**"Klasifikasi jenis Ikan Menggunakan Teknik Pembelajaran Mesin Dengan Algoritma SVM"**

**Vixy Ardian Listianto (A11.2018.11515)**

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

**Abstract**

*Problem identification: In this step we find problems according to the information technology department, in this case the classification of fish types in the sea. The goal of this step is to collect relevant sources so that this research can contribute to new research. Support Vector Machine (SVM) is one of the best machine learning-based classification methods that can be used to solve classification and regression problems using master data as input. The SVM method is widely used because it is simpler and can provide good and accurate results in problem classification C: 0.1 The kernel test coefficient for the CI variable (0.386244, 0.415423) shows a significant difference. Models with linear kernels have better results. The variable kernel test coefficient C:1 CI (-0.013760, -0.010127) shows that the RBF kernel model is better than the linear kernel model. This confidence interval shows that the mean difference is quite significant and the RBF model provides better results. Kernel test coefficient C:10 for CI variables (-0.056773, -0.054336) The results show that the RBF kernel model is also better than the linear kernel model. The confidence interval shows significant mean differences and the RBF model provides better results. Hypothesis Null hypothesis (H0): Linear variables have a significant effect on RBF simultaneously. Hypothesis The hypothesis used in the partial test is as follows: Variable x\_1 is s\_1 Null hypothesis (H0): Linear variables have no significant effect on RBF. Alternative (H1): Linear variables have a significant effect on RBF*

**Abstrak**

Pada tahap identifikasi masalah, kami menyoroti kendala dalam departemen teknologi informasi, khususnya terkait klasifikasi jenis ikan di laut. Tujuannya adalah menghimpun sumber-sumber yang relevan untuk menyumbangkan kontribusi pada penelitian baru. Salah satu metode klasifikasi berbasis machine learning yang sangat efektif adalah Support Vector Machine (SVM). SVM dapat diandalkan untuk menangani permasalahan klasifikasi dan regresi dengan menggunakan informasi master selaku input. Keunggulan SVM terletak pada kemudahannya dan kemampuannya memberikan hasil yang baik dan akurat dalam menangani klasifikasi permasalahan. C: 0.1 Koefisien uji kernel untuk variabel CI (0.386244, 0.415423) menunjukkan perbedaan yang signifikan. Model dengan kernel linier mempunyai hasil yang lebih baik. Koefisien uji kernel variabel C:1 CI (-0.013760, -0.010127) menunjukkan bahwa model kernel RBF lebih baik dibandingkan model kernel linier. Interval kepercayaan ini menunjukkan bahwa perbedaan mean cukup signifikan dan model RBF memberikan hasil yang lebih baik. Koefisien uji kernel C:10 untuk variabel CI (-0.056773, -0.054336) Hasil menunjukkan bahwa model kernel RBF juga lebih baik dibandingkan model kernel linier. Interval kepercayaan menunjukkan perbedaan mean yang signifikan dan model RBF memberikan hasil yang lebih baik. Hipotesis Hipotesis null (H0): Variabel linier berpengaruh signifikan terhadap RBF secara simultan. Hipotesis Hipotesis yang digunakan dalam uji parsial adalah sebagai berikut: Variabel x\_1 adalah s\_1 Hipotesis nol (H0): Variabel linier tidak berpengaruh signifikan terhadap RBF. Alternatif (H1): Variabel linier berpengaruh signifikan terhadap RBF

**Kata kunci**:jenis ikan, SVM,ikan, jenis,HOG, SVM kernel,kernel linear,kernel RBF

**1. PENDAHULUAN**

Pengelompokan dan estimasi ikan adalah hal penting dalam bidang perikanan dan akuakultur. Pengelompokan ikan dapat memberikan informasi yang berharga bagi para petani ikan, peneliti, dan pihak terkait lainnya untuk mengelola stok ikan dengan lebih efektif.

