 0,334
A3	MIN(MAX(0,5*(1-0,331);-0,5);2-0,5) = 0,334251068	0,334 > 0,001 (iterasi berlanjut)

Perhitungan dengan kernel linier dan polynomial telah

memenuhi kondisi pemberhentian sedangkan perhitungan dengan kernel Radial Basis Function (RBF) masih berlanjut karena belum memenuhi kondisi pemberhentian. Pada kernel linier dan polynomial, semua nilai delta alpha menunjukkan angka negatif (<0), sehingga nilainya lebih kecil dari nilai epsilon. Sedangkan pada kernel RBF menunjukkan nilai delta alpha positif (>0), sehingga nilainya masih lebih besar dari nilai epsilon.

4. Menghitung nilai alpha baru

Selanjutnya akan dihitung nilai alpha baru dengan cara menambahkan nilai alpha lama dengan nilai delta alpha.

TABEL 7  
NILAI ALPHA BARU

	$\alpha_{baru}$
A1	0,487811454
A2	0,727448343
A3	0,834251068

5. Menghitung bias

Setelah diperoleh nilai alpha baru, selanjutnya akan dihitung nilai bias (b) menggunakan persamaan  $b = -\frac{1}{2}(\bar{w} \cdot \vec{x}_{-1} > + \bar{w} \cdot \vec{x}_{+1} >)$ . Terlebih dahulu harus dihitung nilai w.  $W_i^+$  adalah bobot dot product data dengan alpha terbesar di kelas positif. Sedangkan  $W_i^-$  adalah bobot dot product data dengan alpha terbesar di kelas negatif. Sehingga untuk nilai w dapat dihitung sebagai berikut:

$$W_i^+ = (0,488 * -1,69 * 0,37) + (0,727 * 0,56 * 0,467) + (0,834 * 0,56 * 0,567) = 0,762$$

$$W_i^- = (0,488 * -1,69 * 3,1) + (0,727 * 0,56 * -0,354) + (0,834 * 0,56 * -0,697) = -3,025$$

$$b = -\frac{1}{2}(0,762 + (-3,025)) = 1,131$$

Sampai dengan perhitungan nilai bias menunjukkan satu iterasi telah selesai. Untuk melanjutkan ke iterasi berikutnya

maka diulangi kembali proses mulai menghitung error ( $E_i$ ), delta alpha dan seterusnya. Langkah-langkah ini akan dilakukan terus-menerus hingga kondisi iterasi maksimum tercapai atau  $\max(|\delta \alpha_i|) < \epsilon$ , yaitu kondisi alpha yang konvergen.

C. Perancangan Proses Testing (Klasifikasi)

Hasil perhitungan (proses pembelajaran) belum bisa digunakan untuk fungsi keputusan karena iterasi masih harus diteruskan sampai mencapai maksimum iterasi atau kondisi konvergen. Namun, dalam penelitian ini akan

digunakan hasil perhitungan tersebut hanya sebagai contoh perhitungan dengan fungsi keputusan saja. Sebagai contoh akan digunakan data uji sebagai berikut:

TABEL 8  
SAMPEL DATA UJI

Att	Data	Att	Data
$x_1$	-0,72	$x_{10}$	0,22
$x_2$	-1,33	$x_{11}$	0,38
$x_3$	-0,5	$x_{12}$	-0,64
$x_4$	2,03	$x_{13}$	-0,2
$x_5$	1,49	$x_{14}$	-1,41
$x_6$	0,22	$x_{15}$	0
$x_7$	0,41	$x_{16}$	1,97
$x_8$	-1,95	$x_{17}$	0,56
$x_9$	-0,88		

Langkah pertama untuk menguji adalah menghitung nilai antara data uji dengan semua data latih dengan fungsi kernel RBF.

