

# TEKNIK *GENETIC MODIFIED K-NEAREST NEIGHBOR* UNTUK ESTIMASI HASIL PRODUKSI GULA TEBU BERDASARKAN NILAI KLOORIFIL DAUN TEBU

Siti Mutrofin<sup>1)</sup>, R. V. Hari Ginardi<sup>2)</sup>, dan Chastine Faticah<sup>3)</sup>

<sup>1)</sup>Sistem Informasi, Fakultas Teknik, Universitas Pesantren Tinggi Darul Ulum Kompleks Ponpes Darul Ulum Rejoso Peterongan Jombang 61485

<sup>2,3)</sup>Teknik Informatika, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember Jl. Teknik Kimia, Gedung Teknik Informatika, Kampus ITS Sukolilo, Surabaya 60111  
e-mail: [mutrofin12@mhs.if.its.ac.id](mailto:mutrofin12@mhs.if.its.ac.id)<sup>1)</sup>, [hari@its.ac.id](mailto:hari@its.ac.id)<sup>2)</sup>, [chastine@cs.its.ac.id](mailto:chastine@cs.its.ac.id)<sup>3)</sup>

## ABSTRAK

*Pada umumnya, taksasi (estimasi) hasil produksi gula didasarkan pada luas area, masa tanam, faktor panjang leng, jumlah batang/leng, tinggi batang, bobot batang, dan rendemen. Faktor rendemen sangat berpengaruh terhadap kualitas gula (manis atau tidak), sedangkan faktor yang lain hanya mempengaruhi kuantitas gula. Salah satu tolak ukur produksi gula maupun rendemen dapat dilihat dari nilai klorofil daun tebu. Klorofil juga memiliki peranan yang sangat penting bagi proses fotosintesis. Untuk itu, dalam penelitian taksasi produksi gula tebu ini didasarkan pada nilai klorofil daun tebu. Dalam penelitian ini, performa algoritma Modified k-Nearest Neighbor (MKNN) ditingkatkan dengan cara mengoptimalkan nilai k dengan menggunakan algoritma genetika, yang selanjutnya algoritma tersebut dinamakan dengan algoritma Genetic Modified K-Nearest Neighbor (GMKNN). Tujuan dari penelitian ini adalah GMKNN sebagai algoritma baru digunakan untuk melakukan estimasi hasil produksi gula tebu berdasarkan nilai klorofil daun tebu. Dari sejumlah uji coba, terbukti bahwa klorofil (Model2) dapat digunakan untuk melakukan taksasi, walaupun memiliki hasil yang kurang baik dibandingkan ketika tidak menggunakan klorofil (Model1). Model2 unggul pada percobaan dengan menggunakan 95 data, Model2 juga terbukti memiliki perbedaan yang signifikan berdasarkan uji T. Kinerja algoritma GMKNN lebih baik dari pada MKNN dengan nilai MSE terkecil sebesar 3737 pada percobaan dengan menggunakan 95 data, dan MSE terbesar sebesar 2053730 pada percobaan dengan menggunakan 201 data, namun GMKNN juga memiliki kekurangan, yaitu komputasi tinggi, terjebak pada optimum lokal, sedangkan kekurangan MKNN adalah penentuan nilai ambang batas dalam mendapatkan nilai similaritas antar data latih yang masih manual pada kasus estimasi.*

**Kata Kunci:** *Genetic Modified K-Nearest Neighbor, klorofil, MKNN, tebu, validitas.*

## I. PENDAHULUAN

Salah satu komoditas pangan utama di Indonesia yang akan ditingkatkan produksinya oleh pemerintah Indonesia adalah gula, sebagai target swasembada untuk industri pada tahun 2014. Sebab, swasembada untuk komoditas gula konsumsi langsung rumah tangga sudah terwujud sejak tahun 2008. Revitalisasi teknologi dan industri hilir termasuk dari strategi Kementerian Pertanian (Kementan) untuk mewujudkan hal tersebut [1].

