

**PERBANDINGAN SUPPORT VECTOR MACHINE DAN DECISION TREE MENGGUNAKAN  
BINARY DRAGONFLY ALGORITHM PADA KLASIFIKASI INDEKS KHUSUS  
PENANGANAN STUNTING (IKPS) KABUPATEN/KOTA DI INDONESIA**

**Andini Eka Saputri Pakaya**

Program Studi S1 Matematika, FMIPA, Universitas Negeri Gorontalo

email : andinipakaya20@gmail.com

**Isran K. Hasan**

Jurusan Matematika, FMIPA, Universitas Negeri Gorontalo

\*email : isran.hasan@ung.ac.id

**Novianita Achmad**

Jurusan Matematika, FMIPA, Universitas Negeri Gorontalo

email : novianita.achmad@ung.ac.id

**Abstrak**

Stunting merupakan kondisi hambatan pertumbuhan dan perkembangan anak akibat defisiensi gizi kronis dan infeksi berulang. Dalam kondisi ini, anak menunjukkan tinggi atau panjang tubuh dibawah ukuran standar yang ditetapkan. Untuk memantau upaya percepatan penurunan stunting yang telah dilakukan di tingkat nasional, provinsi, dan kabupaten/kota, dibentuklah suatu instrumen Indeks Khusus Penanganan Stunting atau IKPS. Penelitian ini menggunakan metode *Support Vector Machine* (SVM) dan *Decision Tree* serta menggunakan *Binary Dragonfly Algorithm* sebagai algoritma seleksi fitur, pada *dataset* Indeks Khusus Penanganan Stunting (IKPS) Kabupaten/Kota tahun 2022 yang bersumber dari *Website* Badan Pusat Statistik, dengan variabel penelitian yang terdiri dari imunisasi ( $X_1$ ), penolong persalinan oleh tenaga kesehatan di fasilitas kesehatan ( $X_2$ ), KB modern ( $X_3$ ), ASI eksklusif ( $X_4$ ), MP ASI ( $X_5$ ), air minum layak ( $X_6$ ), sanitasi layak ( $X_7$ ), Pendidikan Anak Usia Dini ( $X_8$ ), kepemilikan JKN/jamkesda ( $X_9$ ), dan penerima KPS/KKS atau bantuan pangan ( $X_{10}$ ), serta IKPS ( $Y$ ). Penelitian ini bertujuan untuk membandingkan metode *Support Vector Machine* dan *Decision Tree* menggunakan algoritma seleksi fitur *Binary Dragonfly Algorithm* dalam mengklasifikasikan Indeks Khusus Penanganan Stunting (IKPS) Kabupaten/Kota di Indonesia. Hasil penelitian menunjukkan metode SVM memperoleh akurasi klasifikasi sebesar 86%, sedangkan metode *Decision Tree* memperoleh akurasi sebesar 50%.

**Kata Kunci:** IKPS, Klasifikasi Data, *Support Vector Machine*, *Decision Tree*, *Binary Dragonfly Algorithm*

**Abstract**

Stunting is a condition that inhibits the growth and development of children due to chronic nutritional deficiencies and recurrent infections. In this condition, children show height or body length below the standard size set. To monitor efforts to accelerate the reduction of stunting that have been carried out at the national, provincial and district / city levels, a Special Index for Stunting Handling or IKPS instrument was formed. This study uses the Support Vector Machine (SVM) and Decision Tree methods and uses the Binary Dragonfly Algorithm as a feature selection algorithm, on the 2022 District / City Special Index for Stunting Handling (IKPS) dataset sourced from the Central Statistics Agency Website, with research variables consisting of immunization ( $X_1$ ), delivery assistance by health workers in health facilities ( $X_2$ ), modern family planning (contraception methods) ( $X_3$ ), exclusive breastfeeding ( $X_4$ ), complementary feeding ( $X_5$ ), proper drinking water ( $X_6$ ), proper sanitation ( $X_7$ ), early childhood education ( $X_8$ ), JKN/jamkesda ownership ( $X_9$ ), and recipients of KPS/KKS or food assistance ( $X_{10}$ ), and IKPS ( $Y$ ). This study aims to compare the Support Vector Machine and Decision Tree methods using the Binary Dragonfly Algorithm feature selection algorithm in classifying the Special Index for Stunting Handling (IKPS) of Regency / City in Indonesia. The results showed that the SVM method obtained a classification accuracy of 86%, while the Decision Tree method obtained an accuracy of 50%.

**Keywords:** IKPS, Data Classification, *Support Vector Machine*, *Decision Tree*, *Binary Dragonfly Algorithm*

## 1. PENDAHULUAN

Isu stunting telah menjadi perhatian signifikan di berbagai negara, termasuk Indonesia (Syafika & Karisma, 2023). Stunting merupakan kondisi hambatan pertumbuhan dan perkembangan anak akibat defisiensi gizi kronis dan infeksi berulang (Badan Pusat Statistik, 2023). Berdasarkan data dari SSGI (Studi Status Gizi Indonesia), prevalensi stunting pada tahun 2022 sebesar 21.6% (Kementerian Kesehatan RI, 2022). Angka tersebut belum memenuhi standar yang disahkan oleh WHO (*World Health Organization*) yakni sebesar 20% (WHO, 2021).

Sebagai pemantau upaya percepatan penurunan stunting yang telah dilakukan di tingkat nasional, provinsi, dan kabupaten/kota, dibentuklah suatu instrumen Indeks Khusus Penanganan Stunting atau IKPS di tahun 2018 yang disempurnakan pada tahun 2020 (Badan Pusat Statistik, 2023). Untuk meningkatkan efektivitas penanganan stunting, dibutuhkan pengklasifikasian kategori IKPS seperti penelitian yang dilakukan oleh (Syafika & Karisma, 2023) yang mengklasifikasikan IKPS di tingkat provinsi.

