

PEMODELAN PREDIKTIF KELAYAKAN BEASISWA PIP MENGGUNAKAN SUPPORT VECTOR MACHINE: STUDI KASUS PADA SEKOLAH MENENGAH KEJURUAN NEGERI DI CIREBON

Muhammad Syafri Syamsudin¹, Indra Maulana², Suratno³

^{1),2)} Informatics and Computer Engineering Education, Institut Prima Bangsa, Cirebon, Indonesia

³⁾ Universitas Pendidikan Indonesia, Bandung, Indonesia

email : syafri.ice@ipbcirebon.ac.id¹, indramaulana360@gmail.com², suratno@upi.edu³

Abstrak

Memastikan akses pendidikan yang setara bagi siswa dari keluarga berpenghasilan rendah tetap menjadi tantangan strategis dalam pengembangan sumber daya manusia di Indonesia. Program Indonesia Pintar (PIP) merupakan inisiatif pemerintah yang memberikan bantuan keuangan langsung kepada siswa kurang mampu atau yang berisiko putus sekolah karena kesulitan ekonomi. Namun, proses seleksi penerima beasiswa masih menghadapi permasalahan terkait akurasi dan objektivitas. Penelitian ini bertujuan untuk mengembangkan model klasifikasi guna mengidentifikasi calon penerima PIP yang layak menggunakan algoritma Support Vector Machine (SVM), berdasarkan data siswa dari Sekolah Menengah Kejuruan Negeri (SMK Negeri) di Cirebon. Dataset terdiri atas 216 data siswa, mencakup fitur seperti prestasi akademik, pendapatan orang tua, status KETM (keluarga ekonomi tidak mampu), serta variabel demografis lainnya. Model SVM dikembangkan menggunakan kernel linear pada platform R, dengan kinerja divalidasi melalui metrik akurasi, presisi, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model mencapai akurasi klasifikasi sebesar 90% pada data pelatihan dan 85,19% pada data pengujian, dengan interval kepercayaan 95%. Temuan ini menunjukkan bahwa algoritma SVM efektif dalam mengklasifikasikan kelayakan penerima PIP berdasarkan variabel kuantitatif yang tersedia. Penelitian ini berkontribusi pada pengembangan sistem pendukung keputusan berbasis data di sektor pendidikan, khususnya dalam meningkatkan transparansi dan efisiensi distribusi beasiswa yang adil dan tepat sasaran.

Kata kunci: Klasifikasi, Machine Learning, Prediksi, Support Vector Machine, Beasiswa.

Abstract

Ensuring equitable access to education for students from low-income families remains a strategic challenge in the development of human resources in Indonesia. The Program Indonesia Pintar (PIP) is a government initiative that provides direct financial assistance to students who are poor or at risk of dropping out due to economic hardship. However, the selection process for scholarship recipients still faces issues related to accuracy and objectivity. This study aims to develop a classification model for identifying eligible PIP recipients using the Support Vector Machine (SVM) algorithm, based on student data from public vocational high schools (SMK Negeri) in Cirebon. The dataset comprises 216 student records, including features such as academic performance, parental income, KETM (low-income household status), and other demographic variables. The SVM model was developed using a linear kernel on the R platform, with performance validated through metrics including accuracy, precision, recall, and F1-score. Testing results indicate that the model achieved a classification accuracy of 90% on training data and 85.19% on testing data, with a 95% confidence interval. These findings suggest that the SVM algorithm is effective in classifying PIP eligibility based on available quantitative variables. This research contributes to the development of data-driven decision support systems in the education sector, particularly in enhancing transparency and efficiency in fair and targeted scholarship distribution.

Keywords: Classification, Machine Learning, Prediction, Support Vector Machine, Scholarship.

1. PENDAHULUAN

Dalam era globalisasi, pemerataan akses terhadap pendidikan telah menjadi isu sentral dalam kebijakan sosial dan pembangunan berkelanjutan. Salah satu tantangan utama adalah memastikan bahwa siswa dari keluarga berpenghasilan rendah dapat melanjutkan pendidikan mereka secara optimal [1], [2]. Program bantuan seperti Program Indonesia Pintar (PIP) hadir sebagai strategi nasional untuk menjawab tantangan tersebut. PIP memberikan bantuan langsung berupa dana tunai kepada peserta didik yang tergolong miskin atau rentan miskin dengan tujuan mencegah terjadinya putus sekolah [3], [4]. Di sisi lain, peningkatan efisiensi dan transparansi dalam penyaluran beasiswa menjadi kebutuhan yang mendesak. Seiring dengan kemajuan teknologi kecerdasan buatan, metode prediktif berbasis machine learning kini berperan penting dalam mendukung pengambilan keputusan kebijakan yang berorientasi pada data [5]. Salah satu teknik yang menonjol dalam bidang ini adalah Support Vector Machine (SVM), yang dikenal memiliki tingkat akurasi tinggi dalam mengklasifikasikan data non-linear [6]. Dalam konteks pendidikan, SVM memiliki potensi besar untuk memprediksi kelayakan penerima beasiswa berdasarkan variabel sosial-ekonomi dan akademik, sehingga memungkinkan penyaluran bantuan keuangan yang lebih tepat sasaran dan efektif [7], [8].

