

## ANALISIS PERBANDINGAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DAN SUPPORT VECTOR MACHINE (SVM) DALAM KLASIFIKASI DATA PERBANKAN

Rosalia Hadi<sup>1</sup>, Ni Luh Gede Pivin Suwirmayanti<sup>2</sup>, I Gusti Ngurah Ady Kusuma<sup>3</sup>, I Gusti Ayu Desi Saryanti<sup>4</sup>, Putu Devi Novayanti<sup>5</sup>

ITB STIKOM Bali<sup>12345</sup>

Jl. Raya Puputan No 86 Renon, Denpasar

E-mail : rosa@stikom-bali.ac.id<sup>1</sup>, pivin@stikom-bali.ac.id<sup>2</sup>, ady\_kusuma@stikom-bali.ac.id<sup>3</sup>, desi@stikom-bali.ac.id<sup>4</sup>, devinovayanti @stikom-bali.ac.id<sup>5</sup>

### ABSTRAK

Data perbankan memiliki karakteristik yang beragam, mencakup transaksi keuangan, riwayat kredit, detail nasabah, dan lainnya. Kompleksitas data ini menimbulkan tantangan dalam pengelolaan dan analisis data yang efektif. Salah satu tantangannya adalah menilai kelayakan kredit pemohon pinjaman rumah. Dalam proses pemberian pinjaman, beberapa faktor dipertimbangkan, termasuk karakter pemohon, kemampuan pembayaran, jaminan, kondisi ekonomi, usia, dan sumber pendanaan. Analisis risiko kredit dapat dilakukan melalui berbagai metode, termasuk analisis pasar dan teknik pembelajaran mesin big data. Dalam penelitian ini, kami menggunakan dua algoritma populer: K-Nearest Neighbors (KNN) dan Support Vector Machine (SVM). KNN, dikenal karena kesederhanaannya dan kemampuannya dalam menangani data yang beragam, sering digunakan untuk klasifikasi. Namun, KNN menghadapi keterbatasan dalam kecepatan komputasi pada dataset besar dan kerentanan terhadap data yang bising. Sebaliknya, SVM, sebuah algoritma yang lebih canggih, dirancang untuk memaksimalkan pemisahan antara kelas data dan unggul dalam menangani data berdimensi tinggi. Temuan kami menunjukkan bahwa SVM lebih baik daripada KNN dalam mengklasifikasikan risiko kredit pada data perbankan, terutama untuk pinjaman rumah. SVM dengan kernel dot dan ANOVA mencapai akurasi 80,56%, dibandingkan dengan 69,44% untuk KNN dengan k=50. Temuan ini menunjukkan bahwa SVM lebih unggul dalam klasifikasi risiko kredit pada data perbankan, khususnya untuk kredit kepemilikan rumah..

Kata kunci : K-Nearest Neighbor, Support Vector Machine, Perbankan

### ABSTRACTS

*Banking data is characterized by its diverse nature, encompassing financial transactions, credit histories, customer details, and more. This complexity poses challenges in effective data management and analysis. One such challenge lies in assessing the creditworthiness of home loan applicants. In the lending process, several factors are considered, including the applicant's character, repayment capacity, collateral, economic condition, age, and funding source. Credit risk analysis can be undertaken using various methods, including market analysis and big data machine learning techniques. In this study, we employed two popular algorithms: K-Nearest Neighbors (KNN) and Support Vector Machine (SVM). KNN, known for its simplicity and ability to work with diverse data, is frequently used for classification. However, it faces limitations in terms of computational speed on large datasets and susceptibility to noisy data. Conversely, SVM, a more sophisticated algorithm, is designed to maximize the separation between data classes and excels in handling high-dimensional data. Our findings reveal that SVM outperforms KNN in classifying credit risk in banking data, particularly for home loans. SVM with dot and ANOVA kernels achieved an accuracy of 80.56%, compared to 69.44% for KNN with k=50. These findings indicate that SVM is superior in credit risk classification in banking data, especially for home ownership loans.*

*Keywords: K-Nearest Neighbor, Support Vector Machine, Banking.*

### 1.PENDAHULUAN

Data perbankan memiliki ciri khas tersendiri yang mencakup berbagai aspek, seperti transaksi keuangan, riwayat kredit, informasi nasabah, dan banyak lagi. Kompleksitas data ini menciptakan tantangan dalam pengelolaan dan analisis yang efektif. Sebagai contoh, transaksi keuangan dapat memiliki variasi nilai yang signifikan, sementara atribut nasabah dapat mencakup informasi pribadi yang beragam. Oleh karena itu, diperlukan suatu pendekatan yang cermat untuk mengklasifikasikan data ini dengan tepat.

