

Klasifikasi Tulang Tengkorak Berdasarkan Jenis Kelamin Dalam Antropologi Forensik Menggunakan Metode *Support Vector Machine*

Siti Sri Rahayu¹, Iis Afrianty², Elvia Budianita³, Fadhilah Syafria⁴

^{1,2,3,4}Universitas Islam Negeri Sultan Syarif Kasim Riau,

Jl. HR. Soebrantas, Simpang Baru, Kota Pekanbaru, Riau, Indonesia

E-mail: 12050120515@students.uin-suska.ac.id¹, iis.afrianty@uin-suska.ac.id², elvia.budianita@uin-suska.ac.id³, fadhilah.syafria@uin-suska.ac.id⁴

Abstract - Classification of skull bones by sex is part of human biological profile identification in forensic anthropology that aims to determine whether the skeleton belongs to a male or female. The most popular method for determining sex from bones is DNA analysis. However, under some conditions such as burnt, damaged, or very dry skeletal remains, DNA analysis cannot provide accurate results. So forensic anthropology is developing by utilizing the help of machine learning technology. This research shows the performance of Support Vector Machine in classifying skull bones based on gender. The skull parameter data used is data collected by Dr. William Howells from craniometric measurements consisting of male and female data with a total of 2524 data and 82 features, namely bizygomatic breadth, glabello-occipital length and others. In building the skull bone classification model, the Support Vector Machine kernels used are linear, RBF, and polynomial. Based on the test results, the best accuracy was obtained in each kernel function, namely the linear kernel obtained the best accuracy of 88.14% with $C = 2$. For the RBF kernel, the best accuracy was 91.30% at $C = 2$, $\gamma = \text{'auto'}$. For the polynomial kernel, the best accuracy was 88.14% at $C = 1$ and 2 , $\gamma = 1$ and 2 , $d = 1$. The evaluation results show that the Support Vector Machine model with the RBF kernel has proven to be the optimal choice in skull bone classification compared to other kernels, based on accuracy, precision, recall, and CrossValidation measurements reaching values above 90%. These results indicate that the skull bone classification model based on gender using Support Vector Machine is highly recommended in skull bone classification.

Keywords - *Forensic Anthropology, Skull Bones, Sex Classification, Support Vector Machine.*

Intisari – Klasifikasi tulang tengkorak berdasarkan jenis kelamin merupakan bagian dari identifikasi profil biologis manusia pada antropologi forensik yang bertujuan untuk mengetahui kerangka tersebut milik laki-laki atau perempuan. Metode yang paling populer untuk menentukan jenis kelamin dari tulang adalah analisis DNA. Namun beberapa kondisi seperti sisa-sisa kerangka terbakar, rusak, atau sangat kering, maka analisis DNA tidak dapat memberikan hasil yang akurat. Sehingga antropologi forensik berkembang dengan memanfaatkan bantuan teknologi *machine learning*. Penelitian ini menunjukkan kinerja *Support Vector Machine* dalam mengklasifikasikan tulang tengkorak berdasarkan jenis kelamin. Data parameter tengkorak yang digunakan adalah data yang dikumpulkan Dr. William Howells dari pengukuran kranimetri yang terdiri dari data pria dan wanita dengan total data 2524 dan 82 fitur, yaitu *bizygomatic breadth*, *glabello-occipital length* dan lainnya. Dalam membangun model klasifikasi tulang tengkorak, kernel *Support Vector Machine* yang digunakan adalah linear, RBF, dan polynomial. Berdasarkan hasil uji coba, akurasi terbaik diperoleh pada masing-masing fungsi kernel, yaitu kernel linear memperoleh akurasi terbaik sebesar 88,14% dengan $C = 2$. Untuk kernel RBF akurasi terbaik sebesar 91,30% pada $C = 2$, $\gamma = \text{'auto'}$. Untuk kernel polinomial akurasi terbaik sebesar 88,14% pada $C = 1$ dan 2 , $\gamma = 1$ dan 2 , $d = 1$. Hasil evaluasi menunjukkan bahwa model *Support Vector Machine* dengan kernel RBF telah terbukti menjadi pilihan optimal dalam klasifikasi tulang tengkorak dibandingkan dengan kernel lainnya, berdasarkan pengukuran *accuracy*, *precision*, *recall*, dan *CrossValidation* mencapai nilai di atas 90%. Dengan hasil ini menandakan bahwa model klasifikasi tulang tengkorak berdasarkan jenis kelamin menggunakan *Support Vector Machine* sangat direkomendasikan baik dalam klasifikasi tulang tengkorak.

Kata Kunci – Antropologi Forensik, Tulang Tengkorak, Klasifikasi Jenis Kelamin, *Support Vector Machine*.

I. PENDAHULUAN

Antropologi forensik adalah penerapan metode dan teori yang berkaitan dengan pemulihan dan analisis kerangka untuk menafsirkan variasi manusia, biologi kerangka dan biomekanika tulang. Dalam praktik antropologi forensik sering memperkirakan jenis kelamin, garis keturunan, usia, dan perawakan dari sisa kerangka individu yang tidak teridentifikasi [1]. Identifikasi jenis kelamin adalah komponen paling krusial dalam antropologi forensik individu dan merupakan langkah awal dalam proses pengenalan sisa-sisa kerangka [2]. Pengetahuan tentang jenis kelamin dari sisa-sisa kerangka tulang yang tidak dikenal sangat penting untuk membuat estimasi usia yang lebih akurat. Oleh karena itu, penentuan jenis kelamin adalah langkah yang diperlukan untuk mengidentifikasi usia, keturunan, dan perawakan lebih lanjut [3].

Panggul, tulang paha, tulang paha bawah, tulang tibia, tulang humerus, tulang radius, tulang hasta, tulang rahang bawah, dan tulang tengkorak adalah bagian dari tulang yang dapat membantu untuk penentuan jenis kelamin [4]. Diantara semua karakteristik tulang, ciri-ciri panggul memperlihatkan perbedaan yang signifikan antara jenis kelamin laki-laki dan perempuan sehingga dianggap sebagai indikator yang paling andal terkait penentuan jenis kelamin [5]. Tetapi tengkorak karena pengawetannya yang lebih baik, terdiri dari jaringan keras serta memiliki stabilitas fitur dimorfik seksual lebih baik sehingga diyakini sebagai metode kedua paling efektif dalam menentukan jenis kelamin [4], [6]. Tengkorak adalah salah satu komponen kerangka yang paling berguna untuk menentukan jenis kelamin dan merupakan tulang terbaik yang memiliki tingkat akurasi hingga 90% setelah tulang panggul [7].