Salah satu tantangan terbesar dalam penerapan sistem beasiswa berbasis pemerataan adalah ketiadaan mekanisme seleksi yang akurat dan terotomatisasi. Banyak sekolah masih mengandalkan pendekatan manual dan administratif yang rentan terhadap bias serta ketidakefisienan [9]. Proses seleksi sering kali belum mampu mempertimbangkan secara komprehensif kombinasi variabel penting, seperti pendapatan orang tua, prestasi akademik, dan status kemiskinan [10]. Hal ini berisiko menyebabkan beasiswa tidak tepat sasaran, yakni tidak sampai kepada siswa yang benar-benar membutuhkan. Beberapa penelitian telah menunjukkan bahwa teknologi machine learning mampu mengatasi keterbatasan tersebut. Metode Support Vector Machine (SVM), misalnya, terbukti memiliki kinerja tinggi dalam mengklasifikasikan kasus yang kompleks, khususnya dengan dataset yang relatif kecil [11]. Dalam konteks prediksi beasiswa, Ji [12] menemukan bahwa model SVM dapat secara akurat mengidentifikasi penerima yang layak melalui proses evaluasi otomatis yang mampu mengurangi bias manusia. Oleh karena itu, pengembangan sistem pendukung keputusan berbasis SVM menjadi sangat penting untuk meningkatkan transparansi dan akuntabilitas dalam distribusi Program Indonesia Pintar (PIP), khususnya di studi-studi terbaru menunjukkan adanya kemajuan yang signifikan dalam penerapan machine learning untuk sistem pendukung keputusan di bidang pendidikan. Dalam dua hingga tiga tahun terakhir, sejumlah penelitian telah mengimplementasikan Support Vector Machine (SVM) dalam berbagai konteks pendidikan, termasuk prediksi kinerja akademik, pemilihan jurusan, dan kelayakan beasiswa [6], [13]. SVM dianggap unggul karena kemampuannya dalam mengklasifikasikan data dengan margin maksimum serta fleksibilitasnya dalam menangani data non-linear melalui penggunaan kernel trick [14]. Sebagai contoh, Fajardo et al [1] menerapkan SVM pada sistem prediksi penerima beasiswa pemerintah dan memperoleh kinerja yang tinggi, meskipun masih kalah dibandingkan Logistic Regression. Sementara itu, Nikumbh et al. [3] menggunakan SVM sebagai bagian dari sistem rekomendasi beasiswa yang dipersonalisasi berdasarkan profil siswa.

Namun demikian, beberapa penelitian masih menghadapi keterbatasan, seperti ukuran dataset yang kecil, kurangnya keragaman variabel, serta ketergantungan pada prosedur preprocessing yang kompleks [10], [15]. Selain itu, sebagian besar penelitian sebelumnya cenderung lebih menekankan pada pengembangan algoritmik daripada integrasi langsung dengan konteks kebijakan. Hal ini tampak pada penelitian Agustina et al. [16] yang menggunakan metode MAUT untuk menyeleksi penerima beasiswa PIP di SMKN 1 Kefamenanu secara manual, serta pada penelitian Maulina & Kurniadi [17] yang mengembangkan aplikasi berbasis SAW untuk merekomendasikan beasiswa di SMP Plus Madaarikul Ulum tanpa memanfaatkan algoritma prediktif. Aspek motivasional dan karakteristik psikososial penerima beasiswa juga disoroti oleh Airulanang [18], [19], yang menekankan perlunya sistem seleksi yang tidak hanya efisien secara teknis, tetapi juga responsif terhadap kondisi menyeluruh siswa. Studi-studi tersebut menunjukkan bahwa pemanfaatan teknologi cerdas seperti Support Vector Machine (SVM) dalam sistem seleksi beasiswa masih terbatas dan belum sepenuhnya terimplementasi dalam skema nasional seperti Program Indonesia Pintar (PIP). Oleh karena itu, integrasi SVM ke dalam sistem pendukung keputusan beasiswa berbasis sekolah, khususnya di tingkat Sekolah Menengah Kejuruan (SMK), merupakan area inovasi penting yang dapat meningkatkan akurasi, transparansi, dan keadilan dalam distribusi bantuan pendidikan [18], [20], [21].

