

**EVALUASI ALGORITMA SUPPORT VECTOR MACHINE
DENGAN OPTIMASI METODE GENETIC ALGORITHM PADA
KLASIFIKASI BANJIR DI KOTA SAMARINDA**

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar
Sarjana Komputer

DISUSUN OLEH:

YULIANA DILLA EVITASARI

1911102441117



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR**

2023

**Evaluasi Algoritma Support Vector Machine dengan
Optimasi Metode Genetic Algorithm pada Klasifikasi Banjir
Di Kota Samarinda**

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar
Sarjana Komputer

Disusun Oleh:

Yuliana Dilla Evtasari

1911102441117



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
2023**

HALAMAN PENGESAHAN

EVALUASI ALGORITMA *SUPPORT VECTOR MACHINE*
DENGAN OPTIMASI METODE *GENETIC ALGORITHM* PADA
KLASIFIKASI BANJIR DI KOTA SAMARINDA

DISUSUN OLEH:
YULIANA DILLA EVITASARI
1911102441117

Telah melaksanakan ujian skripsi dan dinyatakan lulus,
Pada tanggal 06 Juli 2023

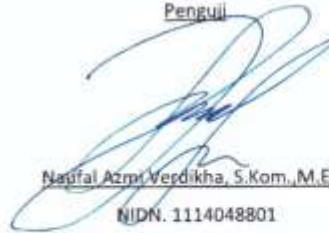
Dosen Pembimbing



Wawan Joko Pranoto, S.Kom, M.TI

NIDN. 1102057701

Penguji



Nasifal Azmi Verdikha, S.Kom, M.Eng

NIDN. 1114048801

Dekan



Prof. Dr. Sarlito, MT., Ph.D.

NIDN. 0610116204

Program Studi



Asas M. Ratipah, S.Kom., M.Cs

NIDN. 1124098902

PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama Mahasiswa : Yuliana Dilla Evtasari

NIM : 1911102441117

Konsentrasi : Sistem Informasi

Menyatakan bahwa Skripsi dengan judul berikut :

Evaluasi Algoritma Support Vector Machine Dengan Optimasi Metode Genetic Algorithm Pada Klasifikasi Banjir Di Kota Samarinda

Dosen Pembimbing : Wawan Joko Pranoto, S.Kom, M.TI

6. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar Sarjana Komputer, baik di Universitas Muhammadiyah Kalimantan Timur (UMKT) maupun di Perguruan Tinggi lainnya
7. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing
8. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
9. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas Muhammadiyah Kalimantan Timur (UMKT)
10. Pernyataan ini SAYA buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi Muhammadiyah.

Samarinda, 06 Juli 2023

Yang Menyatakan,

Yuliana Dilla Evtasari
1911102441117

PRAKATA



Assalamualaikum Warahmatullahi Wabarakatuh

Puji dan syukur dipanjatkan kepada Tuhan Yang Maha Esa, karena atas rahmatnya Skripsi yang berjudul “Evaluasi Peningkatan Akurasi Algoritma Support Vector Machine Dengan Optimasi Metode Genetic Algorithm Pada Klasifikasi Banjir Di Kota Samarinda” ini diajukan sebagai salah satu syarat untuk memenuhi persyaratan mencapai gelar Sarjana Komputer di Universitas Muhammadiyah Kalimantan Timur.

Penulis menyadari bahwa dalam penyelesaian Proposal ini mendapat bantuan dari berbagai pihak, baik langsung maupun tidak langsung. Oleh karena itu pada kesempatan ini penulis menyampaikan terima kasih kepada:

1. Orang tua penulis Bapak Daryanto dan Ibu Juwita yang selalu memberikan doa serta dukungan.
2. Bapak Wawan Joko Pranoto, S.Kom, M.Kom, selaku Dosen Pembimbing yang telah memberikan bimbingan dan masukan kepada penulis.
3. Ibu Asslia Johar Latipah, M.Cs., ketua Program Studi S1 Teknik Informatika.
4. Bapak Prof. Ir. Sarjito, M.T.,Ph.D., selaku Dekan Fakultas Sains & Teknologi.
5. Bapak Prof. Dr. H. Bambang Setiaji, selaku Rektor Universitas Muhammadiyah Kalimantan Timur.
6. Seluruh dosen S1 Teknik Informatika Universitas Muhammadiyah Kalimantan Timur yang telah memberikan ilmunya kepada penulis.
7. Kepada sahabat serta rekan bimbingan seperjuangan saya yaitu Endah Aulia Nisaulatifah, Septia Intan Permata Sari, Suci Mawaddah, Kamirul Hakim, Putra Saymen, Rahmat Ramadhani dan Muhammad Fath Thoriq yang telah banyak membantu dan berjuang Bersama-sama serta memberikan saran dan semangat sehingga dapat menyelesaikan skripsi ini.
8. Kepada teman-teman S1 Teknik Informatika Angkatan 2019.

