

# KLASIFIKASI KERUSAKAN KAWASAN KONSERVASI DENGAN METODE *SUPPORT VECTOR MACHINE* (SVM) MENGGUNAKAN KERNEL GAUSSIAN: STUDI KASUS THE NATURE CONSERVANCY

Syaiful Anwar<sup>1</sup>, M.Syafrullah<sup>2</sup>

<sup>1,2</sup>)Program Studi Magister Ilmu Komputer, Program Pascasarjana, Universitas Budi Luhur  
Jl. Raya Ciledug, Petukangan Utara, Pesanggrahan, Jakarta Selatan 12260  
Telp. (021) 5853753, Fax. (021) 5869225  
<sup>1</sup>aefsyaiiful.anwar@gmail.com, <sup>2</sup>mohammad.syafrullah@budiluhur.ac.id

## ABSTRAK

Untuk dilakukan konservasi alam perlu dilakukan perancangan sebuah kawasan dengan analisa data dari hasil survey lalu diolah dan dipetakan pada aplikasi GIS dan dilakukan penentuan kawasan konservasi dengan melibatkan pakar setiap dilakukan perancangan. SVM merupakan metode analisis data yang digunakan untuk membentuk model yang mendeskripsikan kelas data yang penting, atau model yang memprediksikan trend data. Pada penelitian ini, untuk mengatasi kendala aplikasi metode SVM pada data yang tidak dapat dipisah secara linier digunakan metode kernel berbasis Gaussian. Metode kernel lain yaitu kernel polynomial juga akan dibahas untuk dibandingkan dengan metode kernel Gaussian. Hasil dari penelitian ini menunjukkan bahwa penggunaan kernel Gaussian dapat meningkatkan tingkat akurasi dari klasifikasi data menggunakan SVM dengan pengujian dataset Pulau Sumba dengan menggunakan fungsi kernel Gaussian tingkat akurasi klasifikasi mencapai 98.3% dan menggunakan kernel Polynomial 73,09%.

**Kata Kunci:** Konservasi, data mining, klasifikasi, SVM, kernel gaussian, akurasi

## 1. PENDAHULUAN

Negara Indonesia merupakan salah satu negara kepulauan terbesar di dunia yang memiliki sekitar 17.508 buah pulau yang membentang sepanjang 5.120 km dari timur ke barat sepanjang katulistiwa, dan 1.760 km dari utara ke selatan. Disamping itu dengan panjang pantai sekitar 81.791 km, menambah predikat Indonesia sebagai salah satu negara yang memiliki pantai terpanjang di dunia [1].

Salah satu yang menyebabkan terancamnya sebuah kawasan dengan keanekaragaman hayati adalah dari penduduk itu sendiri maupun pendatang yang mencari sumberdaya alam untuk kepentingan pribadinya. Makin besarnya ancaman makin kecil peluang kawasan tersebut di konservasi dengan pertimbangan mengoptimalkan kawasan yang masih terjaga. Bisa jadi kawasan yang sudah terancam rusak akan di jadikan kawasan industri, wisata, maupun kawasan rehabilitas. Untuk kawasan yang masih terjaga dengan sedikitnya ancaman yang timbul maka akan semakin dilindungi [2].

Sistem aplikasi pemetaan yang digunakan saat ini oleh Program Marine Konservasi diantaranya Aplikasi Geografi Information System (ArcGIS), yang berfungsi memuat informasi pemetaan wilayah konservasi dan untuk penentuan kawasan konservasi berdasarkan penghitungan pakar.

Dengan kebutuhan nilai akurasi data salah satu metode untuk teknik data mining adalah klasifikasi [3]. Salah satu metode klasifikasi menggunakan Support Vector Machine [4]. SVM mulai diperkenalkan pertama kali tahun 1979 oleh Vapnik.

Pada tahun 1992, SVM berkembang dengan pesat. SVM adalah salah satu teknik yang *relative* baru dibandingkan dengan teknik lain, tetapi memiliki performa yang lebih baik di berbagai bidang aplikasi seperti bio informatika, pengenalan tulisan tangan, klasifikasi teks dan lain sebagainya [5]. Kernel Gaussian adalah salah satu kernel yang sering digunakan dalam klasifikasi data SVM. Hal ini disebabkan kernel Gaussian memiliki tingkat akurasi yang paling besar dibanding kernel lain seperti *polynomial* untuk itu kernel menjadi solusi pengolahan data menggunakan metode SVM.

### 1.1. Maksud dan Tujuan

Adapun maksud dan tujuan dari penelitian ini adalah Bagaimana Algoritma *Support Vector Machine* (SVM) dapat digunakan untuk klasifikasi kerusakan kawasan dan Bagaimana pengujian untuk menghasilkan nilai akurasi klasifikasi data menggunakan algoritma *Support Vector Machine* (SVM) dengan kernel Gaussian.