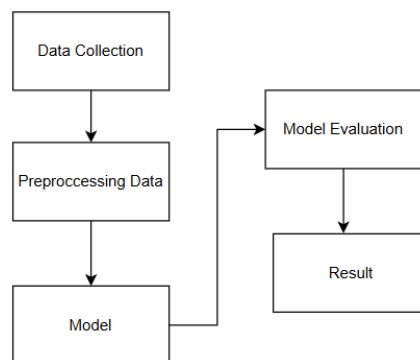
Meskipun beberapa penelitian telah menunjukkan efektivitas SVM dalam prediksi akademik, penelitian yang secara khusus mengintegrasikan algoritma ini untuk menentukan penerima Program

Indonesia Pintar (PIP) di tingkat sekolah masih terbatas [22]. Sebagian besar penelitian hanya berfokus pada validasi algoritma tanpa secara langsung mempertimbangkan konteks implementasi kebijakan pendidikan. Selain itu, eksplorasi terkait integrasi sistem pendukung keputusan dengan platform analitik seperti R juga masih sangat terbatas, yang menunjukkan adanya kesenjangan dalam literatur [9], [23]. Untuk menjawab kesenjangan tersebut, penelitian ini mengusulkan pendekatan berbasis Support Vector Machine (SVM) yang terintegrasi dengan proses data preprocessing serta evaluasi klasifikasi menggunakan confusion matrix. Kebaruan dari penelitian ini terletak pada penerapan model SVM secara khusus dalam konteks distribusi beasiswa PIP di Sekolah Menengah Kejuruan (SMK) melalui pendekatan sistematis dan pengujian empiris. Pendekatan ini diharapkan dapat mendorong pengembangan sistem pendukung keputusan beasiswa yang efisien, akurat, dan dapat direplikasi oleh lembaga pendidikan lainnya [22].

Penelitian ini bertujuan untuk mengembangkan model prediktif dalam mengidentifikasi penerima beasiswa Program Indonesia Pintar (PIP) pada Sekolah Menengah Kejuruan (SMK) dengan menggunakan algoritma Support Vector Machine (SVM). Model ini dirancang untuk meningkatkan akurasi seleksi penerima beasiswa dengan memanfaatkan data akademik dan sosial-ekonomi siswa. Kebaruan utama penelitian ini terletak pada integrasi teknik klasifikasi SVM dengan validasi akurasi menggunakan platform R, sehingga memungkinkan penerapan langsung di lingkungan sekolah. Selain menilai kinerja statistik, penelitian ini juga menekankan pentingnya kontribusi praktis dalam mendukung kebijakan pendidikan berbasis data. Melalui pendekatan ini, proses seleksi beasiswa diharapkan menjadi lebih objektif, transparan, dan efisien, sehingga bantuan pendidikan dapat tepat sasaran kepada siswa yang benar-benar membutuhkan [24].

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif prediktif yang bertujuan untuk mengembangkan model klasifikasi biner dalam memprediksi kelayakan siswa sebagai penerima Program Indonesia Pintar (PIP). Model tersebut dibangun dengan menggunakan algoritma Support Vector Machine (SVM), yang dikenal efektif dalam menangani data berdimensi tinggi serta mampu menghasilkan margin klasifikasi optimal baik pada dataset linear maupun non-linear.



Gambar 1. Metodologi Penelitian

Gambar 1 menggambarkan metodologi penelitian yang terdiri atas beberapa tahapan berurutan: pengumpulan data, data preprocessing, pengembangan model, evaluasi model, serta pembuatan keluaran akhir. Setiap tahapan dilaksanakan secara sistematis untuk membangun model klasifikasi yang akurat dan andal dalam seleksi penerima beasiswa.

2.1. Data Collection (Pengumpulan Data)

Data yang digunakan dalam penelitian ini diperoleh dari dokumentasi resmi yang disediakan oleh salah satu Sekolah Menengah Kejuruan Negeri (SMK Negeri) di Cirebon, dengan total sebanyak 216 data siswa. Setiap data mencakup atribut-atribut sebagai berikut: jenis kelamin, asal sekolah, nilai akademik,

pendapatan orang tua, status keluarga ekonomi tidak mampu (KETM), serta status penerima PIP. Adapun label target, status penerima PIP, dikodekan sebagai variabel biner, di mana nilai 1 menunjukkan penerima, sedangkan nilai 0 menunjukkan bukan penerima.

Tabel 1. Kumpulan Data

No	Name	Gender	School	Value	Income	Label
1	A. S*****	Male	SMP N 1 T*****	184	1.900.000	√
2	F*****	Male	SMP N 4 C*****	208	2.000.000	√
3	Y*****	Female	SMP N 11 C*****	184	4.300.000	X
4	R. A. A*****	Male	SMP N 7 C*****	220	5.700.000	X
5	A. P*****	Male	MTS N 2 C*****	221	7.600.000	X
6	R. A. P*****	Male	SMP N 6 C*****	213	8.400.000	X
7	M. P. S*****	Male	SMP N 8 C*****	205	5.800.000	X
8	A. N*****	Male	SMP N 1 T*****	193	2.600.000	X
9	F. A. P*****	Male	SMP N 1 C*****	255	2.000.000	√
...	...	...	...	...	...	...
216	U. H*****	Female	SMP N 5 C*****	199	3.000.000	X

