

SISTEM KLASIFIKASI UKURAN BAJU DENGAN METODE SUPPORT VECTOR MACHINE (SVM)

Muhammad Avi Majid Kaaffah^{*)}, Achmad Hidayatno dan Yosua Alvin Adi Soetrisno

Departemen Teknik Elektro, Universitas Diponegoro
Jl. Prof. Sudharto, SH, Kampus UNDIP Tembalang, Semarang 50275, Indonesia

^{*)E-mail : avikaaffah@students.undip.ac.id}

Abstrak

Ukuran baju merupakan rujukan utama pengguna saat memilih baju yang cocok bagi tubuhnya. Pengguna memutuskan membeli baju dengan mencoba mencari ukuran yang cocok di kamar pas sehingga membutuhkan waktu yang banyak. Ukuran baju di setiap produsen baju belum tentu sama, sehingga Badan Standarisasi Nasional (BSN) mengeluarkan Standar Nasional Indonesia (SNI) untuk ukuran baju dengan nomor SNI 2161:2010 berjudul Ukuran Pakaian – Kaos Pria. Perancangan sistem menggunakan metode Support Vector Machine (SVM) untuk melakukan klasifikasi ukuran baju. SVM merupakan metode klasifikasi dengan softmargin yang digunakan untuk memisahkan beberapa fitur dengan jarak support garis ke fitur yang paling optimal. Masukan fitur dari sistem adalah tiga parameter ukuran tubuh yaitu lingkaran badan, panjang badan, dan lebar bahu. Dataset dari fitur tersebut diambil dari pengukuran langsung (oleh penjahit profesional) dan pengukuran mesin menggunakan pengolahan citra digital. Ukuran baju yang tercantum pada SNI 2161:2010 menjadi rujukan data latih dan data uji sistem klasifikasi ukuran baju. Keluaran dari sistem klasifikasi adalah rekomendasi ukuran yang cocok berdasarkan ukuran tubuh pengguna. Kinerja sistem dihitung dengan menggunakan metode confusion matrix. Akurasi dari SVM yang dicapai adalah 92%.

Kata kunci : Klasifikasi, ukuran baju, SNI 2161:2010, SVM.

Abstract

Clothing size is the user's main reference when choosing clothes that are suitable for his body. The user decides to buy clothes by trying to find a suitable size in the fitting room so it requires a lot of time. The size of clothes in each clothing manufacturer is not necessarily the same, so the Badan Standarisasi Nasional (BSN) issues the Standar Nasional Indonesia (SNI) for clothes size with SNI number 2161: 2010 entitled Size of Clothing - Men's T-Shirt. The system design uses the Support Vector Machine (SVM) method to classify shirt sizes. SVM is a classification method with soft margin used to separate some features with the line support distance to the most optimal features. The input features of the system are three body size parameters namely body circumference, body length, and shoulder width. The dataset of these features is taken from direct measurements (by professional tailors) and machine measurements using digital image processing. The size of the clothes listed in SNI 2161: 2010 becomes a reference for the training data and the test data for the shirt size classification system. The output of the classification system is a suitable size recommendation based on the user's body size. System performance is calculated using the confusion matrix method. Accuracy of the SVM achieved is 92%.

Keywords : Classification, clothing size, SNI 2161:2010, SVM.

1. Pendahuluan

Ukuran baju merupakan rujukan utama pengguna saat memilih baju yang pas bagi tubuhnya. Saat hendak membeli baju di toko, pemilihan baju dilakukan dengan mencoba memakai baju yang hendak dibeli satu per satu. Percobaan tersebut dilakukan di kamar pas yang biasanya hanya boleh membawa maksimal dua atau tiga baju untuk sekali percobaan. Selain itu banyaknya calon pembeli lain yang juga ingin menggunakan kamar pas sehingga harus mengantri. Tentu proses tersebut membutuhkan waktu yang banyak.

Ukuran baju yang sangat umum digunakan dan sangat sering ditemui di pasaran adalah kelompok ukuran S, M, L, dan XL (selanjutnya disebut SML), walaupun terkadang kita temukan ukuran yang lebih kecil dari S (XS, XXS) atau lebih besar dari XL (XXL, XXXL) dan sebagainya. Namun ukuran tersebut bukanlah ukuran dengan angka parameter (panjang badan, lebar bahu, lingkaran badan) yang tetap, melainkan berbeda pada setiap konveksi atau produsen baju yang mengeluarkan *chart* ukuran mereka sendiri. Selain kelompok ukuran baju SML, ada kelompok ukuran lain yaitu ukuran baju anak dan beberapa baju *slim fit*, yang keduanya memakai ukuran angka [1], [2]. Karena ukuran SML yang berbeda-beda dan juga adanya

kelompok ukuran lain, Badan Standarisasi Nasional (BSN) mengeluarkan Standar Nasional Indonesia (SNI) untuk ukuran baju. SNI tersebut terdapat pada dokumen dengan nomor SNI 2161:2010 berjudul Ukuran Pakaian – Kaos Pria [3]. Ukuran yang tercantum pada standar tersebut menjadi rujukan data latih dan data uji sistem klasifikasi ukuran baju menggunakan metode SVM.

Support Vector Machine (SVM) merupakan salah satu algoritma pada metode klasifikasi [4]. SVM bekerja atas prinsip *Structural Risk Minimization* (SRM) dengan tujuan menemukan hyperplane terbaik. SVM pada dasarnya didesain untuk klasifikasi 2-kelas, namun dalam beberapa makalah disebutkan bahwa kombinasi K SVM dapat digunakan untuk menyelesaikan masalah klasifikasi K-kelas [5]. Maka dari itu, SVM menjadi metode yang tepat untuk menyelesaikan masalah klasifikasi ukuran kaos pria yang mana terdiri dari 4 kelas yaitu S, M, L, dan XL.