$$\text{Data ke-1} = \text{Exp}(-(\text{Abs}((-0,72+1,31)+(-1,33-0,72)+(-0,5-1,08)+(2,03-0,8)+(1,49+0,64)+(0,22-0,22)+(0,41-0,41)+(-1,95-0,49)+(0,88+0,88)+(0,22-0,22)+(0,38-0,38)+(-0,64+0,64)+(0,2+1,54)+(-1,41+1,41)+(0-0)+(1,97-0,32)+(0,56+1,69))^2)/(2*2)^2) = 0,54149336$$

$$\text{Data ke-2} = \text{Exp}(-(\text{Abs}((-0,72-1,67)+(-1,33-0,72)+(-0,5-1)+(2,03+0,43)+(1,49+0,64)+(0,22-0,22)+(0,41-0,41)+(-1,95-0,49)+(-0,88-1,08)+(0,22-0,22)+(0,38-0,38)+(-0,64-1,49)+(-0,2-1,14)+(-1,41+0,18)+(0*0)+(1,97+0,7)+(0,56-0,563))^2)/(2*2)^2) = 0,570291056$$

$$\text{Data ke-3} = \text{Exp}(-(\text{Abs}((-0,72+0,12)+(-1,33-0,72)+(-0,5+1,18)+(2,03+0,43)+(1,49+0,64)+(0,22-0,22)+(0,41-0,41)+(-1,95-0,49)+(-0,88+0,88)+(0,22-0,22)+(0,38-0,38)+(-0,64-1,49)+(-0,2+0,2)+(-1,41+0,46)+(0*0)+(1,97+0,7)+(0,56-0,563))^2)/(2*2)^2) = 0,982004043$$

Selanjutnya dilakukan perhitungan fungsi keputusan dengan persamaan

$$f(x) = \sum_{i=1}^m \alpha_i y_i K(x, x_i) + b$$

Penentuan klasifikasi data uji:

$$F(x) = \text{sign}((0,488 * -1,688 * 0,541) + 1,131) + ((0,727 * 0,563 * 0,570) + 1,131) + ((0,834 * 0,5627 * 0,982) + 1,131) = \text{sign}(3,643) = 1$$

Fungsi keputusan menghasilkan nilai 1, jadi data uji diatas termasuk kedalam kelas positif atau pinjamannya di terima. Proses pengujian klasifikasi ini dapat dilakukan



untuk banyak data uji sebelum kemudian ditentukan besaran pinjaman untuk data yang masuk kelas positif (diterima). Sedangkan data yang masuk kelas negatif, tidak perlu ditentukan besaran pinjamannya.

**D. Penentuan Besaran Pinjaman dengan RLB**

Variabel terikat yang digunakan adalah besaran pinjaman koperasi dan diskalakan menjadi menjadi 4, yaitu (1) Pinjaman rendah, (2) Pinjaman sedang, (3) Pinjaman Tinggi dan (4) Pinjaman Sangat tinggi. Sedangkan variable bebas yang digunakan ada 2, yaitu X<sub>3</sub> (pendapatan/bulan) dan X<sub>4</sub> (level pinjaman).

Model regresi linear berganda disajikan dengan persamaan sebagai berikut

$$Y = \beta_0 + \beta_3 X_3 + \beta_4 X_4$$

Keterangan:

Y = Besaran Pinjaman

X<sub>3</sub> = Pendapatan

X<sub>4</sub> = Level Pinjaman

β<sub>0</sub> = Konstanta

β<sub>i</sub> = Slope (Koefisien estimate)

Contoh digunakan data hasil klasifikasi sebagai berikut:

TABEL 9  
DATA KELAS "LAYAK MENERIMA PINJAMAN"

Y	X <sub>3</sub>	X <sub>4</sub>
0,74	1,49	0,995
0,74	-0,489	-0,3883
0,74	0,806	-0,3883
0,74	-0,489	-0,3883
0,74	-0,824	-0,3883

Untuk mempermudah perhitungan dalam menentukan nilai β<sub>0</sub> dan β<sub>i</sub>, maka dibuat tabel berikut:

TABEL 10  
TABLE PERSIAPAN RLB

Y	X <sub>3</sub>	X <sub>3</sub> Y	X <sub>3</sub> <sup>2</sup>	X <sub>4</sub>	X <sub>4</sub> Y	X <sub>4</sub> <sup>2</sup>
0,74	1,49	1,09	2,21	0,99	0,73	0,99
0,74	-0,48	-0,36	0,24	-0,38	-0,29	0,15
0,74	0,80	0,59	0,65	-0,38	-0,29	0,15
0,74	-0,48	-0,36	0,24	-0,38	-0,29	0,15
0,74	-0,82	-0,61	0,68	-0,38	-0,29	0,15
<b>3,68</b>	<b>0,48</b>	<b>0,36</b>	<b>4,01</b>	<b>-0,55</b>	<b>-0,41</b>	<b>1,59</b>

Selanjutnya akan dihitung nilai untuk β<sub>3</sub> dan β<sub>4</sub> menggunakan persamaan berikut:

$$\beta_i = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}$$

$$\beta_3 = \frac{5(0,36) - (0,489)(3,68)}{5(4,012) - 4,012} = 2,3016$$

$$\beta_4 = \frac{5(-0,411) - (-0,558)(3,68)}{5(1,593) - 1,593} = 3,39016$$

Setelah dihitung nilai β<sub>i</sub> selanjutnya akan dihitung nilai β<sub>0</sub> menggunakan persamaan sebagai berikut:

$$\beta_0 = \frac{1}{n} \sum_{i=1}^n Y_i - b_3 \frac{1}{n} \sum_{i=1}^n X_i - b_4 \frac{1}{n} \sum_{i=1}^n X_i = \bar{Y} - b_3 \bar{X} - b_4 \bar{X}$$

$$\beta_0 = \frac{1}{5}(3,68) - 2,3016 \frac{1}{5}(0,489) - 3,39016 \frac{1}{5}(-0,558) = 0,889655$$

Setelah diketahui semua nilai β maka selanjutnya telah terbentuk persamaan regresi linier berganda sebagai berikut:

$$Y = 0,8897 + 2,3016 X_3 + 3,39016 X_4$$

Persamaan tersebut memiliki arti bahwa pendapatan per bulan (X<sub>3</sub>) dan level pinjaman (X<sub>4</sub>) berpengaruh positif atau sebanding dengan besaran pinjaman pada Koperasi di Banjarmasin. Pendapatan per bulan (X<sub>3</sub>) berpengaruh positif terhadap besaran pinjaman sebesar 2,3016. Level pinjaman (X<sub>4</sub>) berpengaruh positif terhadap besaran pinjaman sebesar 3,39016. Apabila tidak terjadi perubahan pendapatan per bulan (X<sub>3</sub>) dan level pinjaman (X<sub>4</sub>), maka akan terjadi kenaikan besaran pinjaman sebesar 0,8897.

Berdasarkan persamaan linier tersebut maka dapat diperoleh skala prediksi besaran pinjaman pada Koperasi di Banjarmasin. Skala minimal diperoleh dengan cara mengasumsikan X<sub>3</sub>=X<sub>4</sub>=0. Sedangkan skala maksimal diperoleh dengan cara mengasumsikan X<sub>3</sub>=X<sub>4</sub>=1. Sehingga diperoleh skala besaran pinjaman adalah antara 0,8897 sampai 6,58. Skala besaran pinjaman tersebut dibagi menjadi empat skala untuk masing-masing kategori besaran pinjaman. Sehingga masing-masing skala bernilai 1,423.

TABEL 11  
PENENTUAN SKALA KATEGORI BESARAN PINJAMAN

No	Kategori	Skala
1	Pinjaman Rendah	750000 – 2300000
2	Pinjaman Sedang	2300000 – 5000000
3	Pinjaman Tinggi	5000000 – 25000000
4	Pinjaman Sangat Tinggi	>25000000

Dalam hal ini terdapat 5 data yang akan ditentukan besaran pinjamannya. Masing-masing dari 5 data tersebut akan dihitung menggunakan persamaan regresi linier yang telah dibentuk sebelumnya. Hasil perhitungan prediksi besaran pinjaman menggunakan regresi linier berganda ditunjukkan pada tabel berikut.