Klasifikasi data merupakan proses kategorisasi menggunakan *machine learning* dengan penerapan *data mining* (Somantri dkk., 2022). Ada beberapa metode yang diterapkan untuk melakukan klasifikasi, diantaranya adalah *Support Vector Machine* (SVM), suatu metode yang mampu mengidentifikasi *hyperplane* dengan nilai optimal untuk membedakan dua kelas dalam ruang input (Sihombing & Yuliati, 2021). Selain SVM, *Decision Tree* juga termasuk metode yang dapat digunakan untuk klasifikasi. *Decision Tree* adalah metode yang membentuk suatu model struktur pohon, di mana pohon tersebut terdiri dari simpul yang merepresentasikan keputusan yang didasarkan pada nilai-nilai atribut, serta cabang-cabang yang menggambarkan hasil dari keputusan tersebut (Amirudin & Wowor, 2023). Kedua metode tersebut, kerap dibandingkan berdasarkan tingkat akurasi dalam mengklasifikasikan data, sebagaimana yang ditemukan dalam sejumlah penelitian, diantaranya penelitian yang dilakukan oleh (Jaya Purnama dkk., 2020), (Sarumi, 2021), (Oktafiani & Rianto, 2023), (Syahputra dkk., 2023), (Devisetty & Kumar, 2023), dan (Aneke dkk., 2024).

Sebelum menerapkan metode untuk mengklasifikasikan data, diperlukan tahapan seleksi fitur untuk mengurangi atribut yang tidak relevan dalam data yang dapat mempengaruhi hasil kinerja pada *machine learning* (Fitriani dkk., 2020). Algoritma untuk melakukan seleksi fitur diantaranya adalah *Binary Dragonfly Algorithm*, yang merupakan penyempurnaan *Dragonfly Algorithm*. *Dragonfly Algorithm* adalah pendekatan optimasi terbaru yang terinspirasi dari perilaku capung di alam (Chantar dkk., 2021). Penggunaan *Binary Dragonfly Algorithm* untuk seleksi fitur telah diterapkan dalam beberapa penelitian, seperti yang dilakukan oleh (Alhakeem dkk., 2023), (Raouf & Abd, 2023), dan (Guendouz & Amine, 2023). Hasil penelitian-penelitian tersebut menunjukkan bahwa penerapan algoritma ini memberikan kinerja yang lebih baik dibandingkan dengan yang tidak menggunakan *Binary Dragonfly Algorithm*.

Berdasarkan penelitian-penelitian sebelumnya, pada penelitian ini melakukan perbandingan metode *Support Vector Machine* dan *Decision Tree* menggunakan *Binary Dragonfly Algorithm* sebagai penentuan fitur-fitur yang akan digunakan, dengan tujuan memperoleh hasil klasifikasi Indeks Khusus Penanganan Stunting (IKPS) Kabupaten/Kota di Indonesia dengan kedua metode tersebut, dan menganalisis perbandingan keakuratannya, serta diharapkan dapat memberikan informasi kepada pemerintah sebagai dasar pertimbangan dalam merumuskan kebijakan untuk mengatasi permasalahan stunting pada Kabupaten/Kota di Indonesia.

## 2. KAJIAN TEORI

### Indeks Khusus Penanganan Stunting (IKPS)

Indeks Khusus Penanganan Stunting (IKPS) adalah suatu instrumen yang dirancang sebagai evaluasi kinerja implementasi program percepatan untuk mengurangi tingkat stunting skala nasional, provinsi, dan kabupaten/kota. IKPS baik secara nasional, provinsi, dan kabupaten/kota memiliki tujuan untuk dipergunakan dalam membandingkan seperti apa perkembangan lingkup-lingkup intervensi terhadap rumah tangga sasaran. IKPS berperan sebagai penghubung antara kepentingan kementerian/lembaga dalam menjalankan upaya penanganan stunting. IKPS di tingkat kabupaten/kota dibangun dari lima dimensi, yaitu

dimensi kesehatan yang mencakup indikator imunisasi, penolongan persalinan oleh tenaga kesehatan di fasilitas kesehatan, dan Keluarga Berencana (KB) modern, dimensi gizi mencakup indikator ASI eksklusif dan Makanan Pendamping (MP) ASI, dimensi perumahan mencakup indikator air minum layak, dan sanitasi layak, dimensi pendidikan terdiri dari indikator Pendidikan Anak Usia Dini (PAUD), serta dimensi perlindungan sosial yang mencakup indikator kepemilikan JKN/Jamkesda, dan penerima KPS/KKS atau bantuan pangan (Badan Pusat Statistik, 2023).

**Binary Dragonfly Algorithm**

*Binary Dragonfly Algorithm* adalah penyempurnaan dari *Dragonfly Algorithm* yang merupakan pendekatan optimasi yang terinspirasi dari perilaku dinamis dan statis capung di alam, diusulkan oleh Sendedali Mirjalili pada tahun 2016 (Mirjalili, 2016).

Ada tiga prinsip primitif perilaku berkerumun untuk mensimulasikan perilaku capung, yaitu pemisahan (*S*), keselarasan (*A*), dan kohesi (*C*). Selain itu, tujuan utama *Dragonfly Algorithm* adalah membuat setiap kawanan dapat bertahan hidup, sehingga semua individu tertarik pada sumber makanan (*F*) dan mengalihkan perhatian musuh (*E*). Model matematika dari implementasi perilaku capung tersebut, adalah sebagai berikut.