Salah satu tantangan dalam perbankan adalah kredit pemilikan rumah yang diberikan kepada masyarakat. Dalam proses peminjaman ada tahap pengecekan apakah yang mengajukan kredit dapat di setujui atau tidak. Pada proses ini terdapat beberapa atribut yang akan dilakukan analisa seperti penilaian karakter, kapasitas pembayaran, besarnya anggunan atau jaminan, kondisi ekonomi dari peminjam, umur, sumber pendanaan, kemampuan bayar, dan lain sebagainya.

Analisis risiko kredit dapat dilakukan dengan berbagai cara seperti analisis pasar dan data besar melalui pembelajaran mesin [1] seperti menggunakan metode KNN dan SVM. K-Nearest Neighbor (KNN) adalah algoritma machine learning yang bekerja berdasarkan prinsip bahwa data yang memiliki kesamaan fitur akan memiliki klasifikasi yang mirip. Metode KNN yang sederhana dalam penerapannya sering digunakan karena kemampuannya untuk mengelola data dalam berbagai bentuk dan ukuran tanpa memerlukan asumsi khusus tentang distribusi data. Meskipun, KNN memiliki keterbatasan, terutama dalam hal kecepatan perhitungan pada dataset besar dan sensitivitasnya terhadap data yang bising. Beberapa penelitian terbaru telah menyelidiki cara untuk meningkatkan kinerja KNN, seperti melalui pemilihan fitur yang lebih optimal dan penggunaan teknik pra-pemrosesan yang canggih.

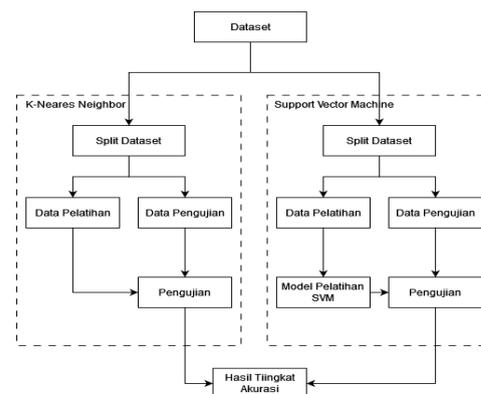
Support Vector Machine (SVM) adalah algoritma yang lebih kompleks dan dirancang untuk menemukan hyperplane yang memisahkan dua kelas data dengan margin maksimum. SVM dikenal efektif dalam

menangani kumpulan data dengan banyak fitur dan dapat memberikan kinerja yang baik meskipun pada data yang tidak dapat dipisahkan secara linear. SVM juga memiliki keunggulan dalam kemampuan penyebaran yang lebih baik dibandingkan KNN. Walau begitu, SVM memerlukan penyetelan parameter yang lebih teliti, seperti pemilihan inti yang sesuai, untuk mencapai hasil yang maksimal.

Penelitian sebelumnya tentang perbankan menggunakan KNN adalah menjelaskan tentang penggunaan metode KNN dalam melakukan klasifikasi kredit pada Bank [2], selain itu terdapat juga metode decision tree dan naïve bayes melakukan klasifikasi menentukan kelayakan pemberian kredit [3] dan penelitian menggunakan metode SVM sebelumnya menjelaskan bahwa memprediksi risiko pinjaman dengan pendekatan machine learning menggunakan beberapa metode seperti Support Vector Machine (SVM) [4]. Sejumlah penelitian sebelumnya telah mencoba mengoptimalkan algoritma klasifikasi, termasuk KNN, pada berbagai jenis data. Namun, masih belum terdapat penerapan normalisasi dan seleksi fitur khususnya pada data perbankan. Dengan demikian, penelitian ini akan memberikan kontribusi penting terhadap literatur dengan menyelidiki dampak integrasi metode normalisasi dan seleksi fitur terhadap performa KNN dan SVM pada dataset perbankan.