9. Semua pihak yang tidak dapat penulis sebutkan satu persatu, terimakasih atas bantuan semangat serta nasehat dan dorongan dalam proses penyusunan skripsi ini.

Penulis menyadari skripsi ini masih jauh dari sempurna, oleh karena itu dengan segala hormat dan kerendahan hati, penulis berharap saran dan kritik yang membangun untuk hasil yang dituangkankan dalam skripsi ini bisa bermanfaat bagi pihak-pihak yang memerlukan.

Samarinda, 20 Maret 2023



Yuliana Dilla Evtasari

1911102441117

ABSTRAK

Banjir merupakan bencana alam yang sering terjadi di Indonesia, terutama di kota Samarinda yang terletak di Kalimantan Timur. Penelitian ini bertujuan untuk meningkatkan akurasi dengan menerapkan metode seleksi fitur menggunakan Genetic Algorithm (GA). Melalui analisis data banjir kota Samarinda, ditemukan bahwa terdapat tiga atribut yang paling berpengaruh terhadap terjadinya banjir, yaitu kelembapan, lamanya penyinaran matahari, dan kecepatan angin. Selanjutnya, penelitian ini menggunakan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan data banjir. Dengan menerapkan seleksi fitur menggunakan GA, hasil pengujian menunjukkan peningkatan akurasi algoritma SVM sebesar 13.45%. Sebelum penerapan seleksi fitur, akurasi SVM hanya mencapai 52,71%, namun setelah penerapan seleksi fitur menggunakan GA, akurasi meningkat menjadi 66,16%. Hasil ini membuktikan bahwa seleksi fitur dengan menggunakan GA efektif dalam meningkatkan akurasi prediksi banjir. Kesimpulan dari penelitian ini adalah seleksi fitur menggunakan GA dapat mengidentifikasi atribut-atribut yang paling berpengaruh terhadap terjadinya banjir di kota Samarinda. Penerapan seleksi fitur ini menghasilkan peningkatan signifikan dalam akurasi algoritma SVM untuk prediksi banjir.

Kata Kunci: Akurasi, *Support Vector Machine*, *Genetic Algorithm*, Seleksi fitur, Optimasi, Banjir.

ABSTRACT

Flooding is a natural disaster that often occurs in Indonesia, especially in the city of Samarinda, which is located in East Kalimantan. This research aims to improve accuracy by applying a feature selection method using Genetic Algorithm (GA). Through the analysis of flood data in Samarinda, it was found that there are three attributes that have the most influence on the occurrence of floods, namely humidity, length of sunshine, and wind speed. Furthermore, this research used the Support Vector Machine (SVM) algorithm to classify the flood data. By applying feature selection using GA, the test results showed an increase in the accuracy of the SVM algorithm by 13.45%. Before the application of feature selection, the accuracy of SVM only reached 52.71%, but after the application of feature selection using GA, the accuracy increased to 66.16%. These results prove that feature selection using GA is effective in improving flood prediction accuracy. The conclusion of this research is that feature selection using GA can identify the attributes that have the most influence on the occurrence of floods in Samarinda. The application of this feature selection results in a significant increase in the accuracy of the SVM algorithm for flood prediction.

Keywords: *Accuracy, Support Vector Machine, Genetic Algorithm, Feature Selection, Optimization, Flood.*

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN JUDUL	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN KEASLIAN SKRIPSI	iv
PRAKATA.....	v
ABSTRAK.....	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian.....	3
1.4. Batasan Masalah	3
1.5. Manfaat Penelitian	4
BAB 2 TINJAUAN PUSTAKA.....	5
2.1. Banjir.....	5
2.2. Data Mining	6
2.2.1. Klasifikasi.....	6
2.2.4. <i>Genetic Algorithm</i>	10
2.2.5. <i>Data Pre-Processing</i>	11
2.3. <i>K-Fold Cross Validation</i>	12
2.4. <i>Confusion Matrix</i>	12
2.5. Penelitian Terkait.....	13
BAB 3 METODOLOGI PENELITIAN	16
3.1. Objek Penelitian	16
3.2. Tahap Pelaksanaan	16