Pemisahan (*separation*) :

$$S_i = -\sum_{j=1}^N X - X_j \tag{1}$$

Keselarasan (*alignment*) :

$$A_i = \frac{\sum_{j=1}^N V_j}{N} \tag{2}$$

Kohesi (*cohesion*) :

$$C_i = \frac{\sum_{j=1}^N X_j}{N} - X \tag{3}$$

Daya tarik (*attraction*) :

$$F_i = F_{location} - X \tag{4}$$

Gangguan (*distraction*) :

$$E_i = E_{location} + X \tag{5}$$

dimana *X* adalah posisi individu capung saat ini, *X<sub>i</sub>* menunjukkan posisi tetangga ke- *j* dari *X*, *V<sub>j</sub>* menunjukkan vektor kecepatan tetangga ke- *j*, dan *N* mewakili jumlah tetangga, *F<sub>location</sub>* merupakan lokasi sumber makanan, serta *E<sub>location</sub>* menunjukkan posisi musuh saat ini (Li dkk., 2020).

Untuk menemukan solusi optimal, *Dragonfly Algorithm* mendefinisikan vektor langkah yang

mengacu pada arah perjalanan capung dirumuskan sebagai berikut :

$\Delta X_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + w \Delta X_t$  (6) dimana *s, a, c, f*, dan *e* diketahui sebagai faktor pembobotan pemisahan (*S<sub>i</sub>*), keselarasan (*A<sub>i</sub>*), kohesi (*C<sub>i</sub>*), daya tarik (*F<sub>i</sub>*), dan gangguan (*E<sub>i</sub>*), masing-masing dari capung sebagai agen pencarian ke- *i*, sedangkan *w* adalah berat inersia.

Serta vektor posisi capung sebagai agen pencarian (*X*), yang dirumuskan sebagai berikut :

$$X_{t+1} = X_t + \Delta X_{t+1} \tag{7}$$

dimana *t* adalah iterasi saat ini.

Versi dasar *Dragonfly Algorithm* ditujukan untuk masalah dalam ruang pencarian berkelanjutan, sedangkan pemilihan fitur adalah masalah optimasi biner, sehingga strategi pembaruan seperti pada persamaan (6) dan (7) tidak dimungkinkan untuk ruang pencarian biner. *Binary Dragonfly Algorithm* menggunakan fungsi transfer berikut untuk mengubah nilai vektor langkah menjadi angka yang dibatasi dalam [0,1].

$$T(\Delta X) = \left\lfloor \frac{\Delta X}{\sqrt{\Delta X^2 + 1}} \right\rfloor \tag{8}$$

Fungsi transfer diatas digunakan untuk mencari probabilitas pembaruan posisi capung dalam kawanan, kemudian posisi capung (agen pencari) diperbarui sebagai berikut.

$$X_{t+1} = \begin{cases} -X_t r < T(\Delta X_{t+1}) \\ X_t r \geq T(\Delta X_{t+1}) \end{cases} \tag{9}$$

Dimana *r* adalah nomor acak dalam interval [0,1] (Chantar dkk., 2021).

**Support Vector Machine**

*Support Vector Machine* merupakan salah satu metode pembelajaran mesin dengan konsep dasarnya adalah memaksimumkan batas *hyperplane* untuk memisahkan dua kelas dalam suatu ruang *feature*. *Hyperplane* adalah penentuan letak pemisah atau pembatas antar kelompok (Kusuma, 2020).

Dalam pemodelan klasifikasi, SVM dapat menyelesaikan permasalahan klasifikasi linier maupun non-linier (Hasanah, 2022). SVM linier merupakan SVM yang menangani masalah klasifikasi biner dari titik data yang ditargetkan dengan membagi dua kelas (Anugrah & Suparwito, 2022). Sedangkan SVM non-linear diterapkan untuk data yang dipisahkan secara non-linier, ini berarti apabila suatu dataset tidak dapat diklasifikasikan dengan menggunakan garis lurus, maka data

tersebut dianggap sebagai data non-linier. SVM non-linier menggunakan konsep kernel untuk mengubah data dalam dimensi tertentu ke dimensi yang lebih tinggi. Berikut ini kernel yang sering diterapkan pada SVM.

1. Linier, berikut persamaannya :

$$K(x_n, x_i) = x_n, x_i \quad (10)$$

2. *Gaussian Radial Basis Function* (RBF), berikut persamaannya :

$$K(x_n, x_i) = \exp(-\gamma \|x_n, x_i\|^2 + C) \quad (11)$$

3. Sigmoid, berikut persamaannya :

$$K(x_n, x_i) = \tanh(\gamma(x_n, x_i) + r) \quad (12)$$

4. Polinomial, berikut persamaannya :

$$K(x_n, x_i) = (\gamma(x_n, x_i) + r)^d \quad (13)$$

dimana  $x_n$  dan  $x_i$  adalah data pada *data training* masing-masing ke-  $n$  dan ke-  $i$ ,  $\gamma$  adalah Gamma,  $C$  adalah *Cost*,  $r$  adalah *coefficient*, dan  $d$  adalah pangkat (*degree*) (Nanda dkk., 2018).

SVM pada dasarnya adalah pemodelan *binary classifier*, dirancang untuk mengklasifikasikan 2 kelas. Namun, SVM juga dapat digunakan untuk klasifikasi multi-kelas melalui dua metode yang berbeda, yaitu metode *one-against-all* yang mengimplikasikan konsep klasifikasi yang membandingkan antara satu kelas dengan semua kelas lainnya dan *one-against-one* melibatkan pembentukan model SVM biner dengan membandingkan satu kelas terhadap satu kelas lainnya (Dr, 2017).