### 2.METODE PENELITIAN

Adapun metode penelitian yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

Penelitian ini diawali dengan pengambilan dataset perbankan berupa data kredit pemilikan rumah yang opensource. Data tersebut akan diproses menggunakan metode K-Nearest Neighbor. Pada masing-masing proses klasifikasi akan dilakukan proses pemisahan dataset menjadi dataset pelatihan dan dataset pengujian. Pada metode KNN dilakukan pengujian langsung menggunakan dataset pengujian berdasarkan dataset pelatihan maka hasil akhir tersebut mendapatkan nilai tingkat akurasi dari metode KNN. Pada metode SVM dilakukan proses pelatihan terlebih dahulu menggunakan dataset pelatihan untuk mendapatkan model pelatihan. Model yang telah didapatkan dilakukan proses pengujian menggunakan dataset pengujian, sehingga hasil dari proses tersebut mendapatkan hasil berupa tingkat akurasi.

### 2.1. K-Nearest Neighbor

Algoritma K-Nearest Neighbor (K-NN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. K-Nearest Neighbor berdasarkan 169erupa 'learning by analogy'. Data learning dideskripsikan dengan atribut numerik  $n$ -dimensi. Tiap data learning merepresentasikan sebuah titik, yang ditandai dengan  $c$ , dalam ruang  $n$ -dimensi. Jika sebuah data query yang labelnya tidak diketahui diinputkan, maka K-Nearest Neighbor akan mencari  $k$  buah data learning yang jaraknya paling dekat dengan data query dalam ruang  $n$ -dimensi. Jarak antara data query dengan data learning dihitung dengan cara mengukur jarak antara titik yang merepresentasikan data query dengan semua titik yang merepresentasikan data learning dengan rumus Euclidean Distance [5]

Pada fase training, 169erupakan ini hanya melakukan penyimpanan 169erupa-vektor fitur dan klasifikasi data training sample. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk testing data (klasifikasinya belum diketahui)[6]. Jarak dari 169erupa yang baru ini terhadap seluruh 169erupa training sample dihitung, dan sejumlah  $k$  buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut. Nilai  $k$  yang terbaik untuk

169erupakan ini tergantung pada data secara umumnya, nilai  $k$  yang tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat 169erupakan antara setiap klasifikasi menjadi lebih kabur. Nilai  $k$  yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan cross-validation. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain,  $k = 1$ ) disebut 169erupakan nearest neighbor [7]

Ketepatan 169erupakan K-NN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Riset terhadap 169erupakan ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur, agar performa klasifikasi menjadi lebih baik.

Algoritma K-Nearest Neighbor menggunakan klasifikasi ketetangaan (neighbor) sebagai nilai prediksi dari query instance yang baru. Algoritma ini sederhana, bekerja berdasarkan jarak terpendek dari query instance ke training sample untuk menentukan ketetanggaannya. Langkah-langkah untuk menghitung metode k-Nearest Neighbor antara lain:

1. Menentukan parameter  $k$
2. Menghitung jarak antara data yang akan dievaluasi dengan semua pelatihan
3. Mengurutkan jarak yang terbentuk
4. Menentukan jarak terdekat sampai urutan  $k$
5. Memasangkan kelas yang bersesuaian
6. Mencari jumlah kelas dari tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelas data yang akan dievaluasi

### 2.2. Support Vector Machine

Support Vector Machine (SVM) 169erupakan suatu 169erupa untuk melakukan prediksi, baik prediksi dalam kasus regresi maupun klasifikasi. Teknik SVM digunakan untuk mendapatkan fungsi pemisah (*hyperplane*) yang optimal untuk memisahkan observasi yang memiliki nilai 169erupakan target yang berbeda. *Hyperplane* ini dapat berupa line pada two dimension dan dapat berupa flat plane pada multiple dimension [8].

