

# Klasifikasi Penyakit Diabetes Menggunakan Particle Swarm Optimaze Pada Algoritma Support Vector Machine

\*Sena Daniswara<sup>1</sup>, Zalfa Ibtisamah Arishandy<sup>2</sup>, Eva Yulia Puspaningrum<sup>3</sup>

<sup>1,2,3</sup> Program Studi Informatika, UPN "Veteran" Jawa Timur, Indonesia

**Artikel Histori:**

Disubmit: Oktober 2025  
Diterima: November 2025  
Diterbitkan: Desember 2025

**DOI**

[10.33005/jifti.v7i2.163](https://doi.org/10.33005/jifti.v7i2.163)



**ABSTRAK**

Penyakit diabetes mellitus menjadi salah satu tantangan kesehatan global dengan prevalensi yang terus mengalami peningkatan dari tahun ke tahun. Penelitian ini bertujuan untuk mengklasifikasikan penyakit diabetes dengan menerapkan algoritma Support Vector Machine (SVM) yang ditingkatkan performanya melalui Particle Swarm Optimization (PSO) serta teknik Synthetic Minority Over-sampling Technique (SMOTE). Dataset pada penelitian ini bersumber dari National Institute of Diabetes and Digestive and Kidney Diseases dan terdiri dari 768 data. Penelitian mencakup tahap praproses data, penanganan ketidakseimbangan kelas dilakukan dengan metode SMOTE, optimasi parameter SVM dilakukan dengan bantuan PSO, dan evaluasi kinerja model dilakukan menggunakan metrik akurasi, presisi, recall, serta F1-score. Hasil pengujian menunjukkan bahwa model SVM-SMOTE-PSO mencapai akurasi sebesar 83,95%, meningkat dibandingkan model SVM-SMOTE tanpa PSO yang hanya mencapai 82,72%. Peningkatan ini terutama terlihat pada prediksi kelas minoritas, di mana PSO membantu mengoptimalkan parameter model SVM. Dengan demikian, metode ini terbukti efektif dalam meningkatkan akurasi dan keseimbangan prediksi klasifikasi penyakit diabetes.

*Kata Kunci:* Diabetes mellitus, klasifikasi, Particle Swarm Optimization, Support Vector Machine, SMOTE.

**How to Cite:**

Sena Daniswara, Zalfa Ibtisamah Arishandy, Eva Yulia Puspaningrum. (2025). Klasifikasi Penyakit Diabetes Menggunakan Particle Swarm Optimaze Pada Algoritma Support Vector Machine. *Jurnal Ilmiah Teknologi Informasi dan Robotika*, 7(2), 20-28. <https://doi.org/10.33005/jifti.v7i2.163>.

**\*Corresponding Author:**

Email : 22081010296@student.upnjatim.ac.id  
Alamat : Jl. Rungkut Madya, Gn. Anyar



This article is published under a Creative Commons Attribution 4.0 International License.

## PENDAHULUAN

Diabetes melitus merupakan salah satu penyakit kronis dengan prevalensi yang terus meningkat secara global, termasuk di Indonesia. Rahmah, dkk (2023) menyatakan sesuai data dari *World Health Organization* (WHO), diabetes menyumbang sekitar 4% dari total kematian dunia, dengan jumlah penderita yang meningkat dari tahun ke tahun. Penyakit ini disebabkan oleh ketidakmampuan tubuh menghasilkan insulin dalam jumlah yang cukup atau pemanfaatan insulin yang kurang optimal. Oleh sebab itu, deteksi dini dan klasifikasi diabetes menjadi hal yang penting untuk mencegah terjadinya komplikasi yang lebih berat sesuai dengan penelitian Putri, dkk (2023).

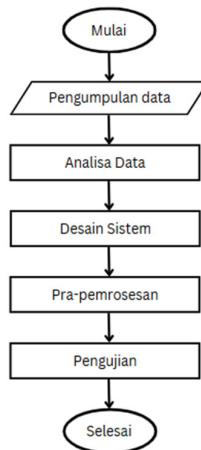
Dalam upaya meningkatkan akurasi klasifikasi diabetes, berbagai metode dan algoritma telah dikembangkan, salah satunya adalah *Support Vector Machine* (SVM). SVM dikenal sebagai salah satu metode pembelajaran mesin yang cukup efektif dalam pengklasifikasian *data non-linear*. Namun demikian, kendala utama dalam penerapan SVM terletak pada pemilihan fitur yang relevan guna meningkatkan performa model seperti penelitian dari Setiawan, dkk (2022). Dalam konteks ini, *Particle Swarm Optimization* (PSO) dapat digunakan sebagai teknik optimasi untuk memilih fitur-fitur terbaik yang akan digunakan dalam model SVM.