TABEL 12  
HASIL PREDIKSI BESARAN PINJAMAN

No	X <sub>3</sub>	X <sub>4</sub>	Y	Besar Pinjaman	Rp
1	1,49	0,995	7,6808	sangat tinggi	32000000
2	-0,49	-0,3883	-1,5522	rendah	20000000
3	0,81	-0,3883	1,4283	rendah	15000000
4	-0,49	-0,3883	-1,5522	rendah	20000000
5	-0,82	-0,3883	-2,3233	rendah	23000000



## V. UJI COBA

Pembagian data training dan testing digunakan teknik Cross validasi. Uji cross validasi bertujuan untuk mengukur akurasi nilai hasil pelatihan dan klasifikasi yang dihasilkan. Pembagian data training dan testing dengan teknik cross validasi digunakan 10 fold. Masing-masing fold akan berisi jumlah 1013 record yang masing-masing fold data akan diatur sedemikian rupa akan terjadi variasi kelas, sehingga

dalam setiap fold data terdapat kelas “Layak: dan juga terdapat kelas “Tidak Layak”. Terdapat 2 jenis pengujian yaitu uji klasifikasi dengan algoritma SVM dan yang kedua adalah uji penentuan besar pinjaman menggunakan Regresi Linier berganda.

Pengujian klasifikasi dengan SVM bertujuan untuk menilai keakurasian hasil klasifikasi data. Percobaan akan dilakukan sebanyak 10 kali sesuai dengan jumlah fold (k=10) dalam cross validation. Proses validasi akan dilakukan pada setiap fold data yang dimiliki. kernel yang dipakai dalam pengujian yaitu linear, polynomial dan RBF. Data yang digunakan sebanyak 10.134 dengan data testing sebanyak 3041 dan data training 7093 artinya data yang dipakai sebagai data testing sebanyak 30% dan data training 70% dan pengujian dilakukan dengan variasi parameter sebanyak 10 kali.

TABEL 13  
HASIL PENGUJIAN KLASIFIKASI DENGAN SVM

Data ke-	Kernel Linear	Kernel Polynomial	Kernel RBF
1	71,69	71,93	71,96
2	72,19	70,56	71,72
3	71,3	70,80	70,73
4	71,66	71,16	72,79
5	70,26	71,16	72,55
6	70,63	71,96	71,23
7	70,40	71,79	69
8	69,87	71,96	68,44
9	70,93	72	68,60
10	70,73	70,16	65,45

Pengujian Penentuan Besaran Pinjaman dengan Regresi Linier Berganda bertujuan untuk mencari Mean Square Error (MSE). Percobaan akan dilakukan sebanyak 10 kali ujicoba dengan data yang sama dengan data yang dipakai SVM dengan parameter SVM bervariasi.

TABEL 14  
HASIL PENGUJIAN PENENTUAN BESARAN PINJAMAN DENGAN MLR

No.	LINEAR	POLY	RBF
	MSE	MSE	MSE
1	0,01145	0,01124	0,01148
2	0,01124	0,01103	0,01194
3	0,01131	0,0115	0,01178
4	0,01115	0,01172	0,0116
5	0,011	0,01138	0,01153
6	0,01154	0,01116	0,01127
7	0,01162	0,01186	0,01127
8	0,01149	0,01135	0,01141
9	0,01112	0,01176	0,01158
10	0,01146	0,01151	0,0116
<b>AVG</b>	<b>0,01134</b>	<b>0,01145</b>	<b>0,01155</b>

Dari keseluruhan uji coba yang telah dilakukan dapat disimpulkan bahwa penentuan besaran pinjaman pada koperasi di Banjarmasin dengan memanfaatkan RLB yang telah dibangun menunjukkan hasil yang cukup baik. Hal ini dibuktikan dengan Mean Square Error ( MSE ) dengan rata-rata 0,011

