

# Estimasi Keberhasilan Siswa dalam Pemodelan Data Berbasis Learning Menggunakan Algoritma Support Vector Machine

Suryani<sup>1,\*</sup>, Mustakim<sup>2</sup>

<sup>1,2</sup>Fakultas Sains dan Teknologi, Prodi Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

<sup>1,2</sup>Puzzle Research Data Technology (Predatech), Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: <sup>1,\*</sup>[12050320389@students.uin-suska.ac.id](mailto:12050320389@students.uin-suska.ac.id), <sup>2</sup>[mustakim@uin-suska.ac.id](mailto:mustakim@uin-suska.ac.id)

Email Penulis Korespondensi: [12050320389@students.uin-suska.ac.id](mailto:12050320389@students.uin-suska.ac.id)

## Abstrak

SMK Negeri 5 Pekanbaru bertujuan menyiapkan lulusan terampil dan mampu bersaing di pasar global. Perwujudan tujuan tersebut dipengaruhi oleh prestasi atau keberhasilan siswa dalam mengikuti pembelajaran di sekolah. Hasil belajar menentukan kemampuan dan kesiapan siswa terjun dan bersaing di dunia kerja. Berdasarkan hasil observasi, diketahui bahwa hasil belajar siswa di SMK Negeri 5 Pekanbaru cenderung rendah. Hal ini turut ditunjukkan oleh data yang berhasil dikumpulkan melalui bagian Kurikulum. Berdasarkan data yang dikumpulkan, dapat dilakukan ekstraksi menggunakan metode supervised learning sehingga menghasilkan model klasifikasi hasil belajar siswa. Algoritma supervised learning yang digunakan adalah Support Vector Machine (SVM). Algoritma SVM diterapkan pada data nilai siswa kelas X semester ganjil pada tahun 2020 yang jumlahnya 160 data. Proses klasifikasi dilakukan dengan penerapan metode GridSearch terlebih dahulu untuk menemukan kernel terbaik yang akan diimplementasikan. Berdasarkan penerapan GridSearch, kernel yang akan digunakan adalah Radial Basis Function (RBF) yang memiliki parameter Cost (C) dan Gamma ( $\gamma$ ). Berdasarkan 16 skenario percobaan dengan nilai input parameter berbeda, hasil klasifikasi terbaik didapatkan pada percobaan menggunakan nilai Cost (C) = 0.1 dan nilai Gamma ( $\gamma$ ) = 0.01 dengan nilai akurasi 94%.

**Kata Kunci:** Hasil Belajar Siswa; Klasifikasi; Radial Basis Function; Supervised Learning; Support Vector Machine

## Abstract

SMK Negeri 5 Pekanbaru aims to prepare competent graduates who can compete in the global market. The realization of these goals is influenced by student achievement at school. Student achievements determine the ability of students to work in certain fields. Based on observations, it is known that student achievement at SMK Negeri 5 Pekanbaru tend to be low. This is also shown by the data that has been collected through the Curriculum section. Based on the data, there can be extraction using the supervised learning method to make a classification model of student achievements. The supervised learning algorithm used in this research is a Support Vector Machine (SVM). The data used in this study are student's data grade X SMK Negeri 5 Pekanbaru in 2020 totaling 160 data. The classification process is carried out by applying the GridSearch method to find the best kernel to be implemented. Based on the implementation of GridSearch, the kernel to be used is Radial Basis Function (RBF) with Cost (C) and Gamma ( $\gamma$ ) parameters. Based on 16 experiments with different parameter values, the best classification results are obtained using the value of Cost (C) = 0.1 and the value of Gamma ( $\gamma$ ) = 0.01, with accuracy values of 94%.

