

IMPLEMENTASI METODE *K-NEAREST NEIGHBOR* UNTUK KLASIFIKASI PENDUDUK MISKIN DI DESA NGEPLAK KIDUL KABUPATEN PATI JAWA TENGAH

Ahmad Zainuddin

Program Studi Teknik Informatika, STMIK Cikarang

E-mail: ahmadzain5@gmail.com

ABSTRAKSI

Penanggulangan kemiskinan dan pemberdayaan masyarakat, merupakan salah satu tugas pemerintah daerah yang harus dijalankan, sesuai dengan amanah peraturan perundangan, sehingga pengklasifikasian atau pendataan yang baik merupakan suatu hal yang harus dilakukan. Untuk mendapatkan hasil tersebut maka dibutuhkan penerapan teknologi masa kini yaitu tentang pemanfaatan program aplikasi. Program aplikasi merupakan salah satu cara yang efektif dalam hal pengklasifikasian penduduk miskin tersebut. Keberadaan program aplikasi tentunya akan memberikan data yang akurat dan tepat, sehingga kebijakan yang harus diambil oleh pemerintah desa Ngeplak Kidul jadi tepat sasaran atau dengan kata lain sesuai dengan kebutuhan masyarakat. Pengklasifikasian data penduduk miskin juga dapat menentukan arah kebijakan bagi pemerintah daerah dalam hal program mengentaskan kemiskinan. Klasifikasi pendataan penduduk miskin dengan memanfaatkan teknologi terkini untuk memenuhi kebutuhan perangkat desa, sebagai suatu pemecahan masalah dari ketidakakuratan, ketepatan data untuk mendapatkan hasil klasifikasi penduduk miskin, proses dan tujuan dari klasifikasi itu sendiri. Tujuan penelitian ini adalah membuat Program aplikasi Klasifikasi Penduduk Miskin, dengan menggunakan metode *K-nearest neighbor* sebagai pengambilan keputusan klasifikasi penduduk miskin. Hasil dari pembuatan program aplikasi klasifikasi penduduk miskin, di desa Ngeplak Kidul kabupaten Pati, dengan metode *k-nearest neighbor* adalah untuk mempermudah perangkat desa, dalam mendapatkan hasil secara tepat sasaran, data akurat secara efisien waktu.

Keyword : Klasifikasi, *K-Nearest Neighbor*, Desa Ngeplak Kidul.

1. Pendahuluan

Desa Ngeplak Kidul adalah salah satu Pemerintah yang berada di Kabupaten Pati. Desa Ngeplak Kidul ini mempunyai kewajiban untuk melayani masyarakat dan selalu ingin memberikan informasi pelayanan terbaik bagi masyarakat, dengan cara selalu memperbaiki sistem yang ada. Permasalahan kemiskinan merupakan salah satu persoalan mendasar yang terus di hadapi sejumlah daerah di Indonesia. Berbagai upaya untuk mengentaskan kemiskinan telah dilakukan oleh pemerintah yang diaplikasikan dalam wujud kebijakan dan program-program baik yang bersifat langsung maupun tidak langsung.

Kebijakan bersifat langsung yaitu berupa program yang langsung diberikan kepada penduduk miskin, contoh: bantuan langsung tunai (BLT), raskin, sedangkan kebijakan tidak langsung, contoh program Jamkesmas, program IDT, BOS. Walaupun telah dilakukan berbagai upaya namun kemiskinan tidak dapat dihilangkan seluruhnya artinya fenomena kemiskinan dengan mudah dapat dijumpai di hampir seluruh wilayah baik di perkotaan maupun di pedesaan.

Program kemiskinan yang saat ini dilakukan baik yang berasal dari pemerintah maupun non pemerintah umumnya hanya sementara, artinya program tersebut akan berjalan selama masih ada anggaran (dana), setelah dana habis maka selesai pula kegiatan program. Dengan kata lain bahwa program-program kemiskinan yang selama ini dilaksanakan merupakan berdasarkan pada pendekatan proyek dan bukan pendekatan program. Tidak heran jika akhirnya program pengentasan kemiskinan tidak berkelanjutan sehingga angka kemiskinan secara absolut di Indonesia tetap saja tinggi.