Banyak peneliti yang telah berperan di bidang revitalisasi teknologi, diantaranya adalah terkait penanganan hama penyakit tebu, peningkatan kualitas bibit tebu, perbaikan perlakuan dalam

bertani tebu, peningkatan produktivitas tebu, maupun tentang taksasi hasil produksi gula. Menurut [2] dan [3] salah satu upaya untuk meningkatkan produktivitas gula atau mewujudkan swasembada gula nasional 2014 adalah dengan cara meningkatkan rendemen tebu. Tolak ukur produksi gula maupun rendemen tebu dapat dilihat dari nilai klorofil daun tebu [3]. Klorofil juga memiliki peranan yang sangat penting bagi proses fotosintesis. Sementara itu, secara umum dalam melakukan taksasi produksi gula, pabrik gula hanya memperhatikan 6 faktor, yaitu: 1) Faktor leng (juring), 2) Jumlah batang/leng, 3) Tinggi batang, 4) Berat batang, 5) Rendemen, 6) Luas lahan dan 7) Masa tanam.

Pentingnya klorofil terhadap hasil produktivitas gula, menjadikan tantangan untuk melakukan penelitian terkait taksasi (estimasi) produksi gula berdasarkan nilai klorofil daun tebu, karena dalam proses taksasi gula belum pernah dilakukan penelitian taksasi produksi gula yang menggunakan nilai klorofil daun tebu. Ada beberapa teknik untuk melakukan estimasi, salah satunya adalah menggunakan algoritma *K-Nearest Neighbor* (kNN) [4] dan [5].

kNN adalah salah satu algoritma estimasi yang mendasar dan sederhana, kNN memiliki beberapa kelebihan, diantaranya: 1) Mudah digunakan, 2) Kuat terhadap data pelatihan yang memiliki derau, dan 3) Efektif jika data pelatihan besar. Namun, kNN juga memiliki kelemahan, yaitu: 1) Komputasi yang cukup tinggi, karena harus menghitung jarak setiap kueri untuk semua sampel pelatihan, 2) Membutuhkan memori yang besar, 3) Tingkat akurasi rendah ketika data pelatihan memiliki multi dimensi, 4) Harus menentukan nilai  $k$  secara manual, 5) Jarak yang berdasarkan pembelajaran tidak jelas, jenis jarak apa yang digunakan, dll [6]. Penelitian terkait untuk perbaikan kNN sudah banyak ditemui, salah satunya adalah *Modified k-Nearest Neighbor* (MKNN), MKNN merupakan salah satu turunan dari algoritma kNN, yang bertujuan untuk meningkatkan akurasi dari algoritma kNN. MKNN mempunyai kinerja yang lebih baik dari pada kNN [7] karena menggunakan bobot dan validitas titik data untuk mengestimasi tetangga terdekat. Namun, masih terdapat kekurangan dari algoritma kNN yang belum terselesaikan oleh algoritma MKNN, yaitu komputasi dengan tingkat kompleksitas yang tinggi [8]. Namun, banyak penelitian tentang MKNN hanya diperuntukkan untuk tugas klasifikasi, sedangkan kasus usulkan dalam penelitian ini adalah estimasi, sehingga perhitungan validitasnya, dalam penelitian ini dilakukan pergantian dengan menggunakan perhitungan nilai MSE dari data latihnya.

Berdasarkan permasalahan tersebut, dalam penelitian yang diusulkan bertujuan untuk melakukan perbaikan MKNN dengan cara mengoptimalkan nilai  $k$  menggunakan algoritma *Genetic Algorithm* (GA) pada kasus estimasi hasil produksi gula tebu berdasarkan nilai klorofil tebu. Karena, GA sudah terbukti dapat digunakan untuk optimalisasi nilai  $k$  dari algoritma kNN [9]. Algoritma tersebut selanjutnya dinamakan sebagai

algoritma *Genetic Modified k-Nearest Neighbor* (GMKNN).