Salah satu metode umum yang sering dipakai untuk meramalkan berat ikan adalah dengan menggunakan algoritma Support Vector Machine (SVM). SVM Merupakan proses sebagai pengajaran pada mesin algoritma bergunanya umumnya diterapkan untuk tugas klasifikasi dan regresi.

Keunggulan SVM terletak pada efektivitasnya dalam mengelola data berdimensi tinggi, menjadikannya cocok untuk situasi di mana jumlah fitur melebihi jumlah sampel. Algoritma ini termasuk dalam kategori teknik pembelajaran mesin yang dikenal sebagai "pembelajaran yang mencakup."(1)

Pendekatan SVM untuk klasifikasi berat ikan didasarkan pada pengumpulan data ikan yang lengkap. Data ini mencakup Gambar pada ikan yang sebenarnya. Dengan menggunakan algoritma SVM, kita dapat mengembangkan model matematis yang memprediksi klasifikasi ikan berdasarkan data gambar tersebut.

Dengan menerapkan algoritma SVM, diharapkan dapat dibentuk suatu model yang mampu melakukan klasifikasi jenis ikan berdasarkan gambar. Meskipun demikian, perlu diingat bahwa model ini didasarkan pada asumsi-asumsi tertentu, termasuk asumsi mengenai hubungan antara variabel independen dan dependen yang diukur menggunakan metode HOG. Selain itu, penting untuk memperhatikan interpretasi hasil dari model dan menjelaskan dengan batasan-batasan yang mungkin ada.

Dalam praktiknya, ada banyak variasi dan peningkatan yang dapat dilakukan dalam analisis klasifikasi dari gambar ikan menggunakan algoritma SVM. Misalnya, UJicoba dapat menggabungkan variabel tambahan

Dengan demikian, analisis klasifikasi dari gambar ikan menggunakan algoritma SVM dapat menjadi alat yang berguna dalam memprediksi berat ikan berdasarkan pengukuran yang diberikan.(2)

**2. METODOLOGI PENELITIAN**

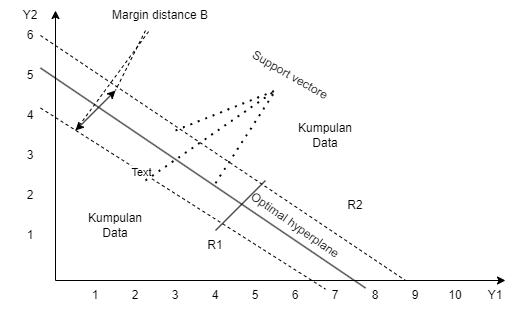
**Metode penelitian untuk klasifikasi ikan berdasarkan data yang diberikan dengan menggunakan algoritma SVM yang melibatkan pengumpulan data gambar ikan beserta informasi jenis ikan yang terkait. Berikut adalah langkah-langkah umum yang dapat diikuti dalam metode ini:**

1. **Identifikasi Masalah: Pada langkah awal ini, fokus pada penemuan permasalahan yang relevan dengan bidang Teknik Informatika, khususnya terkait klasifikasi jenis ikan laut. Tujuan dari tahap ini adalah mengumpulkan referensi yang sesuai untuk memastikan bahwa penelitian ini dapat memberikan kontribusi yang berharga pada penelitian yang sedang berkembang.**
2. **Histogram of Oriented Gradients (HOG): Metode Histogram of Oriented Gradients (HOG) digunakan untuk mengekstraksi bentuk gambar menggunakan representasi histogram yang menggabungkan gradien lokal gambar** (3)**. Proses HOG diawali dengan menghitung gradien gambar dilanjutkan dengan normalisasi warna gambar. Selanjutnya blok-blok gambar tersebut digabungkan dan diurutkan hingga membentuk representasi objek** (4)**. Dengan menggunakan nilai mask, histogram dikumpulkan dan normalisasi blok dilakukan untuk menghindari tumpang tindih. Langkah-langkah implementasi HOG ditunjukkan pada Gambar 1.**