### Decision Tree

*Decision Tree* merupakan metode yang diterapkan untuk konstruksi sebuah model pengambilan keputusan dalam bentuk pohon berdasarkan atribut data *training* (Id, 2021) Struktur *Decision Tree* berisi simpul akar, simpul internal, simpul daun, dan cabang. Simpul akar (*root node*) adalah awal dari sebuah pohon, simpul internal (*internal node*) adalah simpul setelah simpul akar dan sebelum simpul daun, simpul daun (*leaf node*) merupakan simpul yang tidak mengalami pembelahan lebih lanjut, dan cabang (*branch*) adalah penghubung antar simpul (Xu, 2023).

*Decision Tree* memiliki beberapa algoritma populer dalam penyusunannya salah satunya adalah algoritma c4.5. Pada algoritma c4.5, pembentukan pohon keputusan dimulai dengan menentukan simpul akar (*root node*) yang diambil dari atribut atau fitur dengan nilai *gain ratio* tertinggi. Berikut

persamaan-persamaan yang digunakan dalam algoritma c4.5.

1. *Entropy*, berikut persamaannya :

$$Entropy(S) = \sum_{j=1}^n -p_i * \log_2 p_i \quad (14)$$

di mana  $S$  merupakan himpunan kasus,  $n$  adalah jumlah partisi,  $p_i$  adalah proporsisi  $S_i$  terhadap  $S$ .

2. *Gain*, berikut persamaannya :

$$Gain(S, A) = entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * entropy(S_i) \quad (15)$$

3. *Split info*, berikut persamaannya :

$$Split Info(S, A) = - \sum_{j=1}^n \frac{|S_i|}{|S|} * \log_2 \frac{|S_i|}{|S|} \quad (16)$$

4. *Gain ratio*, berikut persamaannya :

$$Gain Ratio = \frac{Gain(S, A)}{Split Info(S, A)} \quad (17)$$

$A$  adalah atribut,  $n$  adalah jumlah partisi atribut  $A$ , dan  $|S_i|$  adalah jumlah kasus pada partisi ke-  $i$ , dan  $|S|$  merupakan jumlah kasus dalam  $s$  (Saleh, 2020).

### 3. METODE

Pada bagian ini terdapat lima tahapan yaitu pra-pemrosesan data, seleksi fitur, *splitting data*, proses klasifikasi, dan evaluasi kinerja klasifikasi.

#### Pra-pemrosesan Data

Penelitian ini menggunakan data yang terdiri dari beberapa variabel dengan skala yang bervariasi. Oleh karena itu, dilakukan pra-pemrosesan data agar semua variabel dapat berkontribusi secara seimbang dan tanpa ada yang mendominasi, diawali dengan proses diskritisasi dengan mengubah data kontinu pada label menjadi data kategori yang bertujuan membagi data ke dalam kelas-kelas, dimana data penelitian ini dibagi menjadi kelas rendah, sedang, dan tinggi. Diskritisasi data IKPS dilakukan dengan ketentuan menghitung selisih nilai maksimum dan nilai minimum pada label dan dibagi dengan jumlah kelas yang sudah ditentukan sebelumnya untuk menentukan interval setiap kelas. Selanjutnya dilakukan *encoding* untuk mengubah format data kategori, yaitu rendah, sedang, dan tinggi yang masing-masing menjadi 0, 1, dan 2. Kemudian pada atribut atau yang menjadi fitur-fitur dalam klasifikasi, dilakukan normalisasi data dengan *min-max normalization* untuk menyelaraskan skala data dalam rentang 0 hingga 1 sehingga algoritma *machine learning* dapat bekerja lebih efektif.

#### Seleksi Fitur

Seleksi fitur pada penelitian ini menggunakan Algoritma *metaheuristic* yaitu *Binary Dragonfly*

*Algorithm.* Algoritma ini mencari solusi optimal dengan menggunakan representasi biner (0 dan 1). Terdapat 10 fitur yang diseleksi berdasarkan korelasi, yang mana digunakan korelasi pearson untuk menilai kekuatan hubungan antara fitur-fitur tersebut terhadap kelas yang telah ditentukan.

Adapun proses dari seleksi fitur diawali dengan mendefinisikan *fitness function*. Selanjutnya dilakukan inialisasi populasi dan arah atau kecepatan populasi dengan nilai acak biner, serta mendefinisikan *Binary Dragonfly Algorithm* yang melibatkan fitur, label atau kelas, jumlah capung dalam populasi, dan jumlah maksimum iterasi, dimana proses iterasi menghitung *fitness* untuk setiap capung dan memperbarui *fitness* terbaik dan *fitness* terburuk. Kemudian dilakukan pengaturan parameter bobot ( $w$ ), laju ( $rate$ ), dan parameter acak pada  $s, a, c, f, e$  dan melakukan iterasi untuk mengidentifikasi dan menyimpan informasi tetangga dari setiap individu dalam populasi, informasi tersebut digunakan untuk perhitungan dengan menerapkan persamaan (1), persamaan (2), persamaan (3), persamaan (4), dan persamaan (5) yang hasilnya digunakan untuk menghitung vektor langkah menggunakan persamaan (8), dari vektor langkah yang diperoleh digunakan untuk memperbarui vektor posisi capung menggunakan persamaan (9), sehingga solusi terbaik diperoleh.

### Splitting Data

Setelah memperoleh fitur terbaik untuk mengklasifikasikan data IKPS, perlu dilakukan *splitting data* untuk membagi data penelitian menjadi *training data* yang diterapkan dalam proses pembentukan model, dan *testing data* yang diterapkan sebagai validasi apakah model tersebut memiliki tingkat akurasi yang baik. Penelitian ini melakukan *splitting data* dengan rasio 80% *training data* dan 20% *testing data*, berdasarkan *proses trial and error* dengan beberapa rasio *splitting data* yang umum digunakan.