Dari penelitian sebelumnya, metode yang paling populer untuk menentukan jenis kelamin dari tulang yang ditemukan adalah analisis *Deoxyribose Nucleic Acid* (DNA) melalui pemeriksaan laboratorium atau pemeriksaan radiologi [3]. Namun dalam beberapa situasi, seperti ketika sisa-sisa kerangka terbakar, rusak, atau sangat kering, maka analisis DNA tidak dapat memberikan hasil yang akurat disebabkan ekstraksi protein DNA menjadi tidak mungkin dalam proses tersebut. Oleh karena itu, bagian penting pada tulang yang berisi informasi tidak dapat diakses [3]. Dalam bidang diagnostik yang merekomendasikan metode laboratorium membutuhkan waktu yang lama serta biaya yang mahal untuk mendapatkan hasil yang akurat, dikarenakan perlu mempersiapkan pekerjaan kerangka, individualitas kerangka harus diperiksa, jejak penyakit dicari dan waktu sejak kematian dinilai [8]. Sehingga diperlukan teknik *machine learning* untuk mengatasi keterbatasan metode DNA [9].

Metode tradisional untuk memperkirakan jenis kelamin tengkorak terutama adalah diskriminasi morfologi dan diskriminasi pengukuran. Diskriminasi morfologi yang bergantung pada pemahaman ahli tentang perbedaan karakteristik morfologi tengkorak laki-laki dan perempuan dari pengamatan kriteria visual. Pemahaman subjektif para ahli dapat menyebabkan tingkat pengenalannya rendah, secara teoritis tidak memadai dan memiliki keandalan yang rendah. Diskriminasi pengukuran merujuk pada proses awal yang melibatkan identifikasi titik-titik fitur pada tengkorak, pengukuran sejumlah indeks fitur khusus jenis kelamin pada tengkorak, dan penerapan fungsi diskriminan menggunakan data geometris ini untuk memprediksi jenis kelamin [10].

Antropologi forensik berkembang untuk meningkatkan identifikasi parameter penting secara keseluruhan dengan memanfaatkan kemajuan ilmu dan teknologi melalui pembelajaran mesin (*Machine Learning*) yang dapat mengotomatiskannya [11]. *Machine Learning* adalah subdisiplin ilmu komputer yang memiliki kemampuan untuk belajar dari data dan melakukan prediksi untuk masa depan berdasarkan data yang diberikan [12]. Penerapan *machine learning* dalam antropologi forensik mampu menghindari kesalahan, memeriksa dan memproses data secara lebih efisien dan memberikan informasi yang lebih terperinci yang dapat mendukung

pengambilan keputusan medis oleh praktisi kesehatan [13]. Selain itu, penerapan *machine learning* dalam antropologi forensik memberikan keunggulan dengan kemampuan untuk memproses dan menganalisis banyak sampel menggunakan berbagai metode, yang sering kali menghasilkan klasifikasi yang 100% akurat [14].

Pendekatan *machine learning* telah digunakan dalam mengidentifikasi jenis kelamin dari sisa-sisa kerangka, seperti yang ditunjukkan dalam penelitian klasifikasi jenis kelamin menggunakan tulang tengkorak dengan metode LVQ1 (*Learning Vector Quantization 1*), LVQ2 (*Learning Vector Quantization 2*), dan LVQ3 (*Learning Vector Quantization 3*). Hasil menunjukkan metode LVQ1 mencapai akurasi tertinggi yaitu 91.39% dengan pengaturan *learning rate* 0.1, 0.4, 0.7, dan 0.9. Metode LVQ2 mencapai akurasi tertinggi yaitu 77.05% dengan *learning rate* 0.9 dan *window* 0.2. Sementara metode LVQ3 mencapai akurasi tertinggi yaitu 80.04% dengan pengaturan *learning rate* 0.7, *window* 0.1, dan *epsilon* 0.3. Hasil yang didapatkan menunjukkan pendekatan *machine learning* dengan metode LVQ1 mampu memberikan akurasi prediksi tertinggi dalam identifikasi jenis kelamin dari data tulang tengkorak. Sehingga pada implemtasi dari ketiga metode tersebut, metode LVQ1 merupakan metode yang mencapai tingkat akurasi terbaik dibanding dengan metode LVQ2 dan LVQ3 [15]. Penelitian lanjutannya adalah menentukan jenis kelamin dari data tulang tengkorak dengan menggunakan metode LVQ1 dengan melakukan seleksi fitur menggunakan *Gain Rasio*. Sehingga menghasilkan akurasi tertinggi dengan *threshold* 0.01 dengan *learning rate* 0.1 sebesar 92.01% [16]. Selain itu penelitian lanjutan lainnya menggunakan metode *Corelation Based Feature Selection* pada LVQ2 menghasilkan akurasi tertinggi yaitu 82,51% dengan pengaturan *learning rate* 0.9 dan *window* 0.3 [17].

Penelitian lainnya yang telah dilakukan dengan menggunakan metode *Support Vector Machine* dalam mengklasifikasikan tulang sacrum berdasarkan jenis kelamin memperoleh hasil bahwa kernel polynomial mencapai tingkat akurasi tertinggi, yakni 85,56% [18]. Pada penelitian lainnya, dalam pembuatan model estimasi jenis kelamin obyektif berbasis tulang tengkorak pada populasi Cina dengan menggunakan transformasi *wavelet* dan transformasi *fourier* dan menggunakan fitur yang diekstraksi diklasifikasikan oleh *Support Vector Machine*. Hasil menunjukkan tingkat akurasi diskriminasi jenis kelamin pria sebesar 90.9% dan wanita sebesar 94.4% [19].

Support Vector Machine menjadi pilihan yang konsisten dalam berbagai studi klasifikasi karena dikenal memberikan tingkat akurasi yang tinggi, kemampuan generalisasi yang kuat, dan memiliki stabilitas yang tinggi [20]. Karena memiliki dasar teori yang baik dan memiliki kapasitas generalisasi yang baik, dalam beberapa tahun terakhir, *Support Vector Machine* telah menjadi salah satu metode klasifikasi yang paling banyak digunakan [21]. Metode *machine learning* seperti *Support Vector Machine* memiliki tingkat akurasi prediksi yang tinggi dan bias yang rendah dalam menentukan jenis kelamin dari kerangka dalam antropologi forensik [12].