**Keywords:** Classification; Students Achievement; Radial Basis Function; Supervised Learning; Support Vector Machine

## 1. PENDAHULUAN

Sekolah Menengah Kejuruan (SMK) adalah lembaga pendidikan formal yang memiliki tujuan menciptakan Sumber Daya Manusia (SDM) terampil dan memiliki keahlian sehingga siap terjun ke dunia kerja atau melanjutkan pendidikan ke jenjang lebih tinggi [1]. Berdasarkan data milik Dinas Pendidikan Provinsi Riau, SMK Negeri 5 Pekanbaru tercatat sebagai SMK Negeri bidang industri teknologi dan rekayasa yang beralamat di Jalan Yos Sudarso Kota Pekanbaru. Sebagai salah satu SMK Negeri tertua di kota Pekanbaru, SMK Negeri 5 Pekanbaru memiliki misi mempersiapkan tamatan yang disiplin, mandiri, kreatif, peduli, bertanggung jawab, profesional, berdaya saing global dan siap mengantisipasi perkembangan IPTEK dalam dunia kerja dengan berbekal iman, taqwa dan akhlak mulia [2]. Dalam upaya mewujudkan tujuan SMK sebagai lembaga pendidikan formal, dibutuhkan kerjasama antar manajemen sekolah, kemitraan dengan dunia kerja, kerjasama orang tua dan keinginan siswa itu sendiri [3]. Elemen ini saling terkait dalam menghasilkan lulusan SMK berprestasi. Prestasi adalah sesuatu yang didapat melalui kegiatan belajar di sekolah yang sifatnya kognitif dan mampu diukur menggunakan rentang nilai [4]. Tidak meratanya sebaran prestasi atau keberhasilan siswa dalam belajar menjadi masalah dalam upaya perwujudan tujuan dan misi SMK Negeri 5 Pekanbaru dalam menghasilkan lulusan siap terjun ke dunia kerja. Upaya yang dapat ditawarkan dalam penyelesaian masalah ini adalah membentuk model klasifikasi menggunakan algoritma data mining.

Data mining adalah metode ekstraksi data dari database berukuran besar yang memanfaatkan gabungan bidang ilmu seperti statistik, matematika, kecerdasan buatan dan *machine learning* dengan maksud menemukan pola dan mengidentifikasi suatu informasi yang berguna [5][6][7]. Salah satu metode pada data mining yang digunakan untuk ekstraksi data adalah klasifikasi. Klasifikasi adalah metode yang dapat memetakan input menjadi output diskrit yang disebut label dan kategori [6]. Algoritma klasifikasi yang diketahui memiliki kemampuan baik menyelesaikan permasalahan klasifikasi adalah *Support Vector Machine* (SVM) [6]. SVM adalah algoritma yang bekerja dengan prinsip *Structural Risk Minimization* (SRM) yang bertujuan menemukan *hyperlane* terbaik untuk memisahkan dua kelas inputan data, dan memaksimalkan margin antara dua kelas [5][8][9]. Margin adalah jarak antara *hyperlane* dengan *pattern* masing-masing kelas data. Kemampuan algoritma SVM dalam menyelesaikan permasalahan klasifikasi ditunjukkan pada penelitian terdahulu dengan memberikan nilai akurasi yang baik. Berdasarkan penelitian yang dilakukan oleh Kesumawati

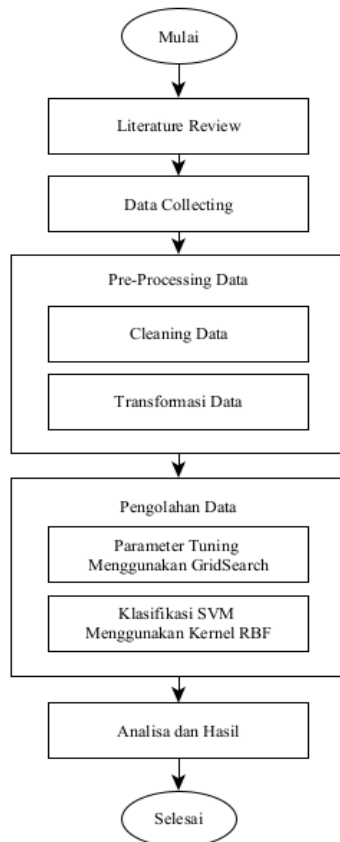
dan Utari (2018), dilakukan perbandingan antara *Naive Bayes Classifier* dan SVM dalam membentuk model prediksi waktu kelulusan mahasiswa. Hasil penelitian tersebut adalah kesimpulan bahwa algoritma SVM memiliki performa lebih baik daripada *Naive Bayes Classifier* dengan akurasi 69.15% [5]. Pratama, dkk (2018) melakukan penelitian serupa dengan menerapkan algoritma SVM dalam membentuk model prediksi ketepatan waktu kelulusan mahasiswa. Skenario pengujian terbaik mendapatkan nilai akurasi sebesar 80,55% menggunakan kernel *Gaussian RBF* dengan nilai variabel  $\gamma = 0,01$ ,  $\kappa = 0,5$ ,  $C = 1$ , IterasiMax = 100, dan epsilon = 0,001 menggunakan perbandingan data latih dan uji 170 dan 18 [10].

