

# Penerapan Algoritma Support Vector Machine Dalam Prediksi Diabetes Melitus Menggunakan Data Pasien RSUD Prof. Dr. Margono Soekarjo

1<sup>st</sup> Rafli Bayu Pratama  
Fakultas Informatika  
Universitas Telkom Purwokerto  
Purwokerto, Indonesia  
raflibayuu@student.telkomuniversity.a  
c.id

2<sup>nd</sup> Aditya Dwi Putro W  
Fakultas Informatika  
Universitas Telkom Purwokerto  
Purwokerto, Indonesia  
adityaw@telkomuniversity.ac.id

3<sup>rd</sup> Dian Kartika Sari  
Fakultas Informatika  
Universitas Telkom Purwokerto  
Purwokerto, Indonesia  
dianks@telkomuniversity.ac.id

**Abstrak** — Diabetes Melitus (DM) adalah penyakit kronis yang semakin meningkat prevalensinya setiap tahun dan menjadi salah satu tantangan utama di bidang kesehatan. Penelitian ini bertujuan membangun model prediksi risiko DM menggunakan algoritma Support Vector Machine (SVM) dengan data pasien dari RSUD Prof. Dr. Margono Soekarjo. Algoritma SVM dipilih karena kemampuannya dalam mengklasifikasikan data non-linear dengan baik. Penelitian ini menggunakan pendekatan data mining dengan tahapan preprocessing data, pemodelan, evaluasi, dan analisis. Data yang digunakan mencakup variabel usia, jenis kelamin, tekanan darah, kadar glukosa, dan indeks massa tubuh (BMI). Hasil penelitian menunjukkan bahwa SVM dengan kernel RBF memberikan performa terbaik dengan akurasi 93%, precision 93%, recall 94%, dan f1-score 93%. Analisis feature importance menunjukkan bahwa variabel glukosa dan BMI memiliki kontribusi terbesar dalam menentukan risiko diabetes, sedangkan tekanan darah, usia, dan jenis kelamin memiliki pengaruh yang lebih kecil. Dari hasil ini disimpulkan bahwa kombinasi SVM dengan kernel RBF dan pendekatan oversampling (SMOTE) merupakan metode paling optimal untuk memprediksi diabetes melitus pada dataset ini. Penelitian ini diharapkan dapat membantu tenaga medis dalam diagnosis dini dan memberikan dasar bagi pengembangan aplikasi prediksi risiko diabetes berbasis teknologi untuk mendukung pengambilan keputusan klinis yang lebih baik.

**Kata kunci**— diabetes melitus, SVM, klasifikasi, SMOTE, prediksi

## I. PENDAHULUAN

*Diabetes Melitus (DM)* merupakan gangguan metabolik yang menjadi tantangan besar dalam kesehatan global. Diabetes terjadi karena adanya gangguan pada organ pankreas yang mengakibatkan kondisi hiperglikemia atau terjadinya penurunannya jumlah insulin [1]. Insulin adalah hormon yang dihasilkan pankreas dan berperan penting dalam pengaturan kadar gula darah. Klasifikasi *DM* terdapat 2 jenis yaitu diabetes tipe 1 dan tipe 2 [2]. Diabetes tipe 1 disebabkan oleh respon *autoimun* pada tubuh menyerang sel-beta di pankreas, sehingga produksi insulin menjadi sangat minim bahkan berhenti sama sekali. Diabetes jenis ini dapat

berkembang pada segala usia, tetapi jenis ini paling sering terjadi pada anak-anak dan dewasa muda. Sedangkan Diabetes tipe 2 ini disebabkan akibat gangguan fungsi insulin dalam tubuh sehingga tubuh tidak merespon insulin secara efektif. Kondisi ini menyebabkan peningkatan kadar glukosa dalam darah dan memicu produksi insulin secara berlebihan. Diabetes tipe 2 ini mencakup 90% dari seluruh dunia [3].