2.2. Preprocessing Data (Pra-pemrosesan Data)

Data yang digunakan dalam penelitian ini diperoleh dari dokumentasi resmi yang disediakan oleh salah satu Sekolah Menengah Kejuruan Negeri (SMK Negeri) di Cirebon, dengan total sebanyak 216 data siswa. Setiap data mencakup atribut-atribut sebagai berikut: jenis kelamin, asal sekolah, nilai akademik, pendapatan orang tua, status keluarga ekonomi tidak mampu (KETM), serta status penerima PIP. Adapun label target, yaitu status penerima PIP, dikodekan sebagai variabel biner, di mana nilai 1 menunjukkan penerima dan nilai 0 menunjukkan bukan penerima.

Tabel 2. Training Data

No	Name	Gender	School	Value	Income	Label
1	A. S*****	Male	SMP N 1 T*****	184	1.900.000	√
2	F*****	Male	SMP N 4 C*****	207	2.000.000	X
3	Y*****	Female	SMP N 11 C*****	184	4.300.000	X
4	A. P*****	Male	MTS N 2 C*****	221	7.600.000	X
...	...	...	...	...	...	...
162	R. A. P*****	Male	SMP N 6 C*****	213	8.400.000	X

Tabel 3. Testing Data

No	Name	Gender	School	Value	Income	Label
1	R. A. A*****	Male	SMP N 7 C*****	200	5.700.000	X
2	M. P. S*****	Male	SMP N 8 C*****	205	5.800.000	X
3	A. N*****	Male	SMP N 1 T*****	193	2.600.000	X
4	F. A. P*****	Male	SMP N 1 C*****	255	2.000.000	√
...	...	...	...	...	...	...
54	U. H*****	Female	SMP N 5 C*****	199	3.000.000	X

2.3. Development Model (Pengembangan Model)

Model klasifikasi dalam penelitian ini dikembangkan dengan menggunakan algoritma Support Vector Machine (SVM), secara khusus menerapkan pendekatan C-classification dengan linear kernel.

Konfigurasi ini dipilih berdasarkan keterpisahan linear yang teramati pada distribusi data awal. Implementasi dilakukan menggunakan paket e1071 pada R versi 4.1.3, dengan parameter cost bawaan (1), karena evaluasi awal menunjukkan kinerja yang memadai tanpa memerlukan regularisasi tambahan. Model dilatih dengan menggunakan tiga fitur utama, yaitu nilai akademik, pendapatan orang tua, dan status Keluarga Ekonomi Tidak Mampu (KETM), untuk memprediksi status beasiswa sebagai variabel biner. Proses pelatihan menghasilkan sebanyak 46 support vector, sebagaimana ditunjukkan pada Gambar 3, yang menunjukkan bahwa sebagian data observasi berkontribusi langsung dalam pembentukan optimal hyperplane.

```
> library("e1071")
Warning message:
package 'e1071' was built under R version 4.1.3
>
> attach(datLatih)
> svm.Lin = svm(Beasiswa~ Nilai+Penghasilan+KETM, data=datLatih, type="C-classification", kernel="linear")
> print(svm.Lin)

Call:
svm(formula = Beasiswa ~ Nilai + Penghasilan + KETM, data = datLatih, type = "C-classification", kernel = "linear")

Parameters:
  SVM-Type:  C-classification
 SVM-Kernel: linear
      cost:  1

Number of Support Vectors:  46

> |
```

Gambar 2. Output Support Vectors

Hasil ini semakin menegaskan potensi SVM sebagai pendekatan yang andal dalam mendukung proses seleksi beasiswa yang objektif, berbasis data, dan berkeadilan di lingkungan pendidikan.

2.4. Evaluation Model (Evaluasi Model)

Algoritma Support Vector Machine (SVM) dengan tipe C-classification dan linear kernel diadopsi dalam penelitian ini berdasarkan temuan awal yang menunjukkan adanya distribusi yang relatif dapat dipisahkan secara linear antara penerima dan non-penerima beasiswa. Proses pelatihan model melibatkan tiga variabel prediktor utama, yaitu kinerja akademik, tingkat pendapatan orang tua, dan keanggotaan dalam kategori Keluarga Ekonomi Tidak Mampu (KETM). Hasil evaluasi menunjukkan bahwa model mencapai tingkat akurasi sebesar 90,74%, sebagaimana ditunjukkan pada Gambar 4, serta menghasilkan nilai Cohen's Kappa sebesar 0,64. Nilai ini mengindikasikan tingkat kesesuaian klasifikasi yang substansial, melampaui kesesuaian yang diharapkan secara acak.

```
Confusion Matrix and Statistics

 datLatih.Prediksi
      0      1
0 130      5
1   10     17

      Accuracy : 0.9074
      95% CI : (0.8519, 0.9472)
 No Information Rate : 0.8642
 P-Value [Acc > NIR] : 0.06276

      Kappa : 0.64

McNemar's Test P-Value : 0.30170

      Sensitivity : 0.9286
      Specificity : 0.7727
      Pos Pred Value : 0.9630
      Neg Pred Value : 0.6296
      Prevalence : 0.8642
      Detection Rate : 0.8025
      Detection Prevalence : 0.8333
      Balanced Accuracy : 0.8506

      'Positive' Class : 0
```

Gambar 3. Output Evaluasi Model

Rincian kinerja model ditunjukkan melalui matriks pada tabel 4.