PSO merupakan algoritma optimasi berbasis populasi yang terinspirasi dari perilaku sosial hewan, seperti burung atau ikan. Sesuai dengan penelitian Patricia (2024) metode ini telah terbukti efektif dalam berbagai permasalahan optimasi, termasuk pada proses seleksi fitur untuk meningkatkan akurasi klasifikasi data medis. Dengan mengintegrasikan PSO ke dalam SVM, diharapkan kinerja klasifikasi diabetes dapat meningkat melalui pengurangan dimensi data serta penekanan pada fitur-fitur paling berpengaruh terhadap proses diagnosis secara lebih tepat dan terarah. Penelitian dari Yusuf, dkk (2021) dan Maulidah, dkk (2020) menunjukkan bahwa kombinasi PSO dengan algoritma klasifikasi lain, seperti *Naive Bayes* dan *Decision Tree*, mampu memberikan peningkatan signifikan pada akurasi klasifikasi diabetes. Sehingga, penelitian yang dilakukan berfungsi agar dapat mengetahui kinerja PSO dalam meningkatkan performa SVM pada klasifikasi penyakit diabetes. Dengan begitu, penelitian yang dilakukan tidak hanya berkontribusi pada pengembangan metode klasifikasi diabetes, tetapi memberikan wawasan tambahan mengenai penerapan teknik optimasi dalam bidang kesehatan, khususnya pada data klinis yang kompleks.

Melalui penelitian yang dilakukan, diharapkan memperoleh model klasifikasi yang lebih akurat dan efisien dalam mendeteksi diabetes melitus. Hasil penelitian juga diharapkan dimanfaatkan sebagai dasar peningkatan kajian sistem pendukung keputusan untuk diagnosis diabetes yang lebih baik dan dapat diandalkan pada masa mendatang.

## METODE PENELITIAN

Menurut Ibnu Sina (2022), metode penelitian merupakan rangkaian proses keilmuan yang dipergunakan untuk mengumpulkan informasi yang valid dengan tujuan menemukan, mengembangkan, atau membuktikan suatu pengetahuan tertentu, sehingga dapat dimanfaatkan untuk memahami, memecahkan, serta mengantisipasi permasalahan dalam suatu kajian. Berdasarkan pemahaman tersebut, penulis menyatakan bahwa metode penelitian adalah langkah terstruktur yang diterapkan untuk menangani persoalan atau memperoleh informasi yang dapat dipertanggungjawabkan. Agar dapat mendukung proses klasifikasi di penelitian, dirancanglah serangkaian alur yang akan dilaksanakan selama kegiatan kejadian berjalan. Gambar 1 memperlihatkan alur proses tersebut.



Gambar. 1 Alur Metodologi Penelitian  
Sumber: Data Diolah

#### A. Pengumpulan Data

Data yang digunakan berbentuk *data sekunder* yang telah ada sebelumnya. Metode ini dilakukan melalui pengambilan data secara tidak langsung dari sumber aslinya, sehingga penulis memanfaatkan dataset yang telah tersedia untuk kebutuhan penelitian. Data yang digunakan berasal dari *National Institute of Diabetes and Digestive and Kidney Diseases*, yang dipublikasikan sebagai dataset pada situs Kaggle. Dataset tersebut memuat beberapa variabel seperti pada tabel 1, yaitu *Pregnancies*, *Glucose*, *Blood Pressure*, *Skin Thickness*, *Insulin*, *BMI*, *Diabetes Pedigree Function*, dan *Age*. Dataset yang digunakan sebagai data penelitian dengan jumlah total 768 data, yang selanjutnya dimanfaatkan dalam proses pengolahan dan pengujian sesuai tahapan yang direncanakan.