Diabetes sering disebut sebagai *silent killer* karena gejala awalnya yang tidak disadari oleh sebagian besar penderita, padahal komplikasi yang ditimbulkannya sangat serius[4]. Penyakit ini menjadi penyebab utama berbagai gangguan kesehatan seperti kebutaan, gagal ginjal, stroke, serangan jantung, hingga amputasi anggota tubuh (Kemenkes, 2016). Kompleksitas penyakit ini menuntut penanganan medis yang berkelanjutan dan pendekatan prediktif untuk deteksi dini yang lebih efektif.

Berdasarkan data dari *IDF* di tahun 2021, diperkirakan 537 juta orang dewasa yang berusia antara 20 hingga 79 tahun diketahui hidup dengan diabetes. Jumlah ini diprediksi akan meningkat hingga 643 juta pada tahun 2030 dan mencapai 783 juta pada tahun 2045. Di Indonesia sendiri, prevalensi diabetes pada tahun 2021 mencapai 10,8% dari total populasi orang dewasa sebanyak 179 juta jiwa, atau sekitar 1 dari 9 orang. Angka ini menunjukkan urgensi untuk mengembangkan metode prediksi dini yang efektif guna menurunkan tingkat kematian dan komplikasi akibat diabetes[5].

Prediksi dini untuk *diabetes melitus* ini memiliki peran krusial dalam memastikan kesehatan masyarakat. Dalam menghadapi kompleksitas data kesehatan, pendekatan menggunakan teknik data mining ini menjadi pilihan yang tepat untuk dilakukannya prediksi yang dimana dapat meningkatkan efisiensi dan akurasi[6]. Dengan digunakannya algoritma prediktif seperti *Support Vector Machine (SVM)* memberikan solusi yang menjanjikan. Algoritma SVM yang merupakan algoritma telah terbukti baik dalam klasifikasi data [7], dengan memanfaatkan informasi klinis seperti riwayat medis, tes laboratorium, dan faktor risiko, algoritma ini bisa mengidentifikasi pola yang rumit dalam data dan memprediksi kemungkinan terjadinya *diabetes melitus*.

Beberapa penelitian terdahulu telah menunjukkan penerapan algoritma Support Vector Machine, seperti pada penelitian oleh Isyami (2024) [8] menggunakan SVM untuk melakukan klasifikasi pada masyarakat yang ada di Desa Cisarua mengenai keluarga mana yang akan menerima bantuan program keluarga harapan, dan hasilnya didapatkan nilai akurasinya mencapai 89,89%. Penelitian lain yang dilakukan oleh Yonal (2024) [9] ini memanfaatkan SVM untuk melakukan klasifikasi pada mata katarak dengan menggunakan 4222 gambar, yang hasilnya model ini mendapatkan nilai akurasi 95,035% dengan menggunakan kernel RBF. Penelitian lainnya yang dilakukan oleh Arifin (2024) [10] ini melakukan klasifikasi penyakit diabetes melitus dengan menggunakan data publik yang diambil pada website Kaggle yang menghasilkan nilai akurasi sebesar 77,27%.

Namun, ketiga studi tersebut terdapat satu penelitian yang mengkaji penerapan algoritma SVM dalam konteks prediksi penyakit diabetes berbasis data klinis tetapi data yang digunakan merupakan data publik. Maka dari itu penelitian ini menggunakan data asli dari RSUD Prof. Dr. Margono Soekarjo sebagai sumber data yang lebih representatif terhadap kondisi pasien, berbeda dengan penelitian sebelumnya yang menggunakan data publik. Selain itu, penelitian ini menerapkan kombinasi teknik *GridSearchCV* untuk optimasi parameter dan SMOTE untuk penyeimbangan kelas pada dataset, yang belum banyak digunakan secara bersamaan dalam studi sejenis. Penelitian ini juga menyertakan analisis *feature importance* untuk mengevaluasi kontribusi masing-masing variabel terhadap prediksi risiko diabetes, sehingga dapat memberikan wawasan tambahan bagi tenaga medis dalam memahami faktor-faktor utama yang berperan dalam diagnosis dini. Dengan pendekatan ini, diharapkan model prediksi yang dibangun dapat lebih akurat dan aplikatif dalam mendukung pengambilan keputusan klinis berbasis data.