## II. PERSIAPAN DATA

Pada penelitian ini data primer didapatkan dari Pabrik Gula (PG) Kebon Agung Malang pada periode tahun 2013 yang bertujuan untuk mendapatkan data klorofil daun tebu, sedangkan data sekunder didapatkan dari PG Pesantren Baru Kediri pada periode tahun 2009-2012 untuk mendapatkan data rendemen. Data dari PG Kebon Agung Malang maupun PG Pesantren Baru Kediri memiliki permasalahan, yaitu banyak terdapat *missing value* (data yang tidak bernilai atau kosong) seperti ditunjukkan pada Tabel I, sehingga untuk menangani hal ini, menurut [10] perlu dilakukan imputasi menggunakan algoritma *k-Nearest Neighbor* (kNN). Proses normalisasi data sebelum membangkitkan data yang bersifat *missing value* menggunakan kNN perlu dilakukan, karena besaran nilai antara data tidak sama, seperti Pada Tabel I, luas lahan memiliki data antara 0,5 sampai dengan 5,3. Sedangkan pada atribut klorofil memiliki data antara 27,78 sampai dengan 38,98. Proses normalisasi diharapkan dapat menghasilkan estimasi data yang bersifat *missing value* mendekati nilai asli dari data yang memiliki karakter data yang sama berdasarkan atribut pada data latih. Metode normalisasi yang digunakan pada penelitian ini adalah *min-max*.

Berdasarkan hasil pemilihan dan pembersihan atribut yang didapatkan dari data sekunder, varietas tebu yang digunakan dalam penelitian ini hanya menggunakan BL (Buluh Lawang), dikarenakan hanya varietas BL yang ada pada PG Kebon Agung Malang dan PG Pesantren Baru Kediri. Pada Tabel I, data no. 1 sampai dengan no. 6 data didapatkan dari PG Kebon Agung Malang, Sedangkan data no. 7 sampai dengan no. 10 data didapatkan dari PG Pesantren Baru Kediri. Tabel I adalah menunjukkan beberapa data yang digunakan dalam penelitian ini, data secara keseluruhan adalah sebanyak 201 data. Pada atribut masa tanam, data berisi waktu tanam, waktu tanam dalam sebulan dapat dilakukan dua kali (dua minggu). Contoh masa tanam, jika waktu tanam dilakukan bulan Januari

TABEL I  
DATA TAKSASI PRODUKSI GULA PADA VARIETAS BL  
(BULUH LAWANG) YANG TELAH DISELEKSI DAN  
DIBERSIHKAN

	L N o	u a	M a s a	Fa k t o r	Ju m l a h	Ti n g g i	B o b o t	Ren d e m e n	Kl o r o f i l	Ha s i l
	La h a n	Ta n a m	Le n g g / H a	Bat a n g	( K m g )	( K g )	(%)			(K u)
1	1,	1							38	11
	5								,9	62
									8	
2	0,	23			2,	0,			35	50
	5				55	5			,0	0
									8	
3	0,	17			3,	0,			31	85
	6				11	35			,0	0
									5	
4	2	17			3,	0,			35	40
					27	27			,1	00
									3	
5	2,	23			1,	0,			35	41
	3				85	5			,1	40
									3	
6	0,	1			2,	0,			27	
	6				97	66			,7	
									8	
7	4	12					10,1			89
							8			2
8	2	13					10,1			70
							8			0
9	5,	11					10,1			81
	3						8			0
10	4,	11					10,1			81
	8						8			6

minggu pertama, maka notasinya adalah 1, jika waktu tanam pada bulan Januari Minggu ketiga, maka notasinya adalah 2, jika waktu tanama adalah bulan Februari minggu pertama, maka notasinya adalah 3, jika waktu tanam pada bulan Februari minggu ketiga, maka notasinya adalah 4, begitu seterusnya. Minggu pertama pada setiap bulan selalu bernilai ganjil, Minggu ketiga pada setiap bulannya

bernilai genap.