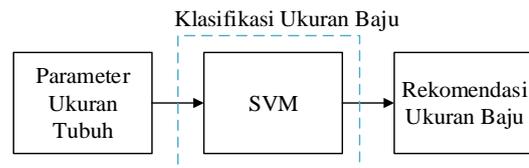
Perancangan sistem menggunakan algoritma SVM untuk melakukan klasifikasi ukuran baju pria dewasa. Fitur atau ciri yang diperlukan pada sistem klasifikasi ini adalah lingkaran badan, panjang badan, dan lebar bahu. Dataset dari fitur tersebut diambil dari pengukuran langsung (oleh penjahit profesional) dan pengukuran oleh mesin menggunakan pengolahan citra digital. Pengukuran oleh mesin menggunakan pengolahan citra digital dibahas di makalah M. Alfian Qadar [6] dan Rizal Fachmi [7]. Terdapat 142 data latih untuk melatih sistem klasifikasi dan 10 data data uji untuk menguji kinerja sistem klasifikasi. *Output* akhir dari sistem klasifikasi adalah rekomendasi ukuran yang cocok berdasarkan ukuran tubuh pengguna. Kinerja sistem dihitung dengan menggunakan metode *confusion matrix*.

2. Metode

Perancangan perangkat sepenuhnya dalam bentuk software. Dengan kata lain perangkat keras yang digunakan hanyalah personal computer dengan OS windows dalam perancangannya. Perangkat dirancang menggunakan bahasa pemrograman Python. Bahasa ini digunakan secara luas pada komputasi saintifik, data *mining*, dan *machine learning*. Tumbuhnya kebutuhan terhadap *developer machine learning* inilah yang meningkatkan popularitas Python [8].

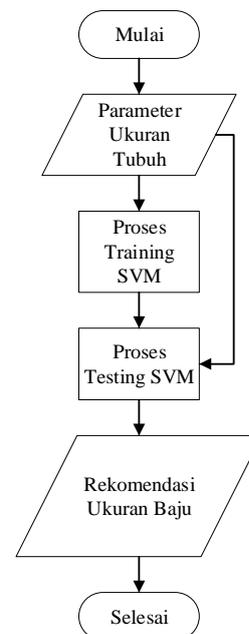
Seperti yang sudah disebutkan sebelumnya, masukkan SVM adalah nilai parameter ukuran tubuh yang merupakan output dari blok identifikasi dan perhitungan ukuran tubuh. Keluaran dari SVM adalah rekomendasi ukuran tubuh yang cocok untuk pengguna. Diagram blok klasifikasi

ukuran baju menggunakan algoritma SVM ditunjukkan pada Gambar 1.



Gambar 1. Diagram blok sistem klasifikasi ukuran baju

Parameter ukuran tubuh terdiri dari 3 parameter yaitu lingkaran badan, panjang badan, dan lebar bahu. Ketiga parameter tersebut diberi label ukuran S, M, L, atau XL. Pemberian label ukuran berdasarkan parameter lingkaran badan karena merujuk pada SNI 2161:2010 bahwa ukuran baju pada standar ini dibuat berdasarkan lingkaran badan.



Gambar 2. Diagram alir algoritma subsistem klasifikasi ukuran baju

Diagram alir sistem klasifikasi ukuran baju diberikan pada Gambar 2. Parameter ukuran tubuh hasil komputasi nilai ukuran tubuh menjadi masukan pada program SVM. Sejumlah 142 data yang terkumpul menjadi data latih untuk melakukan proses *training* pada SVM. Lalu sebuah data uji diinputkan pada SVM yang sudah di-*training* untuk melakukan proses *testing* SVM. Keluaran yang dihasilkan dari proses *testing* SVM adalah rekomendasi ukuran baju yang paling cocok untuk pengguna.

Tabel 1. Syarat ukuran pakaian – kaos pria dewasa

No.	Parameter	Satuan	Ukuran						Toleransi
			S	M	L	XL	XXL	XXXL	
1	Lingkar Badan	cm	88	94	100	106	112	118	+2 / -1
2	Lebar Lingkar Leher								
	a. Pakai kerah	cm	18	18,5	19	19,5	20	20,5	+0,5 / -0,5
	b. Tanpa kerah	cm	18	18,5	19	19,5	20	20,5	+0,5 / -0,5
3	Tinggi lingkar leher								
	a. Pakai kerah	cm	5,5	6	6,5	7	7,5	8	+0,5 / -0,5
	b. Tanpa kerah	cm	7	7,5	8	8,5	9	9,5	+0,5 / -0,5
4	Panjang bahu	cm	11	12	13	14	15	16	+1 / -1,5
5	Ukuran lengan								
	a. Lengan panjang	cm	53	54	55	56	57	58	+1 / -0,5
	b. Lengan pendek	cm	20	21	22	23	24	25	+1 / -0,5
	c. Lingkar pangkal lengan	cm	39,5	42	43,5	46	48,5	50,5	+1 / -0,5
	d. Lingkar ujung lengan	cm	36	37	38	39	40	41	+1 / -0,5
6	Panjang Badan	cm	63	66	67	69	71	73	+1 / -1

Sistem klasifikasi yang dirancang adalah *supervised learning* yang artinya setiap data pada data latih yang telah terhimpun diberikan label ukuran secara manual oleh perancang sesuai referensi SNI 2161:2010 seperti yang telah disebutkan di atas. Sama halnya dengan data latih namun data uji yang telah dilabelkan ukuran bertujuan bukan untuk menjadi masukan SVM melainkan untuk memverifikasi keluaran dari SVM apakah ukuran yang diprediksi benar atau salah.

Data latih bisa jadi tersebar secara non linear. Pada suatu dataset yang tersebar secara linear, SVM yang digunakan cukup dengan menggunakan *kernel* linear. Namun, jika SVM dengan *kernel* linear tidak bisa atau tidak maksimal melakukan klasifikasi, maka klasifikasi harus menggunakan SVM dengan *kernel* non linear (contoh: RBF, Polinomial, Sigmoid).

