

*Corresponding author: Ruliana,
Department of Statistics, Universitas
Negeri Makassar, Makassar, Indonesia

E-mail: ruliana.t@unm.ac.id

RESEARCH ARTICLE

Comparison of k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM) Methods for Classification of Poverty Data in Papua

Fauziah, Muhammad Arif Tiro, & Ruliana*

Department of Statistics, Universitas Negeri Makassar, Makassar, Indonesia

Abstract: Classification is a job of assessing data objects to include them in a particular class from a number of available classes. The classification method used is the k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM) methods. The data used in this study is data on poverty in Papua with the category of the number of low/high level poor people. Of the 29 regencies/cities that were sampled, 15 regencies/cities represent the number of low-level poor people and 14 districts/cities are the number of high-level poor people. The results of the analysis obtained are the k-Nearest Neighbor (k-NN) method with a value of $k=15$ producing an accuracy of 58.62%, while the Support Vector Machine (SVM) method with Parameter cost = 1 using the RBF kernel produces an accuracy value. by 93.1%. The classification criteria to find the best method is to look at the Root Mean Square Error (RMSE) which states that the Support Vector Machine (SVM) method is better than the k-Nearest Neighbor (k-NN) method.

Keywords: Classification, k-Nearest Neighbor, Support Vector Machine, poverty

1. Introduction

Masalah kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian pemerintah di negara manapun. Salah satu aspek penting untuk mendukung strategi penanggulangan kemiskinan adalah tersedianya data kemiskinan yang akurat dan tepat sasaran. Pengukuran kemiskinan yang dapat dipercaya dan dapat menjadi instrumen tangguh bagi pengambil kebijakan dalam memfokuskan perhatian pada kondisi hidup orang miskin. Data kemiskinan yang baik dapat digunakan untuk mengevaluasi kebijakan pemerintah terhadap kemiskinan, membandingkan kemiskinan antar waktu dan daerah, serta menentukan target penduduk miskin dengan tujuan untuk memperbaiki kondisi mereka (BPS, 2019).

Salah satu upaya pemerintah untuk mengentaskan kemiskinan yaitu melalui beberapa program penanggulangan kemiskinan. Dalam hal ini pemerintah sering kesulitan untuk memilah tingkat-tingkat kemiskinan dalam masyarakat. Untuk menunjang keberhasilan pelaksanaan program terutama yang berkaitan dengan penanggulangan kemiskinan, diperlukan suatu penelitian yang dapat membantu pemerintah dalam mengidentifikasi dan mengklasifikasi jumlah penduduk miskin yang mempunyai ciri-ciri atau karakteristik kemiskinan yang hampir sama. Dengan mengetahui informasi mengenai kriteria kemiskinan dari setiap kabupaten/kota diharapkan program kebijakan pemerintah daerah dapat disusun secara lebih terarah sesuai target atau sasaran yang ingin dicapai.

Kecermatan dan ketepatan dalam hal klasifikasi data merupakan hal yang sangat penting. Salah satu upaya untuk mendukung upaya pemerintah yaitu dengan melakukan klasifikasi



indikator kemiskinan. Adapun metode yang sering digunakan yaitu rule based, neural network, Support Vector Machine, naive bayes, dan nearest neighbour. Dalam penelitian ini salah satu metode yang digunakan yaitu metode k-nearest neighbor (k-NN) dan Support Vector Machine (SVM).

Metode k-NN merupakan salah satu metode klasifikasi yang terdapat dalam data mining selain SVM. k-NN juga termasuk dalam kelompok instance-based learning. k-NN dilakukan dengan mencari k objek dalam data training yang paling dekat (mirip) dengan objek pada data testing (Wu & Kumar 2009).

Tujuan dari metode SVM menemukan fungsi pemisah (classifier) yang optimal yang bisa memisahkan dua set data yang berbeda (Vapnik 1995). Fungsi pemisah atau hyperplane terbaik adalah hyperplane yang terletak di tengah-tengah dua objek dari kedua kelas.