Pada umumnya, PG selalu menggunakan 7 atribut dalam melakukan taksasi produksi gula, yang terdiri dari luas lahan, masa tanam, faktor leng, jumlah batang, tinggi, bobot, rendemen. Pada penelitian ini, dari 7 atribut tersebut, dilakukan uji korelasi berdasarkan nilai klorofil. berdasarkan uji korelasi, atribut yang memiliki hubungan dengan nilai klorofil adalah masa tanam, faktor leng, bobot dan rendemen, maka 4 atribut itu digantikan oleh atribut klorofil. Sehingga pada penelitian ini memiliki 2 model dalam melakukan taksasi, yaitu Model1 terdiri dari 7 atribut yang umum digunakan oleh PG, Model2 terdiri dari 4 atribut (luas lahan, jumlah batang, tinggi dan klorofil).

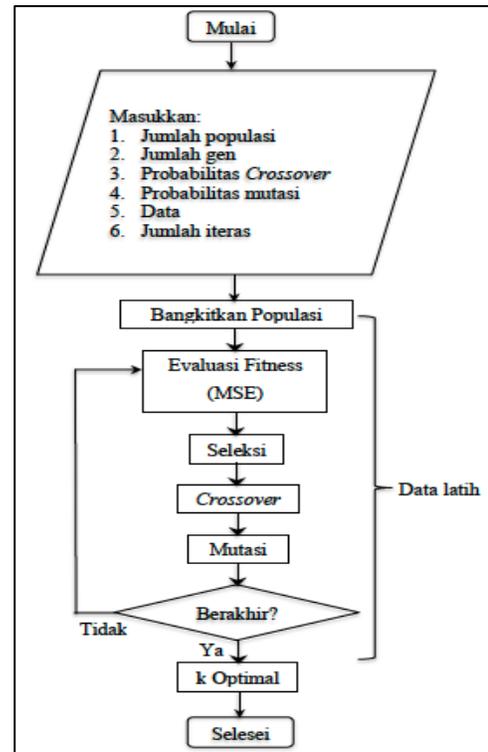
### III. GMKNN: GENETIC MODIFIED K-NEAREST NEIGHBOR

GMKNN adalah algoritma yang bertujuan untuk meningkatkan algoritma MKNN dengan cara mengoptimalkan nilai  $k$  menggunakan algoritma genetika (Gambar 1), dalam penelitian ini GMKNN (Gambar 3) digunakan untuk kasus estimasi (taksasi) produksi gula tebu berdasarkan nilai klorofil daun tebu.

Algoritma genetika memiliki masukan berupa jumlah populasi, jumlah gen, probabilitas *crossover* ( $P_c$ ), probabilitas mutasi ( $P_m$ ), data dan jumlah iterasi. Jumlah populasi merepresentasikan kemungkinan solusi  $k$  ( $k$  adalah salah satu parameter pada MKNN,  $k$  pada algoritma genetika direpresentasikan sebagai kromosom) yang optimal pada MKNN, jika yang diharapkan adalah kemungkinan solusi  $k$  yang optimal adalah sebanyak 10 buah, maka jumlah populasi bernilai 10. Jumlah gen merepresentasikan batasan nilai  $k$  pada MKNN, batasan nilai  $k$  adalah kurang dari data latih, skema pengkodean yang digunakan dalam penelitian ini adalah *binary encoding*. misalkan, data latih yang digunakan adalah sebanyak 75, maka representasi *binary encoding* yang kurang dari 75 adalah sebesar 63 atau 6 bit (Gambar 2) dalam bilangan biner (Tabel II).  $P_c$  pada penelitian ini digunakan sebesar 0,8 [11] sedangkan  $P_m$  yang digunakan adalah 0,1 [12]. Data latih yang digunakan adalah untuk Percobaan1 sebesar 75, untuk Percobaan2 sebesar 130 dan Percobaan3 sebesar 150. Sedangkan data uji untuk Percobaan1 sebesar 20, Percobaan2 sebesar 30 dan

untuk Percobaan3 sebesar 51. Iterasi yang digunakan beberapa variasi sebesar 5, 10, 25, 50, 100, 500 dan 1000. Proses untuk membangkitkan populasi dilakukan secara acak. Evaluasi *fitness* didasarkan pada MSE (*Mean Square Error*), di mana nilai *fitness* berbanding terbalik dengan nilai MSE. Seleksi induk baik untuk pindah silang maupun mutasi menggunakan metode *roulette wheel*. Keluaran dari algoritma genetika ini adalah berupa *k* (kromosom) yang optimal untuk MKNN.