Pada Perancangan Sistem Klasifikasi Ukuran Baju ini menggunakan SVM dengan *kernel* RBF (*Radial Basis Function*). Dalam menggunakan *kernel* RBF diperlukan dua parameter yaitu *C* (*Complexity*) dan *gamma*. Pada parameter *C* semakin besar nilainya semakin ketat batas kelas. Sedangkan pada parameter *gamma* semakin besar nilainya semakin sempit wilayah kelas.

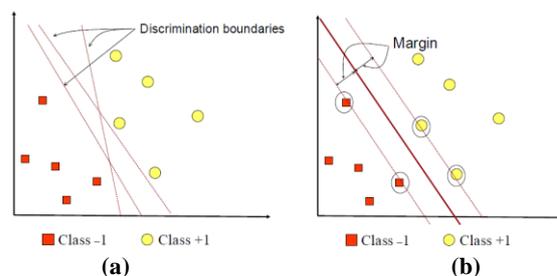
2.1. SNI 2161:2010 Ukuran Pakaian - Kaos Pria Dewasa

Standar Nasional Indonesia yang disusun ini merupakan revisi SNI 08-2161-1991, Ukuran kaos pria dewasa. Revisi ini dimaksudkan untuk menyempurnakan standar ukuran pakaian – kaos pria dewasa yang telah ada, dengan perubahan ukuran dari nomor (80, 85, 90, 95, 100) menjadi huruf (S, M, L, XL, XXL, XXXL) juga karena adanya penyempurnaan acuan normatif, cara pengambilan contoh dan prasyarat mutu yang belum tercantum sesuai dengan prosedur cara uji serta adanya perubahan format penyusunan SNI [3]. Ukuran pakaian-kaos pria dewasa yang dicantumkan dalam standar ini merupakan hasil

pengukuran dari berbagai kaos pria dewasa yang di pasar dan *factory outlet* produksi di dalam negeri yang bermerk (*branded*) maupun tidak bermerk serta telah dilakukan uji pembuatan dan grading pola di Laboratorium Pola dan Pematangan, Sekolah Tinggi Teknologi Tekstil Bandung. Syarat ukuran pakaian – kaos pria dewasa tertera pada Tabel 1 [3].

2.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) dikembangkan oleh Boser, Guyon, Vapnik, dan pertama kali dipresentasikan pada tahun 1992 di *Annual Workshop on Computational Learning Theory*. Konsep dasar SVM sebenarnya merupakan perpaduan harmonis dari teori-teori komputasi yang telah ada puluhan tahun sebelumnya, seperti *margin hyperplane* (Duda & Hart tahun 1973, Cover tahun 1965, Vapnik 1964, dsb.), pada tahun 1950 Aronszajn memperkenalkan *kernel*, suatu metode pada SVM untuk menyelesaikan masalah non-linear, dan demikian juga dengan konsep-konsep pendukung yang lain. Tetapi hingga tahun 1992, belum pernah ada upaya merangkaikan komponen-komponen tersebut [5].



Gambar 3. SVM mencari hyperplane

Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dua buah class pada input space [5]. *Hyperplane* dalam ruang vektor berdimensi *d* adalah *affine subspace*

berdimensi $d-1$ yang membagi ruang vector tersebut ke dalam dua bagian, yang masing-masing berkorespondensi pada class yang berbeda [9]. Contohnya, jika ruang vector berdimensi 2, maka hyperplane-nya berdimensi 1 atau garis.

Gambar 3 (a) [5] memperlihatkan beberapa pattern yang merupakan anggota dari dua buah class : +1 dan -1. Pattern yang tergabung pada class -1 disimbolkan dengan warna merah (kotak), sedangkan pattern pada class +1, disimbolkan dengan warna kuning (lingkaran). Problem klasifikasi dapat diterjemahkan dengan usaha menemukan garis (*hyperplane*) yang memisahkan antara kedua kelompok tersebut. Berbagai alternatif garis pemisah (*discrimination boundaries*) ditunjukkan pada Gambar 3 (a).

Hyperplane pemisah terbaik antara kedua *class* dapat ditemukan dengan mengukur *margin hyperplane* tersebut dan mencari titik maksimalnya. Margin adalah jarak antara *hyperplane* tersebut dengan pattern terdekat dari masing-masing *class*. Pattern yang paling dekat ini disebut sebagai *support vector*. Garis solid pada Gambar 3 (b) [5] menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua *class*, sedangkan titik merah dan kuning yang berada dalam lingkaran hitam adalah *support vector*. Usaha untuk mencari lokasi *hyperplane* ini merupakan inti dari proses pembelajaran pada SVM [5].

Data yang tersedia dinotasikan sebagai $\vec{x}_i \in \mathcal{R}^d$ sedangkan label masing-masing dinotasikan $y_i \in \{-1, +1\}$ untuk $i = 1, 2, \dots, l$, yang mana l adalah banyaknya data. Diasumsikan kedua class -1 dan +1 dapat terpisah secara sempurna oleh hyperplane berdimensi d , yang didefinisikan

$$\vec{w} \cdot \vec{x} + b = 0 \tag{1}$$

Pattern \vec{x}_i yang termasuk class -1 (sampel negatif) dapat dirumuskan sebagai pattern yang memenuhi pertidaksamaan

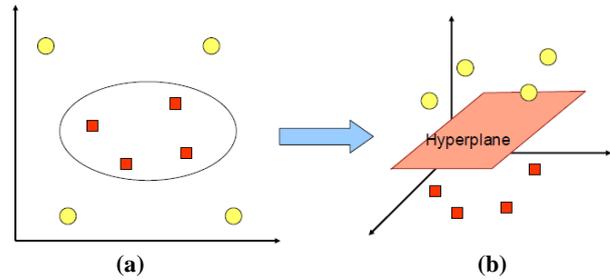
$$\vec{w} \cdot \vec{x} + b \leq -1 \tag{2}$$

sedangkan pattern \vec{x}_i yang termasuk class +1 (sampel positif)

$$\vec{w} \cdot \vec{x} + b \geq +1 \tag{3}$$

2.3. SVM Non-Linear

Pada umumnya untuk masalah yang ada dalam domain dunia nyata, kebanyakan bersifat non linier. Algoritma *Support Vector Machine (SVM)* dimodifikasi dengan cara memasukkan fungsi kernel kedalam non linier SVM, dengan cara yang pertama yaitu data \vec{x}_i dipetakan ke dalam fungsi $\Phi(x)$ ke ruang vector yang memiliki ukuran dimensi tinggi [5], [9].