2. Literature Review

2.1. Data Mining

Data mining adalah pengembangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar (Davies, 2004). Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambahan berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data (Pramudiono, 2007).

2.2. Klasifikasi

Klasifikasi adalah proses mengelompokkan data menjadi kelas tertentu berdasarkan kelas yang sudah ada (Simatupang dkk, 2016). Klasifikasi juga dapat mempermudah mengidentifikasi data dalam jenis dan kelompok, sehingga klasifikasi bisa memprediksi anggota kelompok dari masing-masing data (Bhavsar dkk, 2012).

2.3. k-Nearest Neighbor (k-NN)

Metode k-NN adalah suatu metode klasifikasi yang terdapat dalam data mining selain SVM. k-NN dilakukan dengan mencari k objek dalam data training yang paling dekat (mirip) dengan objek pada data testing (Wu & Kumar 2009). Nilai k dalam k-NN merupakan variabel jumlah tetangga (*Neighbor*) terdekat yang akan diambil untuk proses klasifikasi.

Adapun cara untuk mengukur kedekatan antara data baru dengan data yang lama (data training), diantaranya Euclidean distance dan mahattan distance. Yang paling sering digunakan untuk mengukur kedekatan antar data adalah *euclidean distance* (Bramer 2007) yaitu:

$$\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \quad (1)$$

dimana $a = a_1, a_2, \dots, a_n$ dan $b = b_1, b_2, \dots, b_n$ mewakili n nilai atribut dari dua *record*.

$$D(X, Y) = \sqrt{\sum_{t=1}^p (X_t - Y_t)^2} \quad (2)$$

Dengan:

D(X, Y) : Jarak Eucliden (*Euclidean Distance*)
Xi : Sampel data
Yi : Data uji
p : Dimensi data
i : Variabel data

2.4. Support Vector Machine (SVM)

SVM merupakan metode yang bersifat *non-parametric* dan biasanya digunakan dalam klasifikasi data serta pengolahan citra. Tingkat akurasi pada metode ini diambil dari

parameter dan kernel, pengguna dapat menentukan parameter dan pada kernel setiap parameter akan memiliki dampak yang berbeda (Ustuner dkk, 2015). Konsep SVM adalah mendesain sebuah *hyperplane* yang dapat mengklasifikasikan seluruh data training menjadi dua buah kelas (Aulia, 2015).

Tujuan dari metode SVM menemukan fungsi pemisah (*classifier*) yang optimal yang bisa memisahkan dua set data yang berbeda (Vapnik, 1995). Fungsi pemisah atau *hyperplane* terbaik adalah *hyperplane* yang terletak di tengah-tengah dua objek dari kedua kelas. Adapun fungsi linier dari SVM:

$$g(x) = \text{sign}(f(x)) \tag{3}$$

dengan

$$f(x) = w^T x + b$$

$x, w \in R^n$ dan $b \in R$ masalah klasifikasi ini bisa dirumuskan sebagai berikut: untuk menemukan nilai dari parameter (w, b) maka $\text{sgn}(f(x)) = \text{sgn}(\langle w, x \rangle + b) = y_i$ untuk semua i .

Fungsi kernel yang digunakan untuk mengatasi data *non-linier*. diperoleh fungsi untuk non-linear sebagai berikut:

$$\begin{aligned} f(x) &= \text{sign} \left(\sum_{i=1}^n y_i \hat{\alpha}_i (\phi(x_i), \phi(x_j)) + \hat{b} \right) \\ &= \text{sign} \left(\sum_{i=1}^n y_i \hat{\alpha}_i (K(x_i, x_j)) + \hat{b} \right) \end{aligned}$$

Semua nilai $f(x) < 0$ diberi label "- 1" dan $f(x) > 0$ diberi label "+ 1". Macam-macam kernel diberikan pada Tabel 1 (Prasetyo 2012).