### Proses Klasifikasi

Penelitian ini mengklasifikasikan data IKPS menggunakan dua metode klasifikasi, berupa *Support Vector Machine* dan *Decision Tree* dengan algoritma c4.5. Proses klasifikasi metode SVM dimulai setelah data di *split* dengan membuat model, pada penelitian ini diterapkan menggunakan kernel

polinomial persamaan (13), yang terdiri dari parameter  $d$  untuk menentukan derajat polinomial,  $r$  yang mempengaruhi margin *hyperplane* dan parameter  $\gamma$  sebagai penentuan seberapa jauh pengaruh suatu sampel set data pelatihan. Kemudian menentukan pendekatan multi-kelas yang digunakan, selanjutnya melatih model dengan *training data*. Setelah itu, model dievaluasi dalam memprediksi label atau kelas untuk *testing data* menggunakan model yang telah dilatih.

Adapun proses metode *Decision Tree* dengan algoritma c4.5 dimulai dengan mendefinisikan fungsi untuk menghitung *entropy* (persamaan 14), *gain* (persamaan 15), *split info* (persamaan 16), dan *gain ratio* (persamaan 17) dari masing-masing fitur, dan fungsi untuk membuat diagram pohon. Setelah itu, data di *split* menjadi *training data* dan *testing data*, selanjutnya membangun model berdasarkan hasil *entropy*, *gain*, *split info*, *gain ratio* dan dilatih dengan *training data*. Kemudian, model dievaluasi dalam memprediksi label atau kelas untuk *testing data* menggunakan model yang telah dilatih.

### Evaluasi Kinerja Klasifikasi

Setelah proses klasifikasi, dilakukan evaluasi pada hasil klasifikasi menggunakan *confusion matrix*. *Confusion matrix* merupakan suatu tabel matriks yang memberikan gambaran tentang performa suatu model klasifikasi pada rangkaian data testing yang nilai aktualnya diketahui. Berikut istilah-istilah yang merepresentasikan hasil dari proses klasifikasi dalam *confusion matrix* (Anugrah & Suparwito, 2022).

1. *True Positive* (TP), yaitu data positif yang diprediksi kebenarannya
2. *True Negative* (TN), yaitu data negatif yang diprediksi kebenarannya
3. *False Positive* (FP), yaitu data negatif tetapi diprediksi positif
4. *False Negative* (FN), yaitu data positif tetapi diprediksi negative

Tabel 1. *Confusion Matrix*

Kelas		Nilai Aktual	
		Positif	Negatif
Nilai Prediksi	Positif	TP	FP
	Negatif	FN	TN

Pengujian performa klasifikasi dengan *confusion matrix* adalah sebagai berikut (Oktafiani & Rianto, 2023).

1. *Accuracy*, untuk mengukur sejauh mana kesamaan antara nilai aktual dan nilai yang diprediksi, berikut persamaannya.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \quad (18)$$

2. *Precision*, untuk mengukur tingkat ketepatan informasi yang diprediksi dan respon yang diberikan sistem, berikut persamaannya.

$$Precision = \frac{TP}{TP+FP} \quad (19)$$

3. *Recall*, untuk mengukur sejauh mana sistem berhasil dalam memulihkan informasi, berikut persamaannya.

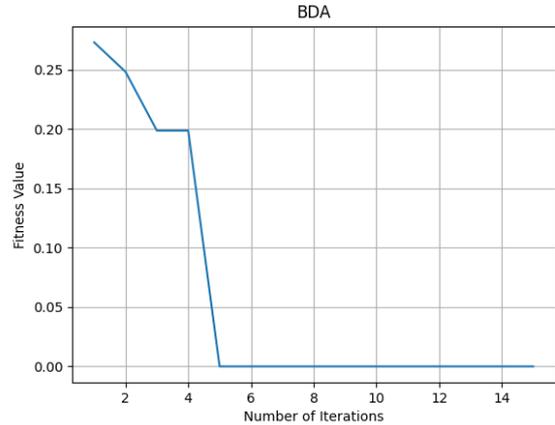
$$Recall = \frac{TP}{TP+FN} \quad (20)$$

#### 4. DATASET

Dataset yang digunakan pada penelitian ini yaitu data Indeks Khusus Penanganan Stunting (IKPS) kabupaten/kota di Indonesia tahun 2022 yang bersumber dari *website* Badan Pusat Statistik, yang terdiri dari 514 baris data, 10 atribut, dan 3 kelas yaitu rendah, sedang, dan tinggi. 10 atribut pada dataset ini meliputi imunisasi( $X_1$ ), penolongan persalinan oleh tenaga kesehatan di fasilitas kesehatan ( $X_2$ ), KB modern ( $X_3$ ), ASI eksklusif ( $X_4$ ), MP ASI ( $X_5$ ), air minum layak ( $X_6$ ), sanitasi layak ( $X_7$ ), PAUD ( $X_8$ ), kepemilikan JKN/jamkesda ( $X_9$ ), dan penerima KPS/KKS atau bantuan pangan ( $X_{10}$ ).