Gambar 4. Mapping Fungsi Φ

Ilustrasi dari konsep ini dapat dilihat pada Gambar 4 [5]. Pada Gambar 4 (a) diperlihatkan data pada class kuning dan data pada class merah yang berada pada input space berdimensi dua tidak dapat dipisahkan secara linear. Selanjutnya Gambar 4 (b) menunjukkan bahwa fungsi Φ memetakan tiap data pada input space tersebut ke ruang vector baru yang berdimensi lebih tinggi (dimensi 3), dimana kedua class dapat dipisahkan secara linear oleh sebuah *hyperplane* [5]. Notasi matematika dari mapping ini adalah pada persamaan (2.4)

$$\Phi: \mathcal{R}^d \rightarrow \mathcal{R}^q \quad d < q \tag{4}$$

Karena umumnya transformasi Φ ini tidak diketahui, dan sangat sulit untuk dipahami secara mudah, maka perhitungan dot product tersebut sesuai teori Mercer dapat digantikan dengan fungsi kernel $K(\vec{x}_i, \vec{x}_j)$ yang mendefinisikan secara implisit transformasi Φ [5]. Hal ini disebut sebagai *Kernel Trick*, yang dirumuskan pada persamaan (2.5)

$$K(\vec{x}_i, \vec{x}_j) = \Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j) \tag{5}$$

Kernel trick memberikan berbagai kemudahan, karena dalam proses pembelajaran SVM, untuk menentukan support vector, kita hanya cukup mengetahui fungsi kernel yang dipakai, dan tidak perlu mengetahui wujud dari fungsi non linear Φ . Berbagai jenis fungsi kernel dikenal, sebagaimana dirangkumkan pada Tabel 2 [5]. Pada perancangan ini, *kernel* yang digunakan adalah Gaussian.

Tabel 2. Kernel umum yang dipakai dalam SVM

Jenis Kernel	Definisi
Polynomial	$K(\vec{x}_i, \vec{x}_j) = (\vec{x}_i \cdot \vec{x}_j + r)^p$
Gaussian	$K(\vec{x}_i, \vec{x}_j) = \exp\left(-\frac{\ \vec{x}_i - \vec{x}_j\ ^2}{2\sigma^2}\right)$
RBF	$K(\vec{x}_i, \vec{x}_j) = \exp(-\gamma\ \vec{x}_i - \vec{x}_j\ ^2)$
Sigmoid	$K(\vec{x}_i, \vec{x}_j) = \tanh(\alpha\vec{x}_i \cdot \vec{x}_j + \beta)$

2.4. Multiclass SVM

Ada dua pilihan untuk mengimplementasikan multiclass SVM yaitu dengan menggabungkan beberapa SVM biner

atau menggabungkan semua data yang terdiri dari beberapa kelas ke dalam sebuah bentuk permasalahan optimal [10]. Namun pada pendekatan yang kedua permasalahan optimasi yang harus diselesaikan jauh lebih rumit. Berikut ini adalah metode yang umum digunakan untuk mengimplementasikan multiclass SVM dengan pendekatan yang pertama:

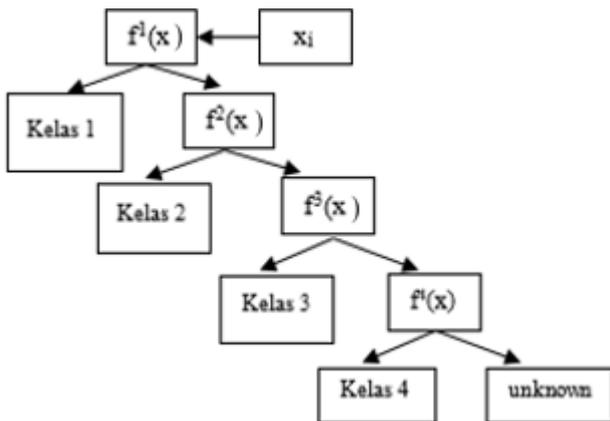
1. Metode *One-against-all*

Dengan menggunakan metode ini, dibangun k buah model SVM biner (k adalah jumlah kelas). Dari setiap klasifikasi ke-i dilatih dengan menggunakan data secara keseluruhan. Misal, terdapat sebuah permasalahan klasifikasi dengan 4 buah kelas yang digunakan pada pelatihan hanya 4 buah SVM biner seperti yang ditunjukkan pada Tabel 3 [10]:

Tabel 3. Contoh empat SVM dengan metode *one-against-all*

$Y_i = 1$	$Y_i = -1$	Hipotesis
Kelas 1	Bukan kelas 1	$f^1(x) = (w^1)x + b^1$
Kelas 2	Bukan kelas 2	$f^2(x) = (w^2)x + b^2$
Kelas 3	Bukan kelas 3	$f^3(x) = (w^3)x + b^3$
Kelas 4	Bukan kelas 4	$f^4(x) = (w^4)x + b^4$

Decision tree pada contoh SVM dengan 4 kelas metode *one-against-all* ditunjukkan pada Gambar 5 [10].