Berdasarkan penelitian oleh Very Riyanto, dkk (2019), dilakukan perbandingan antara algoritma *Neural Network*, SVM dan *Decision Tree* untuk membentuk model prediksi kelulusan mahasiswa. Ketiga algoritma tersebut menunjukkan hasil yang baik dengan akurasi lebih dari 80%. Algoritma *Neural Network* memiliki akurasi 84.68%, *Decision Tree* memiliki akurasi 84.96% dan SVM memiliki akurasi tertinggi yaitu 85.18% [11]. Penelitian selanjutnya dilakukan oleh Haryatmi dan Pramita (2021) yang menerapkan algoritma SVM untuk membentuk model prediksi waktu kelulusan mahasiswa. Dengan skenario pengujian data *training* 90% dan data *testing* 10%, algoritma SVM memberikan akurasi yang baik yakni 94.6% [6]. Selanjutnya, berdasarkan penelitian Mailana, dkk (2021), dilakukan perbandingan antara SVM dan C.45 untuk menemukan performa terbaik dalam memprediksi waktu kelulusan mahasiswa. Hasil penelitian tersebut algoritma SVM diketahui lebih baik daripada C.45 dengan nilai akurasi 85% [12]. Berdasarkan penjabaran masalah dan merujuk pada penelitian terdahulu, maka penelitian yang akan dilakukan adalah eksperimen untuk menguji bagaimana tingkat akurasi algoritma SVM dalam membentuk model klasifikasi keberhasilan siswa SMK Negeri 5 Pekanbaru dalam proses pembelajaran di sekolah. Dengan demikian, model prediksi yang dihasilkan diharapkan mampu membantu manajemen sekolah dalam melakukan evaluasi dan menetapkan kebijakan tepat sasaran dalam upaya meningkatkan keberhasilan siswa dalam proses belajar.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Tahapan penelitian yang diterapkan terdiri dari beberapa proses utama sebagaimana disajikan dalam Gambar 1.



Gambar 1. Tahapan Penelitian

Tahapan penelitian terdiri dari 5 proses utama. Diawali dengan *literature review* dengan tujuan menemukan *state of the art* serta merumuskan perbedaan antara penelitian terdahulu dengan penelitian yang dilakukan. Selanjutnya

dilakukan pengumpulan data, *preprocessing*, klasifikasi menggunakan algoritma, dan analisis untuk merumuskan simpulan dari hasil penelitian.

## 2.2 Pengumpulan Data

Data yang digunakan adalah nilai siswa kelas X semester ganjil program keahlian Desain Pemodelan Informasi Bangunan (DPIB), Teknik Instalasi Tenaga Listrik (TITL), Teknik Tenaga Listrik (TTL), dan Teknik Pemesinan (TP) SMK Negeri 5 Pekanbaru pada tahun 2020. Pengumpulan data dilakukan dengan metode observasi dan studi dokumen pada bagian Kurikulum SMK Negeri 5 Pekanbaru dalam kurun waktu 1 bulan. Data yang berhasil dikumpulkan berjumlah 170 record dengan 10 atribut yaitu ID, Jenis Kelamin, Agama, Jurusan, Pendidikan Agama, Pendidikan Pancasila, Bahasa Indonesia, Matematika, Sejarah, Bahasa Inggris, Seni Budaya, Pendidikan Jasmani Olahraga, Fisika, Kimia, Simulasi Digital, dan variabel target (label), ditunjukkan pada Tabel 1.

Tabel 1. Dataset

ID	Jenis Kelamin	Agama	Jurusan	...	P Pancasila	B Indonesia	Matematika	...	Fisika	Kimia	...	Label
1	Perempuan	Islam	DPIB	...	78	74	76	...	73	81	...	Cukup Baik
2	Laki-Laki	Kristen	DPIB	...	79	74	70	...	71	73	...	Cukup Baik
3	Laki-Laki	Islam	DPIB	...	78	74	78	...	70	80	...	Cukup Baik
4	Perempuan	Islam	DPIB	...	80	82	86	...	83	85	...	Baik
5	Laki-Laki	Kristen	TP	...	78	71	71	...	66	76	...	Cukup Baik
6	Laki-Laki	Kristen	TP	...	78	72	78	...	68	78	...	Cukup Baik
...			...	...			...	...	...	...	...	...
168	Perempuan	Islam	TTL	...	81	78	78	...	72	87	...	Cukup Baik
169	Perempuan	Islam	TTL	...	78	81	74	...	75	81	...	Cukup Baik
170	Laki-Laki	Islam	TITL	...	79	75	74	...	73	72	...	Cukup Baik

## 2.3 Data Mining

Data mining merupakan disiplin ilmu yang bertujuan menemukan keteraturan, pola, atau hubungan yang sifatnya implisit dan belum diketahui sebelumnya dari dataset berukuran besar. Data mining adalah bagian dari *Knowledge Discovery Database* (KDD) yang prosesnya mencakup eksplorasi dan ekstraksi informasi dari dataset dengan formula tertentu sehingga dapat dilakukan pemrosesan lebih detail [9][13][14][5]. Algoritma data mining terbagi menjadi tiga kategori, yaitu *supervised*, *unsupervised* dan *semi-supervised learning*. *Supervised learning* adalah pemodelan prediktif yang mekanisme kerjanya mentraining kumpulan data yang memiliki variabel target, kemudian melakukan klasifikasi berdasarkan tingkat kemiripan antar data [7]. Pemodelan prediktif umumnya dimanfaatkan untuk meramalkan fenomena masa depan berdasarkan kumpulan data di masa lalu [9]. Algoritma data mining yang masuk kategori *supervised learning* adalah klasifikasi [7].