### 1.2. Batasan Permasalahan

Ruang lingkup permasalahan yang akan dibahas pada penelitian ini hanya dibatasi permasalahan sebagai berikut:

1. Klasifikasi kawasan berdasarkan pembagian desa dengan identifikasi kawasan tersebut Layak atau tidak untuk dilakukan konservasi.
2. Metode yang dikembangkan menggunakan metode *Support Vector Machine* (SVM).
3. Data yang diambil untuk dikembangkan dalam pengolahan data dari hasil pemetaan ruang Pulau sumba dari hasil pengawasan kawasan dengan menggunakan ArcGIS.

4. Hasil dari analisa data akan menghasilkan pola grafik berpotensi untuk layak dan tidak layak.

## 2. TINJAUAN PUSTAKA

### 2.1 Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar.

Pengelompokan Data Mining Ada beberapa teknik yang dimiliki data mining berdasarkan tugas yang bisa dilakukan [5], yaitu

#### 1. Estimasi

Estimasi mirip dengan klasifikasi, kecuali variabel tujuan yang lebih kearah numerik dari pada kategori.

#### 2. Prediksi

Prediksi memiliki kemiripan dengan estimasi dan klasifikasi. Hanya saja, prediksi hasilnya menunjukkan sesuatu yang belum terjadi (mungkin terjadi di masa depan).

#### 3. Klasifikasi

Dalam klasifikasi variabel, tujuan bersifat kategorik. Misalnya, kita akan mengklasifikasikan pendapatan dalam tiga kelas, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah atau mengkasifikasi dua kelas seperti lulus atau gagal.

#### 4. Clustering

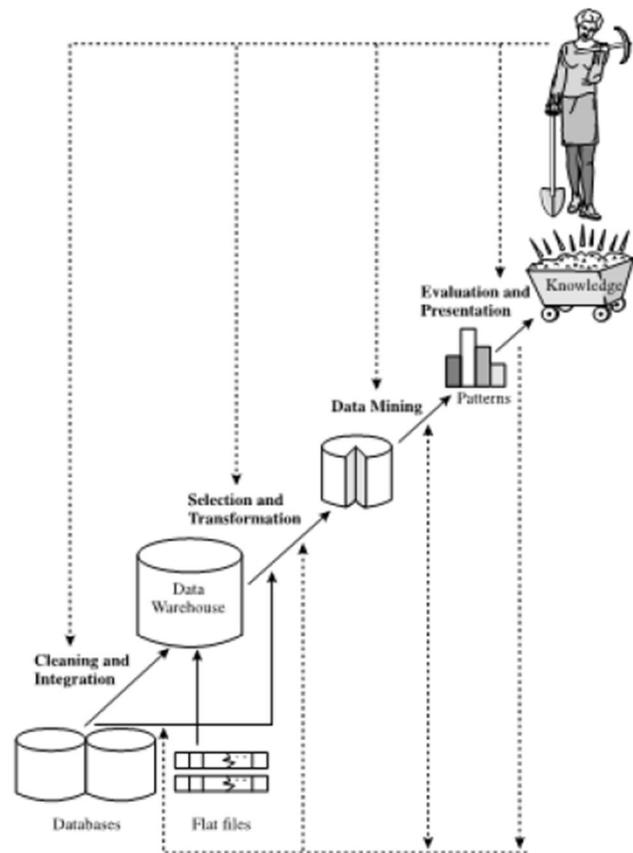
*Clustering* lebih ke arah pengelompokkan record, pengamatan, atau kasus dalam kelas yang memiliki kemiripan.

#### 5. Asosiasi

Mengidentifikasi hubungan antara berbagai peristiwa yang terjadi pada satu waktu.

### 2.2 Tahap-tahap Data Mining

Sebagai suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap proses yang diilustrasikan. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantara *knowledge base* [6].



Gambar 1. Ilustrasi Tahap data mining

Tahap-tahap umum dari data mining yang perlu diingat sebelum seorang analis menerapkan tahapan-tahapan data mining tersebut, sebagai prasyarat penerapan data mining, diperlukan pemahaman terhadap data dan proses diperolehnya data tersebut. Yang lebih mendasar lagi adalah diperlukannya pemahaman mengapa menerapkan data mining dan target yang ingin dicapai. Sehingga secara garis besar sudah ada hipotesa mengenai aksi-aksi yang dapat diterapkan dari hasilnya nanti. Pemahaman-pemahaman tersebut akan sangat membantu dalam mendesain proses data mining dan juga pemilihan teknik data mining yang akan diterapkan. Selain itu, bagi dunia bisnis akan memudahkan untuk melakukan pengukuran return on investment-nya (ROI).

### 2.3 Klasifikasi

Merupakan metode analisis data yang digunakan untuk membentuk model yang mendeskripsikan kelas data yang penting, atau model yang memprediksikan *trend* data. Klasifikasi digunakan untuk memprediksikan kelas data yang bersifat *categorical*, sedangkan prediksi untuk memodelkan fungsi yang mempunyai nilai *continuous*[4].

Contoh:

- Model klasifikasi yang dibangun untuk mengkategorisasikan aplikasi-aplikasi bank sebagai aplikasi yang aman atau beresiko

- Model prediksi yang dibangun untuk memprediksikan pengeluaran konsumen berdasarkan pendapatan dan pekerjaannya.