Gambar 5. Contoh klasifikasi dengan metode *One-against-all*

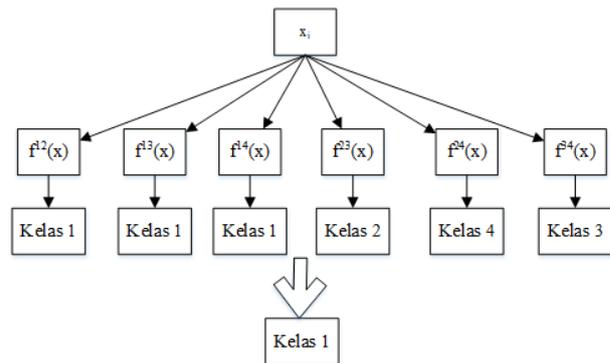
2. Metode *One-against-one*

Dengan menggunakan metode ini, dibangun $k(k-1)/2$ buah model klasifikasi biner (k adalah jumlah kelas). Terdapat beberapa metode untuk melakukan pengujian setelah keseluruhan $k(k-1)/2$ model klasifikasi selesai dibangun, salah satunya adalah metode *voting*. Proses *voting* ditunjukkan pada Tabel 4 [10].

Tabel 4. Contoh empat SVM dengan metode *one-against-one*

$Y_i = 1$	$Y_i = -1$	Hipotesis
Kelas 1	Kelas 2	$f^{12}(x) = (w^{12})x + b^{12}$
Kelas 1	Kelas 3	$f^{13}(x) = (w^{13})x + b^{13}$
Kelas 1	Kelas 4	$f^{14}(x) = (w^{14})x + b^{14}$
Kelas 2	Kelas 3	$f^{23}(x) = (w^{23})x + b^{23}$
Kelas 2	Kelas 4	$f^{24}(x) = (w^{24})x + b^{24}$
Kelas 3	Kelas 4	$f^{34}(x) = (w^{34})x + b^{34}$

Diagram proses *voting* untuk memutuskan kelas yang terpilih ditunjukkan pada Gambar 6 [10]



Gambar 6. Diagram proses *voting* pada metode *One-against-one*

2.4. *Confusion Matrix*

Confusion matrix adalah metode untuk mengukur kinerja suatu sistem klasifikasi. *Confusion matrix* menurut Han dan Kamber [11] dapat diartikan sebagai suatu alat yang memiliki fungsi untuk melakukan analisis apakah *classifier* tersebut baik dalam mengenali tuple dari kelas yang berbeda. Nilai dari *True-Positive* dan *True-Negative* memberikan informasi ketika *classifier* dalam melakukan klasifikasi data bernilai benar, sedangkan *False-Positive* dan *False-Negative* memberikan informasi ketika *classifier* salah dalam melakukan klasifikasi data. Tabel confusion matrix diberikan pada Tabel 5 [11]

Tabel 5. *Confusion Matrix* menampilkan total *positive* dan *negative tuple*

		Predicted class		Total
		yes	no	
Actual Class	Yes	TP	FN	P
	No	FP	TN	N
Total		P'	N'	P+N

Keterangan Tabel 5 [11] *Confusion Matrix* adalah sebagai berikut:

- True Positives (TP)* adalah jumlah *record* data positif yang diklasifikasikan sebagai nilai positif
- False Positives (FP)* adalah jumlah *record* data negatif yang diklasifikasikan sebagai nilai positif
- False Negatives (FN)* adalah jumlah *record* data positif yang diklasifikasikan sebagai nilai negatif
- True Negatives (TN)* adalah jumlah *record* data negatif yang diklasifikasikan sebagai nilai negatif

Nilai yang dihasilkan melalui metode *Confusion Matrix* adalah berupa evaluasi sebagai berikut [12] :

- Accuracy*, merupakan rasio prediksi benar (positif dan negatif) dengan keseluruhan data.

$$Accuracy = (TP + TN) / (TP + FP + FN + TN) \quad (2.6)$$

- Precision*, merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif.

$$Precision = TP / (TP + FP) \quad (2.7)$$

- Recall*, merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif

$$Recall = TP / (TP + FN) \quad (2.8)$$

Sementara itu, pada klasifikasi dengan jumlah keluaran kelas yang lebih dari dua (*multiclass*), cara menghitung *accuracy*, *precision*, dan *recall* dapat dilakukan dengan menghitung rata-rata dari nilai *accuracy*, *precision*, dan *recall* pada setiap kelas. Persamaan (2.9), (2.10), dan (2.11) merupakan formula untuk menghitung nilai akurasi, presisi dan recall dari sistem klasifikasi *multiclass*

$$Accuracy = \frac{\sum_{i=1}^l (TP_i + TN_i) / \sum_{i=1}^l (TP_i + FP_i + FN_i + TN_i)}{l} \quad (9)$$

$$Precision = \sum_{i=1}^l TP_i / \sum_{i=1}^l (TP_i + FP_i) \quad (10)$$

$$Recall = \sum_{i=1}^l TP_i / \sum_{i=1}^l (TP_i + FN_i) \quad (11)$$

Keterangan persamaan (9), (10), dan (11) adalah sebagai berikut:

- TP_i adalah *True Positive*, yaitu jumlah data positif yang terklasifikasi dengan benar oleh sistem untuk kelas ke-*i*.
- TN_i adalah *True Negative*, yaitu jumlah data negatif yang terklasifikasi dengan benar oleh sistem untuk kelas ke-*i*.
- FN_i adalah *False Negative*, yaitu jumlah data negatif namun terklasifikasi salah oleh sistem untuk kelas ke-*i*.
- FP_i adalah *False Positive*, yaitu jumlah data positif namun terklasifikasi salah oleh sistem untuk kelas ke-*i*
- l* adalah jumlah kelas.