**Tabel 1**  
**Variabel yang digunakan**

No	Atribut	Deskripsi
1.	<i>Pregnancies</i>	Tingkat kehamilan
2.	<i>Glucose</i>	Kadar glukosa
3.	<i>Blood Pressure</i>	Tekanan darah
4.	<i>Skin Thickness</i>	Ketebalan kulit
5.	<i>Insulin</i>	Insulin
6.	<i>BMI</i>	Indeks massa tubuh
7.	<i>Diabetes Pedigree Function</i>	Fungsi keturunan diabetes
8.	<i>Age</i>	Usia

#### B. Analisa Data

Ahli statistik John Tukey (1961) mendefinisikan analisis data sebagai serangkaian prosedur untuk mengolah data, teknik dalam menafsirkan hasil yang diperoleh, serta

pendekatan dalam merencanakan proses pengumpulan data agar kegiatan analisis dapat dilakukan secara lebih efektif dan akurat. Analisis data tidak hanya berfokus pada perhitungan statistik semata, tetapi juga mencakup upaya memahami pola, hubungan, dan makna yang ada di dalam data sehingga berfungsi untuk membantu pengambilan keputusan yang lebih baik dan berbasis data.

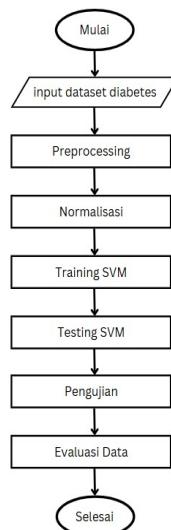
Dalam konteks penelitian ini, analisis data dilakukan dengan tujuan memperoleh hasil yang lebih akurat dan objektif, di mana seluruh proses pengolahan mengikuti kaidah statistik dan matematis yang berlaku untuk memastikan validitas hasil analisis [8]. Data yang dianalisis merupakan dataset diabetes yang diperoleh dari situs Kaggle, yang bersumber dari *National Institute of Diabetes and Digestive and Kidney Diseases*. Proses analisis ini bermaksud agar dapat menggali informasi lebih mendalam mengenai karakteristik dataset yang digunakan, sehingga data tersebut dapat dimanfaatkan secara optimal dalam penelitian klasifikasi penyakit diabetes.

### C. Desain Sistem

Desain sistem dalam penelitian ini dapat diartikan sebagai proses pendefinisian kebutuhan fungsional yang diperlukan untuk membangun sebuah sistem yang mampu mendukung tujuan penelitian. Tahapan ini mencakup perencanaan, perancangan, serta penyusunan kerangka implementasi yang menggambarkan bagaimana setiap komponen sistem saling terintegrasi. Desain sistem juga berfungsi sebagai panduan dalam menyatukan berbagai elemen yang terpisah menjadi satu kesatuan sistem yang utuh, terstruktur, dan dapat berfungsi secara optimal, yang mana ini sesuai dengan penelitian dari Maulidah (2020).

#### 1) Alur Klasifikasi SVM-SMOTE

Dataset diabetes yang digunakan dalam penelitian ini berisi informasi terkait kondisi pasien diabetes dan dijadikan sebagai dasar dalam pembangunan model prediksi. Sebelum dilakukan analisis lebih lanjut, data terlebih dahulu melalui tahap praproses untuk membersihkan data, menangani nilai yang hilang atau tidak konsisten, serta menyesuaikan format data agar sesuai dengan algoritma yang digunakan. Tahapan ini bertujuan untuk meningkatkan kualitas data sehingga model yang dihasilkan dapat bekerja secara lebih optimal. Gambar 2 memperlihatkan secara keseluruhan tahapan yang dilakukan pada alur klasifikasi ini.



Gambar. 2 Alur Klasifikasi SVM-SMOTE  
Sumber: Data Diolah

Langkah selanjutnya adalah proses normalisasi, di mana seluruh data diubah ke dalam skala yang sama untuk memastikan masing-masing fitur punya kontribusi yang seimbang pada tahapan *training* model. Untuk mengatasi permasalahan *imbalance dataset*, digunakanlah metode *Synthetic Minority Over-sampling Technique* (SMOTE). Metode ini digunakan untuk penambahan data buatan pada kelas minoritas (pasien diabetes), sehingga distribusi data antar kelas menjadi seimbang serta potensi bias model untuk kelas yang lebih besar dapat diminimalkan.