Oleh karena itu, penelitian ini bertujuan mengevaluasi kinerja algoritma SVM dalam membangun model prediksi risiko *Diabetes Melitus* yang akurat dan dapat diandalkan untuk mendukung pengambilan keputusan klinis secara data-driven.

## II. KAJIAN TEORI

### A. Diabetes Melitus

*Diabetes Melitus (DM)* merupakan jenis gangguan metabolik kronis yang disebabkan oleh meningkatnya kadar glukosa dalam darah sebagai akibat dari kelainan sekresi insulin, kerja insulin, atau keduanya [1]. *Diabetes Melitus (DM)* adalah salah satu penyakit yang paling umum dijumpai pada masalah Kesehatan dan penyebab utama dari angka morbiditas dan mortalitas secara global [11]. Penyakit Diabetes ini merupakan penyakit yang tidak dapat ditularkan dari orang ke orang [12] dan kerap kali tidak terdeteksi sejak awal karena gejala yang muncul seperti rasa haus berlebihan, sering buang air kecil, dan penurunan berat badan yang tersamarkan. Gejala tersebut khususnya sulit dikenali pada lansia karena perubahan fisik akibat proses penuaan, sehingga kondisi ini sering berkembang menjadi komplikasi serius yang dapat mengancam jiwa [13], sehingga

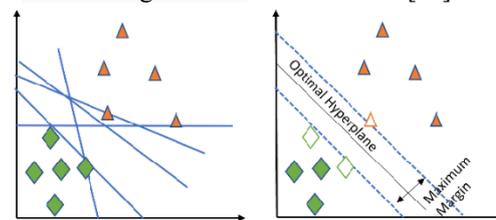
membutuhkan perawatan medis yang berkelanjutan guna menurunkan peluang terjadinya kompleksitas tersebut [14].

Faktor-faktor yang mungkin menyebabkan diabetes diantaranya sebagai berikut [15], [16]:

1. Usia: Pada usia diatas 45 tahun memiliki resiko lebih tinggi menderita *diabetes melitus* karena tingkat sensitifitas terhadap insulin mulai menurun sehingga kadar gula darah yang seharusnya masuk kedalam sel ini tetap berada pada aliran darah yang menyebabkan kadar gula darah meningkat.
2. Faktor genetik: Faktor genetik atau riwayat keluarga menjadi resiko terjadinya *diabetes melitus*, seperti salah satu orang tua mengidap diabetes maka anaknya memiliki kemungkinan untuk terjangkit diabetes juga.
3. Pola makan: Makan yang berlebihan bisa menjadi penyebab *diabetes melitus* karena menyebabkan lemak dan gula menumpuk didalam tubuh.
4. Aktifitas fisik: Aktifitas fisik yang tidak teratur dapat menyebabkan resiko terjadinya *diabetes melitus*.

### B. Support Vector Machine

*Support Vector Machine (SVM)* adalah salah satu algoritma klasifikasi dalam bidang *machine learning* yang termasuk dalam kategori *supervised learning* [17]. SVM bekerja dengan cara mencari *hyperplane* dengan membagi menjadi 2 kelas (kelas positif dilambangkan dengan +1 dan kelas negative dilambangkan dengan -1) dengan tujuan memaksimalkan margin antara kelas tersebut [18].



GAMBAR 1 ILUSTRASI OPTIMASI HYPERPLANE [19]

Berdasarkan kasus pembagian datanya *Support Vector Machine* dibagi menjadi 2 macam, yaitu data yang dapat dipisahkan secara sempurna dan data yang tidak dapat dipisahkan secara sempurna.