Support Vector Machine (SVM) digunakan untuk melatih dan mengklasifikasikan pola terbaik dan menghasilkan nilai yang benar. Ini

melatih model support vector machine (SVM) dari data kereta dan data pengujian. Data latih digunakan untuk melatih model dan data uji digunakan untuk memvalidasi data latih [9]

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Pengolahan Sampel Data

Data yang digunakan pada penelitian ini bersifat 170erupa mengenai persetujuan peminjaman pada pengajuan kredit dibank dengan jumlah baris data sebesar 614 buah data dan 13 buah atribut dimana atribut yang digunakan dapat dilihat pada tabel berikut:

Tabel 1 Atribut Dataset Perbankan

No	Atribut	Keterangan
1	Loan ID	ID Peminjam dengan id nilai
2	Gender	Jenis Kelamin Peminjam diisi Male untuk laki-laki berisi nilai 1, Female untuk perempuan berisi nilai 0
3	Married	Status Pernikahan diisi dengan nilai Yes memiliki nilai 1 dan No memiliki nilai 0
4	Dependents	Tanggungannya dimiliki oleh peminjam nilai yang diisi 0, 1, 2, 3+
5	Education	Tingkat Pembelajaran diisi dengan nilai 0 untuk Not Graduate dan 1 untuk Graduate
6	Self Employed	Pekerjaan dari peminjam apakah Merupakan wiraswata, jika Yes berikan nilai 1 dan Jika No berikan nilai 0
7	Applicant Income	Berisikan nilai dari pendapatan peminjam mulai dari 150 sampai 81000
8	Co Applicant Income	Berisikan nilai dari pendapatan 170erupak dari peminjam mulai dari 0 sampai 41667
9	Loan Amount	Berisikan jumlah dari pinjaman yang diajukan
10	Loan Amount Term	Berisikan jangka waktu peminjaman

11	Credit History	Berisikan apakah memiliki kredit, nilai 1 untuk Yes dan nilai 0 untuk No
12	Property	Berisikan rumah yang dimiliki perkotaan, semi perkotaan dan pedesaan
13	Loan Status	Penerimaan kredit berisikan nilai Y untuk pemberian kredit dan N untuk tidak memberikan kredit

Pada tabel 1 dijelaskan bahwa terdapat 13 atribut dengan definisi dari masing-masing kolom, sehingga dapat ditentukan atribut Loan Status 170erupakan label dari setiap data yang ada. Adapun data dari dataset dapat dilihat pada Gambar 2.

Gambar 2 Dataset perbankan

#### 3.2. Konversi Nilai Atribut Karakter menjadi Angka

Sebelum melakukan proses klasifikasi maka dilakukan proses perubahan nilai pada masing-masing atribut seperti atribut Married dan atribut Self Employed yang masing menggunakan nilai Yes dan No, maka dari itu dilakukan perubahan data dari String ke bilangan. Adapun proses yang dilakukan dapat dilihat pada tabel 2.

Tabel 2

No	Atribut	Perubahan
1	Gender	Nilai Male = 0 dan Female = 1
2	Married	Nilai Yes = 1 dan No = 0
3	Education	Nilai Graduate = 0 dan Not Graduate = 1

4	Self Employed	Nilai Yes = 1 dan No = 0
5	Property	Nilai Urban = 0, Semi Urban 2 dan Rural = 1

Berdasarkan hasil tabel maka dilakukan proses pembagian dataset dimana dataset pelatihan dengan jumlah 70% dan dataset pengujian dengan jumlah 30%.

### 3.3. Hasil Pengujian K-Nearest Neighbor

Berdasarkan data diatas kemudian dilakukan proses perhitungan menggunakan metode KNN untuk melihat. Nilai k yang digunakan pada proses Algoritma knn adalah 5. Sebagai contoh disini akan menggunakan 10 buah data pada baris ke 1 sampai ke 10, dimana data ke 11 sebagai data baru. Maka perhitungan yang dapat diberikan pada metode KNN pada data ke 11 ke masing-masing baris data adalah sebagai berikut :

#### baris data 1

$$= \sqrt{\frac{(-0.404 - -0.138)^2 + (2.493 - -0.028)^2 + (0.686 - -0.208)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(-0.266)^2 + (2.521)^2 + (0.894)^2 + (0)^2 + (0)^2}{0.071 + 6.353 + 0.799 + 0 + 0}} = \sqrt{7.224} = 2.687$$