Tujuannya adalah *record-record* yg sebelumnya tidak terlihat dinyatakan kelasnya seakurat mungkin. Model klasifikasi digunakan untuk pemodelan deskriptif sebagai perangkat penggambaran untuk membedakan objek-objek dari kelas berbeda.

**2.4 Confussion Matrix**

*Confusion matrix* adalah sebuah tabel yang menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan. Contoh *confusion matrix* untuk klasifikasi biner ditunjukkan pada tabel 1.

Tabel 1. Confusion Matrix untuk klasifikasi biner [7]

Cost Matrix	Predicted Class		
	Class = Yes	Class = No	
Observed Class	Class = Yes	TP (True Positif)	FN (False Negatif)
	Class = No	FP (False Positif)	TN (True Negatif)

Keterangan untuk tabel 1 dinyatakan sebagai berikut:

**True Positive (TP)**, yaitu jumlah dokumen dari kelas *Yes* yang benar dan diklasifikasikan sebagai kelas *Yes*.

**True Negative (TN)**, yaitu jumlah dokumen dari kelas *No* yang benar diklasifikasikan sebagai kelas *No*.

**False Positive (FP)**, yaitu jumlah dokumen dari kelas *No* yang salah diklasifikasikan sebagai kelas *Yes*.

**False Negative (FN)**, yaitu jumlah dokumen dari kelas *Yes* yang salah diklasifikasikan sebagai kelas *No*.

Perhitungan akurasi dinyatakan dalam persamaan [7].

Untuk menghitung akurasi digunakan formula :

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{Jumlah data yang diprediksi secara benar}}{\text{jumlah prediksi yang dilakukan}} \\
 &= \frac{f11 + f00}{f11 + f10 + f01 + f00}
 \end{aligned}$$

Untuk menghitung laju *error* (kesalahan prediksi) digunakan formula :

$$\begin{aligned}
 \text{laju error} &= \frac{\text{Jumlah data yang diprediksi secara salah}}{\text{jumlah prediksi yang dilakukan}} \\
 &= \frac{f10 + f01}{f11 + f10 + f01 + f00}
 \end{aligned}$$

Cost Matrix	Predicted Class				
		Class = Yes		Class = No	
Observed Class	Class = Yes	TP =	21	TN =	22
	Class = No	FP =	4	FN =	3

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP+TN}{TP+FN+FP+TN} = \frac{21+22}{21+3+4+22} \\
 &= \frac{43}{50} \\
 &= 0,86 \times 100\% \\
 &= \mathbf{86.00\%}
 \end{aligned}$$

**2.5 Metode Support Vector Machine**

Metode ini dari teori pembelajaran statistik yang hasilnya sangat memberikan hasil yang lebih baik daripada metode yang lain. Ide dasar SVM adalah memaksimalkan batas *hyperplane* (*maximal margin hyperplane*) [8].

SVM adalah metode *learning machine* yang bekerja atas prinsip *Structural Risk Minimization* (SRM) dengan tujuan menemukan *hyperplane* terbaik yang memisahkan dua buah *class* pada *input space*.

Pada awalnya SVM dikembangkan untuk persoalan klasifikasi dua kelas, kemudian dikembangkan kembali untuk klasifikasi multikelas. Dalam klasifikasi kasus multikelas, *hyperplane* yang terbentuk adalah lebih dari satu. Salah satu metode pendekatan yang digunakan adalah satu lawan semua (SLA). Metode SLA untuk kasus klasifikasi k-kelas, menemukan k *hyperplane* dimana k adalah banyak kelas dan ρ adalah *hyperplane*. Dalam metode ini ρ(ℓ) diujikan dengan semua data dari kelas ℓ dengan label +1, dan semua data dari kelas lain dengan label -1. Konsep pada SLA yaitu dimisalkan pada kasus tiga kelas, kelas 1, 2, dan 3. Bila akan diujikan ρ(1), semua data dalam kelas 1 diberi label +1 dan data dari kelas 2 dan 3 diberi label -1. Pada ρ(2), semua data dalam kelas 2 diberi label +1 dan data dari kelas 1 dan 3 diberi label -1. Begitu juga untuk ρ(3), semua data dalam kelas 3 diberi label +1 dan data dari kelas 1 dan 2 diberi label -1. Kemudian dicari *hyperplane* dengan algoritma SVM dua kelas. Maka akan didapat *hyperplane* untuk masing-masing kelas diatas. Kemudian kelas dari suatu data baru x ditentukan berdasarkan nilai terbesar dari *hyperplane* [8].

Meskipun waktu pelatihan SVM kebanyakan lambat, tetapi metode ini sangat akurat karena kemampuannya untuk menangani model-model nonlinier yang kompleks. SVM dapat digunakan untuk prediksi dan klasifikasi.