Penelitian ini menggunakan metode diskriminasi pengukuran dan menerapkan metode *Support Vector Machine* dalam klasifikasi tulang tengkorak berdasarkan jenis kelamin dalam antropologi forensik. Penelitian ini bertujuan untuk menguji kinerja metode *Support Vector Machine* berdasarkan parameter uji dari setiap kernel. Dengan melakukan penerapan bidang keilmuan data mining dalam pengklasifikasian data dengan menggunakan metode *Support Vector Machine* diharapkan dapat memberikan kepada pembaca agar dapat memahami langkah-langkah dalam proses klasifikasi dengan metode *Sipport Vector Machine*.

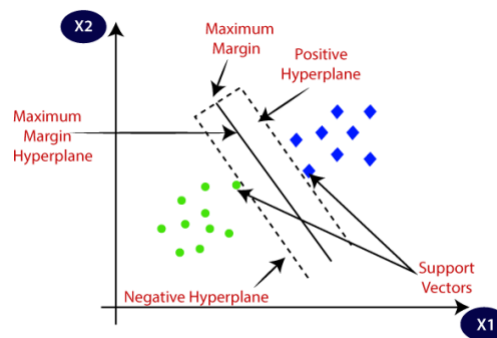
II. SIGNIFIKANSI STUDI

A. *Support Vector Machine*

Algoritma *Support Vector Machine* pertama Kali diperkenalkan pada tahun 1992 di *Annual Workshop on Computational Learning Theory* oleh Vapnik bersama Isabelle Guyon dan

Bernhard Boser. *Support Vector Machine* merupakan salah satu metode klasifikasi yang paling banyak diimplementasikan dalam berbagai studi kasus klasifikasi dua kelas dalam teknik *Machine Learning*. *Support Vector Machine* bertujuan untuk menemukan *hyperplane* yaitu pemisah linier yang bisa memaksimalkan jarak ke individu terdekat dari masing-masing dua kelas yang disebut margin [3]. Algoritma *Support Vector Machine* memiliki aplikasi yang baik untuk memisahkan dua kelas dan memberikan kinerja klasifikasi yang sangat baik [22].

Support Vector Machine adalah salah satu algoritma dalam supervised learning yang digunakan dalam klasifikasi dan regresi. Salah satu keunggulan dari *Support Vector Machine* (SVM) adalah kemampuannya dalam menentukan jarak dengan menggunakan vektor dukungan, yang mengakibatkan proses komputasi menjadi lebih efisien. Selain itu *Support Vector Machine* mempunyai tingkat yang rendah untuk kesalahan generasi dan mudah untuk interpretasi. Dalam penerapan *Support Vector Machine*, pemilihan kernel yang tepat sangat penting untuk mencapai hasil yang optimal. Kernel digunakan untuk mentransformasikan data ke ruang dimensi yang lebih tinggi agar lebih mudah memisahkan pola atau struktur data. Hal ini bertujuan untuk meningkatkan pemisahan antara kelas data sehingga model dapat melakukan klasifikasi dengan lebih baik [23]. *Support Vector Machine* bekerja dengan mencari *hyperplane* yang memisahkan dua kelas data dengan jarak maksimum, sehingga semua titik dari satu kelas berada di satu sisi *hyperplane* dan semua titik dari kelas lainnya berada disisi yang berlawanan [24].



Gambar 1. Ilustrasi *Support Vector Machine*

Ada dua jenis variabel utama dalam dataset. Satu variabel x_i mewakili data, dan variabel lainnya y_i mewakili kelas setiap data dalam dataset. *Support Vector Machine* digunakan untuk melakukan proses memisahkan dua kelas dalam suatu *dataset* dengan menggunakan *hyperplane* yaitu bidang pemisah ruang berdimensi tinggi. Akurasi model hasil penerapan *Support Vector Machine* sangat tergantung pada jenis fungsi kernel yang dipilih [25].

Metode klasifikasi *Support Vector Machine* memiliki parameter kernel diantaranya Linear, RBF, Sigmoid dan Polinomial. Pada penelitian ini penggunaan kernel Linear, RBF, dan polynomial dalam klasifikasi *Support Vector Machine* untuk meningkatkan keterpisahan data dan mencapai akurasi yang lebih tinggi dalam proses klasifikasi [26]. Selain itu, kernel Linear, RBF dan Polynomial banyak digunakan dalam penelitian klasifikasi *Support Vector Machine* karena memiliki performa yang baik dalam pengenalan pola dan kemampuan interpretasi yang kuat, terutama saat digunakan dengan kernel non-linear [27]. Kernel RBF mampu memberikan performa yang konsisten untuk masalah-masalah Dimana komputasi dapat dilakukan secara numerik, sementara kernel polynomial mampu melakukan komputasi lebih cepat dan mengurangi resiko *overfitting* [28]. Pada dasarnya, tujuan utama *Support Vector Machine* adalah mempermudah pengklasifikasian data dengan menemukan *hyperplane* yang mampu memisahkan dataset secara linear dengan hasil tingkat akurasi yang tinggi.

Penelitian ini menggunakan LibSVM sebagai klasifikasi tulang tengkorak. LibSVM adalah sebuah perpustakaan (*library*) berbasis kernel yang menyediakan implementasi algoritma *Support Vector Machine* dalam berbagai bahasa pemrograman, yaitu *Python*. Penggunaan

LibSVM dalam klasifikasi *Support Vector Machine* untuk mempercepat pengembangan penelitian dan menghemat waktu karena tidak perlu menulis kode dari awal [29]. Selain itu, LibSVM digunakan untuk membantu meningkatkan akurasi pengklasifikasi dengan memilih sampel pelatihan yang tepat dan mengoptimalkan parameter.

Berikut variasi fungsi kernel yang digunakan diuraikan pada Tabel I [30].