#### baris data 2

$$= \sqrt{\frac{(-0.404 - -0.417)^2 + (2.493 - -0.604)^2 + (0.686 - -0.208)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(0.01)^2 + (3.097)^2 + (1.664)^2 + (0)^2 + (0)^2}{0 + 9.59 + 2.77 + 0 + 0}} = \sqrt{12.36} = 3.516$$

#### baris data 3

$$= \sqrt{\frac{(-0.404 - -0.49)^2 + (2.493 - 0.297)^2 + (0.686 - -0.307)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(0.086)^2 + (2.196)^2 + (0.994)^2 + (0)^2 + (0)^2}{0.007 + 4.822 + 0.987 + 0 + 0}} = \sqrt{5.817} = 2.412$$

#### baris data 4

$$= \sqrt{\frac{(-0.404 - 0.112)^2 + (2.493 - -0.604)^2 + (0.686 - -0.046)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(-0.516)^2 + (3.097)^2 + (0.73)^2 + (0)^2 + (0)^2}{0.267 + 9.589 + 0.537 + 0 + 0}} = \sqrt{10.393} = 2.224$$

#### baris data 5

$$= \sqrt{\frac{(-0.404 - 0.009)^2 + (2.493 - 0.999)^2 + (0.686 - 1.159)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(-0.414)^2 + (1.494)^2 + (-0.832)^2 + (0)^2 + (0)^2}{0.171 + 2.231 + 0.692 + 0 + 0}} = \sqrt{3.095} = 1.76$$

#### baris data 6

$$= \sqrt{\frac{(-0.404 - -0.535)^2 + (2.493 - -0.025)^2 + (0.686 - -0.618)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(0.131)^2 + (2.517)^2 + (1.304)^2 + (0)^2 + (0)^2}{0.017 + 6.338 + 1.701 + 0 + 0}} = \sqrt{8.056} = 2.838$$

#### baris data 7

$$= \sqrt{\frac{(-0.404 - -0.411)^2 + (2.493 - 0.353)^2 + (0.686 - 0.165)^2}{(0.275 - 0.275)^2 + (1 - 0)^2}}$$

$$= \sqrt{\frac{(0.006)^2 + (2.517)^2 + (1.304)^2 + (0)^2 + (1)^2}{0 + 4.58 + 0.272 + 0 + 1}} = \sqrt{5.852} = 2.42$$

#### baris data 8

$$= \sqrt{\frac{(-0.404 - -0.24)^2 + (2.493 - -0.21)^2 + (0.686 - 0.289)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(-0.165)^2 + (2.514)^2 + (0.398)^2 + (0)^2 + (0)^2}{0.027 + 6.319 + 0.158 + 0 + 0}} = \sqrt{6.504} = 2.55$$

#### baris data 9

$$= \sqrt{\frac{(-0.404 - 1.319)^2 + (2.493 - 3.586)^2 + (0.686 - 2.537)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(-1.723)^2 + (-1.093)^2 + (-1.851)^2 + (0)^2 + (0)^2}{2.97 + 1.195 + 3.425 + 0 + 0}} = \sqrt{7.59} = 2.755$$

#### baris data 10

$$= \sqrt{\frac{(-0.404 - -0.382)^2 + (2.493 - -0.3366)^2 + (0.686 - -0.928)^2}{(0.275 - 0.275)^2 + (1 - 1)^2}}$$

$$= \sqrt{\frac{(-0.022)^2 + (2.829)^2 + (1.615)^2 + (0)^2 + (0)^2}{0.001 + 8.004 + 2.607 + 0 + 0}} = \sqrt{10.612} = 3.258$$

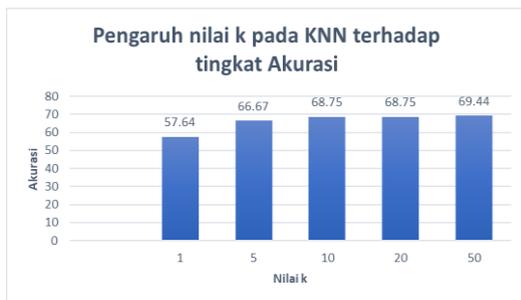
Berdasarkan hasil diatas maka dapat dilakukan proses pengurangan nilai dari jarak terpendek sampai jarak terbesar, seperti yang ditampilkan pada Tabel 3.