3. Pengujian dan Analisis

Klasifikasi ukuran baju diuji dengan *Confusion Matrix*, suatu metode yang sering dipakai untuk menguji kinerja

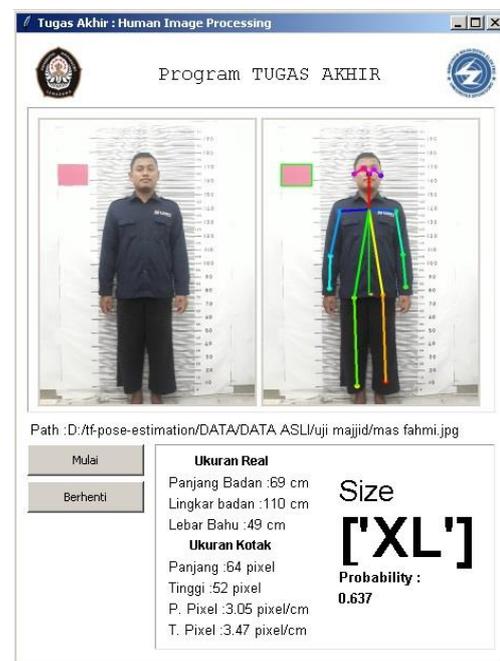
suatu metode klasifikasi. Dari *Confusion Matrix* kemudian dapat dihitung akurasi, presisi, dan *recall*. Ketiganya merupakan parameter yang menunjukkan kinerja sistem klasifikasi.

Klasifikasi yang menggunakan metode SVM untuk menentukan ukuran baju pengguna telah dibangun dengan 142 data latih. Data uji telah terkumpul sejumlah 25 data, 10 data diantaranya adalah data yang juga terdapat di data latih dan 15 data sisanya adalah data baru yang diambil untuk dilakukan pengujian langsung di tempat. Pengujian ini dilakukan dengan dua rujukan aktual yang berbeda yaitu berdasarkan ukuran SNI 2161:2010 dan berdasarkan ukuran subjektif pengguna.

3.1. Pengujian Klasifikasi Ukuran Baju Berdasarkan Ukuran SNI 2161:2010

Pengujian pertama adalah berdasarkan rujukan ukuran SNI 2161:2010. Pengujian hanya dilakukan sekali untuk setiap data uji. Pengujian dilakukan dengan memasukkan satu per satu data uji ke program utama “Sistem Identifikasi Ukuran Tubuh Untuk Menentukan Ukuran Baju”. Keluaran klasifikasi yaitu rekomendasi ukuran baju tercantum pada keluaran program utama yang dalam bentuk GUI bersamaan dengan informasi dari keluaran subsistem lainnya.

Berikut keluaran hasil prediksi klasifikasi berupa rekomendasi ukuran baju ditunjukkan pada Gambar 7.



Gambar 7. Tampilan keluaran klasifikasi pada program utama

Hasil pengujian 10 data uji ditunjukkan pada Tabel 6

Tabel 6. Hasil Pengujian Ukuran Baju vs Ukuran SNI

No.	Nama Pengguna	Ukuran Rekomendasi	Ukuran SNI 2161:2010
1	David	S	S
2	Aldo	S	S
3	Absor	L	L
4	Fahmi	XL	XL
5	Batul	XL	XL
6	Rizqi	M	M
7	Zaki	L	L
8	Irfan	L	L
9	Yere	M	M
10	Faiz	M	M
11	Frenky	S	S
12	Cristian	S	S
13	Gabriel	S	S
14	Arif	L	XL
15	Arinal	M	M
16	Trio	M	M
17	Amri	M	M
18	Rico	S	S
19	Hafidz	M	M
20	Ridho	L	L
21	Bondan	M	M
22	Avi	L	L
23	Rizal	L	L
24	Ibek	S	M
25	Fauzan	L	L

Tabel 7. Confusion Matrix Pengujian Klasifikasi Ukuran Baju Berdasarkan Ukuran SNI 2161:2010

		Prediksi			
		S	M	L	XL
Aktual	S	7	0	0	0
	M	1	7	0	0
	L	0	0	7	0
	XL	0	0	1	2

Dari Confusion Matrix pada tabel di atas didapatkan

- Akurasi

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{Jumlah hasil klasifikasi yang benar}}{\text{Total seluruh data}} \times 100\% \quad (12) \\
 &= \frac{7 + 5 + 8 + 2}{10} \times 100\% \\
 &= \frac{23}{25} \times 100\% \\
 &= 92\%
 \end{aligned}$$

- Presisi

$$\begin{aligned}
 \text{Presisi } S &= \frac{\text{Jumlah hasil prediksi } S \text{ yang benar}}{\text{Total seluruh data yang diprediksi } S} \quad (13) \\
 &= \frac{7}{8} = 0,875
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi } M &= \frac{\text{Jumlah hasil prediksi } M \text{ yang benar}}{\text{Total seluruh data yang diprediksi } M} \quad (14) \\
 &= \frac{7}{7} = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi } L &= \frac{\text{Jumlah hasil prediksi } L \text{ yang benar}}{\text{Total seluruh data yang diprediksi } L} \quad (15) \\
 &= \frac{7}{8} = 0,875
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi } XL &= \frac{\text{Jumlah hasil prediksi } XL \text{ yang benar}}{\text{Total seluruh data yang diprediksi } XL} \quad (16) \\
 &= \frac{2}{2} = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi} &= \frac{\text{Presisi } S + \text{Presisi } M + \text{Presisi } L + \text{Presisi } XL}{4} \times 100\% \quad (17) \\
 &= \frac{0,875 + 1 + 0,875 + 1}{4} \times 100\% \\
 &= 93,8\%
 \end{aligned}$$