Algoritma MKNN adalah algoritma kNN yang telah dilakukan modifikasi dengan cara menambahkan nilai validitas berdasarkan kemiripan dari target data latih, persamaan dari validitas pada kasus estimasi ditunjukkan pada persamaan (1), validitas didasarkan pada similaritas (persamaan (2)), similaritas didasarkan pada jarak terdekat (persamaan (3)) antar data latih.



Gambar 1. Flowchart algoritma genetika untuk mendapatkan *k* optimal dari MKNN

$$Validitas(i) = \frac{1}{k} \sum_{i=1}^k S(target(x), target(N_i(x))), \quad (1)$$

$$S(d_t) = \begin{cases} 1 & d_t < t \\ 0 & d_t \geq t \end{cases} \quad (2)$$

$$d(a, y) = \sqrt{\sum_i (a_i - y_i)^2} \quad (3)$$

Pada persamaan (1) *i* adalah banyaknya data, *k* adalah *k* optimal yang didapatkan dari algoritma genetika, *x* adalah target pada titik *x*, *N* adalah jumlah dari tetangga data latih. Pada persamaan (2) *S* adalah similaritas target antara data latih, *d<sub>t</sub>* adalah jarak Euclidean antara data latih, karena similaritas target ini didasarkan pada jarak Euclidean antar data latih, *t* adalah nilai ambang batas yang didasarkan pada *d<sub>t</sub>*. Pada persamaan (3) *d* adalah jarak Euclidean, persamaan *d* juga dapat digunakan untuk persamaan *d<sub>t</sub>*, hanya saja *d* digunakan untuk jarak antar data uji dengan data latih. *a* adalah target aktual, *y* adalah target hasil estimasi. Pada Persamaan (4)  $\alpha$  adalah *smoothing*, yang bertujuan untuk menghindari penyebut agar tidak dibagi dengan nol.  $\alpha$  pada

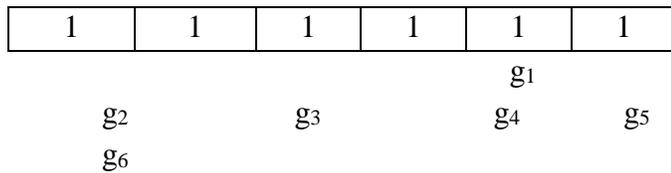
$$W(i) = validitas(i) \times \frac{1}{d + \alpha}, \quad (4)$$

$$\hat{y}_{estimasi} = \frac{\sum_i W_i y_i}{\sum_i W_i} \quad (5)$$

penelitian ini menggunakan 0,5 [6]. Pada persamaan (5)  $\hat{y}_{estimasi}$  adalah target hasil estimasi, *w* adalah bobot, *i* adalah sebanyak *k*, *y* adalah target aktual.

GMKNN memiliki masukan yang terdiri dari data latih, target data latih, data uji, target data uji dan *k* yang optimal. Proses yang pertama kali adalah mendapatkan nilai *d<sub>t</sub>*, selanjutnya digunakan untuk mendapatkan nilai similaritas berdasarkan nilai *t* yang optimal dari *d<sub>t</sub>* yang sudah dilakukan pemeringkatan berdasarkan nilai terkecil sejumlah *k*, jika sudah didapatkan nilai similaritas, maka dapat digunakan untuk perhitungan validitas. Validitas ini akan digunakan untuk mendapatkan bobot. Sebelum menghitung bobot perlu dilakukan perhitungan untuk mendapatkan jarak antara data uji dengan data latih. Semakin kecil jarak semakin baik, begitu juga

sebaliknya, semakin besar bobot, semakin baik. Jika bobot sudah dilakukan, maka bisa dilakukan perhitungan estimasi berdasarkan persamaan (5) [13].