## 2.4 Klasifikasi

Klasifikasi adalah proses menemukan model atau fungsi untuk membedakan kelas data, sehingga bertujuan menemukan kelas dari sekumpulan data yang labelnya belum diketahui [15][16]. Klasifikasi dapat dikatakan sebagai pembelajaran (*training*) yang memetakan item data ke dalam salah satu dari beberapa kelas yang sudah didefinisikan [15]. Beberapa algoritma klasifikasi yang umum digunakan adalah *Bayesian Classification*, *K-Nearest Neighbor*, *Decision Tree Induction*, *Case-Based Reasoning*, *Genetic Algorithm* dan *Support Vector Machine* [17]. Algoritma klasifikasi bekerja berdasarkan 4 komponen utama, yaitu:

1. Label merupakan variabel target yang mewakili objek klasifikasi.
2. Prediktor adalah variabel bebas (atribut) data yang diklasifikasi.
3. Data latih (*training set*) kumpulan data yang dilatih untuk menjalankan fungsi algoritma.
4. Data uji (*testing set*) kumpulan data untuk menghitung akurasi pengujian [18].

## 2.5 Support Vector Machine (SVM)

SVM adalah algoritma *supervised learning* yang memiliki kemampuan baik dalam menganalisis data untuk klasifikasi dan regresi [19][20]. SVM bekerja untuk menemukan *hyperlane* (pemisah) terbaik untuk membagi data ke dalam dua kelas, dan memaksimalkan margin antara dua kelas [10]. Model linier yang digunakan pada SVM untuk menghasilkan *hyperlane* optimal menggunakan persamaan 1 [21] :

$$Y = w^T x_i + b, \quad i = 1, 2, \dots, l \quad (1)$$

Keterangan:

$X_i$  = vektor baris berdimensi k (jumlah fitur)

$Y$  = nilai target dari himpunan data berupa [-1, +1]

$L$  = Jumlah data

$W$  = parameter bobot

$B$  = bias atau error

Margin maksimum dapat diketahui berdasarkan *support vector* yang terbentuk [22]. Perhitungan margin ditunjukkan pada persamaan 2

$$(w \cdot x_i) + b = 0 \quad (2)$$

Data ( $x_i$ ) yang termasuk pada kelas -1 (negatif) dapat dirumuskan menggunakan persamaan 3

$$(w \cdot x_i + b) \leq 1, y_i = -1 \quad (3)$$

Sedangkan data ( $x_i$ ) yang termasuk pada kelas +1 (positif) menggunakan perhitungan pada persamaan 4

$$(w \cdot x_i + b) \geq 1, y_i = 1 \quad (4)$$

Pada beberapa kasus, ditemukan kondisi dimana linier SVM tidak mampu bekerja optimal dalam mengklasifikasi sekumpulan data sehingga mengakibatkan hasil klasifikasi menjadi buruk [10]. Permasalahan demikian dapat diatasi dengan penggunaan kernel non-linier yang memanfaatkan *kernel trick* [20]. *Kernel trick* adalah fungsi yang melakukan *mapping* data input ke *feature space* yang dimensinya lebih tinggi supaya data memiliki struktur yang lebih baik dan membentuk *hyperlane* optimal [20][22]. *Kernel trick* yang dapat digunakan adalah Linier, Sigmoid, *Radial Basis Function* (RBF) dan Polynomial. Persamaan dari masing-masing kernel ditunjukkan pada Tabel 2.

**Tabel 2.** Persamaan Kernel SVM

Kernel	Parameter	Persamaan
Linear	Cost (C)	$K(x_i, x) = x_i^T x$
Polynomial	Cost (C), Gamma, Degree	$K(x_i, x) = (\gamma \cdot x_i^T x + r)^p, \gamma > 0$
<i>Radial Basis Function</i> (RBF)	Cost (C), Gamma	$K(x_i, x) = \exp(-\gamma  x_i - x ^2), \gamma > 0$
Sigmoid	Cost (C), Gamma	$K(x_i, x) = \tanh(\gamma x_i^T x + r)$

Berdasarkan Tabel 2, parameter gamma menyatakan *slope* yaitu perubahan nilai output akibat perubahan variabel input. Parameter C menyatakan intercept dari model yang berfungsi mengontrol optimasi antara margin dan kesalahan klasifikasi [21]. Sedangkan parameter *degree* menyatakan pangkat dari fungsi polynomial [22].