Tabel 4. Matriks Data		
	Prediktion No (0)	Predicted: Scholarship (1)
Actual: No (0)	130 (True Negative)	5 (False Positive)
Actual: Scholarship (1)	10 (False Negative)	17 (True Positive)

Berdasarkan confusion matrix yang diperoleh, beberapa metrik evaluasi utama dihitung untuk mengukur kinerja klasifikasi model. Setiap metrik dihitung menggunakan rumus Sensitivity (Positive Recall) menunjukkan proporsi data kelas positif (penerima beasiswa) yang berhasil diklasifikasikan dengan benar. Dihitung menggunakan rumus:

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Sensitivity} = \frac{17}{17 + 10} = \frac{17}{27} = 62.86\%$$

1. Specificity (Negative Recall)

Menggambarkan proporsi data kelas negatif (bukan penerima beasiswa) yang berhasil diklasifikasikan dengan benar.

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Specificity} = \frac{130}{130 + 5} = \frac{130}{135} = 96.30\%$$

2. Precision (Positive Predictive Value)

Mengukur proporsi prediksi positif yang benar-benar berasal dari kelas positif:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Precision} = \frac{17}{17 + 5} = \frac{17}{22} = 77.27\%$$

3. Negative Predictive Value (NPV)

Menunjukkan kemungkinan bahwa prediksi negatif benar-benar berasal dari kelas negatif:

$$\text{NPV} = \frac{TN}{TN + FN}$$

$$\text{NPV} = \frac{130}{130 + 10} = \frac{130}{140} = 92.86\%$$

4. Balanced Accuracy

Digunakan untuk mengatasi ketidakseimbangan kelas, dihitung sebagai rata-rata dari sensitivity dan specificity:

$$\text{Balanced Accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2}$$

$$\text{Precision} = \frac{62.86\% + 96.30\%}{2} = 79.60\%$$

Nilai specificity yang tinggi menunjukkan bahwa model sangat efektif dalam mengidentifikasi siswa yang bukan penerima beasiswa, sementara nilai sensitivity yang moderat mengindikasikan masih adanya ruang untuk perbaikan dalam mendeteksi siswa yang seharusnya menerima beasiswa. Nilai Balanced Accuracy sebesar 79,60% mencerminkan kinerja model yang relatif seimbang pada kedua kelas, sehingga sesuai untuk digunakan pada data dengan distribusi kelas yang tidak sepenuhnya seimbang Selain itu, uji McNemar menghasilkan nilai p-value sebesar 0,3017, yang menunjukkan tidak terdapat perbedaan yang signifikan secara statistik antara distribusi klasifikasi aktual dan prediksi. Hal ini mengindikasikan bahwa model tidak menunjukkan bias sistematis terhadap salah satu kelas.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Klasifikasi Model SVM

Penelitian ini menghasilkan model klasifikasi untuk memprediksi kelayakan penerima Beasiswa Program Indonesia Pintar (PIP) pada Sekolah Menengah Kejuruan (SMK) Negeri di Cirebon dengan menggunakan algoritma Support Vector Machine (SVM) tipe C-classification dan linear kernel. Dataset yang digunakan terdiri atas 216 data siswa, dengan atribut yang mencakup nilai akademik, pendapatan

orang tua, status keluarga ekonomi tidak mampu (KETM), serta variabel demografis lainnya. Data tersebut dibagi menjadi data latih (75%) dan data uji (25%) menggunakan teknik stratified sampling untuk menjaga proporsi distribusi kelas secara seimbang.

3.2. Evaluasi Data Uji

Model kemudian diuji pada 54 data uji untuk menilai kemampuan generalisasi terhadap data baru. Hasil klasifikasi menunjukkan tingkat akurasi sebesar 85,19%, yang mengindikasikan adanya sedikit penurunan dibandingkan dengan dataset pelatihan, namun tetap berada pada kategori sangat baik. Penurunan ini dapat dipahami mengingat kompleksitas dataset uji serta kemungkinan adanya noise, tetapi tetap menunjukkan kestabilan kinerja model. Selain akurasi, model juga memperlihatkan distribusi klasifikasi yang seimbang pada nilai precision dan recall, yang mengindikasikan bahwa model tidak bias terhadap salah satu kelas target.