Gambar 2. Skema pengkodean kromosom pada GA yang merepresentasikan  $k$  pada kNN (Suyanto, 2005)

TABEL II  
KONVERSI BINER KE DESIMAL

	$2^8$	$2^7$	$2^6$	$2^5$	$2^4$	$2^3$	2	2	2
Konversi Biner ke Desimal	25	12	64	3	1	8	4	2	1
	6	8		2	6				

Biner	1	1	1	1	1	1	1	1	1
Desimal	50	25	12	6	3	1	7	3	1
1	3	7	7	3	1	5			

#### IV. HASIL UJI COBA

Uji coba dilakukan dengan beberapa kali percobaan dengan berbagai macam variasi, baik variasi  $k$  secara manual pada MKNN,  $t$  (0,001; 0,01; 0,1; dan 1), data latih (75, 130 dan 150), maupun data uji (20, 30, dan 51). Berdasarkan hasil uji coba (Tabel III)  $t$  optimal memiliki variasi yang berbeda-beda, hal ini dikarenakan setiap data memiliki karakteristik yang berbeda, sama halnya dengan  $k$  optimal yang berbeda-beda. Pada Tabel III, data yang tercetak merah berarti MKNN unggul, sedangkan jika berwarna ungu maka GMKNN unggul. Berdasarkan Tabel III GMKNN terjebak pada optimum lokal, dikarenakan  $k$  maksimal hanya 63 pada saat penentuan gen terdiri dari 6 bit, sedangkan  $k$  yang optimal adalah 74. Namun, GMKNN juga memiliki keunggulan (Tabel IV), di mana GMKNN memiliki nilai MSE sebesar 3737, sedangkan MKNN memiliki MSE sebesar 3847. GMKNN selain memiliki kelemahan pada konvergensi dini (terjebak pada optimum lokal), juga memiliki komputasi yang besar, karena memiliki

perhitungan yang kompleks. Berdasarkan Tabel III dan Tabel IV, Model2 tidak begitu unggul dibandingkan menggunakan Model1.

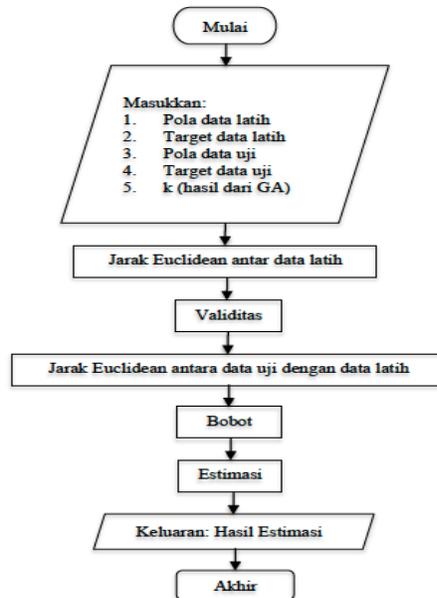
Salah solusi yang dapat digunakan adalah pemilihan skema pengkodean yang cocok dengan kasus algoritma MKNN, karena ketika pada penelitian ini menggunakan skema pengkodean biner, akan terjebak pada optimum lokal pada kasus optimasi nilai  $k$  dari MKNN. Untuk itu pada penelitian kedepan, optimasi nilai  $k$  pada MKNN dapat dilakukan perbaikan baik dengan menggunakan perlakuan tertentu, atau dengan cara pemilihan algoritma optimasi yang lebih baik dari pada algoritma genetika. Berdasarkan uji t dapat diambil kesimpulan bahwa Model 2 memiliki perbedaan yang signifikan.