**2.6 Support Vector Machine Metode Kernel**

Kinerja SVM sangat tergantung pada pilihan yang memadai dari nilai-nilai parameter, termasuk misalnya, kernel dan parameter regularisasi. Pemilihan parameter SVM umumnya sebagai masalah optimasi di mana teknik pencarian digunakan untuk menemukan konfigurasi parameter yang memaksimalkan kinerja SVM [9]

Metode kernel adalah solusi yang digunakan dalam aplikasi SVM ketika data yang akan diklasifikasi sulit atau bahkan tidak mungkin untuk diklasifikasikan dengan bidang pembatas linier. Metode kernel untuk SVM diperkenalkan tahun 1992 oleh Boser, Guyon, dan Vapnik [7]. Penggunaan metode kernel

menyebabkan suatu data  $x$  dalam ruang input dipetakan ke ruang fitur  $F$  dengan dimensi yang lebih tinggi melalui map  $\phi$  sebagai berikut:  $\phi: x \rightarrow \phi(x)$ . Pemetaan ini dilakukan dengan tetap menjaga karakteristik atau topologi data.

### 2.7 Kernel Gaussian

Diantara beberapa kernel SVM yang sering digunakan yaitu linier, *polynomial* dan *Gaussian*, kernel *Gaussian* adalah kernel yang mempunyai tingkat keakuratan yang paling baik. Fungsi kernel Gaussian [7] adalah:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right),$$

dimana  $\sigma^2$  adalah variansi. Seringkali  $\frac{1}{2\sigma^2}$  dituliskan sebagai variabel  $\gamma$  seperti berikut ini:

$$\gamma = \frac{1}{2\sigma^2},$$

Sehingga persamaan menjadi:

$$K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2), \gamma > 0,$$

### 2.8 Kernel Polynomial

Dalam *machine learning*, kernel polinomial adalah kernel yang biasa digunakan dengan mesin dukungan vektor (SVM) dan model *kernelized* lainnya, yang mewakili kesamaan vektor (sampel pelatihan) di ruang fitur lebih *polinomial* variabel asli. Secara intuitif, kernel *polinomial* terlihat tidak hanya pada fitur yang diberikan sampel masukan untuk menentukan kesamaan mereka, tetapi juga kombinasi dari ini. Dalam konteks analisis regresi, kombinasi tersebut dikenal sebagai fitur interaksi. The (implisit) ruang fitur dari kernel *polinomial* setara dengan regresi *polinomial*, tapi tanpa blowup kombinasi dalam jumlah parameter yang akan dipelajari. Ketika fitur input biner bernilai (*boolean*), maka fitur sesuai dengan konjungsi logis dari fitur input.

Polinomial:  $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$

## 3. RANCANGAN SISTEM DAN APLIKASI

### 3.1 Metode Penelitian

Penelitian Data Mining untuk klasifikasi Ekosistem Kawasan konservasi alam di Pulau Sumba Menggunakan metode *Support Vector Machine* dari data hasil pengolahan GIS. Dengan metode SVM proses uji kelayakan kawasan konservasi dilakukan perhitungan dengan berdasarkan data testing dan pelatihan. Hasil perhitungan akan dianalisis untuk disimpulkan tingkat kualitasnya. Kesemua proses pada penelitian ini bertujuan untuk menghasilkan klasifikasi kualitas batubara dengan metode *support vector machine*. Dalam penelitian ini akan menerapkan teori pengembangan kernel Gaussian menggunakan Bahasa pemrograman C#.

### 3.2 Metode Pemilihan Sampel

Untuk Pemilihan sampel berdasar pada metode Quota Sampling karena sampel yang diambil berdasar pada kebutuhan penelitian ini dan telah ditentukan kriteria atau kondisinya. Sampel dalam penelitian adalah data dari sebuah keadaan kawasan di Pulau sumba berdasarkan hasil survey selama beberapa bulan kemudian diolah dan dipetakan di ArcGIS maka diambil 353 data. Adapun parameter yang diambil sebagai pengukur kelayakan sebuah kawasan untuk di konservasi adalah :

1. Habitat : tempat suatu makhluk hidup tinggal dan berkembang biak. Pada dasarnya, habitat adalah lingkungan di sekeliling populasi suatu spesies yang memengaruhi dan dimanfaatkan oleh spesies tersebut.
2. *Species*/Binatang : Binatang yang sering muncul di suatu kawas dan memerlukan perhatian penuh agar kelangsungan hidupnya tidak terancam.
3. Ancaman : Keadaan dimana sebuah kawasan terancam dan ancaman ini sebagian besar muncul dari ulah manusia dan memerlukan penanganan atau memberlakukannya hukum, contoh : Bom Ikan, pendirian pelabuhan, perusakan habitat dll.
4. Sosial Ekonomi : keadaan ekosistem antara manusia dan alam yang dimana harus diberikan arahan dan bimbingan kepada penduduk untuk tidak merusak alam.
5. Alur : sebuah jalur yang dapat merusak ekosistem contoh : Penarikan kabel FO, telepon dll.
6. Kondisi Pendukung Konservasi : sebuah keadaan yang membantu pelestarian alam sebagai contoh pembangunan kawasan diving, snorkeling oleh atau atas seizin pemerintah setempat