### C. Confusion Matrix

*Confusion Matrix* merupakan proses yang dipakai untuk menganalisa nilai akurasi dari model klasifikasi yang digunakan untuk menganalisis data dengan kelas yang berbeda. Dengan mengukur tingkat keakurasian dari data maka dapat diketahui performa dari suatu model klasifikasi yang telah dibuat [20].

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Keterangan :

True Positive (TP) = memiliki nilai prediksi yang benar positif

False Positive (FP) = memiliki nilai salah positif  
 False Negative (FN) = memiliki nilai salah negative  
 True Negative (TN) = memiliki nilai prediksi benar negative

Berikut ini merupakan cara untuk mencari nilai dari *accuracy*, *precision*, *Recall*, dan *F1-score* [21] yang dituliskan pada persamaan (1), (2), (3), dan (4)

a. *Accuracy* : Total seberapa sering model memiliki nilai benar dalam melakukan klasifikasi secara keseluruhan.

$$Accuracy = \frac{TP+TN}{TOTAL} \quad (1)$$

b. *Precision* : Seberapa sering model melakukan prediksi yang bernilai positif dan nilai tersebut benar.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

c. *Recall* : Seberapa sering model melakukan prediksi yang bernilai benar atau positif pada setiap kelas aktualnya positif.

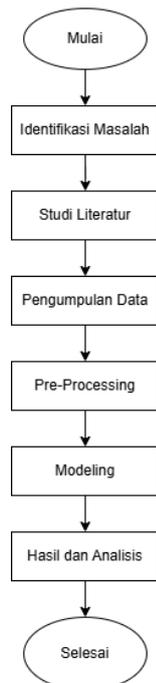
$$Recall = \frac{TP}{TP+FN} \quad (3)$$

d. *F1-score* : Nilai rata-rata dari precision dan recall.

$$F1 - score = \frac{(2 \times recall \times Precision)}{Recall + Precision} \quad (4)$$

### III. METODE

Penelitian ini melakukan klasifikasi terhadap dataset Diabetes melitus yang akan dilakukan pengujian terhadap seberapa besar ketepatan model dalam melakukan klasifikasi terhadap data. Pengujian model dilakukan dengan menggunakan algoritma *Support Vector Machine* dengan alur pengerjaan yang bisa dilihat pada gambar .



GAMBAR 2 ALUR PENELITIAN

#### 1. Pengumpulan data

Penelitian ini menggunakan dataset yang diperoleh melalui data pasien penyakit diabetes melitus pada Ruang

Rekam Medis di RSUD Prof. Dr. Margono Soekarjo pada tanggal 11 Juni hingga 11 Juli 2024.

#### 2. Preprocessing

Proses *preprocessing* data ini dilakukan data cleaning yang mana beberapa tahap seperti penghapusan kolom atribut yang tidak digunakan, encoding kolom, dan menangani nilai missing value.

#### 3. Modelling

Pada tahap modeling ini akan menggunakan 2 jenis model yaitu *support vector machine* dan *random forest*. Pada algoritma *Support Vector Machine* ini dilakukan dengan menggunakan beberapa kernel seperti linear, RBF, poly, dan sigmod. Pada masing-masing kernel digunakan *GridSearchCV* dengan *cross-validation* yang digunakan sebanyak 5. Parameter yang diuji meliputi *C* atau kompleksitas regulasi, *gamma* yang digunakan pada RBF, poly, dan sigmoid, dan *degree* untuk kernel poly.

### IV. HASIL DAN PEMBAHASAN

#### A. Pengumpulan data

Tahap awal penelitian dilakukan yaitu pengumpulan dataset. Pada gambar 2 merupakan hasil dari pengumpulan dataset yang akan digunakan. (Hasil pengambilan data tersebut mendapatkan total 830 kolom dengan 14 atribut data, yang mana atributnya mencakup NO. REG, NO. RM, JK, Usia, Nadi, Tensi (Tekanan Darah), Tinggi Badan (CM), TB(M), Berat Badan, Glukosa, Kolesterol, BMI, dan Outcome.