Setelah data diproses dan diseimbangkan, algoritma *Support Vector Machine* (SVM) dilatih menggunakan data tersebut untuk menemukan hyperplane atau batas pemisah terbaik antara kelas pasien diabetes dan non-diabetes. Model SVM yang telah dilatih kemudian diuji menggunakan data baru yang belum pernah digunakan sebelumnya guna mengukur kemampuan prediksi model secara objektif. Tahap pengujian ini dapat dilakukan melalui validasi silang (*cross-validation*) untuk memastikan kinerja model yang lebih andal dan stabil.

Hasil pengujian selanjutnya dievaluasi menggunakan metrik kinerja seperti akurasi, *presisi*, *recall*, dan *F1-score*. Evaluasi ini memberikan gambaran menyeluruh mengenai kemampuan model dalam memprediksi kelas positif (diabetes) maupun kelas negatif (*non-diabetes*) secara tepat dan konsisten. Dengan demikian, alur klasifikasi SVM-SMOTE yang diterapkan diharapkan mampu menghasilkan model klasifikasi diabetes yang lebih akurat, efisien, dan dapat diandalkan sebagai dasar pengambilan keputusan di bidang kesehatan.

## 2) Alur Klasifikasi SVM-SMOTE-PSO

Dataset diabetes yang digunakan dalam penelitian ini memuat informasi mengenai pasien diabetes dan dijadikan sebagai dasar dalam membangun model prediksi. Secara keseluruhan tahapan ini dapat digambarkan seperti pada gambar 3. Tahap awal yang dilakukan adalah *preprocessing*, yaitu pembersihan data untuk menangani data yang hilang atau tidak konsisten, serta melakukan penyesuaian format data agar sesuai dengan algoritma yang akan digunakan. Selanjutnya, dilakukan normalisasi untuk memastikan seluruh fitur berada pada skala yang sama sehingga kontribusi setiap fitur dalam proses pembelajaran menjadi lebih seimbang.

Untuk mengatasi permasalahan *data imbalance*, diterapkanlah metode SMOTE (*Synthetic Minority Over-sampling Technique*). Metode ini menggenerate data buatan untuk kelas minoritas (pasien diabetes) sehingga komposisi data antar kelas tidak *imbalance* dan model tidak cenderung bias kepada kelompok yang lebih banyak. Metode optimasi berbasis populasi seperti *Particle Swarm Optimization* (PSO) juga digunakan untuk meningkatkan kinerja evaluasi akhir. Proses PSO diawali dengan diawali partikel secara acak, yang mana pada masing-masing partikel memiliki letak dan kecepatan yang merepresentasikan solusi, serta mengalami pembaruan parameter pada setiap iterasi. PSO memperbarui posisi dan kecepatan partikel secara iteratif berdasarkan evaluasi fungsi *fitness*, sehingga partikel dengan nilai *fitness* yang lebih baik dapat memengaruhi partikel lainnya untuk menemukan solusi optimal. Dalam penelitian ini, fungsi *fitness* diimplementasikan untuk menghitung akurasi model untuk memperkirakan kemungkinan diabetes.

Setelah proses optimasi selesai, algoritma klasifikasi yang digunakan adalah model *Support Vector Machine* (SVM). SVM melakukan penentuan *hyperplane* yang optimal untuk mengelompokkan data antara kelas positif (diabetes) dan kelas negatif (*non-diabetes*). Model yang selesai di-*train* kemudian diuji menggunakan data baru untuk mengukur performa serta memastikan kemampuan generalisasi. Proses pengujian dapat melibatkan *cross-validation* agar estimasi kinerja model lebih andal. Selanjutnya, hasil pengujian diuji dengan pengujian kinerja seperti akurasi, *presisi*, *recall*, dan *F1-score*, sehingga dapat memperkirakan hasil evaluasi model untuk memprediksi kelas positif maupun negatif.

Tahap evaluasi ini membantu memastikan model telah berfungsi dengan baik dalam menyelesaikan permasalahan klasifikasi diabetes.



Gambar. 3 Alur Klasifikasi SVM-SMOTE-PSO  
Sumber: Data Diolah

#### D. Pra-Pemrosesan

Tahap awal dalam pra-pemrosesan adalah normalisasi. Tahapan ini, dataset dinormalisasi dengan menyesuaikan variabel numerik ke dalam jangka nilai yang seragam. Tujuannya adalah mengurangi perbedaan skala antar atribut, sehingga setiap atribut dapat memberikan kontribusi yang seimbang terhadap model.