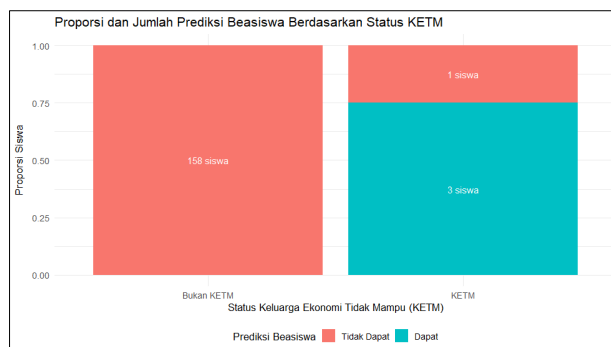
3.3. Diskusi

Hasil penelitian ini menegaskan bahwa algoritma Support Vector Machine (SVM) dengan linear kernel dapat membangun model klasifikasi yang efektif untuk memprediksi kelayakan siswa sebagai penerima beasiswa Program Indonesia Pintar (PIP) berdasarkan fitur utama seperti kinerja akademik, pendapatan orang tua, dan status sosial-ekonomi (KETM). Temuan ini sejalan dengan hasil penelitian sebelumnya, seperti oleh Wahyuni et al. (2021) dan Prasetyo & Santosa (2020), yang menunjukkan ketangguhan SVM dalam tugas klasifikasi di bidang pendidikan karena kemampuannya dalam menangani data berdimensi tinggi serta memaksimalkan margin antar kelas. Bahkan dengan dataset yang relatif kecil, model ini menunjukkan kemampuan generalisasi yang menjanjikan, dengan tingkat akurasi klasifikasi keseluruhan sebesar 85,19% pada data uji serta skor Cohen's Kappa sebesar 0,64. Meskipun kinerjanya menjanjikan, analisis terhadap hasil prediksi (Gambar 5 dan 6) menunjukkan bahwa model cenderung underpredict penerima beasiswa, khususnya pada siswa dari keluarga ekonomi tidak mampu. Sebagai contoh, meskipun terdapat empat siswa dari kategori KETM pada data uji, hanya tiga yang berhasil diklasifikasikan dengan benar sebagai penerima, sementara satu lainnya tetap salah klasifikasi. Hal ini mengindikasikan adanya masalah ketidakseimbangan kelas, di mana kelas minoritas (penerima beasiswa) kurang terwakili sehingga menyebabkan nilai sensitivity lebih rendah meskipun akurasi keseluruhan tinggi. Permasalahan serupa juga dilaporkan dalam penelitian Syakur et al. (2020), yang menyoroti sensitivitas SVM terhadap distribusi kelas yang tidak seimbang serta merekomendasikan mitigasi melalui teknik oversampling atau cost-sensitive learning.

```
> Prediksi = predict(svm.Lin, newdata=datBaru[-8])
> datBaru$Beasiswa = Prediksi
> datBaru
```

No	NamaSiswa	JenisKelamin	AsalSekolah	Nilai	Penghasilan	KETM	Beasiswa
1	ABDULLAH FATTAH	Laki - Laki	SMP NEGERI 1 TALUN	-0.15032367	-0.9984924871	ada	0
2	ADHZARIA JUNYANISA NUR FARDIAH	Laki - Laki	SMP NEGERI 4 CIREBON	-0.96664045	-1.1136436967	ada	0
3	ALIYA TASYADEVITA SHIDQIN	Perempuan	SMP NEGERI 11 CIREBON	0.73909610	-0.8065738044	tidak	0
4	ALVIAN AZZAHRA	Laki - Laki	SMP NEGERI 7 CIREBON	-0.86916979	-0.7298063313	tidak	0
5	ANDHIKA BAGUS SAPUTRA	Laki - Laki	MTS NEGERI 2 CIREBON	-0.61330931	-0.2308177564	tidak	0
6	ANGELINA PUTRI RAMADANI	Laki - Laki	SMP NEGERI 8 CIREBON	-0.06503685	-1.1904111697	tidak	0
7	ANGGA ADITYA	Laki - Laki	SMP NEGERI 6 CIREBON	-1.22250093	0.9590790761	tidak	0
8	ARIEF ADHITYA RAMADHAN	Laki - Laki	SMP NEGERI 11 CIREBON	1.04369191	-1.3554612368	ada	0
9	AURAN FATA MOHAMMAD	Perempuan	SMP NEGERI 10 CIREBON	-0.43055182	-1.2671786428	tidak	0
10	BAYU AJI	Laki - Laki	SMP NEGERI 8 CIREBON	0.64162544	-0.1924340199	tidak	0

Gambar 5. Output Prediksi SVM untuk 10 Siswa Uji. Terlihat Bahwa Seluruh Hasil Prediksi Menghasilkan Kelas 0 (tidak menerima beasiswa).