### 3.3 Metode Pengumpulan Data

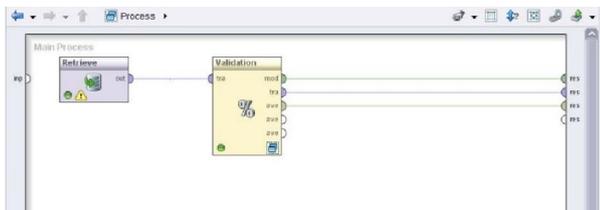
Metode pengumpulan data yang digunakan dalam penelitian ini adalah:

1. Metode Wawancara : Wawancara dengan pihak-pihak yang berkaitan dengan penelitian. Teknik wawancara dilakukan dengan wawancara berstruktur. Dalam wawancara tersebut peneliti telah menyiapkan daftar pertanyaan yang berkaitan dengan pengembangan sistem informasi terintegrasi Biro Perniagaan. Responden dalam wawancara ini adalah Marine Konservasi Planer dan Data Spatial Manager. Pertanyaan-pertanyaan untuk mendapatkan data yang terkait dengan perancangan sebuah kawasan konservasi dan sebuah sistem pada tiap bagian yang sedang berjalan saat ini.
2. Metode Observasi. Observasi atau pengamatan langsung terhadap profil obyek penelitian. Teknik observasi dilakukan dengan observasi berstruktur dengan menyiapkan daftar kebutuhan data dan sumber data. Proses observasi dilakukan untuk mempelajari aplikasi sistem informasi pada tiap bagian, tujuan dan struktur organisasi, proses program, ketersediaan sistem yang berjalan, dan pemanfaatan data survey dan data GIS.

### 3.4 Metode Pemilihan Model

Model data mining dalam penelitian ini dibuat dengan menggunakan *software* Rapid Miner 5.2. yang mana model data ini salah satu tahapan penentuan metode yang akan digunakan dalam fase implementasi. Tahapan ini komparasi metode antara metode *Support Vector Machine*, *Naïve Bayes*, *Neural Network* dan *Decision Tree* yang mana Algoritma *Support Vector Machine* (SVM) adalah metode dengan nilai akurasi terbaik dengan data shapefile Pulau Sumba untuk klasifikasi dan dijalankan dengan menggunakan perangkat lunak tersebut.

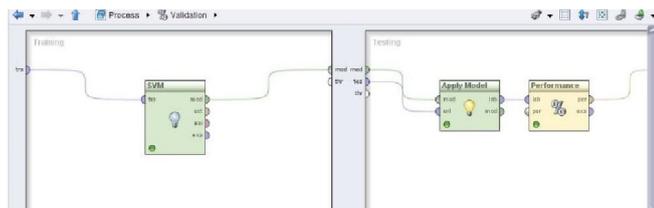
Model dibuat dengan sumber data dari Dataset. Dengan cara yang sama, berikut adalah model data mining menggunakan Rapid Miner 5.3.



Gambar 2. Model Data Mining di Rapidminer

Dari gambar 2 diatas, adalah proses input data ke *Rapidminer* untuk mendapatkan model data lalu di *Validation* untuk menentukan model apa yang akan dibuat.

Langkah selanjutnya adalah penentuan model algoritmanya, dalam hal ini adalah *Support Vector Machine* (SVM) dan *Naïve Bayes* sebagai komparasi nilai akurasi yang lebih baik. Pada tahap ini sekaligus ditambahkan tools untuk mengetahui performansi dari model. Proses ini dilakukan untuk masing-masing dataset yang telah dibuat.



Gambar 3. Training dan Testing SVM

Gambar 3. adalah proses pemilihan model yang dipilih yang mana diatas adalah model SVM yang akan diuji coba menjadi data Training dan menghasilkan data testing dan dijalankan dengan RapidMiner dan dikomparasi.

Dari hasil pengujian data menggunakan metode *Neural Network* memiliki nilai akurasi 78.18%. Dari hasil di atas yang dijalankan dalam aplikasi Rapidminer maka di dapat nilai akurasi dengan menggunakan metode SVM dengan hasil lebih baik yaitu 98.89 % dan di komparasikan menggunakan metode *naïve Bayes* dengan nilai akurasi 93.79% *Decision Tree* 95.15 dan menggunakan *Neural Network* dengan nilai akurasi 78.18% dan setelah dikomparasi maka dalam analisis ini akan diterapkan algoritma dengan metode *Support Vector Machine* (SVM) sebagai model terpilih dengan nilai akurasi 98.89% dengan pemilihan model menggunakan RapidMiner.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Pengolahan Data

Data yang digunakan harus ditransformasi terlebih dahulu sebelum diuji menggunakan aplikasi dengan metode SVM. Transformasi data menggunakan metode max min.