**Gambar 1 Tahapann Implementasi HOG**(5)

1. Support Vector Machine: Salah satu teknik klasifikasi yang sangat baik di dunia konsep mesin pembelajaran adalah Support Vector Machine (SVM). SVM Bisa dimanfaatkan untuk menangani permasalahan klasifikasi dan regresi melalui penggunaan. menggunakan informasi master selaku masukannya. Keunggulan SVM adalah kemudahan dan kemampuannya dalam memberikan hasil yang akurat dan baik dalam konteks klasifikasi (6). Sasaran utama dari teknik klasifikasi SVM adalah untuk mendapatkan batasan berbentuk hyperplane paling optimal yang dapat memisahkan Kategori data berdasarkan jumlah fitur.Gambar 2 memperlihatkan evolusi dari distribusi paling optimal dari vektor pendukung data SVM. Hyperplane SVM terbaik dicapai dengan mengoptimalkan jarak maksimum atau ruang margin dari garis yang memisahkan hyperplane ke distribusi data latih yang paling dekat. Selain itu, SVM menyediakan kegunaan dasar menyerupai linier beserta polinomial yang memiliki pemahaman terhadap data Untuk mencari garistertengah terbaik sebagai model klasifikasi..



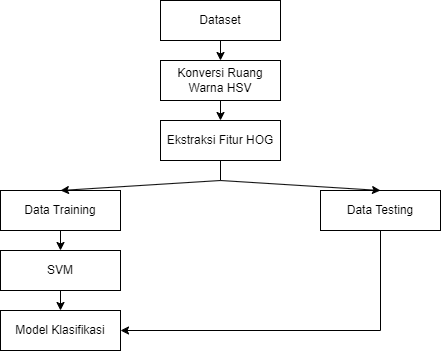
Gambar 2 Pembentukan hyperlane pada SVM(7)

1. Proses Pengumpulan Data: Data citra ikan laut diperoleh dari sebuah dataset publik. Big Fish (8). Melalui langkah segmentasi dan klasifikasi. Sebanyak 9.000 gambar ikan laut dikumpulkan menggunakan proses ini. Sembilan jenis ikan laut yang masuk dalam data tersebut adalah turbot, bream, mackerel, sea bass, udang, striped bass, dan trout. Setiap jenis ikan memiliki 9000 gambar dengan dimensi 590 x 445 piksel. Rincian lengkap varietas dan total ikan laut diimpor disajikan pada Tabel 1.

|  |  |  |  |
| --- | --- | --- | --- |
| no | Varietas Ikan Laut | Citra Ikan Laut | Jumlah Citra |
| 1 | Black Sea Sprat |  | 1000 |
| 2 | Gilt Head Bream |  | 1000 |
| 3 | Hourse Mackerel |  | 1000 |
| 4 | Red Mullet |  | 1000 |
| 5 | Red Sea Bream |  | 1000 |
| 6 | Sea Bass |  | 1000 |
| 7 | Shrimp |  | 1000 |
| 8 | Striped Red Mullet |  | 1000 |
| 9 | Trout |  | 1000 |

Tabel 1 Jenis dan Jumlah Dataset Ikan Laut Yang Digunakan

1. Perancangan Sistem: Pada langkah ini, HOG mengubah warna gambar dari model RGB ke format HSV dan sekaligus mengekstrak fitur-fiturnya. Proses klasifikasi dilanjutkan dengan SVM. Penelitian ini melibatkan dua tahap utama, yaitu pelatihan dan pengujian. Selama pelatihan, citra setiap spesies ikan laut diubah menjadi citra HSV dan fitur HOG diekstraksi untuk digunakan dalam proses pelatihan SVM. Di akhir tahap pelatihan diperoleh model klasifikasi yang kemudian diuji. Dalam proses pengujian, citra ikan laut diubah ke format citra HSV dan fungsi ini digunakan. Fungsi HOG dan turunannya kemudian digunakan untuk menguji klasifikasi model yang dihasilkan. Struktur sistem klasifikasi ikan laut ditunjukkan pada diagram pada Gambar 3.