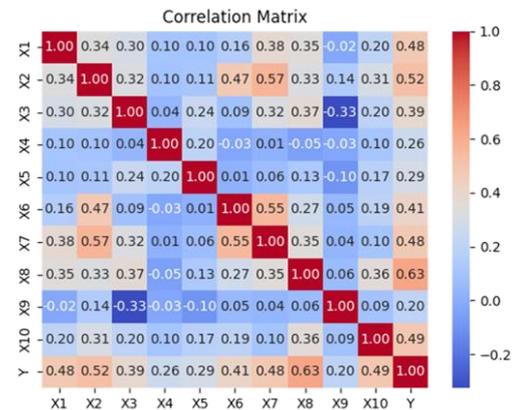
#### 5. HASIL DAN PEMBAHASAN

Pada penelitian ini, dilakukan seleksi fitur menggunakan *Binary Dragonfly Algorithm* dengan mendefinisikan *fitness function* berdasarkan ketentuan nilai korelasi  $\geq 0,40$ , sesuai dengan pedoman yang menunjukkan bahwa nilai korelasi antar variabel cukup kuat (Burhanuddin & Indriyani, 2021). Serta dengan jumlah capung yang digunakan adalah 8 dan maksimum iterasinya adalah 15, diperoleh hasil iterasi dalam kurva konferegensi berikut.



Gambar 1. Iterasi Seleksi Fitur

Dapat dilihat pada gambar, dari iterasi ke-5 hingga ke-15 *fitness value* yang dihasilkan telah mencapai keadaan stabil. Serta memperoleh 6 fitur yang akan digunakan dalam klasifikasi yaitu imunisasi ( $X_1$ ), penolongan persalinan oleh tenaga kesehatan di fasilitas kesehatan ( $X_2$ ), air minum layak ( $X_6$ ), sanitasi layak ( $X_7$ ), PAUD ( $X_8$ ), dan penerima KPS/KKS atau bantuan pangan ( $X_{10}$ ). Untuk membuktikan apakah fitur-fitur tersebut memiliki korelasi yang cukup kuat dengan kelas, dilakukan perhitungan korelasi pearson, sebagai berikut.



Gambar 2. Matriks Korelasi

Ketentuan korelasi yang cukup kuat dengan kelas memiliki nilai korelasi  $\geq 0,40$ . Dilihat pada matriks korelasi di atas, fitur-fitur yang sesuai dengan ketentuan tersebut adalah fitur  $X_1$ ,  $X_2$ ,  $X_6$ ,  $X_7$ ,  $X_8$ , dan  $X_{10}$ , yang masing-masing adalah imunisasi, penolongan persalinan oleh tenaga kesehatan di fasilitas kesehatan, air minum layak, sanitasi layak, PAUD, dan penerima KPS/KKS atau bantuan pangan. Ini berarti fitur-fitur terpilih dari proses seleksi fitur, benar memiliki korelasi yang cukup kuat dengan kelas.

Fitur-fitur yang telah diseleksi, digunakan untuk mengklasifikasikan IKPS kabupaten/kota di Indonesia dengan metode *Support Vector Machine* (SVM) dan *Decision Tree*. Pada SVM menerapkan kernel polinomial dengan parameter  $\gamma = 1, r = 3$ , dan  $d = 2$ , kernel dan parameter tersebut, dipilih setelag melalui serangkaian *trial and error* yang menunjukkan bahwa kombinasi kernel dan parameter ini memberikan akurasi tertinggi dalam membedakan kelas-kelas data, serta menggunakan pendekatan *one-against-one*. Berikut disajikan tabel *confusion matrix* hasil klasifikasi dari metode SVM.

Tabel 2. *Confusion Matrix* metode SVM

Kelas		Nilai Aktual		
		Tinggi	Sedang	Rendah
Nilai Prediksi	Tinggi	44	7	0
	Sedang	6	41	0
	Rendah	0	1	4

Berdasarkan Tabel di atas keseluruhan data dengan kelas tinggi adalah 50 data, tetapi hanya 44 data yang diklasifikasikan dengan benar sebagai kelas tinggi, dan 6 data sisanya diklasifikasikan pada kelas sedang. Pada kelas sedang dengan total data 49, terdapat 41 data diklasifikasi dengan benar sebagai kelas sedang, 7 data sebagai kelas tinggi, dan 1 data sebagai kelas rendah. Sedangkan terdapat 4 data pada kelas rendah, yang semua datanya diklasifikasi dengan benar sebagai kelas rendah. Sehingga diketahui bahwa terdapat 14 data yang tidak dapat diklasifikasi dengan benar.

Adapun metode *Decision Tree* dengan algoritma c4.5 yang memisahkan data pada setiap simpul (*node*) berdasarkan kriteria *gain ratio*. Dari proses klasifikasi dengan metode ini menggunakan bantuan *software phyton* diperoleh hasil *entropy*, *gain*, *split info*, dan *gain ratio* dari masing-masing fitur yang disajikan dalam tabel berikut.

Tabel 3. Hasil Perhitungan Algoritma C4.5

Fitur	<i>Entropy</i>	<i>Gain</i>	<i>Split Info</i>	<i>Gain Ratio</i>
Imunisasi ( $X_1$ )	0,28	0,86	7,86	0,11
Penolongan persalinan oleh tenaga Kesehatan di fasilitas Kesehatan ( $X_2$ )	0,40	0,74	7,35	0,10
Air minum layak ( $X_6$ )	0,46	0,68	7,57	0,09
Sanitasi layak ( $X_7$ )	0,36	0,78	7,91	0,10
PAUD ( $X_8$ )	0,31	0,84	7,98	0,10
Penerima KPS/KKS atau bantuan pangan ( $X_{10}$ )	0,24	0,90	8,02	0,11

Sehingga diperoleh diagram pohon dengan simpul akar (*root node*) adalah fitur imunisasi ( $X_1$ ), sedangkan fitur penolongan persalinan di fasilitas kesehatan ( $X_2$ ) dan sanitasi layak ( $X_7$ ) sebagai simpul internal (*internal node*), serta label atau kelas ( $Y$ ) sebagai simpul daun (*leaf node*). Berikut disajikan tabel *confusion matrix* hasil klasifikasi dari metode *Decision Tree*.