Gambar 6. Jumlah Prediksi Penerima Beasiswa berdasarkan Status KETM

Dari sisi implikasi kebijakan, model yang dikembangkan saat ini berfungsi sebagai komponen dasar dalam membangun sistem pendukung keputusan berbasis data untuk membantu proses seleksi beasiswa. Sistem semacam ini dapat mengurangi subjektivitas dan meningkatkan transparansi dalam distribusi beasiswa, khususnya di lingkungan pendidikan vokasi. Namun, untuk penerapan yang lebih luas dan tingkat keadilan yang lebih tinggi, diperlukan penyempurnaan lebih lanjut. Penelitian selanjutnya disarankan untuk mengeksplorasi metode resampling (misalnya SMOTE), penyesuaian ambang batas (threshold tuning), atau mengintegrasikan SVM ke dalam kerangka ensemble seperti bagging atau boosting guna meningkatkan kinerjanya pada dataset yang tidak seimbang. Selain itu, replikasi model ini di lembaga pendidikan lain dengan data demografis dan akademik yang serupa dapat memvalidasi skalabilitas serta kegunaannya dalam berbagai konteks.

4. KESIMPULAN

Penelitian ini menunjukkan efektivitas algoritma Support Vector Machine (SVM) dalam mengklasifikasikan calon penerima Program Indonesia Pintar (PIP) pada sekolah menengah kejuruan negeri di Cirebon. Berdasarkan hasil pelatihan model menggunakan 216 data siswa yang mencakup atribut sosial ekonomi dan akademik, algoritma SVM dengan linear kernel mencapai tingkat akurasi sebesar 90%. Pemrosesan data dilakukan menggunakan platform R dan dievaluasi dengan metrik kinerja seperti akurasi, presisi, recall, dan F1-score. Tingginya akurasi pada data pelatihan menunjukkan bahwa model mampu mengenali pola data dengan baik serta memberikan prediksi yang andal dalam konteks seleksi beasiswa berbasis data.

Pada tahap pengujian, menggunakan 54 data uji yang sebelumnya tidak dikenal, model mencapai akurasi sebesar 85,19%. Meskipun terjadi sedikit penurunan dibandingkan dengan data pelatihan, model tetap menunjukkan kinerja yang stabil dan seimbang antara kelas penerima dan bukan penerima beasiswa. Temuan ini membuktikan bahwa model SVM dengan linear kernel dapat diandalkan sebagai sistem pendukung keputusan dalam proses seleksi beasiswa yang menuntut objektivitas, efisiensi, dan akuntabilitas. Keberhasilan ini membuka peluang untuk mengembangkan pendekatan serupa di lembaga pendidikan lain, guna meningkatkan pemerataan akses pendidikan bagi siswa dari keluarga kurang mampu secara lebih sistematis dan berbasis bukti.

DAFTAR PUSTAKA

- [1] R. C. A. Fajardo, F. B. Yara, R. F. Ardeña, M. K. L. Hernandez, and J. C. T. Arroyo, "A Data-Driven Approach in Predicting Scholarship Grants of a Local Government Unit in the Philippines Using Machine Learning," vol. 72, no. 6, pp. 74–81, 2024, doi: 10.14445/22315381/IJETT-V72I6P108.
- [2] I. R. W. Astuti et al., "Meningkatkan Motivasi Siswa untuk Melanjutkan Pendidikan ke Perguruan Tinggi," ADMA J. Pengabd. dan Pemberdaya. Masy., vol. 5, no. 2, pp. 387–406, 2025, doi: 10.30812/adma.v5i2.4421.
- [3] P. Nikumbh, K. Guladhe, A. Mishra, R. Wankar, R. Yadav, and J. Avhad, "Scholarspot: Scholarship Recommendation System using Machine Learning," 2024. doi: 10.1109/ICCUBE61740.2024.10775136.
- [4] M. D. T. Wahyuni and I. P. A. Prabawati, "Transparansi dan Ketepatan Sasaran Dalam Penyaluran Program Indonesia Pintar: Studi Kasus Sekolah Dasar di Kecamatan Kuta Utara," Socius J. Penelit. Ilmu ..., vol. 02, no. June, pp. 257–262, 2025.
- [5] F. Hooshmand and M. Nikoomanesh, "Integration of machine constraint learning methods within optimization models," Appl. Math. Model., vol. 146, p. 116184, 2025, doi: <https://doi.org/10.1016/j.apm.2025.116184>.
- [6] P. Castañeda, "Predictive Analysis of Student Dropout in Higher Education," 2025, pp. 303–309. doi: 10.1145/3724504.3724554.
- [7] P. Xuan Lam, P. Q. H. Mai, Q. H. Nguyen, T. Pham, T. H. H. Nguyen, and T. H. Nguyen, "Enhancing educational evaluation through predictive student assessment modeling," Comput. Educ. Artif. Intell., vol. 6, p. 100244, 2024, doi: <https://doi.org/10.1016/j.caeai.2024.100244>.
- [8] J. Li, "Prediction of University Students' Scholarship Grade for Excellent Students Based on Machine Learning," 2023, pp. 416–420. doi: 10.1109/ICEACE60673.2023.10442348.
- [9] N. Nurhayati and K. Lee, "Ensemble learning techniques to improve the accuracy of predictive model performance in the scholarship selection process," vol. 4, no. 3, pp. 264–275, 2023, doi: 10.47738/jads.v4i3.112.