Tabel 1. Jenis Kernel

Nama Kernel	Definisi Fungsi
Linier	$K(x, y) = x \cdot y$
Polinomial	$K(x, y) = (x \cdot y + c)^d$
Gasussian RBF	$K(x, y) = \exp\left(\frac{-\ x-y\ ^2}{2\sigma^2}\right)$
Sigmoid (tangen hiperbolik)	$K(x, y) = \tanh(\sigma(x, y) + c)$
Invers Multikuadrik	$K(x, y) = \frac{1}{\sqrt{\ x-y\ ^2 + c^2}}$

2.5. Pengukuran Akurasi Klasifikasi

Untuk melakukan prediksi akurasi pengklasifikasian seringkali membantu dalam mengukur performa klasifikasi (Bramer, 2007). Berikut ada empat peluang yang dapat diformulasikan dalam tabel kontingensi 2 x 2 untuk menganalisis keakuratan klasifikasi:

Tabel 2. *Confusion Matrix*

		Diklasifikasikan sebagai	
		+	-
Klasifikasi yang Benar	+	Benar Positif (BP)	Salah Positif (SP)
	-	Salah Negatif (SN)	Benar Negatif (BN)

Keterangan untuk Tabel 2.2 dinyatakan sebagai berikut:



BP: jumlah data dari Kelas + yang benar dan diklasifikasikan sebagai kelas +
BN: jumlah data dari kelas - yang benar dan diklasifikasikan sebagai kelas -
SP: jumlah data dari kelas + yang benar dan diklasifikasikan sebagai kelas -
SN: jumlah data dari kelas - yang benar dan diklasifikasikan sebagai kelas +

Penghitungan akurasi klasifikasi dihitung dengan rumus berikut:

$$Akurasi = \frac{BP+BN}{BP+SP+BN+SN} \times 100\% \quad (4)$$

2.6. Pemilihan Kriteria Klasifikasi Terbaik

Dalam teknik klasifikasi untuk melihat tingkat akurasi terbaik dari metode k-NN dan SVM yaitu dengan menggunakan metode *Root Mean Square Error* (RMSE). RMSE merupakan salah satu cara untuk mengevaluasi model regresi linear dengan mengukur tingkat akurasi hasil perkiraan suatu model, RMSE adalah hasil dari akar kuadrat *Mean Square Error* (MSE). Keakuratan metode estimasi kesalahan pengukuran ditandai dengan adanya nilai RMSE yang kecil. Metode estimasi yang mempunyai RMSE lebih kecil dikatakan lebih akurat daripada metode estimasi yang mempunyai RMSE lebih besar. Berikut adalah rumus yang digunakan untuk menghitung nilai klasifikasi terbaik:

$$RMSE = \left(\frac{\sum (y_i - \hat{y}_i)^2}{n} \right)^{1/2} \quad (5)$$

Keterangan:

RMSE = Nilai Root Mean Square Error

y = Nilai hasil observasi

\hat{y} = Nilai hasil prediksi

i = Urutan data pada database

n = Jumlah data

2.7. Kemiskinan

BPS mendefinisikan kemiskinan merupakan ketidakmampuan individu dalam memenuhi kebutuhan dasar minimal untuk hidup layak. yaitu kondisi yang berada dibawah garis nilai standar kebutuhan minimum yang disebut garis kemiskinan (*poverty line*) atau batas kemiskinan (*poverty threshold*). Nilai garis kemiskinan yang digunakan untuk menentukan kemiskinan mengacu pada kebutuhan minimum 2.100 kalori per orang per hari ditambah dengan kebutuhan minimum non pangan. Individu yang pengeluarannya lebih rendah dari garis kemiskinan tersebut dikategorikan miskin (BPS, 2008).

3. Research Method

3.1. Jenis Penelitian

Penelitian yang dilakukan merupakan penelitian dengan pendekatan kuantitatif yaitu dengan mengambil data yang diperlukan dan menganalisisnya dengan menggunakan metode k-NN dan SVM untuk klasifikasi kemiskinan di Papua.

3.2. Sumber Data

Data yang digunakan dalam penelitian ini yaitu data kemiskinan di Papua tahun 2019. Data tersebut adalah data sekunder yang diambil dari BPS di Papua yang merupakan data tahunan.