NO. REG	NO. RM	JK	Usia	Nadi	Tensi (Tekanan Darah)	Tinggi Badan (CM)	Berat Badan (kg)	Glukosa (mg/dl)	Kolesterol (mg/dl)	BMI	Outcome
215110006	772036	L	61	82	166/92	174	178	70	188	22,09	1
248101042	2001088	L	62	83	88/41	41	161	1,81	56	20,9	1
248101044	2220196	P	53	69	130/14	14	100	1,6	56	15,2	1
248101047	2270340	L	36	107	108/122	122	170	1,7	86	17,1	1
248101048	2160303	L	60	63	112/20	20	160	1,6	56	18,53	1
248101050	2048101	L	53	84	121/85	85	0	0	0	28,6	1
248101120	208056	L	58	112	113/102	102	160	1,6	56	13,9	1
248101120	1082544	P	60	82	105/90	90	160	1,6	66	21,5	1
248101146	2086074	P	65	100	134/70	70	150	1,5	56	22,22	1
248101146	101706	L	58	77	130/80	80	160	1,6	78	27,2	1
248101261	2118203	P	56	89	135/86	86	160	1,6	66	31,5	1
248101262	2227060	P	66	105	125/95	95	0	0	0	18,9	1
248101354	2271077	L	45	115	105/61	61	167	1,67	66	21,51	1
248101362	2217060	P	60	98	160/62	62	160	1,6	66	19,9	1
248101364	2271088	P	60	78	126/83	83	160	1,6	66	26,38	1
248101362	2057468	P	58	88	161/54	54	172	1,72	66	17,3	1
248101478	2271168	P	50	80	161/50	50	159	1,59	61	15,7	1
248101801	2271125	P	63	98	160/61	61	147	1,47	48	25,4	1
248101802	810207	P	62	89	159/78	78	155	1,55	63	20,22	1
248101802	1064204	P	60	89	134/85	85	155	1,55	55	20,9	1
248101806	1048181	L	55	102	80/61	61	160	1,6	66	21,44	1
248101805	1038146	L	77	89	153/78	78	165	1,65	70	15,71	1
248101808	1018117	P	68	87	128/81	81	160	1,6	66	22,22	1
248101802	503085	L	70	80	181/97	97	160	1,6	66	25,38	1

Gambar 3 Dataset Diabetes

#### B. Preprocessing

##### 1. Pemilihan atribut data

Pada dataset memiliki 14 total atribut, tetapi pada penelitian ini akan menggunakan 6 atribut seperti JK (Jenis Kelamin), Usia, Tekanan Darah, Glukosa, BMI (*Body Mass Index*), dan Outcome, yang mana pada dataset ini terdapat atribut Tinggi Badan dalam format cm dan m lalu Berat Badan yang mana atribut tersebut sudah dimasukan kedalam bentuk BMI.

Data columns (total 6 columns):				
#	Column	Non-Null Count	Dtype	
0	JK	830 non-null	object	
1	Usia	830 non-null	int64	
2	Tekanan Darah	828 non-null	float64	
3	Glukosa	827 non-null	float64	
4	BMI	809 non-null	object	
5	Outcome	830 non-null	int64	

GAMBAR 4 VARIABEL DATASET YANG DIGUNAKAN

##### 2. Analisis missing value

Pada dataset yang digunakan tidak banyak memiliki nilai missing value atau *NULL* seperti yang bisa dilihat

pada gambar 5. Maka nilai dari masing-masing *missing value* akan diisi dengan nilai median.

```
data.isnull().sum()
```

	0
JK	0
Usia	0
Tekanan Darah	2
Glukosa	3
BMI	21
Outcome	0

GAMBAR 5 NILAI MISSING VALUE

Pada tahap mengisi nilai NULL dengan median digunakan source code dan hasil dari imputer pada gambar 6.

```
[ ] # isi missing values dengan median
imputer = SimpleImputer(strategy='median')
df = pd.DataFrame(imputer.fit_transform(data), columns=data.columns)

[ ] df.isnull().sum()
```