TABEL I
FUNGSI KERNEL

Nama Kernel	Fungsi Kernel
Linear	$K(x_i, x) = X_i^T x$
RBF	$K(x, y) = \exp \left\{ -\frac{ x - y ^2}{2\gamma^2} \right\}$
Polynomial	$K(x_i, x) = (\gamma(x_i^T x) + r)$

B. Dataset

Data yang dipergunakan dalam studi ini adalah kumpulan data yang terdiri dari serangkaian informasi untuk klasifikasi tulang tengkorak berdasarkan jenis kelamin. Dataset tulang tengkorak ini terdiri dari pengukuran yang telah dikumpulkan oleh Dr. William W Howells antara tahun 1965 dan 1980. Dataset ini terdiri dari pengukuran kranimetri yang diambil dari 2.524 data tengkorak manusia yang terdiri dari 1368 data tengkorak laki-laki dan 1156 merupakan data tulang tengkorak perempuan. Data sekunder pada studi ini di peroleh dari laman website <https://web.utk.edu/~auerbach/HOWL.htm>. Dataset tengkorak yang digunakan dalam penelitian ini sebanyak 82 fitur yang digunakan sebagai metrik atau ukuran untuk menentukan jenis kelamin. Pada penelitian ini, terdapat dua label klasifikasi yaitu *Male* dan *Female*.

Parameter yang dikodekan untuk setiap fitur tengkorak dan pengukuran tulang tengkorak pada masing-masing variabel akan dijelaskan pada Tabel II dan III.

TABEL II
FITUR PARAMETER TENGGORAK

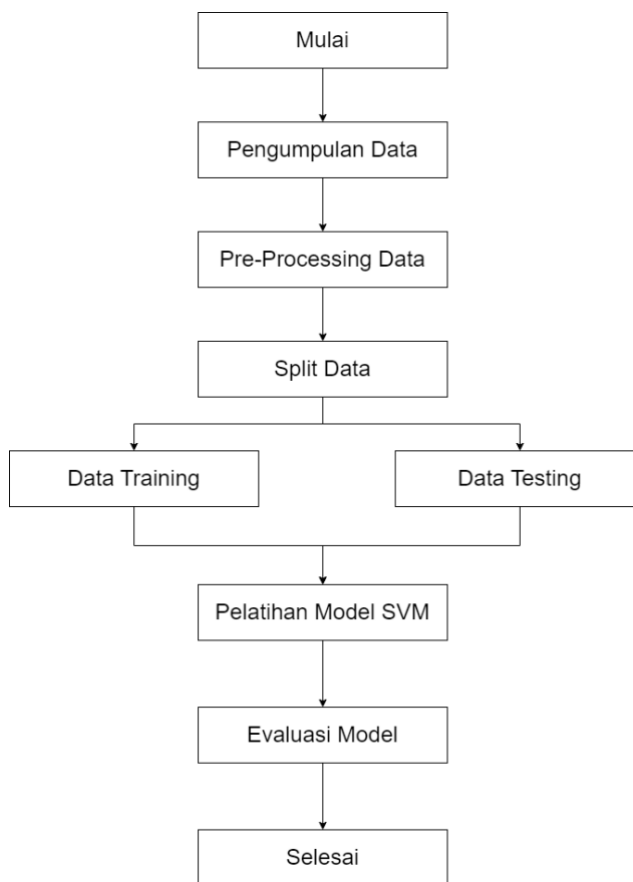
Kode	Fitur Parameter Tulang Tengkorak
AUB	<i>Biauricular breadth</i>
PAS	<i>Parietal Subtense</i>
FRS	<i>Frontal Subtense</i>
PAC	<i>Parietal Chord</i>
NPH	<i>Nasion-prosthion height</i>
PAF	<i>Parietal fraction</i>
BNL	<i>Basion-nasion length</i>
NLH	<i>Nasal height</i>
ZMB	<i>Bimaxillary breadth</i>
BPL	<i>Basion-prosthion length</i>
...	...
TBA	<i>Thiobarbituric acid</i>

TABEL III
PENGUKURAN TULANG TENGGORAK

CLASS	GOL	ZYB	AUB	RFA	OCA	...	TBA
M	189	133	119	0	117	...	0
M	182	137	125	0	119	...	0
M	191	134	125	0	111	...	0
M	178	129	121	0	111	...	0
F	159	117	106	61	132	...	154
F	156	113	102	63	133	...	151
...	
F	160	117	112	60	131	...	156

C. Metode Penelitian

Penelitian ini menggunakan metode *Support Vector Machine* dengan beberapa jenis kernel, yaitu linear, RBF, dan polynomial untuk memprediksi tulang tengkorak berdasarkan jenis kelamin. Dua label kategori yang digunakan adalah "Male" dan "Female". Fokus utama penelitian ini adalah membuat model *Support Vector Machine* yang dapat mengklasifikasikan tulang tengkorak dengan hasil akurasi tinggi. Adapun tahapan pelatihan model yang diterapkan dalam penelitian ini mencakup beberapa tahapan sebagai berikut:



Gambar 2. Tahap Pelatihan Model *Support Vector Machine*

Tahapan pelatihan model *Support Vector Machine* dilakukan dengan langkah pertama adalah melakukan pengumpulan dataset yang akan digunakan pada penelitian. Data yang telah diperoleh dilakukan proses *pre-processing* data meliputi pembersihan data dari nilai yang hilang (*missing values*), menghapus kolom yang tidak diperlukan dalam proses membangun model pembelajaran, dan menghilangkan *noise*. Pada penelitian ini terdapat dua kolom fitur yang dihapus yaitu fitur 'Popnum', dan 'Population' karena hanya memberikan informasi tentang nama populasi dan jumlah populasi saja.

Pada tahapan *pre-processing* data, dilakukan juga proses normalisasi. Normalisasi data hanya dengan satu metode dapat mengakibatkan hasil kinerja yang kurang optimal. Jika dataset pada penelitian yang digunakan berukuran besar dengan memiliki banyak fitur dapat menimbulkan fitur dominan yang berakibat menurunkan kinerja klasifikasi algoritma pembelajaran. Fitur dengan nilai numerik lebih besar mendominasi fitur dengan nilai fitur numerik lebih kecil sekaligus membedakan pola dari data [31]. Metode penskalaan mencoba membuat fitur-fitur sebanding. Pada penelitian ini menggunakan Fitur *Scaling* untuk mengubah setiap fitur sehingga memiliki rata-rata nol dan simpangan baku satu sehingga 82 fitur dalam yang digunakan memiliki kontribusi yang setara dengan. Hal ini dilakukan dengan mengurangi rata-rata dari setiap fitur kemudian membaginya dengan simpangan baku dari fitur tersebut. Fitur *Scaling* cenderung lebih tahan terhadap outlier dan lebih baik menangani fitur-fitur yang memiliki distribusi tidak normal[32].