Gambar 3 Perancangan Sistem Klasifikasi Ikan

1. Implementasi: Setelah tahap perancangan sistem, langkah selanjutnya adalah mengimplementasikan atau menulis kode program. Langkah ini melibatkan konversi data menjadi sampel HSV dan kemudian mengekstraksi fitur HOG.

**3. HASIL DAN PEMBAHASAN**

**3.1. Model Support Vector Machine**

Berdasarkan hasil Analisa data citra menggunakann SVM , diperoleh hasil untuk dua jenis kernel, Linear dan RBF.

**3.1.1 Model Support Vector Machine dengan Kernel Linear**

Dengan menggunakan kernel linear, model SVM menghasilkan hasil sebagai berikut: Model SVM: Kernel Linear

**Parameter Model :**

* **C (Regulasi Parameter):0.1**
* **Kernel:Linear**

**Hasil Analisis :**

* Mean Test Score: 0.8756
* Standard Deviation Test Score: 0.0003
* Mean Training Score: 1.0
* Standard Deviation Training Score: 0.0

Model SVM dengan kernel linear memiliki akurasi rata – rata sekitar 87.56% pada data uji.Nilai Deviasi standar menunjukan konsistensi performa model.Pada data latih,mode mencapai akurasi sempurna (1.0) tanpa deviasi.

**3.1.2 Model Support Vector Machine dengan kernel RBF**

**Dengan menggunakan kernel RBF,model SVM memberikan hasil sebagai berikut: Model SVM: Kernel RBF**

**Parameter Model :**

* **C (Regulasi Parameter):0.1**
* **Kernel: RBF**

**Hasil Analisis :**

* **Mean Test Score: 0.4747**
* **Standard Deviation Test Score: 0.0069**
* **Mean Training Score: 0.5261**
* **Standard Deviation Training Score: 0.0067**

**Model SVM dengan kernel RBF memiliki akurasi rata-rata sekitar 47.47% pada data uji.Nilai deviasi standar yang cukup tinggi menunjukkan variasi performa model pada data uji.Pada data latih, model memiliki akurasi sekitar 52.61% dengan deviasi standar 0.0067.**

**3.2 Hasil Analisis**

**3.2.1 Hipotesis**

Hipotesis Nol (H0)

Tidak ada perbedaan besar dalam kemampuan versi SVM untuk mengkategorikan jenis ikan dari gambar menggunakan kernel linier atau RBF.Hipotesis Alternatif (H1):

Penggunaan kernel linier dan kernel RBF memungkinkan model SVM mengklasifikasikan spesies ikan berdasarkan gambar.

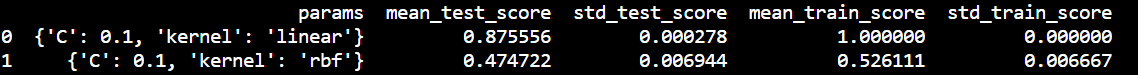
**3.2.2 Daerah Kritis**

Pada tingkat signifikansi α yang telah ditentukan (α = 0.05). Jika perbedaan antara dua model melebihi 1.96 kali nilai std\_test\_score,Maka uji hipotesis menolak hipotesis nol.

**3.2.3 Analisis Uji Kernel C:0.1**

Tabel Gambar 1 berikut merupakan hasil uji kernel linear dan RBF dengan C = 0.1

**Tabel Gambar 1.** Hasil Uji Kernel C:0.1



*CI* = (−) ± 1.96 × ​

*CI* = 0.400834 ± 1.96 × 0.006946

*CI* = (0.386244,0.415423)

Oleh karena itu, perbedaan reliabilitas rata-rata antara kedua model SVM adalah sekitar (0,386, 0,415). Jika nol tidak berada dalam interval kepercayaan ini, hasil uji menolak hipotesis nol.