Peranan data kemiskinan menjadi sangat penting dalam

keberhasilan pelaksanaan program tersebut. Berbagai indikator untuk mengukur tingkat kemiskinan dan menghitung jumlah penduduk miskin telah ditetapkan, dengan harapan upaya mengatasi kemiskinan akan lebih tepat sasaran.

Permasalahan yang sering terjadi dalam program-program pengentasan kemiskinan adalah tidak selarasnya antara program yang dicanangkan oleh Pemerintah Daerah dengan masalah yang dihadapi oleh masyarakat, bahkan mungkin sama sekali tidak tersentuh, sehingga hasil yang diperoleh sudah barang tentu jauh dari harapan masyarakat. Hal ini pada umumnya terjadi karena data penduduk miskin tidak tertata dengan baik, sehingga program pengentasan kemiskinan belum tepat objek dan sasaran yang berhak menerimanya. Berdasarkan hal tersebut, maka dirancang dan dibangun sebuah sistem informasi yang dapat digunakan untuk mengolah data penduduk miskin di Desa Ngeplak Kidul dengan metode *K-Nearest Neighbor*.

2. Landasan Kepustakaan

Penulis memulai penelitian ini dengan terlebih dahulu melakukan studi kepustakaan dari penelitian-penelitian dan sumber-sumber lain. Dari studi kepustakaan itu, penulis menemukan beberapa penelitian yang mendorong untuk mengangkat tema seperti diatas. Penelitian tersebut membahas tentang topik yang terkait dengan penelitian penulis, antara lain adalah penelitian mengenai algoritma yang digunakan penulis yang akan diangkat oleh penulis.

Tabel 1. Penelitian Terkait

No	Judul	Metode (Proses)	Hasil
1	Implementasi Metode K-Nearest Neighbor untuk Rekomendasi Keminatan pada Studi (Studi Kasus: Jurusan Teknik Informatika Universitas Brawijaya)	K-Nearest Neighbor (KNN)	Implementasi metode K-Nearest Neighbor untuk rekomendasi keminatan pada studi mahasiswa teknik informatika dapat digunakan sebagai sarana bagi mahasiswa yang masih bingung untuk menentukan keminatan yang akan diambil. Berdasarkan pengujian yang telah dilakukan dari 30 data uji dengan membandingkan antara hasil sistem dengan data yang diperoleh dari pakar maka diperoleh hasil akurasi sebesar 76,66%.
2	MODEL ALGORITMA K-NEAREST NEIGHBOR (K-NN) UNTUK PREDIKSI KELULUSAN MAHASISWA	K-Nearest Neighbor (KNN)	Dalam penelitian ini dilakukan pengujian model dengan menggunakan algoritma K-Nearest Neighbor (K-NN) dengan menggunakan data kelulusan mahasiswa yang tepat dan terlambat. Model yang dihasilkan diuji untuk mendapatkan nilai accuracy dan AUC dari

			algoritma klasifikasi data mining sehingga didapat pengujian data mahasiswa dengan klatering data k=1 dengan menggunakan K-Nearest Neighbor (K-NN) didapat nilai accuracy adalah 82,25% dan nilai AUC adalah 0.500, dengan cluster data k=2 accuracy adalah 79,45% dan nilai AUC adalah 0.826, dengan cluster data k=3 accuracy adalah 83,95% dan nilai AUC adalah 0.853, dengan cluster data k=4 accuracy adalah 82,62% dan nilai AUC adalah 0.874, dengan cluster data k=5 accuracy adalah 85,15% dan nilai AUC adalah 0.888
--	--	--	--

2.1 Klasifikasi

Klasifikasi adalah sebuah metode untuk mengelompokkan data secara sistem menurut aturan dan kaidah yang telah ditetapkan. Klasifikasi juga dapat diartikan pengelompokkan data atau objek baru berdasarkan variabel yang diamati dengan tujuan untuk memprediksi suatu objek dari yang masih belum diketahui kelas atau kategorinya (Tri Halomoan Simanjuntak, 2014).