## 2.6 Confusion Matrix

*Confusion Matrix* merupakan tools visualisasi yang digunakan sebagai evaluasi mode klasifikasi pada algoritma *supervised learning* [23][24]. *Confusion Matrix* memiliki kolom dan baris, dimana kolom menyajikan kelas prediksi, dan baris mewakili kejadian yang sebenarnya [24]. Evaluasi *Confusion Matrix* adalah membandingkan hasil kelas prediksi dengan hasil kelas data sebenarnya menggunakan persamaan pada Tabel 3 [20]

**Tabel 3.** Persamaan Akurasi

Evaluasi	Persamaan
Akurasi	$\frac{TP + TN}{P + N}$
Keseluruhan	$\frac{P + N}{P + N}$

Persamaan tersebut digambarkan sebagai *Confusion Matrix* yang diilustrasikan pada Tabel 4.

**Tabel 4.** Persamaan *Confusion Matrix*

Actual Class	Predicted Class		Total
	Positif	Negatif	
Positif	TP	FN	TP + FN
Negatif	FP	TN	FP + TN
Total	TP + FP	FN + TN	P+N

Berdasarkan Tabel 4, *True Positif* (TP) menunjukkan kelas positif yang terklasifikasi dengan benar. *True Negatif* (TN) menunjukkan kelas negatif yang terklasifikasi dengan benar. *False Positif* (FP) menunjukkan kelas positif yang terklasifikasi negatif, dan *False Negatif* (FN) adalah kelas negatif yang terklasifikasi negatif [23].

### 3. HASIL DAN PEMBAHASAN

#### 3.1 *Preprocessing*

Sebelum dilakukan klasifikasi menggunakan algoritma *Support Vector Machine* (SVM), data yang berhasil dikumpulkan harus melalui tahap *preprocessing* terlebih dahulu. *Preprocessing* yang dilakukan meliputi cleaning dan tranformasi. Data yang digunakan memiliki 10 atribut yaitu ID, Jenis Kelamin, Agama, Jurusan, Pendidikan Agama, Pendidikan Pancasila, Bahasa Indonesia, Matematika, Sejarah, Bahasa Inggris, Seni Budaya, Pendidikan Jasmani Olahraga, Fisika, Kimia, Simulasi Digital, dan variabel target (label). Hasil *preprocessing* ditunjukkan pada Tabel 5.

**Tabel 5.** Hasil *Preprocessing*

ID	Jenis Kelamin	Agama	Jurusan	...	P Pancasila	B Indonesia	Matematika	...	Fisika	Kimia	...	Label
1	1	0	0	...	78	74	76	...	73	81	...	Cukup Baik
2	0	1	0	...	79	74	70	...	71	73	...	Cukup Baik
3	0	0	0	...	78	74	78	...	70	80	...	Cukup Baik
4	1	0	0	...	80	82	86	...	83	85	...	Baik
5	0	1	3	...	78	71	71	...	66	76	...	Cukup Baik
6	0	1	3	...	78	72	78	...	68	78	...	Cukup Baik
...				...				...			...	...
168	1	0	2	...	81	78	78	...	72	87	...	Cukup Baik
169	1	0	2	...	78	81	74	...	75	81	...	Cukup Baik
170	0	0	1	...	79	75	74	...	73	72	...	Cukup Baik

*Preprocessing* yang dilakukan meliputi *data cleaning* untuk menghapus *outlier* yang berpotensi mengganggu kinerja klasifikasi, kemudian *data transformation* untuk merubah nilai atribut menjadi numerik sesuai kebutuhan algoritma. Hasil *preprocessing* didapatkan 160 data yang dapat digunakan dalam pengujian sebagaimana ditunjukkan pada Tabel 5.

#### 3.2 Klasifikasi *Support Vector Machine* (SVM)

Klasifikasi SVM dilakukan menggunakan bahasa pemrograman Python. Dataset sebelumnya dibagi terlebih dahulu menjadi *training* dan *testing* set dengan ratio 0.2. Artinya, 80% data untuk *training set*, dan 20% adalah *testing set*. Skenario pengujian dilakukan dengan menemukan kernel terbaik terlebih dahulu menggunakan metode *GridSearch*.

Kernel yang dibandingkan adalah *Radial Basis Function* (RBF), *Polynomial* dan *Sigmoid*. Hasil perbandingan menggunakan metode GridSearch menunjukkan kernel dengan performa terbaik adalah RBF. Berdasarkan hasil perbandingan tersebut, selanjutnya dilakukan perhitungan klasifikasi menggunakan algoritma SVM kernel RBF. Proses klasifikasi diawali dengan memberikan nilai input pada masing-masing parameter. Parameter yang digunakan dalam perhitungan menggunakan kernel RBF adalah Cost (C) dan Gamma ( $\gamma$ ). Nilai input parameter ditunjukkan pada Tabel 6.