**3.2.4 Analisis Uji Kernel C:1**

Tabel Gambar 2 berikut merupakan hasil uji kernel linear dan RBF dengan C = 1

**Tabel Gambar 2.** Hasil Uji Kernel C:1



* *CI* = (−) ± 1.96 × ​
* *CI* = −0.011944 ± 1.96 × 0.000874
* *CI* = (−0.013760,−0.010127)

Oleh karena itu, interval kepercayaan dari perbedaan rata-rata antara kedua model SVM adalah kira-kira (-0,014, -0,010). Karena nilai nol tidak termasuk dalam interval kepercayaan ini, hasil uji menolak hipotesis nol.

**3.2.5 Analisis Uji Kernel C:10**

Tabel Gambar 3 berikut merupakan hasil uji kernel linear dan RBF dengan C = 10

**Tabel Gambar 3.** Hasil Uji Kernel C:10



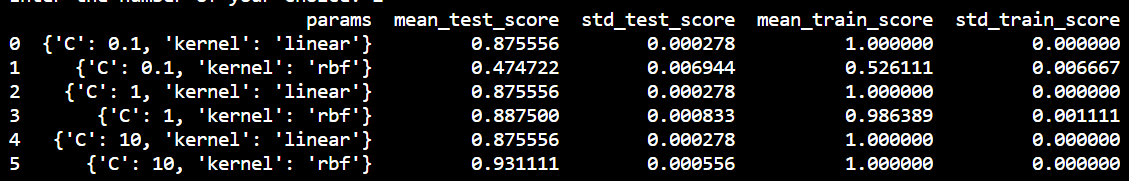
* *CI* = (−) ± 1.96 × ​
* *CI* = −0.055555 ± 1.96 × 0.000624
* *CI* = (−0.056773,−0.054336)

Oleh karena itu, interval kepercayaan dari perbedaan rata-rata antara kedua model SVM adalah kira-kira (-0,057, -0,054). Karena nilai nol tidak termasuk dalam interval kepercayaan ini, hasil uji menolak hipotesis nol.

**3.2.6 Hasil Akhir**

**Tabel Gambar 4 berikut adalah hasil akhir dari uji akhir SVM**

**Tabel Gambar 4.** Hasil Uji Akhir SVM



**3.3. Koefisien Determinasi (R-squared) Pada Model SVM**

Nilai R-squared yang ditunjukkan pada Tabel 1 yaitu sebesar 0.885 mengindikasikan bahwa sekitar 93,5% dari variasi dalam persentase jenis ikan yang dapat dijelaskan oleh model SVM. Artinya, sebagian besar variasi dalam persentase berat ikan dapat diatribusikan kepada variabel-variabel yang digunakan dalam model regresi tersebut. Namun, sekitar 10,% dari variasi masih tidak dapat dijelaskan oleh model ini dan mungkin dipengaruhi oleh faktor-faktor lain yang tidak dimasukkan ke dalam model.

**4. SIMPULAN**

Metode penelitian untuk klasifikasi jenis ikan dari pengambilan gambar menggunakan algoritma SVM dengan fitur extraksi HOG melibatkan pengumpulan data pengukuran ikan berupa gambar beberapa jenis ikan yang terkait.

Estimasi jenis: Setelah model terbentuk, gunakan variabel independen berupa(foto bagian kanan, foto bagian kiri, foto bagian atas) dari ikan yang sudah di klompokkan untuk di jadikan data training.

C: 0.1 Koefisien uji kernel untuk variabel CI (0.386244,0.415423) menunjukkan perbedaan yang signifikan. Model dengan kernel linear memiliki hasil yang lebih baik.