3.3. Definisi Operasional Peubah



Peubah yang digunakan dalam penelitian ini terdiri dari peubah terikat (Y) terdiri dari dua kategori yaitu kategori 1 adalah banyaknya penduduk miskin tingkat rendah dan kategori 2 yaitu banyaknya penduduk miskin tingkat tinggi. Adapun peubah bebas (X) yang digunakan dalam penelitian ini.

1) Banyaknya Penduduk Miskin (Y)

Banyaknya penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan.

2) Produk Domestik Regional Bruto (X1)

Produk domestik regional bruto adalah jumlah nilai tambah bruto yang dihasilkan seluruh unit usaha dalam wilayah tertentu atau merupakan jumlah akhir yang dihasilkan oleh seluruh unit ekonomi.

3) Pengangguran (X2)

Pengangguran adalah suatu keadaan di mana seseorang yang tergolong dalam angkatan kerja ingin mendapatkan pekerjaan tetapi belum dapat memperolehnya.

4) Indeks Pembangunan Manusia (X3)

Indeks pembangunan manusia (IPM) menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan, dan sebagainya.

3.4. Prosedur Penelitian

- 1) Melakukan pengumpulan data sekunder, yaitu banyaknya penduduk miskin, distribusi PDRB ADH, pengangguran, dan IPM di Papua tahun 2019.
- 2) Melakukan pengkodean.
- 3) Melakukan pengklasifikasian k-NN dengan algoritma sebagai berikut:
 - a. Menentukan jumlah k untuk k-NN yaitu 5, 10, dan 15.
 - b. Menghitung akurasi klasifikasi
- 4) Melakukan pengklasifikasian SVM dengan algoritma sebagai berikut:
 - a. Menentukan fungsi kernel untuk pengklasifikasian, dalam penelitian ini menggunakan fungsi kernel *Radial Basis Function* (RBF)
 - b. Menentukan parameter kernel dan *Parameter cost* =1 untuk optimasi
 - c. Menghitung ketetapan klasifikasi
- 5) Membandingkan ketepatan klasifikasi yang diperoleh dari metode k-NN dan SVM dengan melihat RMSE.
- 6) Membuat kesimpulan dan saran.

4. Results and Discussion

For Results.

4.1. Analisis Deskriptif

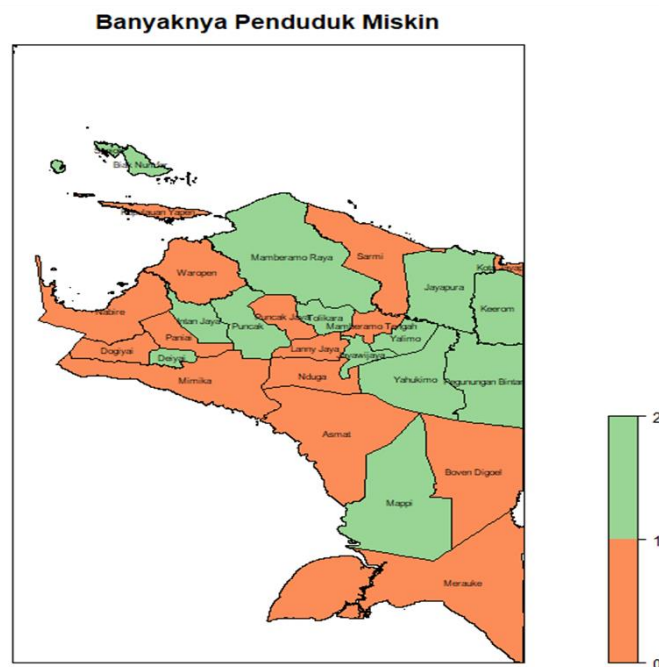
Berdasarkan variabel penelitian yang digunakan dari Data kemiskinan di Papua tahun 2019, diketahui banyaknya sampel yaitu 29 yakni di Kabupaten/Kota yang ada di Papua. Adapun analisis deskriptif yang dapat menggambarkan setiap variabel dapat dilihat dari Tabel 1.