Tahap berikutnya adalah penanganan ketidakseimbangan data (*data imbalance*). Tahapan ini berupaya untuk meng-handle distribusi kelas yang tidak seimbang pada dataset diabetes yang digunakan, dengan tujuan meningkatkan akurasi model klasifikasi yang akan diterapkan. Penerapan *SMOTE* (*Synthetic Minority Over-sampling Technique*) dilakukan untuk menangani data yang bersifat imbalance.

Tahap ketiga adalah pembagian dataset. Dataset diabetes yang telah diproses menggunakan SMOTE kemudian dibagi menjadi 2 kelompok, yaitu data pelatihan (*train*) serta data pengujian (*test*). Hasil pembagian dataset menghasilkan 537 data untuk train serta 231 data untuk test.

#### E. Pengujian

Tahapan pengujian dilakukan untuk model sebelum penerapan SMOTE dan setelah penerapan SMOTE. Pengujian dilakukan untuk mengetahui pengaruh SMOTE terhadap distribusi kelas data dan juga performa model klasifikasi yang dibangun. Dan juga,

pengujian yang dilakukan diharapkan dapat memastikan kinerja model mampu melakukan prediksi secara konsisten dan memiliki kemampuan sintesis yang baik untuk data baru yang mana belum pernah ada di *data train* atau data yang ada.

Untuk memperoleh hasil yang lebih andal, proses pengujian dapat melibatkan validasi silang (*cross-validation*) sehingga estimasi performa model menjadi lebih stabil. Hasil pengujian kemudian dievaluasi menggunakan metrik seperti akurasi, presisi, *recall*, dan *F1-score* sebagai dasar dalam menilai kemampuan model dalam mengklasifikasikan kelas positif (diabetes) maupun kelas negatif (non-diabetes).

Selanjutnya, untuk memperjelas dampak penerapan SMOTE, dilakukan perbandingan distribusi kelas data pada kondisi sebelum dan sesudah proses SMOTE. Perbandingan ini ditampilkan pada Gambar 4, yang menunjukkan distribusi kelas sebelum proses SMOTE, serta perubahan yang terjadi setelah dilakukan penyeimbangan data. Secara distribusi kelas data sebelum SMOTE (*Synthetic Minority Over-sampling Technique*) yaitu Kelas 0 memiliki 237 sampel, lalu Kelas 1 memiliki 138 sampel. Masalahnya data tidak seimbang, karena jumlah sampel kelas 1 lebih sedikit dibandingkan dengan kelas 0.

```
==== Before SMOTE ====
Class distribution in training set: Counter({0: 237, 1: 138})
```

Gambar. 4 Sebelum Proses SMOTE

Sumber: Data Diolah

Pada Gambar. 5 menunjukkan distribusi kelas data setelah SMOTE. Kelas 0 memiliki 237 sampel, lalu Kelas 1 memiliki 237 sampel. Hasilnya SMOTE telah menambahkan sampel sintetis ke kelas minoritas (kelas 1), sehingga distribusi kelas menjadi seimbang.

```
==== After SMOTE ====
Class distribution in training set: Counter({0: 237, 1: 237})
```

Gambar. 5 Sesudah Proses SMOTE

Sumber: Data Diolah

## HASIL DAN PEMBAHASAN

Pada pengujian model *Support Vector Machine* (SVM) untuk menangani data tidak seimbang, teknik oversampling SMOTE digunakan sebagai upaya untuk meningkatkan performa prediksi, khususnya pada kelas minoritas. Pengujian dilakukan dengan beberapa kondisi, yaitu kondisi tanpa dan menggunakan tuning parameter pada *Particle Swarm Optimization* (PSO). Hasil pemrosesan menampilkan bahwa penggunaan SMOTE memperbaiki kemampuan model untuk memprediksi kelas minoritas, sementara penerapan PSO memberikan peningkatan lebih lanjut pada performa keseluruhan model. Evaluasi metrik akurasi, *precision*, *recall*, dan *F1-score*, memperlihatkan bahwa *tuning* parameter menggunakan PSO tidak hanya meningkatkan akurasi, selain itu memperbaiki keseimbangan prediksi kelas mayoritas dan minoritas. Hasil ini mengindikasikan bahwa kombinasi SMOTE dan PSO merupakan pendekatan yang andal dalam mengatasi tantangan data tidak seimbang.