#### V. KESIMPULAN

Taksasi produksi gula berdasarkan nilai Klorofil daun dapat dilakukan dengan cara mengganti atribut masa tanam, juring, bobot dan rendemen berdasarkan uji korelasi, sedangkan untuk mendapatkan nilai klorofil daun dapat menggunakan SPAD meter. Hasil dari taksasi produksi Gula menggunakan nilai klorofil daun tebu juga kurang begitu bagus dibandingkan dengan model yang tidak menggunakan nilai Klorofil daun Tebu, walaupun memiliki pengaruh yang signifikan. Metode GMKNN dapat digunakan untuk melakukan taksasi produksi Gula lebih baik, dari pada MKNN seperti pada percobaan dengan data sebanyak 95 data pada Model2, dengan nilai MSE terbaik sebesar 3737. Metode GMKNN memiliki performa yang kurang baik pada kasus komputasi, karena membutuhkan waktu yang lebih lama, namun memiliki kemampuan dalam menentukan nilai  $k$  secara optimal dan otomatis. Selain permasalahan komputasi, kelemahan GMKNN adalah terjebak pada optimum lokal karena bawaan dari algoritma genetika, dan GMKNN juga memiliki kelemahan dalam penentuan nilai ambang antar data latih ( $d_i$ ) pada kasus estimasi.

Penelitian kedepan pekerjaan yang bisa dilakukan adalah menggunakan algoritma *thresholding* dalam menentukan similaritas antar data latih ( $d_i$ ) pada kasus estimasi atau prediksi dari tugas *datamining*, sehingga memiliki ambang batas yang optimal, serta memiliki nilai MSE yang lebih baik. Algoritma GMKNN kedepannya dapat dilakukan perbaikan dalam menangani permasalahan

konvergensi dini, agar tidak terjebak pada optimum lokal.



Gambar 3. Flowchart algoritma MKNN

TABEL III  
PERBANDINGAN MKNN DENGAN GMKNN UNTUK MELAKUKAN ESTIMASI PADA MODEL1

Data	Data Latih	Data Uji	Variasi k	Jumlah Atribut	Metode	<i>k</i> Optimal	<i>t</i> Optimal	MSE	Waktu (detik)
95	75	20	20	7	MKNN	74	1	<b>3908</b>	<b>0.0628</b>
					GMKNN	63	1	3937	0.2672
130	100	201	21		MKNN	5	0.1	8024	<b>0.0259</b>
					GMKNN	5	0.1	8024	0.2894
201	150	51	22		MKNN	2	0.001	2036463	<b>0.0223</b>
					GMKNN	2	0.001	2036463	0.3368

TABEL IV  
PERBANDINGAN MKNN DENGAN GMKNN UNTUK MELAKUKAN ESTIMASI PADA MODEL2

Data	Data Latih	Data Uji	Variasi k	Jumlah Atribut	Metode	<i>k</i> Optimal	<i>t</i> Optimal	MSE	Waktu (detik)
95	75	20	20	7	MKNN	55	0.1	3847	0.0536
					GMKNN	61	1	<b>3737</b>	0.2701
130	100	201	21		MKNN	1	0.001	12739	0.0204
					GMKNN	1	0.001	12739	0.2834
201	150	51	22		MKNN	4	0.001	2053700	0.025
					GMKNN	4	0.001	2053730	0.308