Tabel 1. Statistika Deskriptif Data Kemiskinan di Papua tahun 2019

Variabel	Minimum	Maksimum	Rata-rata	Simpangan Baku
Banyaknya penduduk miskin (Y)	5,80	83,23	31,94	20,47
Produk domestik regional bruto (X1)	0,69	22,26	3,44	4,67
Pengangguran (X2)	0	14.597	1.961	3.024
Indeks pembangunan manusia (X3)	30,75	80,16	57,30	11,41

4.2. Mengkategorikan variabel Y

Nilai tengah dari banyaknya penduduk miskin di Papua yaitu 27,19 digunakan untuk mencari kategori dari variabel Y dengan pengkategorian banyaknya penduduk miskin rendah dengan angka 1 dan banyaknya penduduk miskin tinggi dengan angka 2.



Gambar 1. Pengkategorian variabel Y di Papua

Berdasarkan Gambar 1 dapat diketahui bahwa daerah dengan simbol 1 (Orange) menggambarkan daerah yang merepresentasikan banyaknya penduduk miskin tingkat rendah. Adapun daerah dengan simbol 2 (Hijau) menggambarkan daerah yang merepresentasikan banyaknya penduduk miskin tingkat tinggi. Dari Gambar 4.1 juga diketahui bahwa ada 15 Kabupaten/Kota di Papua yang berkategori penduduk miskin tingkat rendah diantaranya Merauke, Jayapura, Kepulauan Yapen, Boven Digoel, Mappi, Asmat, Pengunungan Bintang, Sarmi, Keerom, Waropen, Supiori, Mamberamo Raya, Mamberamo Tengah, Yalimo, serta Intan Jaya, selanjutnya dari Gambar 4.1 juga dapat dikatakan bahwa ada 14 Kabupaten/Kota di Papua yang berkategori banyaknya penduduk miskin tingkat tinggi diantaranya Jayawijaya, Nabire, Biak Numfor, Paniai, Puncak Jaya, Mimika, Yahukimo, Tolikara, Nduga, Lanny Jaya, Puncak, Dogiyai, Deiyai, dan Kota Jayapura. Penjelasan dalam Gambar 4.1 diatas juga diberikan dalam Tabel 2.

4.3. k-Nearest Neighbor (k-NN)

Pada metode k-NN dalam penelitian ini penentuan nilai k dimotivasi oleh asumsi bahwa angka-angka tersebut yang diduga merupakan angka yang mewakili banyaknya jumlah data yang paling dekat dengan objek yang pertama. Adapun nilai k yang digunakan yaitu nilai k = 5, k = 10 dan k = 15 untuk menentukan confusion matriks dengan mencari nilai akurasi, sensitivity, dan specificity yang dapat dilihat pada Tabel 3.

Tabel 2. Hasil Kategori dari Variabel Banyaknya Penduduk Miskin

Kategori	Kabupaten/Kota
1	Merauke, Jayapura, Kepulauan Yapen, Boven Digoel, Mappi, Asmat, Pengunungan Bintang, Sarmi, Keerom, Waropen, Supiori, Mamberamo Raya, Mamberamo Tengah, Yalimo, serta Intan Jaya,
2	Jayawijaya, Nabire, Biak Numfor, Paniai, Puncak Jaya, Mimika, Yahukimo, Tolikara, Nduga, Lanny Jaya, Puncak, Dogiyai, Deiyai, dan Kota Jayapura.