	0
JK	0
Usia	0
Tekanan Darah	0
Glukosa	0
BMI	0
Outcome	0

GAMBAR 6 SOURCE CODE MENGGISI NILAI NULL DENGAN MEDIAN

### 3. Encoding data

Tahap *Encoding encoding* pada kolom JK yang awalnya pada dataset memiliki tipe data object dengan nilai 'L' dan 'P' yang bisa dilihat pada gambar 4.1 kemudian dilakukan encoding untuk diubah menjadi tipe data numerik dengan nilai 0 untuk 'L' dan nilai 1 untuk 'P' yang source codenya bisa dilihat pada gambar 7 dengan hasil encoding bisa dilihat pada gambar 8.

```
[ ] #Mengubah tipe kolom data JK dari Object ke bentuk Numerik

if 'JK' in data.columns:
    label_encoder = LabelEncoder()
    data['JK'] = label_encoder.fit_transform(data['JK']) # L=0, P=1
```

GAMBAR 7 SOURCE CODE ENCODING

#	Column	Non-Null Count	Dtype
0	JK	830 non-null	float64

GAMBAR 8 HASIL ENCODING KOLOM

### 4. Standarisasi data

Tahap ini dilakukan konversi nilai desimal pada kolom BMI yang mana pada dataset awal terdapat data yang menggunakan koma ',' atau juga menggunakan titik '.' sebagai pemisah desimal, untuk mempermudah model dalam melakukan *modelling* maka dilakukan normalisasi dengan menyamaratakan bentuk pemisah angka desimal tersebut menjadi bentuk titik '.' yang source codenya bisa dilihat pada gambar 9.

```
[ ] #Menyamarkan data pada kolom BMI menjadi menggunakan Titik

for col in ['BMI']:
    if col in data.columns:
        data[col] = data[col].astype(str).str.replace(',', '.', regex=False)
        data[col] = pd.to_numeric(data[col], errors='coerce')
```

GAMBAR 9 SOURCE CODE KONVERSI NIAI DIGITAL

### 5. Normalisasi data

Tahap normalisasi data pada penelitian ini menggunakan *StandardScaler* yang berfungsi untuk menyamakan skala antar fitur numerik. *StandardScaler* mengubah nilai setiap fitur agar memiliki rata-rata 0 dan standar deviasi 1 yang source codenya bisa dilihat pada gambar 10.

```
# Scaling fitur
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
```

GAMBAR 10 SOURCE CODE NORMALISASI DATA

### C. Modeling

Model *support vector machine* menggunakan teknik *resampling Oversampling SMOTE* untuk menyeimbangkan datanya, hingga didapatkan yang mana dari keempat jenis kernel yang digunakan: RBF, Linear, Poly, dan Sigmoid, dari hasil tuning yang dilakukan kernel RBF memiliki performa terbaik dengan hasil akurasi sebesar 93%, *precision* sebesar 93%, *recall* sebesar 94%, dan nilai *f1-score* sebesar 93%. Hasil evaluasi seluruh kernel dapat dilihat pada tabel 1.

TABEL 1 EVALUASI TERBAIK TIAP KERNEL

	Label	Linear	RBF	Poly	Sigmoid
<b>Accuracy</b>		86%	93%	88%	86%
<b>Precision</b>	0	85%	94%	88%	83%
	1	87%	93%	89%	90%
<b>Recall</b>	0	88%	93%	90%	91%
	1	85%	94%	87%	82%
<b>F1-Score</b>	0	86%	93%	89%	87%
	1	86%	93%	88%	85%

Berdasarkan hasil evaluasi pada model, algoritma *Support Vector Machine* dengan kernel RBF menunjukkan performa terbaik dalam memprediksi diabetes melitus dibandingkan model lainnya, dengan nilai akurasi sebesar 93%, *precision* 93%, *recall* 94%, dan *f1-score* 93%. Kinerja ini mengungguli kernel SVM lainnya seperti *linear*, *polynomial*, dan *sigmoid*. Berdasarkan perbandingan ini, dapat disimpulkan bahwa algoritma SVM dengan kernel RBF merupakan metode paling optimal untuk klasifikasi risiko *diabetes melitus* pada dataset yang digunakan dalam penelitian ini.