Dari kriteria data dengan 6 Kriteria (Habitat, Species, Kondisi Mendukung Konservasi, Ancaman, Sosial Ekonomi dan alur). Proses ini menggunakan perhitungan berdasarkan rumus menggunakan metode normalisasi Min Max dan dihasilkan data yang sudah di transformasi. Berikut Rumus metode normalisasi Min Max beserta penjelasannya.

$$X_{\text{new}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$$

$X_{\text{new}}$  : nilai x setelah normalisasi

$X_{\text{min}}$  : nilai x terkecil

$X_{\text{max}}$  : nilai x terbesar

Setelah di normalisasi dengan metode Max Min maka data siap di klasifikasikan untuk dilihat nilai akurasi dan pengujian menggunakan metode SVM berikut data yang sudah dinormalisasi beserta hasil pengujian menggunakan Confusion Matric. Berikut data yang sudah di transformasikan menggunakan Metode Normalisasi Min Max.

Selain dinormalisasi data untuk diukur dalam phase implementasi harus dicari nilai Y yaitu dengan penghitungan :

$$x_i \cdot w + b \geq +1 \text{ for } y_i = +1$$

$$x_i \cdot w + b \leq -1 \text{ for } y_i = -1$$

### 4.2 Penerapan Algoritma

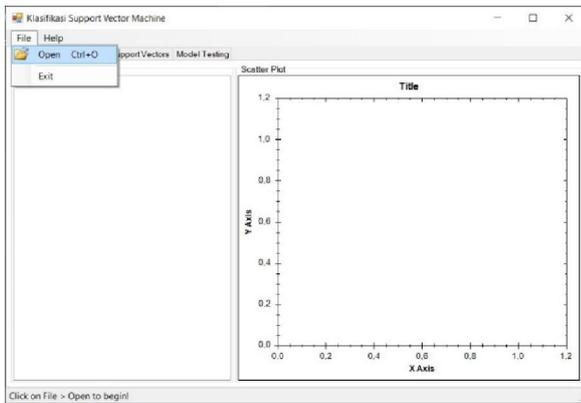
Prototipe aplikasi Klasifikasi kerusakan kawasan konservasi yang dibuat dengan menggunakan bahasa pemrograman C# dengan algoritma terpilih yaitu SVM dengan menggunakan Microsoft Visual Studio.

### 4.3 Implementasi

Implementasi antarmuka yang menggambarkan tampilan prototype yang dibangun yaitu implementasi antarmuka Klasifikasi kerusakan kawasan konservasi menggunakan metode SVM dengan 2 kernel sebagai perbandingan yaitu kernel Gaussian dan kernel Polynomial sebagai Machine Creation dengan hasil yang berbeda dan ditampilkan dengan grafik yang berbeda pula. Berikut ini adalah implementasi antarmuka dari prototype yang dibangun untuk mengukur tingkat akurasi data hasil prediksi yang dibangun menggunakan Bahasa pemrograman C# :

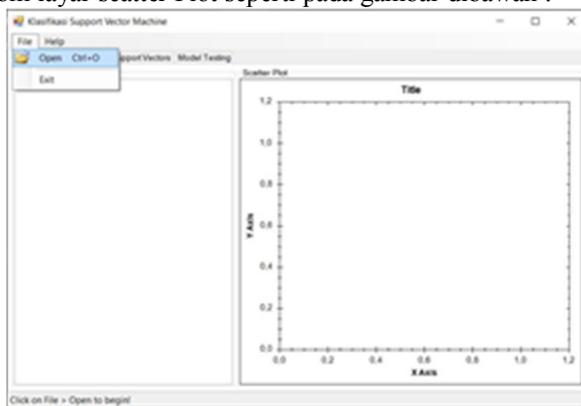
1. Layar Menu Input data

Untuk tampilan awal setelah *Running* aplikasi makan akan tampil Layar Input data, dimana data yang sudah di normalisasi di *upload* ke aplikasi untuk diukur tingkat akurasinya. Pada menu ini *user* akan dibawa ke pencarian file yang mana file tersebut adalah sebuah data yang sudah valid yang sudah sesuai dengan ketentuan program. Selain itu type file yang akan diproses data adalah format *xlsx* setelah dipilih data yang akan dioleh maka *user* akan dimita memilih file tersebut berada di sheet berapa baru data akan diproses oleh *system*. *Dataset* dalam pengujian ini menggunakan satu *dataset* yaitu *dataset* Pulau Sumba. Layar input data ditunjukkan pada gambar 4. Berikut



Gambar 4. Layar menu Insert

Pada layar tampilan ini *user* memasukan data yang akan dioleh menjadi *data training* dan akan di tampilkan data tersebut pada pada kolom *data training*. Setelah di tekan Open pada menu tampilan *Insert* maka akan dibawa ke *directory* untuk memasukan data yang sudah di oleh kedalam aplikasi. Setelah di pilih data yang akan diolah maka data akan ditampilkan dilayar *Data training* dan akan ditampilkan pada kolom layar scatter Plot seperti pada gambar dibawah :



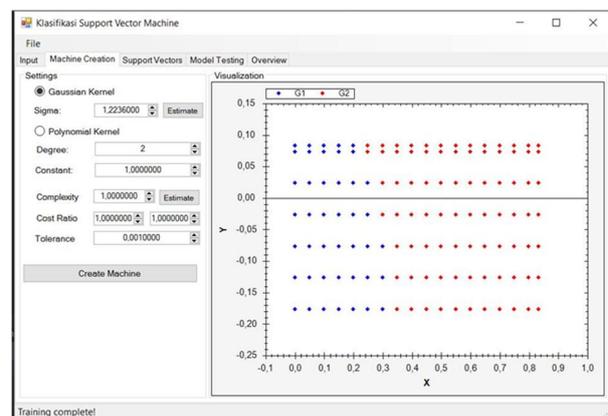
Gambar 5. Tampilan Data Training

Setelah input data maka akan tampil data yang sudah di input dan menghasilkan grafik serta *data training* yang selanjutnya akan di proses menjadi *data testing*.