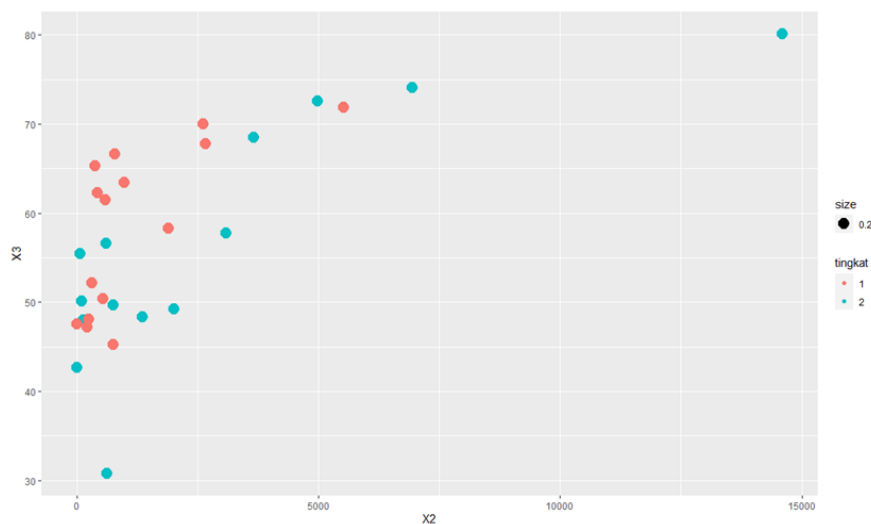
Tabel 3. Hasil Akurasi Metode k-NN yang di tinjau dari Nilai Parameter k

Nilai K	<i>Sensitivity</i>	<i>Specificity</i>	Akurasi
K=5	0,53	0,50	0,51
K=10	0,47	0,33	0,44
K=15	0,56	0,75	0,58

Berdasarkan hasil dari Tabel 3 dengan metode k-NN. Metode ini dilakukan dengan menggunakan parameter k yang diambil adalah k = 5, 10, dan 15. metode k-NN menggunakan nilai k = 5 menghasilkan nilai akurasi sebesar 51% dengan nilai *sensitivity* 53% dan nilai *specificity* sebesar 50%. Untuk metode k-NN menggunakan nilai k = 10 menghasilkan nilai akurasi sebesar 44% dengan nilai *sensitivity* 47% dan nilai *specificity* sebesar 33%. Untuk metode k-NN menggunakan nilai k= 15 menghasilkan nilai akurasi sebesar 58% dengan nilai *sensitivity* 56% dan nilai *specificity* sebesar 75%. Adapun nilai akurasi yang tinggi diantara nilai k = 5, k = 10, dan nilai k = 15 berada pada nilai k = 15 yaitu 58 %. Berdasarkan nilai k yang digunakan, hasil akurasi tertinggi terlihat pada nilai k = 15.

4.4. Plot Data Klasifikasi SVM

Melakukan plot data untuk mengetahui penyebaran data yang akan diklasifikasi, apakah data yang digunakan memiliki plot yang *linear-separable* atau *nonlinear-separable*. Berikut adalah hasil plot data variabel berdasarkan tingkat banyaknya penduduk miskin.



Gambar 2. Plot Data Pengangguran (X2) Dan IPM (X3)

Berdasarkan Gambar 2 diatas, penyebaran data tingkat banyaknya penduduk miskin tidak dapat dipisahkan atau diklasifikasikan secara linear. Oleh karena itu, data perlu ditransformasi ke dimensi yang lebih tinggi yang dalam penelitian ini menggunakan kernel RBF sebagai fungsi untuk mengatasi data yang tidak dapat dipisahkan secara linear (*nonlinear-separable*).

4.5. Support Vector Machine (SVM)

Pada metode SVM menggunakan kernel RBF dengan nilai *Parameter cost* = 1 dengan hasil dapat dilihat pada Tabel 4.

Tabel 4. Hasil Akurasi metode SVM yang di tinjau dari Nilai *Parameter cost*

Kernel	<i>Parameter cost</i>	<i>sensitivity</i>	<i>specifity</i>	akurasi
RBF	1	0,88	1	0,93

Berdasarkan Tabel 4 hasil akurasi dari SVM dengan menggunakan kernel RBF adalah 0,93 atau 93% dengan nilai *Parameter cost* = 1, Adapun nilai *sensitivity* adalah 0,88 atau 88% dengan nilai *specifity* sebesar 1.