## V. KESIMPULAN

Berdasarkan hasil penelitian model *Support Vector Machine* (SVM) untuk melakukan klasifikasi resiko diabetes melitus algoritma SVM dengan kernel *Radial Basis Function* (RBF) menunjukkan performa terbaik. Model ini menghasilkan nilai akurasi sebesar 93%, *precision* 93%, *recall* 94%, dan *F1-score* 93%, lebih tinggi dibandingkan kernel lain seperti *Polynomial* (88%), *Linear* (86%), dan *Sigmoid* (86%). Dengan demikian, SVM kernel RBF merupakan model yang paling optimal dan stabil dalam memprediksi risiko *Diabetes Melitus* berdasarkan dataset yang digunakan.

## REFERENSI

- [1] Lestari, Zulkarnain, and Aisyah Sijid, "Diabetes Melitus: Review Etiologi, Patofisiologi, Gejala, Penyebab, Cara Pemeriksaan, Cara Pengobatan dan Cara Pencegahan," 2021. [Online]. Available: <http://journal.uin-alauddin.ac.id/index.php/psb>
- [2] K. R. Widiyari, I. Made, K. Wijaya, and P. A. Suputra, "Diabetes Melitus Tipe 2: Faktor Risiko, Diagnosis, Dan Tatalaksana," *Ganesha Medicina Journal*, vol. 1, no. 2, pp. 114–120, 2021, doi: <https://doi.org/10.23887/gm.v1i2.40006>.
- [3] International Diabetes Federation, "IDF Diabetes Atlas 10th," International Diabetes Federation. [Online]. Available: [www.diabetesatlas.org](http://www.diabetesatlas.org)
- [4] D. Kurniawan, T. T. Wulansari, M. R. Ibrahim, and R. M. Fajar, "Model Berbasis Logika Fuzzy untuk Mengukur Risiko Menderita Diabetes Melitus," *Jurnal Ilmiah Komputer*, vol. 20, no. 1, pp. 337–347, 2024, doi: [10.35889/progresif.v20i1.1587](https://doi.org/10.35889/progresif.v20i1.1587).
- [5] International Diabetes Federation, "IDF Diabetes Atlas Tenth Edition 2021," International Diabetes Federation. [Online]. Available: <https://www.diabetesatlas.org/data/en/>
- [6] T. Hidayat, "Data Mining untuk Meningkatkan Efisiensi dan Prediksi Produk Garmen Menggunakan Algoritma K-Nearest Neighbor di PT Mas Silueta," *Jurnal Kridatama Sains dan Teknologi*, vol. 06, no. 1, pp. 160–173, 2024, doi: <https://doi.org/10.53863/kst.v6i01.1085>.
- [7] Y. Refo and S. Rostianingsih, "Penerapan SVM untuk Klasifikasi Sentimen pada Review Comment Berbahasa Indonesia di Online Shop," 2021.
- [8] Moch. N. Isyam, D. Indrayana, and W. Apriandari, "KLASIFIKASI PENERIMA BANTUAN PROGRAM KELUARGA HARAPAN MENGGUNAKAN SUPPORT VECTOR MACHINE," *Jurnal Mahasiswa Teknik Informatika (JATI)*, vol. 8, no. 5, pp. 10416–10421, 2024.
- [9] Y. Supit, "Penerapan Algoritma Support Vector Machine (SVM) untuk Deteksi Penyakit Mata Katarak," 2024. [Online]. Available: <https://ejournal.unjaya.ac.id/index.php/ijds>
- [10] S. Arifin, R. Satria, and Trwoto, "Pengembangan Algoritma Support Vector Machine (SVM) untuk Mengklasifikasi Penyakit Diabetes," *JITE*, 2024.