Tahapan selanjutnya dataset dipisahkan menjadi data *training* atau data latih dan data *testing* atau data uji dengan perbandingan 90:10. Pemisahan ini dilakukan sebab metode yang digunakan merupakan *supervised learning* sehingga data terlebih dahulu diajarkan atau melalui proses pembelajaran untuk mengenali pola dalam data input sehingga model klasifikasi yang dihasilkan dapat digunakan untuk prediksi dengan hasil akurat dan dapat dikembangkan. Sehingga pada penelitian ini terbentuk data *training* atau data latih sebanyak 2.271 data yang digunakan untuk melatih model dan 253 data testing yang digunakan untuk menguji performa model yang telah dilatih. Setelah membagi data, dilakukan pelatihan menggunakan tiga model kernel *Support Vector Machine* yaitu linear, rbf dan polynomial.

III. HASIL DAN PEMBAHASAN

Percobaan dilakukan dengan menggunakan 2524 tulang tengkorak dengan 82 variabel yang telah diimplementasikan menggunakan pemrograman *Python* menggunakan tools dari *Google Colab*. 82 variabel ini merupakan hasil dari proses *cleanning* data yang dilakukan pada tahap *preprocessing* data. Eksperimen dilakukan dengan menggunakan tiga kernel *Support Vector Machine* yaitu linear, rbf dan polynomial. Dari ketiganya, model kernel *Support Vector Machine* yang dapat mengevaluasi model dengan baik, dapat dilihat dari hasil akurasi menggunakan *Confusion Matrix*. Adapun hasil akurasi yang diperoleh dari kernel *Support Vector Machine* yang berbeda dapat dilihat pada Tabel

Kernel *Support Vector Machine* linear adalah jenis fungsi kernel yang efektif saat data dapat dibagi secara linier. Kernel ini ideal digunakan ketika data sudah secara alami terpisah secara linear. Kernel linear sangat sesuai untuk data dengan dimensi fitur yang tinggi karena melakukan pemetaan ke dimensi ruang yang lebih tinggi tidak memberikan peningkatan kinerja yang signifikan. Parameter C adalah parameter penalti untuk kesalahan klasifikasi. C mengontrol keseimbangan antara kehalusan batas keputusan dan klasifikasi yang tepat terhadap titik-titik latihan.. Parameter yang digunakan adalah C= 1, 2, dan 3. Nilai akurasi yang diperoleh dapat dilihat pada Tabel IV.

TABEL IV
PERBANDINGAN AKURASI (%) KERNEL LINEAR

Linear	Akurasi
C = 1	86,96
C = 2	88,14
C = 3	87,35

Kernel RBF adalah fungsi yang populer karena dapat digunakan jika data tidak terpisah secara linear. RBF memiliki parameter *cost* (C) dan *gamma* (γ). *Gamma* parameter menentukan seberapa jauh jangkauan pengaruh suatu contoh pelatihan, dengan nilai rendah berarti 'jauh' dan nilai tinggi berarti 'dekat'. Jika *gamma* ukurannya sangat kecil, modelnya terlalu terbatas dan tidak dapat menangkap kompleksitas atau bentuk datanya. Sedangkan jika *gamma* terlalu besar, radius pengaruh vektor tumpuan hanya mencakup vektor tumpuan itu sendiri dan tidak ada jumlah regularisasi nilai *cost* dapat mencegah *overfitting* pada data pelatihan .

Parameter γ ='auto' menyesuaikan nilai *gamma* secara otomatis berdasarkan skala data. Ini berarti bahwa nilai *gamma* akan disesuaikan secara dinamis berdasarkan jumlah fitur dalam dataset dan variansinya. Ini memungkinkan *Support Vector Machine* untuk menyesuaikan dengan baik terhadap berbagai skala dan kompleksitas data. Koefisien *gamma* diatur ke 'auto', yang berarti hasil bagi dari satu dan jumlah (n) fitur [33]. Hasil akurasi dari kernel RBF *Support Vector Machine* dengan *gamma* yang disetel ke 'auto' dapat lebih tinggi dibandingkan dengan *gamma* = 1,2, 3 tergantung pada *dataset* tertentu dan karakteristiknya. Nilai 'auto' untuk *gamma* adalah faktor penskalaan yang secara otomatis ditentukan oleh scikit-learn berdasarkan jumlah fitur dalam data. Ini setara dengan pengaturan $\gamma = 1 / (n_features * X.var())$. Parameter yang digunakan adalah $C=1, 2, \text{ dan } 3$ $\gamma = 1, 2, \text{ dan } 3$. Nilai akurasi yang diperoleh dapat dilihat pada Tabel V dan VI.

TABEL V
PERBANDINGAN AKURASI (%) KERNEL RBF

Cost	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$
C = 1	54,15	54,15	54,15
C = 2	54,15	54,15	54,15
C = 3	54,15	54,15	54,15

TABEL VV
PERBANDINGAN AKURASI (%) KERNEL RBF

Cost	γ ='auto'
C = 1	90,51
C = 2	91,30
C = 3	90,51

Kernel polynomial adalah jenis fungsi kernel yang sangat berguna ketika kita memiliki dataset yang cukup besar, karena kernel ini memetakan vektor sampel pelatihan ke dalam ruang fitur dengan cara yang mempertahankan kesamaan antar vektor. Oleh karena itu, kernel polynomial seringkali merupakan pilihan yang tepat untuk menyelesaikan masalah klasifikasi pada dataset yang telah dinormalisasi. Kernel polynomial pada dasarnya memiliki pendekatan

yang mirip dengan kernel linear. Didalam kernel polynomial tidak hanya bergantung pada satu fitur tertentu dari sampel input, tetapi juga pada kombinasi dalam menentukan kesamaan [34]. Polynomial memiliki parameter *cost* (C), *gamma* (γ), dan *degree* (d). *Degree* atau derajat polynomial yang digunakan untuk menemukan hyperplane yang dapat memisahkan data dengan baik dalam konteks fungsi kernel polinomial pada *Support Vector Machine*. Nilai akurasi yang diperoleh dapat dilihat pada Tabel VII, VIII dan IX.