C: 1 Koefesien uji karnel untuk variable CI (−0.013760,−0.010127) menunjukkan bahwa model kernel RBF lebih baik dibandingkan model kernel linier. Interval kepercayaan ini menunjukkan bahwa perbedaan mean cukup signifikan dan model RBF memberikan hasil yang lebih baik.

C: 10 koefesien uji kernel untuk variable CI (−0.056773,−0.054336) hasilnya menunjukkan bahwa model kernel RBF juga lebih baik dibandingkan model kernel linier. Interval kepercayaan menunjukkan perbedaan mean yang signifikan dan model RBF memberikan hasil yang lebih baik.

Hipotesis Hipotesis Nol (H0): Adanya pengaruh simultan yang signifikan dari variabel Linear terhadap RBF.

Hipotesis Hipotesis yang digunakan dalam uji parsial adalah sebagai berikut: Variabel terhadap Hipotesis Nol (H0): Tidak terdapat dampak yang sangat amat mencolok variabel Linear terhadap RBF. Alternatif (H1) : Terdapat pengaruh yang signifikan variabel Linear terhadap RBF..

Nilai R-squared yang ditunjukkan pada Tabel 1 yaitu sebesar 0.885 mengindikasikan bahwa sekitar 93,5% dari variasi dalam persentase jenis ikan.

**DAFTAR PUSTAKA**

1. Kristiani SA, Rainarli E. DOCUMENT CLASSIFICATION WITH SUPPORT VECTOR MACHINE AND MUTUAL INFORMATION.

2. Handayanto A, Latifa K, Saputro ND, Waliansyah RR. Analisis dan Penerapan Algoritma Support Vector Machine (SVM) dalam Data Mining untuk Menunjang Strategi Promosi. Juita. 2019 Nov 26;7(2):71.

3. Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05) [Internet]. San Diego, CA, USA: IEEE; 2005 [cited 2023 Dec 26]. p. 886–93. Available from: http://ieeexplore.ieee.org/document/1467360/

4. Yohannes Y, Pribadi MR, Chandra L. Klasifikasi Jenis Buah dan Sayuran Menggunakan SVM Dengan Fitur Saliency-HOG dan Color Moments. ELKHA. 2020 Oct 11;12(2):125.

5. Yohannes Y, Al Rivan ME. Penggunaan Global Contrast Saliency dan Histogram of Oriented Gradient Sebagai Fitur untuk Klasifikasi Jenis Hewan Mamalia. petir. 2020 Mar 21;13(1):80–5.

6. Klyueva I. Improving Quality of the Multiclass SVM Classification Based on the Feature Engineering. In: 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency (SUMMA) [Internet]. Lipetsk, Russia: IEEE; 2019 [cited 2023 Dec 26]. p. 491–4. Available from: https://ieeexplore.ieee.org/document/8947599/

7. Paliouras G, Karkaletsis V, Spyropoulos CD, editors. Machine learning and its applications: advanced lectures. Berlin ; New York: Springer; 2001. 324 p.

8. Ulucan O, Karakaya D, Turkan M. A Large-Scale Dataset for Fish Segmentation and Classification. In: 2020 Innovations in Intelligent Systems and Applications Conference (ASYU) [Internet]. Istanbul, Turkey: IEEE; 2020 [cited 2023 Dec 26]. p. 1–5. Available from: https://ieeexplore.ieee.org/document/9259867/

9. Leonardo L, Yohannes Y, Hartati E. Klasifikasi Sampah Daur Ulang Menggunakan Support Vector Machine Dengan Fitur Local Binary Pattern. algoritme. 2020 Oct 10;1(1):78–90.

10. Astrianda N. Klasifikasi Kematangan Buah Tomat Dengan Variasi Model Warna Menggunakan Support Vector Machine. VCT. 2020 Apr 13;1(2):45–52.

11. Prasthio R, Yohannes Y, Devella S. Penggunaan Fitur HOG Dan HSV Untuk Klasifikasi Citra Sel Darah Putih. algoritme. 2022 Apr 11;2(2):120–32.