-
- [10] R. Jha, "A Systematic Study on Student Performance Prediction from the Perspective of Machine Learning and Data Mining Approaches," 2023, pp. 1336–1342. doi: 10.1109/ICCSE57224.2023.10192866.
- [11] S. Al-Hagree et al., "A Comparison of Machine Learning Techniques for Yemeni Universities Admission Examinations Predictions," 2024. doi: 10.1109/ICETI63946.2024.10777278.
- [12] J. Ji, "An automatic scholarships evaluation method based on machine learning," vol. 17, no. 39, pp. 3.1-3.4, 2016, doi: 10.5013/IJSSST.a.17.39.03.
- [13] J.-W. Lee, K.-S. Park, Y.-J. Jun, and S.-H. Eom, "A study on the advancement of building energy simulations: Utilizing ChatGPT to enhance occupancy rate predictions and sustainability outcomes," *Int. J. Air-Conditioning Refrig.*, vol. 33, no. 1, 2025, doi: 10.1007/s44189-025-00077-z.
- [14] K. Swetha, S. Thaila Shree, M. Soundaravalli, and A. Lincy, "Intelligent Enquiry Chatbot Using AI," 2025, vol. 3204, no. 1. doi: 10.1063/5.0245932.
- [15] M. Arashpour et al., "Predicting individual learning performance using machine-learning hybridized with the teaching-learning-based optimization," vol. 31, no. 1, pp. 83–99, 2023, doi: 10.1002/cae.22572.
- [16] A. T. Bere, Y. P. K. Kelen, and F. R. Bobu, "Penentuan Kelayakan Penerima Bantuan Beasiswa Berbasis Sistem Pendukung Keputusan Menggunakan Metode Multy Attribute Utility Theory (Maut) Pada Smkn 1 Kefamenanu," *TeknoIS J. Ilm. Teknol. Inf. dan Sains*, vol. 15, no. 1, pp. 62–73, 2025, doi: 10.36350/jbs.v15i1.292.
- [17] W. S. Maulina and D. Kurniadi, "Perancangan Aplikasi Rekomendasi Calon Penerima Beasiswa Menggunakan Metode Simple Additive Weighting dan Rapid Throwaway Prototyping Model," pp. 527–539, 2025, doi: 10.33364/algorithm/v.22-1.1778.
- [18] A. L. Febriansyah and V. Ratnawati, "Analisis Korelasional Motivasi Belajar dengan Kepercayaan Diri pada Siswa Penerima Beasiswa Progam Indonesia Pintar di MAN 1 Kota Kediri," pp. 997–1004, 2025.
- [19] A. L. Febriansyah and Atrup, "Analisis Motivasi Belajar Pada Penerima Beasiswa Progam Indonesia Pintar Di MAN 1 Kota Kediri Tahun Ajaran 2024/2025," pp. 233–240, 2025.
- [20] M. A. Shafii, I. Zakiya, D. Fitriyani, A. Arkundato, and A. G. Abdullah, "Neutron flux distribution of slab reactor core using one-dimensional multi-group diffusion equation in the thermal energy region," in *AIP Conference Proceedings*, 2022, vol. 2663. doi: 10.1063/5.0108170.
- [21] M. R. Nihan Kristiyani, Azainil, Haerudin, "PENERAPAN SISTEM INFORMASI INDONESIA PINTAR ENTERPRISE DALAM PENYALURAN BEASISWA PROGRAM INDONESIA PINTAR," *Sustain.*, vol. 11, no. 1, pp. 1–14, 2025.
- [22] N. B. Sitepu, M. Zarlis, S. Efendi, and H. W. Dhany, "Analysis of Decision Tree and Smooth Support Vector Machine Methods on Data Mining," 2019, vol. 1255, no. 1. doi: 10.1088/1742-6596/1255/1/012067.
- [23] S. Cherkezova, "Students' mastery goal orientation, academic achievement and education-to-employment transition attitudes. Evidence from 30 countries," *Environ. Soc. Psychol.*, vol. 9, no. 5, 2024, doi: 10.54517/esp.v9i5.2313.
- [24] S. Maniyan, R. Ghousi, and A. Haeri, "Data mining-based decision support system for educational decision makers: Extracting rules to enhance academic efficiency," *Comput. Educ. Artif. Intell.*, vol. 6, p. 100242, 2024, doi: <https://doi.org/10.1016/j.caeai.2024.100242>.