**Tabel 6.** Nilai Input Parameter

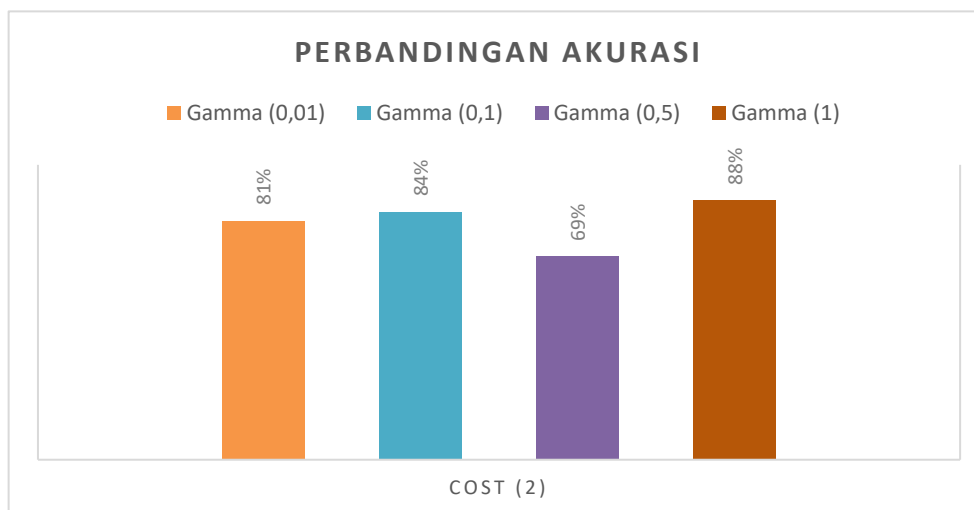
Parameter	Nilai Input			
Cost (C)	0.1	1	1.5	2
Gamma ( $\gamma$ )	0.01	0.1	0.5	1

Berdasarkan nilai input parameter tersebut dilakukan percobaan untuk mengetahui akurasi yang dihasilkan. Hasil akurasi pada masing-masing percobaan ditunjukkan pada Tabel 7.

**Tabel 7.** Hasil Akurasi

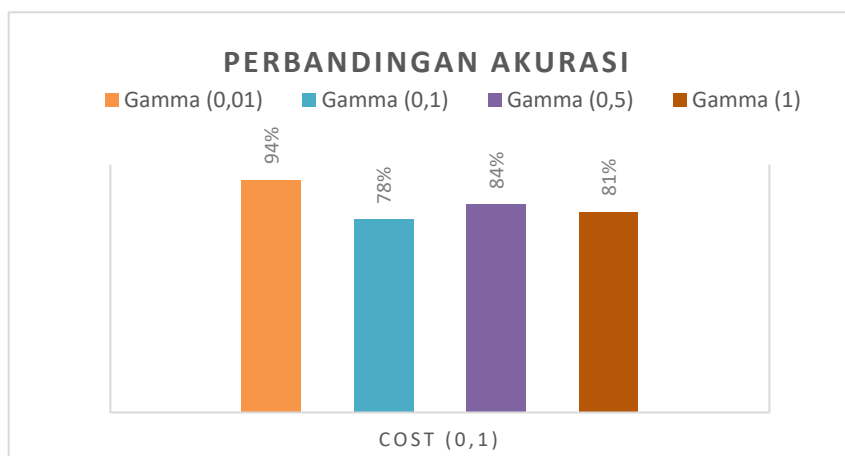
Cost (C)	Gamma ( $\gamma$ )			
	0.01	0.1	0.5	1
0.1	<b>94%</b>	78%	84%	81%
1	88%	91%	84%	75%
1.5	78%	88%	91%	81%
2	81%	84%	<b>69%</b>	88%

Setelah menyelesaikan 16 skenario percobaan, didapatkan akurasi terendah adalah 69% pada penggunaan nilai Cost (C) = 2 dan Gamma ( $\gamma$ ) = 0.5 seperti yang ditunjukkan pada Gambar 2.



**Gambar 2.** Perbandingan Akurasi (1)

Sedangkan akurasi terbaik adalah 94% pada penggunaan nilai Cost (C) = 0.1 dan Gamma ( $\gamma$ ) = 0.01 sebagaimana ditunjukkan pada Gambar 3.