- Recall

$$\begin{aligned}
 \text{Recall } S &= \frac{\text{Jumlah hasil prediksi } S \text{ yang benar}}{\text{Total seluruh data aktual } S} \quad (18) \\
 &= \frac{7}{7} = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall } M &= \frac{\text{Jumlah hasil prediksi } M \text{ yang benar}}{\text{Total seluruh data aktual } M} \quad (19) \\
 &= \frac{7}{8} = 0,875
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall } L &= \frac{\text{Jumlah hasil prediksi } L \text{ yang benar}}{\text{Total seluruh data aktual } L} \quad (20) \\
 &= \frac{7}{7} = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall } XL &= \frac{\text{Jumlah hasil prediksi } XL \text{ yang benar}}{\text{Total seluruh data aktual } XL} \quad (21) \\
 &= \frac{2}{3} = 0,667
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall} &= \frac{\text{Recall } S + \text{Recall } M + \text{Recall } L + \text{Recall } XL}{4} \times 100\% \quad (22) \\
 &= \frac{1 + 0,875 + 1 + 0,667}{4} \times 100\% \\
 &= 88,5\%
 \end{aligned}$$

Hasil dari pengujian klasifikasi ukuran baju berdasarkan ukuran SNI 2161:2010 menunjukkan kinerja klasifikasi yang baik. Akurasi klasifikasi adalah 92%, presisi klasifikasi adalah 93,8% dan recall klasifikasi adalah 88,5%. Pada pengujian ini, terjadi kesalahan klasifikasi pada 2 data uji sedangkan 23 data uji sisanya diklasifikasikan dengan benar sesuai ukuran pada SNI 2161:2010.

3.2. Pengujian Klasifikasi Ukuran Baju Berdasarkan Ukuran Subjektif Pengguna

Pengujian kedua adalah berdasarkan rujukan ukuran subjektif pengguna. Pengujian ini sama dengan pengujian pertama, namun hasil rekomendasi ukuran baju data uji dibandingkan dengan ukuran subjektif pengguna bukan berdasarkan SNI 2161:2012.

Berikut hasil pengujian 10 data uji berdasarkan ukuran subjektif pengguna ditunjukkan pada Tabel. 8

Tabel 8. Hasil Pengujian Ukuran Baju vs Ukuran Subjektif Pengguna

No.	Nama Pengguna	Ukuran Rekomendasi	Ukuran Subjektif Pengguna
1	David	S	S
2	Aldo	S	S
3	Absor	L	L
4	Fahmi	XL	XL
5	Batul	XL	XL
6	Rizqi	M	M
7	Zaki	L	L
8	Irfan	L	L
9	Yere	M	L
10	Faiz	M	M
11	Frenky	S	S
12	Cristian	S	S
13	Gabriel	S	S
14	Arif	L	XL
15	Arinal	M	M
16	Trio	M	M
17	Amri	M	M
18	Rico	S	S
19	Hafidz	M	M
20	Ridho	L	L
21	Bondan	M	L
22	Avi	L	L
23	Rizal	L	L
24	Ibek	S	M
25	Fauzan	L	L

Confusion Matrix pengujian tersebut ditunjukkan pada Tabel 9

Tabel 9. Confusion Matrix Pengujian Klasifikasi Ukuran Baju Berdasarkan Ukuran Subjektif Pengguna

		Prediksi			
		S	M	L	XL
Aktual	S	6	0	0	0
	M	2	5	0	0
	L	0	2	7	0
	XL	0	0	1	2

Dari Confusion Matrix pada tabel di atas didapatkan

- Akurasi

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{Jumlah hasil klasifikasi yang benar}}{\text{Total seluruh data}} \times 100\% \quad (23) \\
 &= \frac{6 + 5 + 7 + 2}{25} \times 100\% \\
 &= \frac{20}{25} \times 100\% \\
 &= 80\%
 \end{aligned}$$

- Presisi

$$\begin{aligned}
 \text{Presisi } S &= \frac{\text{Jumlah hasil prediksi } S \text{ yang benar}}{\text{Total seluruh data yang diprediksi } S} \quad (24) \\
 &= \frac{6}{8} = 0,75
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi } M &= \frac{\text{Jumlah hasil prediksi } M \text{ yang benar}}{\text{Total seluruh data yang diprediksi } M} \quad (25) \\
 &= \frac{5}{7} = 0,714
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi } L &= \frac{\text{Jumlah hasil prediksi } L \text{ yang benar}}{\text{Total seluruh data yang diprediksi } L} \quad (26) \\
 &= \frac{7}{8} = 0,875
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi } XL &= \frac{\text{Jumlah hasil prediksi } XL \text{ yang benar}}{\text{Total seluruh data yang diprediksi } XL} \quad (27) \\
 &= \frac{2}{2} = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi} &= \frac{\text{Presisi } S + \text{Presisi } M + \text{Presisi } L + \text{Presisi } XL}{4} \times 100\% \quad (28) \\
 &= \frac{0,75 + 0,714 + 0,875 + 1}{4} \times 100\% \\
 &= 83,5\%
 \end{aligned}$$

- Recall

$$\begin{aligned}
 \text{Recall } S &= \frac{\text{Jumlah hasil prediksi } S \text{ yang benar}}{\text{Total seluruh data aktual } S} \quad (29) \\
 &= \frac{6}{6} = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall } M &= \frac{\text{Jumlah hasil prediksi } M \text{ yang benar}}{\text{Total seluruh data aktual } M} \quad (30) \\
 &= \frac{5}{7} = 0,714
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall } L &= \frac{\text{Jumlah hasil prediksi } L \text{ yang benar}}{\text{Total seluruh data aktual } L} \quad (31) \\
 &= \frac{7}{9} = 0,778
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall } XL &= \frac{\text{Jumlah hasil prediksi } XL \text{ yang benar}}{\text{Total seluruh data aktual } XL} \quad (32) \\
 &= \frac{2}{3} = 0,667
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall} &= \frac{\text{Recall } S + \text{Recall } M + \text{Recall } L + \text{Recall } XL}{4} \times 100\% \quad (33) \\
 &= \frac{1 + 0,714 + 0,778 + 0,667}{4} \times 100\% \\
 &= 79\%
 \end{aligned}$$

Hasil dari pengujian klasifikasi ukuran baju berdasarkan ukuran subjektif pengguna menunjukkan kinerja klasifikasi cukup baik. Akurasi klasifikasi adalah 80%. Sistem mengklasifikasikan 20 data uji dengan benar dan 5 data uji yang salah. Presisi klasifikasi adalah 83,5% dan recall klasifikasi adalah 79%. Artinya, pada pengujian ini, keluaran klasifikasi ukuran baju yaitu ukuran rekomendasi adalah masih memiliki galat dalam mengklasifikasikan ukuran baju berdasarkan ukuran aktual subjektif pengguna.