Tabel 4. *Confusion Matrix* metode *Decision Tree*

Kelas		Nilai Aktual		
		Tinggi	Sedang	Rendah
Nilai Prediksi	Tinggi	41	38	2
	Sedang	9	10	1
	Rendah	0	1	1

Berdasarkan table di atas, dari total 50 data pada kelas tinggi, hanya 41 data yang diklasifikasikan dengan benar sebagai kelas tinggi, dan 9 data sisanya diklasifikasikan pada kelas sedang. Sedangkan pada kelas sedang dengan total data 49, hanya 10 data diklasifikasi dengan benar sebagai kelas sedang, 38 data sebagai kelas tinggi, dan 1 data sebagai kelas rendah. Serta terdapat 4 data pada kelas rendah, tetapi hanya 1 data yang diklasifikasi dengan benar sebagai kelas rendah, 2 data sebagai kelas tinggi, dan 1 data sebagai kelas sedang. Sehingga diketahui bahwa ada 51 data yang tidak dapat diklasifikasi dengan benar.

Dari hasil klasifikasi dengan kedua metode tersebut, diperoleh perbandingan kinerja klasifikasi yang disajikan pada tabel berikut.

Tabel 5. Perbandingan Kinerja Klasifikasi

Metode	Accuracy	Precision	Recall
SVM	86%	84%	91%
Decision Tree	50%	50%	42%

Dapat dilihat pada tabel di atas, metode SVM memiliki akurasi sebesar 86%, ini menunjukkan metode SVM berhasil mengklasifikasikan kategori kelas rendah, sedang, dan tinggi pada data IKPS dengan tingkat keakuratan mencapai 86%, serta memperoleh *precision* sebesar 84%, dan *recall* sebesar 91%. Sedangkan pada metode *Decision Tree* dengan menggunakan algoritma c4.5 menghasilkan akurasi sebesar 50%, yang artinya metode ini mampu mengklasifikasikan kategori kelas rendah, sedang, dan tinggi pada data IKPS dengan tingkat keakuratan mencapai 50%, serta memperoleh *precision* sebesar 50%, dan *recall* sebesar 42%. Sehingga, dari kedua metode tersebut diketahui bahwa hasil yang dicapai dengan metode SVM memiliki kemampuan yang lebih baik dibandingkan *Decision Tree*.

## 6. PENUTUP

### SIMPULAN

Seleksi fitur dengan menerapkan *Binary Dragonfly Algorithm* berdasarkan korelasi antar fitur dengan kelas pada Indeks Khusus Penanganan Stunting (IKPS) kabupaten/kota di Indonesia, diperoleh 6 fitur dari total 10 fitur, yaitu fitur imunisasi, penolong persalinan oleh tenaga kesehatan di fasilitas kesehatan, air minum layak, sanitasi layak, PAUD, dan penerima KPS/KKS atau bantuan pangan yang akan digunakan dalam proses klasifikasi.

Hasil akurasi klasifikasi Indeks Khusus Penanganan Stunting (IKPS) kabupaten/kota di Indonesia menggunakan metode *Support Vector Machine* sebesar 86%, dan menggunakan metode *Decision Tree* sebesar 50%, menunjukkan bahwa penggunaan metode *Support Vector Machine* memperoleh kinerja yang lebih baik dibandingkan dengan metode *Decision Tree*, dengan tingkat akurasi mencapai 86%.

## SARAN

Dari hasil analisis dan pembahasan yang telah dilakukan pada penelitian ini, saran yang dapat diberikan untuk penelitian selanjutnya adalah menerapkan metode seleksi fitur selain *Binary Dragonfly Algorithm*, baik dengan metode algoritma *metaheuristic* atau yang lainnya, dan menggunakan metode-metode klasifikasi lain, yang menghasilkan akurasi lebih baik lagi.

## DAFTAR PUSTAKA

- Alhakeem, Z. M., Hakim, H., Hasan, O. A., Laghari, A. A., Jumani, A. K., & Jasm, M. N. (2023). Prediction of diabetic patients in Iraq using binary dragonfly algorithm with long-short term memory neural network. *AIMS Electronics and Electrical Engineering*, 7(3), 217–230. <https://doi.org/10.3934/electreng.2023013>
- Amirudin, M., & Wowor, A. D. (2023). Analisis Perbandingan Klasifikasi Balita Beresiko Stunting Menggunakan Metode Support Vector Machine dan Decision Tree. *CENTIVE*, 3(1), 581–591.
- Aneke, C. S., Udoiwod, E. N., & Mfonobong, U. (2024). Prediction Of Breast Cancer Using Support Vector Machine And Decision Tree Machine Learning Approaches. *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, 11(2), 16915–16924. [www.jmest.org](http://www.jmest.org)
- Anugrah, D. P., & Suparwito, H. (2022). Analisis Sentimen Bantuan Langsung Tunai Covid-19 menggunakan Algoritma Support Vector Machine. *Seminar Nasional Sanata Dharma Berbagi*, 72, 72–84.
- Badan Pusat Statistik. (2023). *Laporan Indeks Khusus Penanganan Stunting Kabupaten/Kota 2021-2022* (Vol. 2).
- Burhanuddin, H., & Indriyani. (2021). Pengaruh Penggunaan Lahan terhadap Kualitas Air Permukaan di Sungai Kampar. *Bandung Conference Series: Urban & Regional Planning*, 1(1), 21–29.
- Chantar, H., Tubishat, M., Essgaer, M., & Mirjalili, S. (2021). Hybrid Binary Dragonfly Algorithm with Simulated Annealing for Feature Selection. *SN Computer Science*, 2(4), 295–301. <https://doi.org/10.1007/s42979-021-00687-5>
- Devisetty, G., & Kumar, N. S. (2023). Prediction of Bradycardia using Decision Tree Algorithm and Comparing the Accuracy with Support Vector Machine. *E3S Web of Conferences*, 399. <https://doi.org/10.1051/e3sconf/202339909004>