TABEL VII
PERBANDINGAN AKURASI (%) KERNEL POLYNOMIAL

$\gamma = 1$	$d = 1$	$d = 2$	$d = 3$
$C = 1$	86,96	64,43	84,98
$C = 2$	88,14	64,43	84,98
$C = 3$	87,35	64,43	84,98

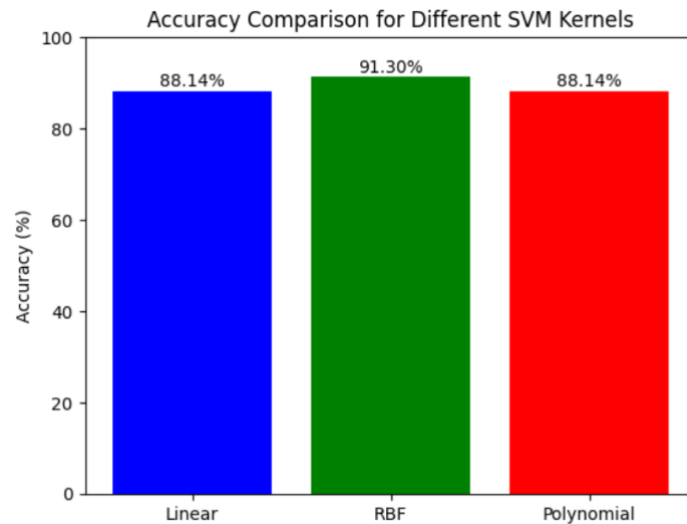
TABEL VIII
PERBANDINGAN AKURASI (%) KERNEL POLYNOMIAL

$\gamma = 1$	$d = 1$	$d = 2$	$d = 3$
$C = 1$	88,14	64,43	84,98
$C = 2$	87,35	64,43	84,98
$C = 3$	87,35	64,43	84,98

TABEL IX
PERBANDINGAN AKURASI (%) KERNEL POLYNOMIAL

$\gamma = 1$	$d = 1$	$d = 2$	$d = 3$
$C = 1$	87,35	64,43	84,98
$C = 2$	87,35	64,43	84,98
$C = 3$	86,16	64,43	84,98

Berkaitan dengan hasil eksperimen yang telah dilakukan, akurasi terbaik yang didapatkan pada masing-masing kernel yaitu kernel linear memperoleh akurasi terbaik sebesar 88,14% dengan $C = 2$. Untuk kernel RBF akurasi terbaik sebesar 91,30% pada $C = 2$, $\gamma = \text{'auto'}$. Untuk kernel polynomial akurasi terbaik sebesar 88,14% pada $C = 1$ dan 2 , $\gamma = 1$ dan 2 , $d = 1$. Gambar 3 berisi akurasi keseluruhan yang diperoleh untuk model yang diusulkan dengan fungsi kernel yang berbeda untuk *Support Vector Machine*.



Gambar 3. Perbandingan Hasil Akurasi Fungsi Kernel *Support Vector Machine*

Setelah model dengan parameter terbaik didapatkan kemudian dilakukan evaluasi dengan cara menghitung nilai *accuracy*, *precision*, dan *recall* pada ketiga kernel yang hasilnya dapat dilihat pada Tabel X.

TABEL X
EVALUASI MODEL KLASIFIKASI (%)

Kernel	Accuracy	Precision	Recall
Linear	86,96	86,97	86,96
RBF	90,51	90,64	90,51
Polynomial	86,56	86,7	86,56

Pada hasil evaluasi yang telah didapatkan, ketiga kernel *Support Vector Machine* yaitu linear, rbf, dan polynomial berhasil memisahkan data dengan sempurna pada dataset yang digunakan pada penelitian ini. Terbukti dengan capaian akurasi diatas 85% untuk *accuracy*, *precision* dan *recall*. Hal ini menandakan bahwa ketiga kernel tersebut menunjukkan kualitas yang sangat baik dalam tugas klasifikasi tulang tengkorak berdasarkan jenis kelamin.

Pada Tabel XI. Menunjukkan hasil kinerja yang diperoleh melalui *cross-validation*. Skor validasi silang adalah nilai rata-rata akurasi untuk setiap kernel yang menggunakan 10 lipatan.

TABEL XI
KINERJA MENGGUNAKAN CROSS VALIDATION (%)

Kernel Linear	Kernel RBF	Kernel Polynomial
88,08	97,98	88,03

Tabel diatas mengindikasikan bahwa kernel Linear, RBF dan Polynomial menunjukkan performa yang sangat baik. Hal ini disebabkan kemampuan dalam memproyeksi titik-titik kedalam ruang dimensi yang lebih tinggi, sehingga meningkatkan keterpisahan data. Pada kernel RBF yang diusulkan dengan nilai *gamma* otomatis memberikan kinerja yang lebih baik. Dengan nilai *gamma* yang otomatis, kernel RBF melakukan penyesuaian parameter secara otomatis untuk pola nonlinier, yang meningkatkan akurasi dan stabilitas hasilnya.

IV. KESIMPULAN

Pendekatan klasifikasi telah banyak digunakan di berbagai bidang, termasuk dalam antropologi forensik. Penelitian ini menggunakan tulang tengkorak untuk klasifikasi jenis kelamin. Eksperimen yang dilakukan dengan menggunakan RBF, polynomial, dan linear menggunakan tiga fungsi yang berbeda. Dari penyajian data dan proses yang telah dijalankan, hasil percobaan menunjukkan akurasi terbaik pada setiap fungsi kernel. Kernel linear memperoleh akurasi terbaik sebesar 88,14% dengan $C = 2$. Untuk kernel RBF akurasi terbaik sebesar 91,30% pada $C = 2$, $\gamma = \text{'auto'}$. Untuk kernel polynomial akurasi terbaik sebesar 88,14% pada $C = 1$ dan 2 , $\gamma = 1$ dan 2 , $d = 1$. Dengan hasil 88,14% yang ditemukan, penelitian ini memiliki akurasi yang tinggi lebih dari 80%.