### A. Hasil Pengujian SVM-SMOTE

Model SVM dengan teknik oversampling SMOTE tanpa PSO seperti pada gambar 6 menghasilkan akurasi sebesar 82,72%. Berdasarkan *confusion matrix*, model mampu memprediksi kelas mayoritas (kelas 0) dengan baik, dengan nilai precision 89% dan recall 86%, yang menunjukkan kemampuan model dalam mengenali data non-diabetes secara konsisten.

0s	[14]	== SVM + SMOTE (without PSO) ==
		Accuracy: 0.8272
		Confusion Matrix:
		[[95 16] [12 39]]
		Classification Report:
		precision recall f1-score support
	0	0.89 0.86 0.87 111
	1	0.71 0.76 0.74 51
		accuracy 0.83 162
	macro avg	0.80 0.81 0.80 162
	weighted avg	0.83 0.83 0.83 162

Gambar. 6 Hasil Pengujian SVM-SMOTE

Sumber: Data Diolah

Pada kelas minoritas (kelas 1), Pengujian menghasilkan *precision* sebesar 71% dan *recall* 76%, yang menandakan bahwa SMOTE cukup membantu meningkatkan sensitivitas model terhadap data diabetes meskipun masih terdapat *error* pada klasifikasi. Secara keseluruhan, nilai *F1-score* berbobot mencapai 83%, yang mencerminkan keseimbangan *precision* dan *recall* yang ada pada kedua kelas tersebut.

Meskipun hasil ini menunjukkan performa yang cukup baik, nilai *recall* pada kelas minoritas masih relatif rendah dibandingkan kelas mayoritas. Hal ini mengindikasikan bahwa model memiliki keterbatasan dalam mengenali seluruh pola data diabetes, sehingga diperlukan pendekatan lanjutan untuk meningkatkan performa klasifikasi pada kelas minoritas.

#### B. Hasil Pengujian SVM-SMOTE

Model SVM yang dikombinasikan dengan SMOTE dan optimasi parameter menggunakan PSO menunjukkan akurasi yaitu 83,95%, yang menunjukkan kinerja lebih baik dibandingkan model tanpa PSO. Pada kelas mayoritas (kelas 0), model mencapai *precision* 90% dan *recall* 86%, menandakan bahwa kemampuan prediksi terhadap data non-diabetes tetap terjaga dengan baik setelah proses optimasi. Untuk kelas minoritas (kelas 1), nilai *precision* meningkat menjadi 73% dan *recall* menjadi 78%, yang menunjukkan adanya perbaikan signifikan dibandingkan model SVM-SMOTE tanpa PSO. Peningkatan ini mengindikasikan bahwa PSO berhasil membantu model menemukan parameter yang lebih optimal sehingga sensitivitas terhadap data diabetes meningkat.

Secara keseluruhan, seperti digambarkan pada gambar 7 nilai *F1-score* berbobot meningkat menjadi 84%, yang mencerminkan keseimbangan prediksi yang lebih baik antara kelas mayoritas dan minoritas. Hasil ini membuktikan bahwa integrasi PSO tidak hanya meningkatkan akurasi model, tetapi juga memperbaiki kinerja klasifikasi pada data tidak seimbang, khususnya dalam mendeteksi kelas minoritas.

0s	[14]	== SVM + SMOTE + PSO ==
		Accuracy: 0.8395
		Confusion Matrix:
		[[96 15] [11 40]]
		Classification Report:
		precision recall f1-score support
	0	0.90 0.86 0.88 111
	1	0.73 0.78 0.75 51
		accuracy 0.84 162
	macro avg	0.81 0.82 0.82 162
	weighted avg	0.84 0.84 0.84 162