- Dr, S. (2017). Data mining untuk klasifikasi dan klusterisasi data. *Bandung : Informatika Bandung*.
- Fitriani, I., Basuki, S., & Minarno, A. E. (2020). Seleksi Fitur Relieff Pada Klasifikasi Malware Android Menggunakan Support Vector Machine(SVM). *REPOSITOR*, 2(11), 1529–1534.
- Guendouz, M., & Amine, A. (2023). A New Feature Selection Method Based on Dragonfly Algorithm for Android Malware Detection Using Machine Learning Techniques. *International Journal of Information Security and Privacy*, 17(1). <https://doi.org/10.4018/IJISP.319018>
- Hasanah, S. H. (2022). Classification Support Vector Machine in Breast Cancer Patiens. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 16(1), 129–136. <https://doi.org/10.30598/barekengvol16iss1p129-136>
- Id, I. D. (2021). *MACHINE LEARNING : Teori, Studi Kasus, dan Implementasi Menggunakan Phytion* (I. D. Id, Ed.; 1 ed.). UR PRESS. <https://books.google.co.id/books?id=JvBPEAAQBAJ>
- Jaya Purnama, J., Mahmud Nawawi, H., Rosyida, S., Ridwansyah, & Risnandar. (2020). Klasifikasi Mahasiswa HER Berbasis Algoritma SVM dan Decision Tree. *Jurnal Teknologi Informasi dan Ilmu Komputer(JTIK)*, 7(6), 1253–1260. <https://doi.org/10.25126/jtiik.202073080>
- Kementerian Kesehatan RI. (2022). *BUKU SAKU Hasil Survei Status Gizi Indonesia (SSGI) 2022*. Kementrian Kesehatan RI.
- Kusuma, D. P. (2020). *MACHINE LEARNING : Teori, Program, dan Studi Kasus* (D. P. Kusuma, Ed.; Mei 2020). CV BUDI UTAMA.
- Li, J., Kang, H., Sun, G., Feng, T., Li, W., Zhang, W., & Ji, B. (2020). IBDA: Improved Binary Dragonfly Algorithm with Evolutionary Population Dynamics and Adaptive Crossover for Feature Selection. *IEEE Access*, 8, 108032–108051. <https://doi.org/10.1109/ACCESS.2020.3001204>
- Mirjalili, S. (2016). Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural computing and applications*, 27, 1053–1073.
- Nanda, M. A., Seminar, K. B., Nandika, D., & Maddu, A. (2018). A Comparison Study of Kernel Functions in the Support Vector Machine and Its Application for Termite Detection. *Information*, 9(1), 5–19. <https://doi.org/10.3390/info9010005>
- Oktafiani, R., & Rianto, R. (2023). Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree untuk Sistem Rekomendasi Tempat Wisata. *Jurnal Nasional Teknologi dan Sistem Informasi*, 9(2), 113–121. <https://doi.org/10.25077/teknosi.v9i2.2023.113-121>
- Raouf, Z. T., & Abd, D. H. (2023). Feature Selection for Binary Dataset using Dragonfly Algorithm. *2023 16th International Conference on Developments in eSystems Engineering (DeSE)*, 480–485. <https://doi.org/10.1109/DeSE60595.2023.1046922>
- Saleh, H. (2020). Analisa Faktor Penyebab Stunting Menggunakan Algoritma C4.5. *ScientiCO: Computer Science and Informatics Journal*, 3(1), 11–17.
- Sarumi. (2021). Fraud Detection in Customers' Electricity Consumption in Nigeria Using Support Vector Machine and C4.5 Decision Tree Algorithms. *Computing, Information System, Development Informatics & Allied Research Journal*, 12(1), 43–64. <https://doi.org/10.22624/AIMS/CISDI/V12N1P6>
- Sihombing, P. R., & Yuliati, I. F. (2021). Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia. *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 20(2), 417–426. <https://doi.org/10.30812/matrik.v20i2.1174>
- Somantri, O., Nugroho, W. E., & Supriyono, A. R. (2022). Penerapan Feature Selection Pada Algoritma Decision Tree Untuk Menentukan Pola Rekomendasi Dini Konseling. *Jurnal Sistem Komputer dan Informatika (JSON)*, 4(2), 272–279. <https://doi.org/10.30865/json.v4i2.5345>
- Syafika, V. A. N., & Karisma, R. D. L. N. (2023). Implementasi Support Vector Machine (SVM) dalam Penentuan Klasifikasi Indeks Khusus Penanganan Stunting di Indonesia (On Implementation of Support Vector Machine (SVM) in Determining the Classification of Stunting-Specific Intervention Index in Indonesia). *Seminar Nasional Official Statistics*, 2(1), 267–276.
- Syahputra, H., Naibaho, S. I., Maulana, M. A., Zulfahmi, I., & Sinaga, E. P. (2023). Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree Untuk Deteksi Tingkat Depresi Mahasiswa. *BINA INSANI ICT JOURNAL*, 10(1), 52–61.
- WHO. (2021). *Levels and trends in child malnutrition*.

Xu, J. (2023). Systematic Analysis and Application Prospect of Decision Tree. *Highlights in Science, Engineering and Technology*, 71, 163–170.