## DAFTAR PUSTAKA

- [17] K. Pertanian. (Desember 2009). Rancangan Rencana Strategis Kementerian Pertanian Tahun 2010-2014. Kementerian Pertanian, Jakarta, Indonesia [Online]. Tersedia: [www.deptan.go.id/renbangtan/rancangan%20renstra%20depta%202010-2014%20lengkap.pdf](http://www.deptan.go.id/renbangtan/rancangan%20renstra%20depta%202010-2014%20lengkap.pdf).
- [18] P3GI. (2008). Konsep Peningkatan Rendemen Untuk Mendukung Program Akselerasi Industri Gula Nasional. P3GI, Indonesia [Online]. Tersedia: [www.sugarresearch.org/wp-content/uploads/2008/12/konsep-peningkatan-rendemen.pdf](http://www.sugarresearch.org/wp-content/uploads/2008/12/konsep-peningkatan-rendemen.pdf).
- [19] Soemarno. M S. (2010). Bagaimana Meningkatkan Rendemen tebu. Malang, Indonesia [Online]. Tersedia: <http://marno.lecture.ub.ac.id/files/2012/01/BAGAIMANA-MENINGKATKAN-Rendemen-tebu.doc>.
- [20] H. Reese, M. Nilsson, P. Sandstrom, et al. (2002). Applications using Estimates of Forest Parameters Derived From Satellite and Forest Inventory Data. *Computers and Electronics in Agriculture*. [Online]. volume 37, hal. 1-20. Tersedia: [http://pub.epsilon.slu.se/3332/1/Reese\\_et\\_al\\_080626.pdf](http://pub.epsilon.slu.se/3332/1/Reese_et_al_080626.pdf).
- [21] J. H. Jung, J. Heo, S. Yoo, et al. "Estimation of Aboveground Carbon using kNN Algorithm: A Case Study in Danyang, Korea", dalam *Proceedings of the Asian Conference on Remote Sensing*, 2010. [Online]. Tersedia: <http://a-a-rs.org/aars/proceeding/ACRS2010/Papers/Oral%20Presentation/TS29-5.pdf>.
- [22] K. Parvin, H. Alizadeh, B. M. Bidgoli. "MKNN: Modified K-Nearest Neighbor". *Proceedings of the World Congress on Engineering and Computer Science*, San Francisco, USA, 2010.
- [23] C. S. Oedegard. (2010). Brief Introduction to the k-Nearest Neighbour Classifier. [Online]. Tersedia: <http://ukult.net/ai/introkn.pdf>.
- [24] N. Bhatia dan Vandana. (2010). Survey of Nearest Neighbor Techniques. *International Journal of Computer Science and Information Security*. [Online]. volume 8(2), hal. 302-305. Tersedia: <http://arxiv.org/pdf/1007.0085.pdf>.
- [25] N. Suguna dan K. Thanushkodi. (2010). An Improved k-Nearest Neighbor Classification Using Genetic Algorithm. *International Journal of Computer Science Issues*. [Online]. volume 7(4), hal. 18-21. Tersedia: <http://ijcsi.org/papers/IJCSI-Vol-7-Issue-4-No-2.pdf>.
- [26] A. Izzah, dan N. Hayatin. (2013). Imputasi Missing Data Menggunakan Metode k-Nearest Neighbor Dengan Optimasi Algoritma Genetika. *Melek IT*. [Online]. volume 1, hal. 1-7. Tersedia: [www.researchgate.net/publication/257603905\\_IMPUTASI\\_MISSING\\_DATA\\_MENGGUNAKAN\\_METODE\\_K-NEAREST\\_NEIGHBOUR\\_DENGAN\\_OPTIMASI\\_ALGORITMA\\_GENETIKA](http://www.researchgate.net/publication/257603905_IMPUTASI_MISSING_DATA_MENGGUNAKAN_METODE_K-NEAREST_NEIGHBOUR_DENGAN_OPTIMASI_ALGORITMA_GENETIKA)
- [27] Coelho, L. D. S. Bora, T. C. dan Klein, C. E. (2014). A genetic programming approach based on Levy flight applied to nonlinear identification of a poppet valve. *Applied Mathematical Modelling*, Vol. 38, No. 5, hal. 179-1736.
- [28] A. Izzah, R. K. Dewi, S. Mutfrofin. "Hybrid Artificial Bee Colony: Penyelesaian Baru Pohon Rentang Berbatas Derajat". Prosiding dari Seminar Nasional Teknologi Informasi dan Multimedia, Yogyakarta, Indonesia, 2013, hal. 18-39 s/d 18-44.
- [29] D. T. Larose. (2005). *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley and Sons, Hoboken, New Jersey.