Klasifikasi merupakan teknik data mining yang melihat atribut dari kelompok data yang sudah didefinisikan sebelumnya. Sehingga dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang sudah diklasifikasi dan menggunakan hasilnya untuk memberikan beberapa aturan. Aturan-aturan tersebut dipakai pada data-data baru untuk selanjutnya diklasifikasi (Siti Hardiyanti, 2014).

Tujuan dari klasifikasi yaitu supaya record yang tidak diketahui pada kategori mana sebelumnya dapat dikelompokkan kelasnya secara akurat (Sofa Zainuddin, 2014).

2.2 K-Nearest Neighbor

K-Nearest Neighbor (KNN) merupakan metode yang biasa digunakan pada klasifikasi data. Algoritma ini digunakan untuk mengklasifikasikan terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. KNN merupakan suatu metode yang menggunakan algoritma supervised dengan hasil dari query instance yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN. Tujuan dari algoritma ini ialah mengklasifikasikan objek baru berdasarkan atribut dan training sample (Larose, 2005).

Prinsip umum dari algoritma ini adalah menemukan k data training untuk menentukan k-nearest neighbor berdasarkan ukuran jarak. Selanjutnya mayoritas dari k tetangga terdekat akan menjadi dasar untuk memutuskan kategori dari sample berikutnya (Yu, 2010). Selain itu algoritma ini sendiri sering digunakan untuk klasifikasi pada teknik data mining meskipun dapat digunakan untuk estimasi dan prediksi data. Metode ini adalah contoh dari instance based learning dimana data training di simpan, sehingga proses klasifikasi dari data yang belum diklasifikasi dapat dengan mudah ditemukan dengan membandingkan data tersebut dengan data yang paling mirip di data training yang ada (Larose, 2005).

Dalam *K-Nearest Neighbor*, diperlukan tiga hal yang penting :

- Kumpulan record yang tersimpan
- Matrik jarak untuk menghitung jarak antar record
- Nilai k jumlah tetangga terdekat yang dicari.

Sehingga untuk mengklasifikasi record yang belum diketahui :

- Menghitung jarak ke data latih lain
- Mengidentifikasi k tetangga terdekat
- Menggunakan kelas label dari tetangga terdekat untuk menentukan label kelas dari record yang belum diketahui (misal : dengan mengambil suara terbanyak)

Adapun langkah-langkah dalam metode KNN adalah sebagai berikut :

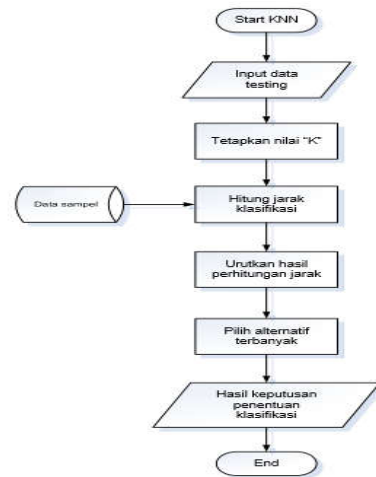
1. Menghitung jarak *Euclidean*, rumus jarak *euclidean*:

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Gambar 2.6 Rumus *Euclidean Distance*

2. Mengurutkan berdasarkan nilai *Euclidean Distance*
3. Menentukan k record klasifikasi terdekat
4. Target output merupakan kelas yang mayoritas



Gambar 2.7 Alur proses klasifikasi metode *K-Nearest Neighbor*

3. Rancangan Sistem dan Aplikasi

Dalam pembahasan ini terdapat dataset sebagai data testing untuk menghitung nilai *euclidean distance* terhadap data baru. Dataset yang mempunyai kelas/target/label nominal.