## VI. Kesimpulan

1. Aplikasi yang dibuat terbukti dapat melakukan klasifikasi terhadap data pemohon pinjaman pada koperasi di Banjarmasin. Hal ini dibuktikan dengan pembahasan yang telah dilakukan pada proses pengujian menunjukkan hasil data testing masuk ke dalam kelas positif atau layak menerima pinjaman.
2. Dilakukan uji coba pelatihan dengan algoritma Support Vector Machine sebanyak 10 kali percobaan dengan variasi parameter. Dari uji coba ini menunjukkan bahwa hasilnya cukup baik yaitu akurasi hasil klasifikasi sebesar 72.79%.
3. Pada uji coba penentuan besaran pinjaman dengan menggunakan Regresi Linier Berganda menunjukkan Mean Square Error (MSE) sebesar 0,011
4. Dalam menjalankan aplikasi untuk penentuan kelayakan dan besaran pinjaman pada koperasi di Banjarmasin ini diperlukan spesifikasi hardware yang cukup tinggi. Hal ini disebabkan karena proses training dan testing melakukan perulangan (iterasi) yang cukup banyak dalam proses pembangkitan matriks dan persamaan langrage.

## DAFTAR PUSTAKA

- [1] L. Changjian and H. Peng, “Credit Risk Assessment for Rural Credit Cooperatives Based on Improved Neural Network,” 2017 Int. Conf. Smart Grid Electr. Autom., pp. 227–230, 2017.
- [2] P. Wongchinsri and W. Kuratach, “SR-based binary classification in credit scoring,” ECTI-CON 2017 - 2017 14th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol., vol. 0, no. c, pp. 385–388, 2017.

- [3] A. K. I. Hassan and A. Abraham, "Modeling consumer loan default prediction using ensemble neural networks," Proc. - 2013 Int. Conf. Comput. Electr. Electron. Eng. 'Research Makes a Differ. ICCEEE 2013, pp. 719-724, 2013.
- [4] A. Gahlaut and P. K. Singh, "Prediction analysis of risky credit using Data mining classification models," 2017.
- [5] Y. Ma and H. Liu, "Research of SVM applying in the risk of bank's loan to enterprises," 2nd Int. Conf. Inf. Eng. Comput. Sci. - Proceedings, ICIECS 2010, no. 3, pp. 1-5, 2010.
- [6] A. Pratama, R. C. Wihandika, and D. E. Ratnawati, "Implementasi Algoritme Support Vector Machine ( SVM ) untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa," vol. 2, no. 4, pp. 1704-1708, 2018.
- [7] N. D. S, "Penerapan Algoritma Support Vector Machine untuk Prediksi Harga Emas," pp. 10-19, 2015.
- [8] S. Nurhayati, E. T. Luthfi, and U. Y. Papua, "Prediksi Mahasiswa Drop Out Menggunakan Metode Support Vector," Prediksi menggunakan SVM, vol. 3, no. 6, pp. 82-93, 1978.
- [9] Bertalya, Konsep Data Mining Klasifikasi: Pohon Keputusan, Universitas Gunadarma, 2009
- [10] Mohammed, Mohssen, Khan, Muhammad Badruddin , dan Bashier, Eihab Bashier Mohammed. Machine Learning: Algorithms and Applications. CRC press. New York. 2017.
- [11] Gorunescu, F. Intelligent Systems Reference Library. Gorunescu, Ed. 2011
- [12] Nugroho, Anto Satriyo, Arief Budi Witarto dan Dwi Handoko. Suport Vector Machines : Teori Aplikasinya dalam Bioinformatika. ilmukomputer.com. 2003
- [13] Kurniawan, Robert. Analisis Regresi Dasar dan Penerapannya dengan R. Prenada Media. Jakarta. 2016
- [14] BigsmiLe, Mr. Mengenal Teknologi Machine Learning. Codepolitan.com.2016