4.6. Support Vector

Support Vector merupakan titik data terdekat dengan *hyperplane* masing-masing kelas.

Tabel 5. *Support Vector*

Kernel	<i>Parameter cost</i>	<i>Support vector</i>
RBF	1	20

Berdasarkan Tabel 5 jumlah *support vector* sebanyak 20 data yang terdapat pada *Parameter cost* = 1. Hal ini berarti terdapat 20 data yang terletak didekat masing-masing *hyperplane* kelompok banyaknya penduduk miskin tingkat rendah dan banyaknya penduduk miskin tingkat tinggi.

4.7. Perbandingan k-NN dan SVM

Nilai *Parameter cost* dan nilai k dijadikan input untuk proses klasifikasi sehingga menghasilkan tingkat akurasi. Berikut perbandingan akurasi terbaik berdasarkan *Parameter cost* untuk SVM dan nilai k untuk k-NN dengan hasil dapat dilihat pada Tabel 6.

Tabel 6. Perbandingan k-NN dan SVM

Metode	RMSE	Akurasi
k-NN	0,69	0,58
SVM	0,26	0,93

Berdasarkan Tabel 6 k-NN dengan akurasi 0,58 atau 58% merupakan k-NN dengan nilai k = 15, sedangkan SVM dengan akurasi 0,93 atau 93% merupakan *Parameter cost* = 1. Adapun nilai RMSE pada metode k-NN sebesar 0,69, sedangkan nilai RMSE pada metode SVM sebesar 0,26. Jadi akurasi terbaik adalah metode SVM dengan *Parameter cost* = 1.

5. Conclusion

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan untuk mengklasifikasikan data kemiskinan di Papua dengan metode k-NN dan SVM dapat disimpulkan bahwa metode k-NN dengan nilai k=15 menghasilkan akurasi sebesar 0,58, sedangkan metode SVM dengan *Parameter cost* = 1 menghasilkan nilai akurasi sebesar 0,93. Kedua metode tersebut jika dibandingkan metode SVM dengan *Parameter cost* = 1 merupakan metode paling tepat dibandingkan dengan metode k-NN dalam data kemiskinan di Papua tahun 2019.

References

- Aulia, S. (2015). *Analisis Perbandingan KNN dengan SVM Klasifikasi Penyakit Diabetes Retinopati berdasarkan Citra Eksudat dan Mikroaneurisma*. Program Studi D3 Teknik Telekomunikasi Fakultas Ilmu Terapan Telkom University: Jurnal Elkomika.



- Bhavsar, H., & Panchal, M. H. (2012). *A Review on Support Vector Machine for Data Classification. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*.
- BPS. (2008). *Analisis dan Penghitungan Tingkat Kemiskinan Tahun 2008*. Badan Pusat Statistik: Jakarta.
- BPS. (2019). *Perhitungan dan Analisis Kemiskinan Makro Indonesia Tahun 2019*. Badan Pusat Statistik: Jakarta.
- Bramer M. (2007). *Principles of Data Mining*. Springer: London.
- Davies, & Paul B. (2004). *Database System Third Edition*. Palgrave Macmillan: New York.
- Pramudiono, I. (2007). *Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data*. ilmukomputer :Spinger
- Simatupang, F. J., Wuryandari, T., & Suparti. (2016). *Klasifikasi Rumah Layak Huni di Kabupaten Brebes dengan Menggunakan Metode Learning Vector Quantization dan Naive Bayes*. Gaussian.
- Ustuner, M., Samli, F. B., & Dixon, B. (2015). *Application of Support Vector Machines for Landuse Classification Using High-Resolution RapidEye Images. A Sensitivity Analysis: European Journal of Remote Sensing*.
- Vapnik, V dan Cortes, C. (1995). *Support Vector Networks. Machine Learning*, 20, 273-297.
- Wu X & Kumar V. (2009). *The Top Ten Algorithms in Data Mining*. Taylor & Francis Group: USA.