Gambar. 7 Hasil Pengujian SVM-SMOTE-PSO

Sumber: Data Diolah

## **SIMPULAN**

Berdasarkan Berdasarkan hasil penelitian mengenai klasifikasi penyakit diabetes dengan algoritma *Support Vector Machine* (SVM) yang dioptimalkan dengan *Particle Swarm Optimization* (PSO) serta teknik SMOTE (*Synthetic Minority Over-sampling Technique*), dapat disimpulkan bahwa penerapan SMOTE berhasil mengatasi *data imbalance* pada dataset diabetes dengan menambahkan data yang telah digeneralisasi pada kelas minoritas, menghasilkan performa model dalam memprediksi pasien diabetes menjadi lebih baik. Kombinasi PSO dalam proses optimasi parameter bersama SVM dan SMOTE juga terbukti mampu meningkatkan akurasi klasifikasi. Model SVM-SMOTE tanpa PSO memperoleh akurasi sebesar 82,72%, sedangkan setelah diterapkan PSO akurasinya meningkat menjadi 83,95%. Selain itu, model SVM-SMOTE-PSO menunjukkan peningkatan pada prediksi kelas minoritas dengan nilai *precision* dan *recall* yang lebih baik dibandingkan SVM-SMOTE tanpa PSO. Dengan demikian, kombinasi SVM, SMOTE, dan PSO dinilai efektif untuk meningkatkan akurasi sekaligus menjaga keseimbangan kinerja model pada data yang *imbalance*. Hasil penelitian ini juga menunjukkan kontribusi pada pengembangan bidang pendukung keputusan untuk membantu diagnosis diabetes yang lebih baik dan efisien di masa datang.

## **DAFTAR PUSTAKA**

- Astriana, R. N., Sepriyanti, M. H. Zikri, I. Ambarani, dan M. Y. B. Shahar, "Implementation of Support Vector Machine and Random Forest for Heart Failure Disease Classification," *PREDATECS: Public Research Journal of Engineering, Data Technology and Computer Science*, vol. 1, no. 1, pp. 1–10, 2023.
- Wulandari, N. D., Aprilia, W. K., Mayangsari, A., & Nugraha, B. S. P. (2022). *PEMANFAATAN TEKNOLOGI TEPAT GUNA UNTUK MENGEMBANGKAN POTENSI MASYARAKAT MELALUI KEMAJUAN UMKM*. <https://new-conference.unisma.ac.id/index.php/KOPEMAS/article/view/1060>.
- Setiawan, D., Putri, R. N., & Suryanita, R. (2019, June 9). *Perbandingan Algoritma Genetika dan Backpropagation pada Aplikasi Prediksi Penyakit Autoimun*. Setiawan | Khazanah Informatika : Jurnal Ilmu Komputer Dan Informatika. <https://journals.ums.ac.id/khif/article/view/7173>.
- Fauzi, A., & Yunial, A. H. (2022). Optimasi Algoritma Klasifikasi Naive Bayes, Decision Tree, K – Nearest Neighbor, dan Random Forest menggunakan Algoritma Particle Swarm Optimization pada Diabetes Dataset. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 8(3), 470. <https://doi.org/10.26418/jp.v8i3.56656>.
- Yusuf, A. L., Misdram, M., & Alamsyah, M. (2024). Prediksi Persediaan Bahan Baku Makanan pada Rumah Makan Menggunakan Metode Apriori (Studi Kasus Rumah Makan Soto Ayam Kampung Pasar Ranggeh). Yusuf | COREAI: Jurnal Kecerdasan Buatan, Komputasi Dan Teknologi Informasi. <https://doi.org/10.33650/coreai.v5i1.6604>.
- Maulidah, N., Abdilah, A., Nurlelah, E., Gata, W., & Hasan, F. N. (2020). Seleksi fitur klasifikasi penyakit diabetes menggunakan Particle swarm Optimization (PSO) pada algoritma Naive Bayes. *journal.stekom.ac.id*. <https://doi.org/10.51903/elkom.v13i2.227>.
- Sina, I. (2022, September 18). *METODOLOGI PENELITIAN*. <https://repository.penerbitwidina.com/publications/556926/metodologi-penelitian>.
- Diabetes Dataset*. (2020, August 5). <https://www.kaggle.com/datasets/mathchi/diabetes-data-set>
- Ayuningtyas, Y., & Suartana, I. M. (2023). Klasifikasi Penyakit Stroke Menggunakan Support Vector Machine (SVM) dan Particle Swarm Optimization (PSO). *ejournal.unesa.ac.id*. <https://doi.org/10.26740/jinacs.v4n04.p451-457>.
- Rayadhani, W. A., & Rahardi, M. (2025). Comparative analysis of random Forest, SVM, and Naive Bayes for cardiovascular disease prediction. *Journal of Applied Informatics and Computing*, 9(6), 3234–3243. <https://doi.org/10.30871/jaic.v9i6.11451>.
- Goel, L., & Nagpal, J. (2022). A Systematic review of recent machine learning techniques for plant disease identification and classification. *IETE Technical Review*, 40(3), 423–439. <https://doi.org/10.1080/02564602.2022.2121772>.