Tabel 3.1
Tabel Data Training

No	X1	X	X3	X4	X5	X6	Keterangan
1	7	150	4	900	1	4000	TM
2	4	100	3	900	1	2000	M
3	9	120	3	1300	2	5000	TM
4	5	100	2	450	1	3000	M
5	8	100	4	900	2	3000	TM
6	4	150	3	450	1	2000	M
7	10	150	3	1300	2	6000	?

Keterangan variabel x yaitu

1. X1 = Jenis Dinding (Untuk jenis dinding penilaian ekuivalen 1-10),
2. X2 = Luas Bangunan (Untuk parameter ekuivalen 1-3),
3. X3 = Jumlah Anggota Keluarga,
4. X4 = Daya Listrik (Untuk parameter ekuivalen 1-4),
5. X5 = Penggunaan Air (1. Sumur Bor, 2. PAM),
6. X6 = Pendapatan Keluarga (Untuk parameter ekuivalen 1-4),
7. Keterangan : TM (Tidak Miskin), M (Miskin).

Tabel 3.2
Tabel Data Indikator

No	Nama Indikator	Satuan	Deskripsi	Jenis Input	Pilihan Input
1	Jenis Dinding		ada 10 Jenis	Select	1) Dinding Bambu, 2) Dinding Papan Triplek, 3) Dinding Kayu, 4) Dinding Bata Tanpa Plaster, 5) Dinding Bata Plaster, 6) Dinding Bata Plaster Tanpa Cat, 7) Dinding Bata Plaster Cat, 8) Dinding Beton, 9) Dinding Beton Plaster, 10) Dinding Beton Paster Cat
2	Luas Bangunan	m ²	Luas Bangunan Rumah	Input	
3	Jumlah Anggota Keluarga	Orang	Jumlah Anggota Keluarga	Input	
4	Daya listrik	watt		Input	
5	Penggunaan Air		Penggunaan Air	Select	1) Sumur Bor, 2) PAM
6	Pendapatan Keluarga	Rupiah	Pendapatan Keluarga	Input	

Data sample pada tabel 3.1 tersebut akan di implementasikan dengan formula *K-Nearest Neighbor*. Adapun cara kerja dari proses perhitungan *K-Nearest Neighbor* yaitu sebagai berikut : Tahapan diawali dengan mengambil data sample atau contoh data seperti pada table 3.1 Kemudian dari tabel diatas dapat di hitung dengan menggunakan formula *K-Nearest Neighbor*, adapun cara kerjanya adalah sebagai berikut :

1. Menentukan parameter K, disini dalam menentukan parameter k=3.
2. Menghitung jarak euclid (*Jarak euclidean*) masing-masing Objek terhadap data sample yang diberikan
3. Urutkan jarak tersebut dan tentukan tetangga mana yang terdekat berdasarkan jarak minimum ke-k.
4. Menentukan kategori dari tetangga terdekat.
5. Gunakan kategori mayoritas yang sederhana dari tetangga yang terdekat tersebut sebagai nilai prediksi dari data yang baru.

Menghitung jarak antara data baru dan semua data yang ada di data training. Misal digunakan square distance dari jarak antara data baru dengan semua data yang ada di data training. Contoh perhitungan numerik (Perhitungan dengan cara manual).

$$\begin{aligned}
 1 &= \sqrt{(7-10)^2 + (150-150)^2 + (4-3)^2 + (900-1300)^2 + (1-2)^2 + (4000-6000)^2} \\
 &= \sqrt{9+0+1+16000+1+4000000} \\
 &= \sqrt{4016010} \\
 &= 2039,610502 \\
 2 &= \sqrt{(8-10)^2 + (100-150)^2 + (3-3)^2 + (900-1300)^2 + (1-2)^2 + (2000-6000)^2} \\
 &= \sqrt{36+2500+0+16000+1+16000000} \\
 &= \sqrt{16018537} \\
 &= 4020,265787 \\
 3 &= \sqrt{(9-10)^2 + (120-150)^2 + (3-3)^2 + (1300-1300)^2 + (2-2)^2 + (5000-6000)^2} \\
 &= \sqrt{1+900+0+0+0+10000000} \\
 &= \sqrt{10009901} \\
 &= 1000,450399 \\
 4 &= \sqrt{(5-10)^2 + (100-150)^2 + (2-3)^2 + (450-1300)^2 + (1-2)^2 + (3000-6000)^2} \\
 &= \sqrt{25+2500+1+722500+1+9000000} \\
 &= \sqrt{9725025} \\
 &= 3118,497555 \\
 5 &= \sqrt{(8-10)^2 + (100-150)^2 + (4-3)^2 + (900-1300)^2 + (2-2)^2 + (3000-6000)^2} \\
 &= \sqrt{4+2500+1+16000+0+9000000} \\
 &= \sqrt{9018505} \\
 &= 3026,962999 \\
 6 &= \sqrt{(4-10)^2 + (150-150)^2 + (3-3)^2 + (450-1300)^2 + (1-2)^2 + (2000-6000)^2} \\
 &= \sqrt{36+0+0+722500+1+16000000} \\
 &= \sqrt{16722537} \\
 &= 4089,31987
 \end{aligned}$$