Tabel 10. Perbandingan Ukuran Rekomendasi, Ukuran SNI 2161:2010, dan Ukuran Subjektif Pengguna

No.	Nama Pengguna	Ukuran Rekomendasi	Ukuran SNI 2161:2010	Ukuran Subjektif Pengguna
1	David	S	S	S
2	Aldo	S	S	S
3	Absor	L	L	L
4	Fahmi	XL	XL	XL
5	Batul	XL	XL	XL
6	Rizqi	M	M	M
7	Zaki	L	L	L
8	Irfan	L	L	L

Tabel 10. Lanjutan

9	Yere	M	M	L
10	Faiz	M	M	M
11	Frenky	S	S	S
12	Cristian	S	S	S
13	Gabriel	S	S	S
14	Arif	L	XL	XL
15	Arinal	M	M	M
16	Trio	M	M	M
17	Amri	M	M	M
18	Rico	S	S	S
19	Hafidz	M	M	M
20	Ridho	L	L	L
21	Bondan	M	M	L
22	Avi	L	L	L
23	Rizal	L	L	L
24	Ibek	S	M	M
25	Fauzan	L	L	L

4. Kesimpulan

Berdasarkan perancangan, pengujian dan analisis yang telah dilakukan maka dapat disimpulkan bahwa sistem klasifikasi ukuran baju menggunakan metode *Support Vector Machine* (SVM) untuk menentukan rekomendasi ukuran baju berdasarkan ukuran tubuh yang telah didapatkan pada identifikasi dan perhitungan ukuran tubuh. Pada pengujian berdasarkan ukuran SNI 2161:2010, akurasi senilai 92%, presisi senilai 93,8%, dan *recall* senilai 88,5%. Sedangkan pada pengujian berdasarkan ukuran subjektif pengguna menghasilkan akurasi senilai 80%, presisi senilai 83,5%, dan *recall* senilai 79%. Kesalahan dalam prediksi ukuran baju berdasarkan ukuran subjektif pengguna dikarenakan ukuran baju yang pengguna pakai adalah ukuran baju yang pengguna pikir adalah yang terbaik bagi tubuhnya, namun tidak sesuai dengan ukuran berdasarkan SNI 2161:2010.

Referensi

- [1]. Harnas Uniform. *Standar Ukuran Kemeja dan Celana Kantor*. [Online]. Tersedia: <https://seragam-kantor.com/info-seragam/standar-ukuran/>. Diakses: Agt. 28, 2019 .
- [2]. Steven Home. *Piyama Anak*. [Online]. Tersedia: <https://www.bukalapak.com/p/fashion-anak/anak-unisex/piyama-1501/>. Diakses: Agt. 28, 2019.
- [3]. Ukuran Pakaian – Kaos Pria Dewasa, SNI 2161, 2010.
- [4]. N. R. S. Huaturuk *dkk.*, “Komparasi Akurasi Naïve Bayes dan Support Vector Machine (SVM) untuk Rekomendasi Produk in Fashion Dress”, dalam *Proc. on Conference on Electrical Engineering, Telematics, Industrial Technology, and Creative Media*, 2018, hal. 169 -174
- [5]. A. S. Nugroho, A. B. Witarto, dan D. Handoko, “Support Vector Machine Teori dan Aplikasinya dalam Bioinformatika”, dalam *Proc. Indonesian Scientific Meeting, Gifu-Japan, 2003*.
- [6]. M. A. Qadar, “Aplikasi Deteksi Kotak Untuk Menentukan Ukuran Tubuh Menggunakan Algoritma HSV”, Laporan Tugas Akhir, Departemen Teknik Elektro, Universitas Diponegoro, Semarang, Indonesia, 2019.
- [7]. R. Fachmi, “Identifikasi dan Perhitungan Ukuran Tubuh”, Tugas Akhir, Departemen Teknik Elektro, Universitas Diponegoro, Semarang, Indonesia, 2019.
- [8]. Makers Institute. *Bahasa Pemrograman yang Populer dan Dibutuhkan*. [Online]. Tersedia: <https://medium.com/@makersinstitute/bahasa-pemrograman-yang-populer-dan-dibutuhkan-49d592b690c8>. Diakses: Sept. 10, 2019.
- [9]. N. Cristianini, J. Shawe-Taylor, “*An introduction to support vector machines and other kernel-based learning methods*”, Cambridgeshire, Britania Raya: Cambridge University Press, 2000.
- [10]. N. Muna, “Identifikasi Diabetic Retinopathy Menggunakan Support Vector Machine”, Laporan Tugas Akhir, Universitas Islam Negeri (UIN) Maulana Malik Ibrahim, Malang, Indonesia, 2013.
- [11]. J. Han, M. Kamber, “Data Mining: C. D. H. Concepts and Techniques”, Oxford University Press, 2012.
- [12]. R. Arthana, *Mengenal Accuracy, Precision, Recall dan Specificity serta yang diprioritaskan dalam Machine Learning*. [Online]. Tersedia: <https://medium.com/@rey1024/mengenal-accuracy-precision-recall-dan-specificity-septa-yang-diprioritaskan-b79ff4d77de8>. Diakses: Sep. 18, 2019.