Tabel 3. Hasil jarak perhitungan k-NN

NO	Baris Data Ke-	Jarak	Label
1	5	1.759	Y
2	3	2.412	Y
3	7	2.419	N
4	8	2.55	Y
5	1	2.688	N
6	9	2.755	N
7	6	2.838	Y
8	4	3.224	Y
9	10	3.258	Y
10	2	3.516	Y

Dari tabel diatas maka dilakukan proses pemilihan ketetanggaan sebanyak 5 buah, yakni dari urutan nomer 1 sampai ke nomer 5. Berdasarkan hal tersebut maka mendapatkan

hasil label Y sebanyak 3 buah dan label N sebanyak 2 buah. Sehingga pada data ke 11 memiliki cluster Y. Hasil dari pengujian menggunakan nilai k sebesar 5 mendapatkan nilai tingkat akurasi sebesar 66,67%. Berdasarkan hasil tersebut maka dilakukan pengujian menggunakan parameter dari metode KNN yaitu nilai k. Pengujian parameter yang dilakukan dengan nilai 1,5,10,20 dan 50. Adapun hasil pengujian tersebut dapat dilihat pada grafik berikut:

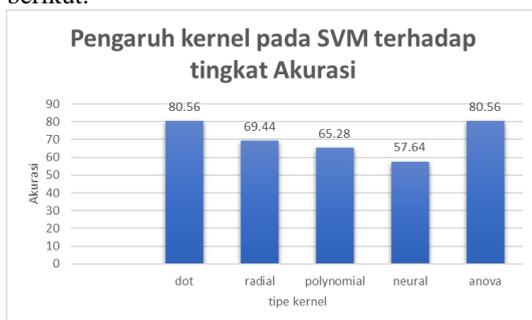


Gambar 3 Pengaruh nilai k pada KNN terhadap tingkat akurasi

Berdasarkan gambar diatas dapat dilihat bahwa nilai k sebesar 1 memiliki tingkat akurasi paling kecil yaitu sebesar 57.64% dan nilai k sebesar 50 memiliki tingkat akurasi paling tinggi yaitu sebesar 69,44% dibandingkan dengan nilai k yang lainnya.

### 3.4. Hasil Pengujian Support Vector Machine

Berdasarkan data diatas kemudian dilakukan proses perhitungan menggunakan metode SVM untuk melihat. Pada metode SVM digunakan beberapa kernel seperti dot, radial, polynomial, naural dan anova untuk melakukan pengujian sehingga hasil pengujian dapat dilihat pada tabel berikut:



Gambar 4 Pengaruh kernel pada SVM terhadap tingkat akurasi

Berdasarkan gambar diatas dapat dilihat bahwa pengujian yang dilakukan menggunakan beberapa kernel memberikan beberapa hasil yaitu kernel dot memberikan nilai tingkat akurasi sebesar 80,56%, kernel radial memberikan nilai tingkat akurasi sebesar 69,44%, kernel polynomial memberikan nilai tingkat akurasi sebesar 65,28%, kernel nural memberikan nilai tingkat akurasi sebesar 57,64% dan kernel anova memberikan nilai tingkat akurasi sebesar 80,56%. Dari hasil tersebut maka kernel neural memberikan hasil nilai tingkat akurasi paling kecil yaitu 57,64% dan kernel dot dan anova memberikan nilai tingkat akurasi sebesar 80,56%.

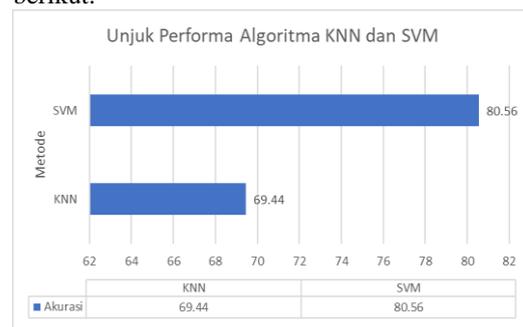
### 3.5. Unjuk Performa Algoritma KNN dan SVM

Berdasarkan hasil pengujian yang diperoleh dari metode KNN dan SVM maka dapat dilihat hasil perbandingan sebagai berikut :

Tabel 4. Unjuk Performa Algoritma KNN dan SVM

Metode	Akurasi
KNN	69.44
SVM	80.56

Dari tabel diatas maka dapat dilihat grafik pengujian Algoritma KNN dan SVM sebagai berikut:



Gambar 5 Grafik Performa Algoritma KNN dan SVM

Hasil pengujian pada gambar 1 menunjukkan bahwa metode SVM memiliki nilai tingkat akurasi yang lebih tinggi yaitu sebesar 80,56 dibandingkan dengan metode KNN dengan nilai tingkat akurasi sebesar 69,44%.