**Gambar 3.** Perbandingan Akurasi (2)

Melalui 16 skenario percobaan, akurasi terbaik ditunjukkan pada penggunaan nilai Cost (C) = 0.1 dan  $\Gamma$  (gamma) = 0.01. Pada skenario tersebut, dilakukan evaluasi menggunakan *confusion matrix*. Hasil evaluasi menggunakan *confusion matrix* ditunjukkan pada Tabel 8.

**Tabel 8.** Evaluasi *Confusion Matrix*

Actual Class	Predicted Class	
	Cukup Baik	Baik
Cukup Baik	30	2
Baik	0	0

Berdasarkan evaluasi menggunakan *confusion matrix* yang ditunjukkan pada Tabel 8, diketahui terdapat 30 data yang terklasifikasi benar sebagai Cukup Baik. Serta terdapat 2 data yang terklasifikasi Baik padahal kelas aktualnya Cukup Baik.

#### 4. KESIMPULAN

Berdasarkan 16 skenario pengujian dengan nilai input parameter yang berbeda-beda, rata-rata akurasi yang dihasilkan adalah 80%. Hal ini menunjukkan bahwa algoritma *Support Vector Machine* (SVM) memiliki kemampuan baik dalam menyelesaikan permasalahan klasifikasi hasil belajar siswa Sekolah Menengah Kejuruan (SMK) Negeri 5 Pekanbaru. Berdasarkan pengujian, hasil klasifikasi terbaik ditunjukkan pada penggunaan nilai Cost (C) = 0.1 dan nilai  $\Gamma$  (gamma) = 0.01 dengan nilai akurasi 94%. Kombinasi nilai parameter Cost (C) dan  $\Gamma$  (gamma) diketahui memberikan pengaruh signifikan terhadap performa klasifikasi. Nilai akurasi terendah dalam pengujian ini adalah 69% pada kombinasi nilai parameter Cost (C) = 2 dan nilai  $\Gamma$  (gamma) = 0.5.

#### REFERENCES

- [1] M. Prasajo and J. Triwidianti, "Prediksi Prestasi Siswa SMK Masuk Pasar Kerja Menggunakan Teknik Data Mining (Studi Kasus SMKN 1 Kota Agung Timur Tanggamus, Lampung)," Seminar Nasional Hasil Penelitian dan Pengabdian Masyarakat, pp. 134–150, 2021.
- [2] "SMK Negeri 5 Pekanbaru – Official Web." <https://smkn5pekanbaru.sch.id/> (accessed Jul. 28, 2022).
- [3] J. Triwidianti, F. Y. Alfian, and M. Prasajo, "Perbandingan Metode Data Mining Untuk Prediksi Prestasi Siswa Tingkat Pendidikan Menengah Kejuruan Pada Sekolah Menengah Kejuruan Negeri (SMKN 1) Gadingrejo Pringsewu Lampung," Seminar Nasional Hasil Penelitian dan Pengabdian Masyarakat, pp. 126–133, 2021.
- [4] S. Marpaung, I. Sekolah Tinggi Ilmu Komputer Stikom Tunas Bangsa Jl Sudirman, S. Barat, K. Pematang Siantar, and S. Utara, "PENERAPAN METODE NAÏVE BAYES DALAM MEMREDIKSI PRESTASI SISWA DI SMA NEGERI 1 PANOMBELIAN PANELI," Jurnal Sistem Informasi dan Ilmu Komputer Prima, vol. 4, no. 2, 2021.
- [5] A. Kesumawati and D. T. Utari, "Predicting patterns of student graduation rates using Naïve bayes classifier and support vector machine," in AIP Conference Proceedings, Oct. 2018, vol. 2021. doi: 10.1063/1.5062769.
- [6] E. Haryatmi and S. Pramita Hervianti, "Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 5, no. 2, pp. 386–392, Apr. 2021, doi: 10.29207/resti.v5i2.3007.
- [7] S. Widaningsih, "PERBANDINGAN METODE DATA MINING UNTUK PREDIKSI NILAI DAN WAKTU KELULUSAN MAHASISWA PRODI TEKNIK INFORMATIKA DENGAN ALGORITMA C4,5, NAÏVE BAYES, KNN DAN SVM," Jurnal Tekno Insentif, vol. 13, no. 1, pp. 16–25, Apr. 2019, doi: 10.36787/jti.v13i1.78.
- [8] S. Dewi Purba, L. Harahap, J. Franky, and R. Panggabean, "PREDICTION OF STUDENTS DROP OUT WITH SUPPORT VECTOR MACHINE ALGORITHM," 2021.
- [9] W. S. Dharmawan, "KOMPARASIALGORITMA KLASIFIKASISVM-PSO DAN C4.5-PSO DALAMPREDIKSI PENYAKIT JANTUNG," Jurnal Informatika, Manajemen dan Komputer, vol. 13, no. 2, pp. 31–41, 2021.
- [10] A. Pratama, R. Cahya Wihandika, and D. E. Ratnawati, "Implementasi Algoritme Support Vector Machine (SVM) untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa," 2018. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [11] V. Riyanto, A. Hamid, and R. Ridwansyah, "Prediction of Student Graduation Time Using the Best Algorithm," Indonesian Journal of Artificial Intelligence and Data Mining, vol. 2, no. 1, pp. 1–9, Mar. 2019, doi: 10.24014/ijaidm.v2i1.6424.
- [12] A. Mailana, A. A. Putra, S. Hidayat, and A. Wibowo, "Comparison of C4.5 Algorithm and Support Vector Machine in Predicting the Student Graduation Timeliness," Jurnal Online Informatika, vol. 6, no. 1, pp. 11–16, Jun. 2021, doi: 10.15575/join.v6i1.608.
- [13] R. H. Sukarna and Y. Ansori, "IMPLEMENTASI DATA MINING MENGGUNAKAN METODE NAIVE BAYES DENGAN FEATURE SELECTION UNTUK PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU," 2022.
- [14] Gunawan, Hanes, and Catherine, "C4.5, K-Nearest Neighbor, Naïve Bayes and Random Forest Algorithms Comparison to Predict Students' On Time Graduation," Indonesian Journal of Artificial Intelligence and Data Mining (IJAIMD), vol. 4, no. 2, pp. 62–71, 2021, doi: 10.24014/ijaidm.v4i2.10833.
- [15] F. Elfaladonna and A. Rahmadani, "ANALISAMETODE CLASSIFICATION-DECISION TREE DAN ALGORITMA C.45 UNTUK MEMREDIKSI PENYAKIT DIABETES DENGAN MENGGUNAKAN APLIKASI RAPID MINER," Scienc and Information Technology Journal, vol. 2, no. 1, pp. 10–17, 2019.
- [16] R. Febryani and T. Arifin, "OPTIMASI NAÏVE BAYES MENGGUNAKAN PSOUNTUK TINGKATKEBERHASILAN CRYOTHERAPY PADA PENYAKIT KUTIL," JURNAL RESPONSIF, vol. 3, no. 2, 2021, [Online]. Available: <http://ejournal.ars.ac.id/index.php/jti>