Berdasarkan hasil evaluasi model menggunakan *confusion matrix*, kernel RBF menghasilkan nilai yang tertinggi yaitu 90% dibandingkan kernel linear dan polynomial untuk *accuracy*, *precision*, dan *recall*. Sementara untuk kernel linear dan polynomial berada pada nilai 86%. Selain Kernel RBF juga menunjukkan performa yang unggul dalam cross-validation dengan nilai rata-rata akurasi sebesar 97,98%. Hasil ini menandakan bahwa model secara konsisten mampu melakukan pemisahan yang baik antara kelas-kelas yang ada dalam dataset. Pada penelitian ini, kernel RBF sebagai pilihan terbaik untuk diimplementasikan pada klasifikasi tulang tengkorak berdasarkan jenis kelamin. Hal ini disebabkan penggunaan gamma dengan nilai default tetap dan nilai *gamma* yang di set 'auto'. Nilai *gamma* otomatis menghasilkan akurasi prediksi yang lebih tinggi dan generalisasi yang lebih baik dalam model klasifikasi tulang tengkorak berdasarkan jenis kelamin dalam antropologi forensik. Hal ini menjadikan *Support Vector Machine* sangat direkomendasikan dalam klasifikasi tulang tengkorak.

Dalam penelian di masa depan, klasifikasi jenis kelamin dari tulang tengkorak dapat ditingkat dengan menerapkan teknik lain yang spesifik untuk meningkatkan akurasi dan kinerja analisis teknik klasifikasi seperti melakukan optimalisasi parameter kernel *Support Vector Machine* parameter kernel menggunakan *Particle Swarm Optimization*.

REFERENSI

- [1] N. V. P. E. J. B. Angi M. Christensen, *Forensic Anthropology: Current Methods and Practice*, Second Edition. London, United Kingdom: Stacy Masucci, 2019.
- [2] J. Bewes, A. Low, A. Morphett, F. D. Pate, and M. Henneberg, "Artificial intelligence for sex determination of skeletal remains: Application of a deep learning artificial neural network to human skulls," *J Forensic Leg Med*, vol. 62, pp. 40–43, Feb. 2019, doi: 10.1016/j.jflm.2019.01.004.
- [3] D. Nasien, M. H. Adiya, I. Afrianty, N. A. Ali, A. A. Samah, and Y. Rahayu, "Determination of Sex and Race in Forensic Anthropology: A Comparison of Artificial Neural Network and Support Vector Machine," in *2021 4th International Conference of Computer and Informatics Engineering (IC2IE)*, IEEE, Sep. 2021, pp. 51–55. doi: 10.1109/IC2IE53219.2021.9649182.
- [4] B. Ramamoorthy, M. M. Pai, L. V. Prabhu, B. V. Muralimanju, and R. Rai, "Assessment of craniometric traits in South Indian dry skulls for sex determination," *J Forensic Leg Med*, vol. 37, pp. 8–14, Jan. 2016, doi: 10.1016/j.jflm.2015.10.001.
- [5] J. Tetteh, A. K. Appiah, C. S. Abaidoo, and C. Adjei-Antwi, "The forensic use of percutaneous femur length in height and sex estimation among Ghanaians," *Forensic Science International: Reports*, vol. 4, p. 100234, Nov. 2021, doi: 10.1016/j.fsir.2021.100234.
- [6] İ. Güven and F. Şimşir, "Demand forecasting with color parameter in retail apparel industry using artificial neural networks (ANN) and support vector machines (SVM)

- methods,” *Comput Ind Eng*, vol. 147, p. 106678, Sep. 2020, doi: 10.1016/j.cie.2020.106678.
- [7] Arthy, Rohit Goel, and Sreenivas M, “Determination of Sex by Osteometry of Third Metatarsal,” *Indian Journal of Forensic Medicine & Toxicology*, vol. 14, no. Vol. 14 No. 3 (2020): Indian Journal of Forensic Medicine & Toxicology, Jul. 2020, doi: 10.37506/ijfmt.v14i3.10315.
- [8] P. Mesejo, R. Martos, Ó. Ibáñez, J. Novo, and M. Ortega, “A Survey on Artificial Intelligence Techniques for Biomedical Image Analysis in Skeleton-Based Forensic Human Identification,” *Applied Sciences*, vol. 10, no. 14, p. 4703, Jul. 2020, doi: 10.3390/app10144703.
- [9] P. Dixit and G. I. Prajapati, “Machine Learning in Bioinformatics: A Novel Approach for DNA Sequencing,” in *2015 Fifth International Conference on Advanced Computing & Communication Technologies*, IEEE, Feb. 2015, pp. 41–47. doi: 10.1109/ACCT.2015.73.
- [10] W. Yang, M. Zhou, P. Zhang, G. Geng, X. Liu, and H. Zhang, “Skull Sex Estimation Based on Wavelet Transform and Fourier Transform,” *Biomed Res Int*, vol. 2020, pp. 1–10, Jan. 2020, doi: 10.1155/2020/8608209.
- [11] A. L. Dallora, P. Anderberg, O. Kvist, E. Mendes, S. Diaz Ruiz, and J. Sanmartin Berglund, “Bone age assessment with various machine learning techniques: A systematic literature review and meta-analysis,” *PLoS One*, vol. 14, no. 7, p. e0220242, Jul. 2019, doi: 10.1371/journal.pone.0220242.
- [12] E. Nikita and P. Nikitas, “On the use of machine learning algorithms in forensic anthropology,” *Leg Med*, vol. 47, p. 101771, Nov. 2020, doi: 10.1016/j.legalmed.2020.101771.
- [13] N. Liu, X. Li, E. Qi, M. Xu, L. Li, and B. Gao, “A Novel Ensemble Learning Paradigm for Medical Diagnosis With Imbalanced Data,” *IEEE Access*, vol. 8, pp. 171263–171280, 2020, doi: 10.1109/ACCESS.2020.3014362.
- [14] S. D. Ousley, “Forensic Classification and Biodistance in the 21st Century,” in *Biological Distance Analysis*, Elsevier, 2016, pp. 197–212. doi: 10.1016/B978-0-12-801966-5.00010-X.
- [15] DARMILA, IIS AFRIANTY, SUWANTO SANJAYA, RAHMAD ABDILLAH, IWAN ISKANDAR, and FADHILAH SYAFRIA, “EVALUASI PERBANDINGAN PERFORMANSI LVQ 1, LVQ 2, DAN LVQ 3 DALAM KLASIFIKASI JENIS KELAMIN MENGGUNAKAN TULANG TENGGORAK,” *Jurnal INSTEK (Informatika Sains dan Teknologi)*, vol. 7, no. 2, pp. 344–353, Oct. 2022, doi: 10.24252/instek.v7i2.32659.
- [16] Y. Harni, I. Afrianty, S. Sanjaya, R. Abdillah, F. Yanto, and F. Syafria, “Performance Analysis of LVQ 1 Using Feature Selection Gain Ratio for Sex Classification in Forensic Anthropology,” *Building of Informatics, Technology and Science (BITS)*, vol. 5, no. 1, Jun. 2023, doi: 10.47065/bits.v5i1.3625.
- [17] Surya Aditya GD, Iis Afrianty, and Suwanto Sanjaya, “Perbandingan Performansi Dengan Metode Corelation Based Feature Selection Pada LVQ 2,” *Jurnal INSTEK (Informatika Sains dan Teknologi)*, vol. 8, no. <https://journal.uin-alauddin.ac.id/index.php/instek/issue/view/2039>, pp. 170–179, Apr. 2023.
- [18] I. Afrianty, D. Nasien, and H. Haron, “Performance Analysis of Support Vector Machine in Sex Classification of The Sacrum Bone in Forensic Anthropology,” *JURNAL TEKNIK INFORMATIKA*, vol. 15, no. 1, pp. 63–72, Jun. 2022, doi: 10.15408/jti.v15i1.25254.
- [19] W. Yang, M. Zhou, P. Zhang, G. Geng, X. Liu, and H. Zhang, “Skull Sex Estimation Based on Wavelet Transform and Fourier Transform,” *Biomed Res Int*, vol. 2020, pp. 1–10, Jan. 2020, doi: 10.1155/2020/8608209.