## 4.KESIMPULAN DAN SARAN

Hasil eksperimen menunjukkan bahwa penggunaan Algoritma K-Nearest Neighbor dan Algoritma Support Vector Machine dapat menghasilkan tingkat akurasi yang baik terkait dengan risiko kredit dan parameter keuangan lainnya. Unjuk kerja metode KNN memberikan hasil tingkat akurasi sebesar 69,44 dengan nilai parameter k sebesar 50, sedangkan unjuk kerja metode SVM dapat memberikan nilai tingkat akurasi sebesar 80,56 dengan menggunakan kernel dot dan kernel anova. Berdasarkan hasil tersebut maka dapat dikatakan metode SVM memberikan nilai tingkat akurasi yang lebih baik untuk penanganan kredit pada dataset perbankan.

#### DAFTAR PUSTAKA

- [1] D. T. Wilujeng et al., “Indonesian Journal of Applied Statistics Analisis Risiko Kredit Perbankan Menggunakan Algoritma K-Nearest Neighbor dan Nearest Weighted K-Nearest Neighbor,” vol. 5, no. 2, pp. 142–148, 2023, doi: 10.13057/ijas.v5i2.58426.
- [2] J. Sains, D. Teknologi, W. Supriana, M. A. Raharja, and P. W. Gunawan, “Sistem Informasi Prediksi Penilaian Kredit Perbankan Menggunakan Algoritma K-Nearest Neighbor Classification,” 2019.
- [3] T. T. Muryono, A. Taufik, and I. Irwansyah, “PERBANDINGAN ALGORITMA K-NEAREST NEIGHBOR, DECISION TREE, DAN NAIVE BAYES UNTUK MENENTUKAN KELAYAKAN PEMBERIAN KREDIT,” *Infotech: Journal of Technology Information*, vol. 7, no. 1, pp. 35–40, Jun. 2021, doi: 10.37365/jti.v7i1.104.
- [4] B. Pernama, H. Dwi Purnomo, and K. Satya Wacana, “Analisis Risiko Pinjaman dengan Metode Support Vector Machine, Artificial Neural Network dan Naive Bayes,” *Jurnal Teknologi Informasi dan Komunikasi*, vol. 7, no. 1, p. 2023, 2023, doi: 10.35870/jti.
- [5] N. A. Anggraini and N. Fadillah, “Analisis Deteksi Emosi Manusia dari Suara Percakapan Menggunakan Matlab dengan Metode KNN,” *InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)*, vol. 3, no. 2, pp. 176–179, 2019, doi: 10.30743/infotekjar.v3i2.1041.
- [6] S. H. Zulaikhah, A. Aziz, and W. Harianto, “OPTIMASI ALGORITMA K-NEAREST NEIGHBOR (KNN) DENGAN NORMALISASI DAN SELEKSI FITUR UNTUK KLASIFIKASI PENYAKIT LIVER,” 2022. [Online]. Available: <https://archive.ics.uci.edu/ml/index.php>
- [7] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, “Efficient kNN classification with different numbers of nearest neighbors,” *IEEE Trans Neural Netw Learn Syst*, vol. 29, no. 5, pp. 1774–1785, May 2018, doi: 10.1109/TNNLS.2017.2673241.
- [8] B. U. Umar, J. Agajo, O. M. Olaniyi, A. Aliyu, O. S. Owolabi, and J. G. Kolo, “Human Detection Using Speeded-Up Robust Features and Support Vector Machine from Aerial,” 2017.
- [9] E. Puerwandono and I. Maulana, “Penerapan Algoritma Svm Untuk Klasifikasi Citra Daun Sirih APPLICATION OF THE SVM ALGORITHM TO IMAGE CLASSIFICATION OF BETEL LEAF,” *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 6, no. 2, 2023.