- [17] R. Thaniket and E. Taufik Luthf, “PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE,” 2020.
- [18] O. Tias Mugi Rahayu, B. Arnawisuda Ningsi, and I. Arofah, “KLASIFIKASI KETEPATAN WAKTU KELULUSAN MAHASISWA DENGAN METODE NAÏVE BAYES,” vol. 15, no. 8, pp. 4993–5000, 2021, [Online]. Available: <http://ejurnal.binawakya.or.id/index.php/MBI>
- [19] G. Sailasya and G. L. Aruna Kumari, “Analyzing the Performance of Stroke Prediction using ML Classification Algorithms,” 2021. [Online]. Available: [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)
- [20] H. C. Husada and A. S. Paramita, “Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM),” *Teknika*, vol. 10, no. 1, pp. 18–26, Feb. 2021, doi: 10.34148/teknika.v10i1.311.
- [21] N. Fitriyah, B. Warsito, D. Asih, and I. Maruddani, “ANALISIS SENTIMEN GOJEK PADA MEDIA SOSIAL TWITTER DENGAN KLASIFIKASI SUPPORT VECTOR MACHINE (SVM),” *JURNAL GAUSSIAN*, vol. 9, no. 3, pp. 376–390, 2020, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>
- [22] N. Rachmalia Feta and A. Rahmat Ginanjar, “KOMPARASI FUNGSI KERNEL METODE SUPPORT VECTOR MACHINE UNTUK PEMODELAN KLASIFIKASI TERHADAP PENYAKIT TANAMAN KEDELAI,” *BRITech (Jurnal Ilmiah Ilmu Komputer, Sains dan Teknologi Terapan)*, vol. 1, no. 1, 2019.
- [23] A. Kesumawati and D. T. Utari, “Predicting patterns of student graduation rates using Naïve bayes classifier and support vector machine,” in *AIP Conference Proceedings*, Oct. 2018, vol. 2021. doi: 10.1063/1.5062769.
- [24] E. Supriyadi and D. I. Sensuse, “OPTIMASI ALGORITMA SUPPORT VECTOR MACHINE DENGAN PARTICLE SWARM OPTIMIZATION DALAM MENDETEKSI KETEPATAN WAKTU KELULUSAN MAHASISWA : STUDI KASUS POLTEK LP3I JAKARTA ‘KAMPUS DEPOK,’” 2015. [Online]. Available: <http://www.nusamandiri.ac.id>,