- [20] O. A. Alimi, K. Ouahada, A. M. Abu-Mahfouz, and S. Rimer, "Power system events classification using genetic algorithm based feature weighting technique for support vector machine," *Heliyon*, vol. 7, no. 1, p. e05936, Jan. 2021, doi: 10.1016/j.heliyon.2021.e05936.
- [21] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, Sep. 2020, doi: 10.1016/j.neucom.2019.10.118.
- [22] I. Afrianty, D. Nasien, M. R. A. Kadir, and H. Haron, "Back-Propagation Neural Network for Gender Determination in Forensic Anthropology," 2015, pp. 255–281. doi: 10.1007/978-3-319-11017-2_11.
- [23] N. Rachmalia Feta, A. Rahmat Ginanjar, I. Teknologi dan Bisnis Bank Rakyat Indonesia, J. R. Harsono No, P. Minggu, and J. Selatan, "KOMPARASI FUNGSI KERNEL METODE SUPPORT VECTOR MACHINE UNTUK PEMODELAN KLASIFIKASI TERHADAP PENYAKIT TANAMAN KEDELAI COMPARISON OF THE KERNEL FUNCTION OF SUPPORT VECTOR MACHINE METHOD FOR MODELING CLASSIFICATION OF SOYBEAN PLAT DISEASE Program Studi Sistem Informasi 2)," *Sains dan Teknologi Terapan*, vol. 1, no. 1, 2019.
- [24] W. Budiharto, *Machine Learning & Computational Intelligence / Widodo Budiharto*. Yogyakarta: Penerbit Andi, 2016.
- [25] I. M. Parapat, M. T. Furqon, and S. Sutrisno, "Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak," *Jurnal Pengembangan Teknologi Informasi Ilmu Komputer*, vol. 2, no. 10, pp. 3163–3169, Oct. 2018.
- [26] S. V. S. Prasad, T. S. Savithri, and I. V. M. Krishna, "PERFORMANCE EVALUATION OF SVM KERNELS ON MULTISPECTRAL LISS III DATA FOR OBJECT CLASSIFICATION," *International Journal on Smart Sensing and Intelligent Systems*, vol. 10, no. 4, pp. 1–16, Jan. 2017, doi: 10.21307/ijssis-2018-022.
- [27] V. Van Belle, B. Van Calster, S. Van Huffel, J. A. K. Suykens, and P. Lisboa, "Explaining Support Vector Machines: A Color Based Nomogram," *PLoS One*, vol. 11, no. 10, p. e0164568, Oct. 2016, doi: 10.1371/journal.pone.0164568.
- [28] R. Izmailov, V. Vapnik, and A. Vashist, "Multidimensional splines with infinite number of knots as SVM kernels," in *The 2013 International Joint Conference on Neural Networks (IJCNN)*, IEEE, Aug. 2013, pp. 1–7. doi: 10.1109/IJCNN.2013.6706860.
- [29] A. Abdiansah and R. Wardoyo, "Time Complexity Analysis of Support Vector Machines (SVM) in LibSVM," *Int J Comput Appl*, vol. 128, no. 3, pp. 28–34, Oct. 2015, doi: 10.5120/ijca2015906480.
- [30] A. Z. Praghakusma and N. Charibaldi, "Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi)," *JSTIE (Jurnal Sarjana Teknik Informatika) (E-Journal)*, vol. 9, no. 2, p. 88, Jun. 2021, doi: 10.12928/jstie.v9i2.20181.
- [31] D. Singh and B. Singh, "Feature wise normalization: An effective way of normalizing data," *Pattern Recognit*, vol. 122, p. 108307, Feb. 2022, doi: 10.1016/j.patcog.2021.108307.
- [32] A. Tedyyana, O. Ghazali, and O. W. Purbo, "Machine learning for network defense: automated DDoS detection with telegram notification integration," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, no. 2, p. 1102, May 2024, doi: 10.11591/ijeecs.v34.i2.pp1102-1109.
- [33] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY: Springer New York, 2000. doi: 10.1007/978-1-4757-3264-1.

- [34] B. Yang, R. Gong, L. Wang, and S. Yang, "WITHDRAWN: Support vector machine in image recognition of nursing methods for critically ill blood purification," *Microprocess Microsyst*, p. 103398, Nov. 2020, doi: 10.1016/j.micpro.2020.103398.

UCAPAN TERIMA KASIH

Penelitian ini merupakan tugas akhir seorang mahasiswa di Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau. Penulis ingin menyampaikan terima kasih kepada orang tua penulis yang selalu mendoakan dan mendukung penulis dalam menyelesaikan penelitian ini. Penulis juga ingin menyampaikan terima kasih kepada Ibu Iis Afrianty, S.T., M.Sc., sebagai pembimbing yang telah memberikan banyak bantuan dalam penyusunan tugas akhir dan memberikan masukan berharga selama penelitian ini dilakukan.