3.3.	Pengumpulan Data	17
3.4.	Data <i>Pre-Processing</i>	18
3.5.	Pembagian Data	19
3.6.	Modelling.....	19
3.6.1.	Pemodelan Algoritma SVM Dengan Python.....	19
3.6.2.	Pemodelan Algoritma SVM + GA Menggunakan Python	21
3.7.	Evaluasi.....	27
BAB 4 HASIL DAN PEMBAHASAN		28
4.1.	Hasil Pengumpulan Data	28
4.2.	Hasil <i>Data Pre-Processing</i>	30
4.2.1.	<i>Data Integration</i>	30
4.2.2.	<i>Data Selection</i>	30
4.2.3.	<i>Data Transformation</i>	30
4.2.4.	Data Cleaning.....	31
4.2.5.	Data Balancing	31
4.3.	Hasil Pembagian Data.....	32
4.4.	Hasil Modelling dan Evaluation	33
4.4.1.	Pemodelan Algoritma SVM Dengan Python.....	33
4.4.2.	Pemodelan SVM + GA Dengan Python	34
4.5.	Pembahasan	35
BAB 5 KESIMPULAN DAN SARAN		37
5.1.	Kesimpulan	37
5.2.	Saran.....	37
LAMPIRAN		42

DAFTAR TABEL

Tabel 2. 1 Nilai Confusion Matrix	13
Tabel 2. 2 Penelitian Terdahulu	14
Tabel 2. 3 Penelitian Terdahulu (Lanjutan)	15
Tabel 3. 1 Atribut Data BMKG dan BPBD	17
Tabel 3. 2 Parameter Pembagian Data Training dan Testing.....	20
Tabel 3. 3 Parameter Mencari Nilai Akurasi SVM	21
Tabel 3. 4 Parameter Evaluasi GA	22
Tabel 3. 5 Parameter Definisi Fungsi Individu dan Populasi dan Fungsi Operasi	23
Tabel 3. 6 Parameter Inisialisasi Populasi dan Definisi Statistik	24
Tabel 3. 7 Parameter GA	24
Tabel 3. 8 Parameter GA (Lanjutan).....	25
Tabel 3. 9 Parameter Mendapatkan Individu Terbaik	25
Tabel 3. 10 Parameter Fitur Terbaik	26
Tabel 3. 11 Parameter Nilai Akurasi SVM + GA.....	26
Tabel 4. 1 Data BMKG	28
Tabel 4. 2 Data BPBD.....	29
Tabel 4. 3 Hasil Data Selection	30
Tabel 4. 4 Hasil Tranformasi Data	31
Tabel 4. 5 Hasil Data Cleaning.....	31
Tabel 4. 6 Hasil Data Duplikat Pada Kelas 1	32
Tabel 4. 7 Hasil Model SVM	33
Tabel 4. 8 Rata-Rata Akurasi SVM.....	33
Tabel 4. 9 Hasil Akurasi SVM + GA	34
Tabel 4. 10 Rata-rata Akurasi SVM + GA	34

DAFTAR GAMBAR

Gambar 2. 1 (a) hyperlane non optimal (b) hyperlane optimal	8
Gambar 2. 2 Persamaan kernel SVM	9
Gambar 3. 1 Alur Penelitian	16
Gambar 3. 2 Import Dataset	19
Gambar 3. 3 Pembagian Data Training dan Testing.....	20
Gambar 3. 4 Mencari Nilai Akurasi SVM	21
Gambar 3. 5 Import Library	22
Gambar 3. 6 Definisi Fungsi GA.....	22
Gambar 3. 7 Inisialisasi Objek Toolbox	23
Gambar 3. 8 Definisi Fungsi Operasi GA	23
Gambar 3. 9 Inisialisasi Populasi dan Statistik	24
Gambar 3. 10 Menjalankan Algoritma GA	24
Gambar 3. 11 Mendapatkan Individu Terbaik	25
Gambar 3. 12 Fitur Terbaik Terpilih	25
Gambar 3. 13 Akurasi SVM + GA	26
Gambar 4. 1 Hasil Data Balancing	32
Gambar 4. 2 Perbandingan Hasil Dua Skema Algoritma.....	36

DAFTAR LAMPIRAN

Lampiran 1. Riwayat Hidup	43
Lampiran 2. Data Awal (BMKG & BPBD)	44
Lampiran 3. Hasil Data Pre-Processing.....	46
Lampiran 4. Source Code	48
Lampiran 5. Surat Ijin Penelitian Prodi.....	51
Lampiran 6. Surat Ijin Penelitian LPPM	52
Lampiran 7. Surat Balasan Kesbangpol	53
Lampiran 8. Lembar Bimbingan Skripsi.....	54
Lampiran 9. Uji Plagiasi	56
Lampiran 10. Hasil Uji Plagiasi.....	57