- [11] S. Delfina, I. Carolita Maharani, S. Habsah, and ta Ayatillahi, "Literature Review : Analisis Determinan Faktor Risiko Kejadian Diabetes Mellitus Tipe 2 Pada Usia Produktif," *Jurnal Kesehatan Tambusai*, vol. 2, no. 4, 2021.
- [12] Suratun, Dewi Pujiana, and Mareta Sari, "PENCEGAHAN DIABETES MELITUS DI PALEMBANG," *Jurnal Masker Medika*, vol. 11, no. 1, 2023, doi: <https://doi.org/10.52523/maskermedika.v11i1.507>.
- [13] Heru Laksono, Hendri Heriyanto, and Risma Apriani, "Determinan Faktor Kejadian Komplikasi Pada Penderita Diabetes Melitus Di Kota Bengkulu Tahun 2021," *Journal of Nursing and Public Health*, vol. 10, no. 1, 2022.
- [14] Fida Maisa Hana, "Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma Decision Tree C4.5," *Jurnal Sistem komputer dan Kecerdasan Buatan*, vol. 4, no. 1, pp. 32–39, 2020, doi: [10.47970/siskom-kb.v4i1.173](https://doi.org/10.47970/siskom-kb.v4i1.173).
- [15] K. Irjayanti P, S. Zaenal, and Suhartatik, "Faktor-Faktor Yang Mempengaruhi Terjadinya Peningkatan Diabetes Melitus Tipe 2," *Jurnal Ilmiah Mahasiswa & Penelitian Keperawatan*, vol. 1, no. 6, pp. 805–813, 2022.
- [16] C. C. Rani and N. S. Mulyani, "Faktor-faktor yang berhubungan dengan kejadian diabetes mellitus tipe-II pada pasien rawat jalan," *Jurnal SAGO Gizi dan Kesehatan*, vol. 2, no. 2, p. 122, Sep. 2021, doi: [10.30867/gikes.v2i2.258](https://doi.org/10.30867/gikes.v2i2.258).
- [17] I. S. Aisah, B. Irawan, and T. Suprpti, "Algoritma Support Vector Machine (SVM) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur'an Digital," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 6, pp. 3759–3765, 2023.
- [18] P. Ayuningtyas et al., "Perbandingan Metode Word2vec Model Skipgram pada Ulasan Aplikasi Linkaja menggunakan Algoritma Bidirectional LSTM dan Support Vector Machine," *Justin (Jurnal Sistem dan Teknologi Informasi)*, vol. 12, no. 1, 2024, doi: [10.26418/justin.v12i1.72530](https://doi.org/10.26418/justin.v12i1.72530).
- [19] S. Y. Prasetyo and G. Z. Nabiilah, "Perbandingan Model Machine Learning Pada Klasifikasi Tumor Otak Menggunakan Fitur Discrete Cosine Transform," *Jurnal Teknologi Terpadu*, vol. 9, no. 1, pp. 29–34, 2023, [Online]. Available: <https://journal.nurulfikri.ac.id/index.php/jtt>
- [20] Y. Afrillia, L. Rosnita, and D. Siska, "Analisis Sentimen Ciutan Twitter Terkait Penerapan Permendikbudristek Nomor 30 Tahun 2021 Menggunakan TextBlob dan Support Vector Machine," *G-Tech: Jurnal Teknologi Terapan*, vol. 6, no. 2, pp. 387–394, Oct. 2022, doi: [10.33379/gtech.v6i2.1778](https://doi.org/10.33379/gtech.v6i2.1778).
- [21] Rifat Wigar Adanendra, "Analisis Sentimen Pada Aplikasi DUOLINGO Menggunakan Metode Naive Bayes Multinomial Dengan Fitur Akstraksi Countinvestorizer," 2024.