Hasil perhitungan menunjukkan nilai euclidean distance data training yang telah diuji satu persatu.

Tabel 3.3
Nilai Euclidean Distance Data Test

Rangking	Euclidean distance	Keterangan
2	2039,610502	TM
5	4020,265787	M
1	1000,450399	TM
4	3118,497555	M
3	3026,962999	TM
6	4089,31987	M

Parameter K yang telah ditentukan K=3. Maka diambil data nilai terdekat yang telah di rangking sejumlah 3 data training.

Tabel 3.4
Rangking Euclidean Diastance

Rangking	Euclidean distance	Keterangan
1	1000,450399	TM
2	2039,610502	TM
3	3026,962999	TM

Dari data hasil perhitungan dengan sejumlah K yg telah ditentukan. Maka dapat disimpulkan dari kelompok yang

mayoritas adalah Tidak miskin (TM) jadi data baru termasuk “Tidak Miskin (TM).

Tabel 4.5
Penentuan Kategori Data Baru

No	X1	X	X3	X4	X5	X6	Euclidean Distance	Rangking	Keterangan
1	7	150	4	900	1	4000	2039,610502	2	TM
2	4	100	3	900	1	2000	4020,265787	5	M
3	9	120	3	1300	2	5000	1000,450399	1	TM
4	5	100	2	450	1	3000	3118,497555	4	M
5	8	100	4	900	2	3000	3026,962999	3	TM
6	4	150	3	450	1	2000	4089,31987	6	M
7	10	150	3	1300	2	6000	-	-	TM

3.1 Tampilan Aplikasi

Tampilan Menu Login



Gambar 3.1
Tampilan Form Login

Tampilan Menu Utama/Dashbord



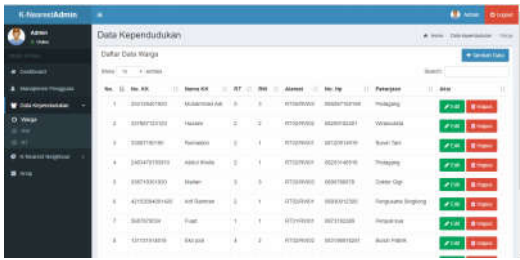
Gambar 3.2
Tampilan Menu Utama/Dashboard