2. Layar Menu *Machine Creation*

Pada menu ini setelah data dimasukan ke aplikasi dan sudah menjadi data training, pada tab selanjutnya menampilkan *data machine creation* yang mana data akan dioleh sesuai kernel yang akan dipilih, ada 2 pilihan kernel yang tersedia sebagai komparasi pemilihan kernel yaitu kernel Gaussian dan kernel Polynomial.

Dalam menu *machine creation* kernel Gaussian terdapat *estimate Sigma*, yaitu parameter dalam klasifikasi yang akan mempengaruhi nilai akurasi yang terbentuk semakin tinggi dimensi suatu ruang vektor, semakin besar kemungkinan menemukan bidang yang memisahkan data dua buah kelas secara linear atas dasar.



Gambar 6. *Machine Creation* Gaussian Kernel

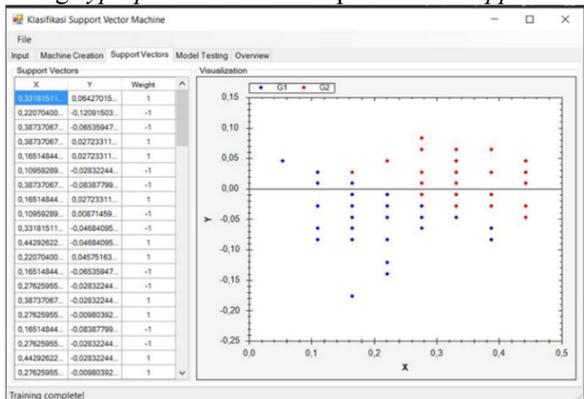
Setelah memilih *machine creation gaussian* lalu klik *estimate* untuk mengukur nilai rata-rata data dan akan terbentuk *support vector* maka akan diketahui nilai *w* setiap kawasan dan akan dibunakan dalam data testing.

3. Tampilan Menu *Support Vector*

Pada menu ini akan di terbentuk *data training* yang sudah di test dengan kernel yang dipilih. Terdapat informasi *Wight* pada tampilan menu ini, *Weights* adalah nilai koefisien sebenarnya tidak lain merepresentasikan posisi *hyperplane* itu pada ruang vektor. Proses *training* pada *perceptron* berjalan dengan mengoreksi nilai *weights* secara berulang, sedemikian hingga kedua buah kelas dapat dipisahkan secara sempurna oleh *hyperplane* itu. *Weight* ini adalah hasil akhir dari proses pembelajaran pada *perceptron*. Adapun pada SVM, tujuan dari proses *training* adalah untuk mencari posisi optimal dari *hyperplane* itu di ruang vektor. Dalam hal ini parameter yang dipakai adalah "margin" yaitu jarak dari *separating hyperplane* ke dua buah kelas.

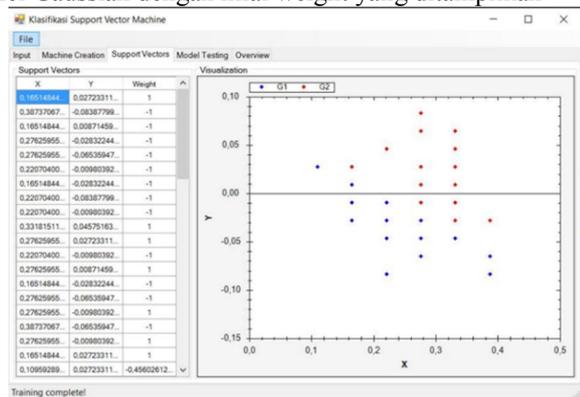
Hasil perhitungan menunjukkan, bahwa *hyperplane* yang terbaik dicapai dengan memaksimalkan nilai margin. Posisi ini tercapai jika *hyperplane* itu terletak tepat di tengah-tengah, memisahkan kedua buah kelas. Pada SVM, fungsi klasifikasi dinyatakan dalam dual form, yaitu dinyatakan sebagai fungsi dari *data training*, bukan *weights*. Proses training SVM sebagaimana dijelaskan diatas, bertujuan mencari data mana dari *training set* yang "paling informatif". Subset dari *training*

set inilah hasil akhir dari training tsb. dan disebut dengan *Support Vektor*. Di ruang *vektor*, *support vektor* ini adalah data dari kedua buah kategori yang terletak paling dekat dengan separating *hyperplane*. Berikut tampilan menu *support Vector*



Gambar 7. Support Vector Gaussian

Pada gambar diatas membentuk pola dan hasil w dari kernel Gaussian dengan nilai weight yang ditampilkan



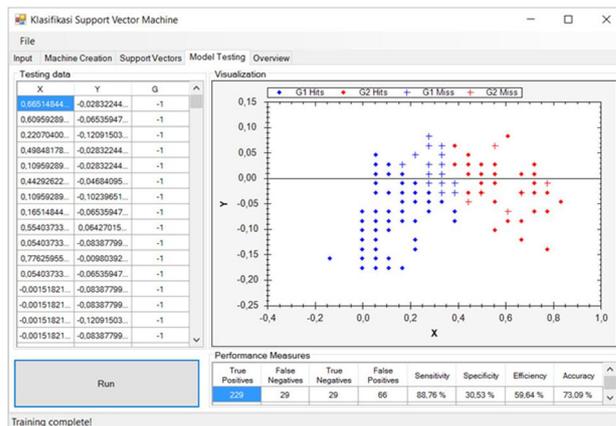
Gambar 8. Support Vector Polynomial

4. Tampilan Menu Model Testing

Pada menu ini adalah menu dimana tingkat yang paling penting yaitu menampilkan nilai akurasi dari sebuah prediksi berdasarkan pengukuran komparasi 2 kernel. Kernel Gaussian dan kernel Polynomial.



Gambar 9. Model Testing Kernel Gaussian



Gambar 10. Model Testing Kernel Polynomial

Berdasarkan hasil pengujian performa pada gambar diatas dengan menggunakan data shapefile di pulau sumba sebanyak 353 dengan matrik :

Tabel 2. Hasil Data Testing

Hasil	Kernel Gaussian	Kernel Polynomial
True Positif	80	229
True Negatif	267	29
False Positif	3	29
False Negatif	3	66
Sensitifity	96.39%	88.76%
Spesificity	98.89%	30.53%
Efficiency	97.64%	59.64%
<b>Accuracy</b>	<b>98.30%</b>	<b>73.09%</b>

5. KESIMPULANAN

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan bahwa:

1. Dengan menggunakan tools rapidminer dapat dibuat model datamining yang menjadi dasar dari pembuatan aplikasi klasifikasi kerusakan kawasan konservasi dibuat dengan menggunakan Bahasa pemrograman Microsoft Visual C#.
2. Klasifikasi dengan metode SVM dimulai dengan melakukan data training untuk memperoleh fungsi pemisah yang membagi data ke dalam dua kelas. Setelah menemukan fungsi pemisah, proses selanjutnya adalah melakukan klasifikasi terhadap data uji untuk melihat tingkat akurasi model SVM.
3. Pengolahan data untuk klasifikasi, algoritma Support Vector Machine (SVM) dapat diterapkan untuk klasifikasi kerusakan kawasan sebagai penentu kawasan konservasi.
4. Hasil pengujian dengan implementasi pembangunan aplikasi menggunakan algoritma SVM dengan Bahasa pemrograman C# di peroleh nilai accuracy, sensitivity, specificity, Efficiency diperoleh nilai dengan hasil terbaik menggunakan kernel Gaussian dengan nilai sensitivity 96.39 specificity 98.89 Efficiency 97.64 Accuracy 98.30%.

## DAFTAR PUSTAKA

- [1] Bhardwaj, B.K., 2011. Data Mining : A prediction for performance improvement using classification. (IJCSIS) International Journal of Computer Science and Information Security,, 9(4).
- [2] Chandani, V., Wahono, R.S. & Purwanto, ., 2015. Komparasi Algoritma Klasifikasi Machine Learning Dan Feature Selection pada Analisis Sentimen Review Film. Journal of Intelligent Systems, 1(1), pp.55–59. Available at: <http://journal.ilmukomputer.org/index.php/jis/article/view/10>.
- [3] Fadli, A., 2011. Konsep Data Mining. , pp.1–9. Available at: [www.ilmukomputer.com](http://www.ilmukomputer.com).
- [4] Furey, T.S. et al., 2000. Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics, 16(10), pp.906–914. Available at: <http://bioinformatics.oxfordjournals.org/content/16/10/906>.
- [5] Hadi, S. & Rusdiana, O., 2015. Kawasan Taman Nasional Ujung Kulon Provinsi Banten ( Potential Analysis of Ecotourism Landscape in the Buffer Zone of Ujung Kulon National Park Banten Province ). , (5), pp.135–144.
- [6] Handayani, N.T., 2013. Model pembobotan untuk penentuan kesesuaian kawasan konservasi terumbu karang di pulau kodingarenglompo kota makassar.
- [7] Jacobus, A. & Winarko, E., 2014. Penerapan Metode Support Vector Machine pada Sistem Deteksi Intrusi secara Real-time. Ijccs, 8(1), pp.13–24.
- [8] Kabakchieva, D., 2012. Student Performance Prediction by Using Data Mining Classification Algorithms. International Journal of Computer Science and Management Research, 1(4), pp.686–690
- [9] Indriani, A. & Nbc, D., 2014. Klasifikasi Data Forum dengan menggunakan Metode Naïve Bayes Classifier. , pp.5–10.