Tampilan Menu Pengguna



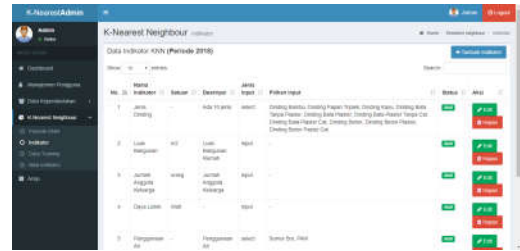
Gambar 3.3
Tampilan Menu Pengguna

Tampilan Menu Data Warga



Gambar 3.4
Tampilan Menu Data Warga

Tampilan Menu Indikator



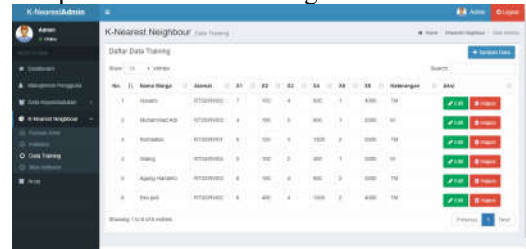
Gambar 3.8
Tampilan Menu Indikator

Tampilan Menu Data RW



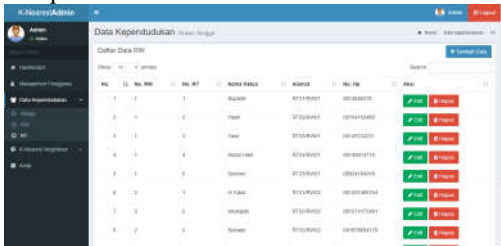
Gambar 3.5
Tampilan Menu Data RW

Tampilan Menu Data Training



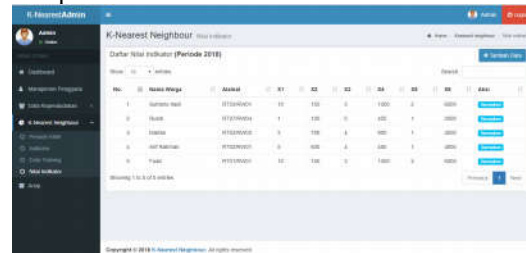
Gambar 3.9
Tampilan Menu Data Training

Tampilan Menu Data RT



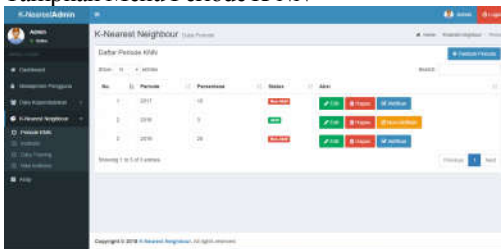
Gambar 3.6
Tampilan Menu Data RT

Tampilan Menu Nilai Indikator



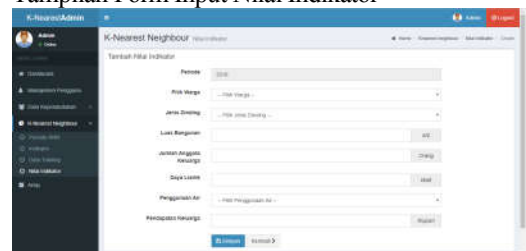
Gambar 3.10
Tampilan Menu Nilai Indikator

Tampilan Menu Periode K-NN



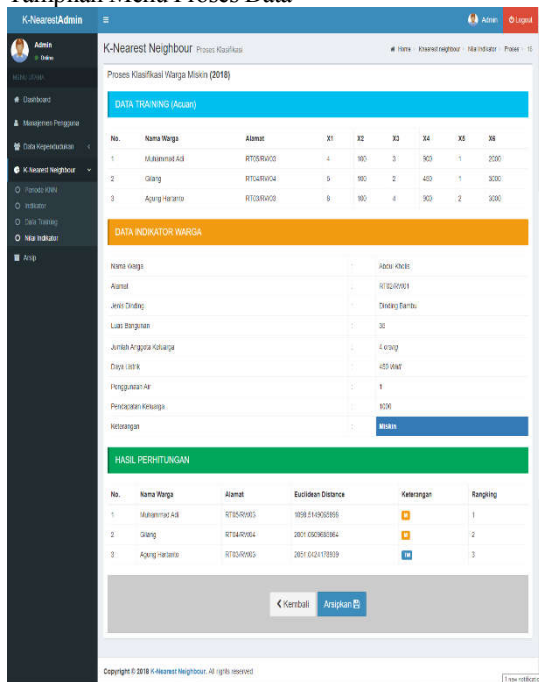
Gambar 3.7
Tampilan Menu Periode K-NN

Tampilan Form Input Nilai Indikator



Gambar 3.11
Tampilan Form Input Nilai Indikator

Tampilan Menu Proses Data



Gambar 3.12

Tampilan Menu Proses Data

Tampilan Menu Arsip



Gambar 3.13

Tampilan Menu Arsip

Tampilan Output Arsip



Gambar 3.14

Tampilan Output Arsip

4. Kesimpulan

Berdasarkan hasil analisa dan perancangan aplikasi program implementasi metode *K-Nearest neighbor* untuk klasifikasi penduduk miskin di desa Ngeplak kidul kabupaten pati jawa tengah, maka dapat diambil kesimpulan sebagai berikut:

1. *K-Nearest Neighbor* dapat digunakan untuk pengklasifikasian penduduk miskin dengan obyek baru berdasarkan atribut. Dan dalam pengklasifikasian penduduk miskin pada setiap priode nilai K yang berubah tidak mempengaruhi hasil akhir dan konsisten dari klasifikasi tersebut. Algoritma KNN menggunakan klasifikasi ketetangaan sebagai nilai prediksi dari query instance yang baru.
2. Sistem Klasifikasi Penduduk Miskin di Desa Ngeplak kidul Kabupaten Pati dengan Menggunakan Metode *K-Nearest Neighbor* dibuat agar memudahkan beban proses kerja perangkat desa Ngeplak kidul selama ini, pada sistem ini perangkat desa hanya menginputkan data baru yang telah masuk untuk mengklasifikasikan penduduk miskin secara tepat, dan akurat dengan indikator yang ditentukan tanpa menyurvei satu persatu rumah warga satu desa.

5. Saran-Saran

Kedepannya penulis mengharapkan akan ada pengembangan positif dari aplikasi program klasifikasi yang telah dibuat seperti :

1. Adanya pelaporan proses dan hasil klasifikasi secara transparan terhadap warga agar warga mengetahui hasil penilaian dari perhitungan indikator, yang diharapkan mampu menjadi panduan pembagian bantuan kemiskinan secara tepat.
2. Pengembangan aplikasi menjadi lebih kompleks sesuai kebutuhan seiring perkembangan masa
3. Dilakukan penelitian lebih mendalam untuk menggali berbagai macam indikator klasifikasi sehingga mencakup segala aspek kondisi sosial di masing-masing daerah.

DAFTAR PUSTAKA

Anshori Luthfi, Putri Rekyan Regasari Mardi, Tibyani. 2018. Implementasi Metode *K-Nearest Neighbor* Untuk Rekomendasi Keminatan Studi (Studi Kasus : Jurusan Teknik Informatika Universitas Brawijaya) e-ISSN: 2548-964X

Hardiyanti, Siti. 2014. Implementasi Metode Modified *K-Nearest Neighbor* (MKNN) Pada Penentuan Keminatan Sekolah Menengah Atas (SMA) (Studi Kasus: SMA Negeri 1 Seririt). Universitas Brawijaya. Malang. ISSN: 2548-964X

Leidiyana, H., 2010. Penerapan Algoritma *K-Nearest Neighbor* Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor. *Jurnal Penelitian*

- Ilmu Komputer, System Embedded & Logic*, Vol : 1. STMIK Nusa Mandiri. 2010
- Mulfikhah Laili, Ratnawati Dian Eka, Putri Rekyan Regasari Mardi. *Data Mining* 2018, ubpress, 2018
- Mustakim, Oktafiani Giantika. 2016. *Algoritma K-Nearest Neighbor Clasification* Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa. ISSN 2407-0939
- Supono dan Putratama Viridiandry. *Pemrograman Web dengan Menggunakan PHP dan Framework Codeigniter 2018*, Deepublish, CV Budi Utama, 2018.
- Tri Halomoan Simanjuntak, Wayan Firdaus Mahmudy dan Sutrisno. 2014. Implementasi Modified K-Nearest Neighbor dengan Otomatisasi Nilai K Pada Pengklasifikasian Penyakit Tanaman Kedelai. e-ISSN : 2548-964X
- Wardana. *Menjadi master PHP dengan framework codeigniter 2010*, PT.Elex media komputindo kelompok Gramedia, Anggota IKAPI, Jakarta, 2010