

# Optimasi Parameter K pada Algoritma KNN untuk Deteksi Penyakit Kanker Payudara

Sattriedi Wahyu Binabar, Ivandari

STMIK Widya Pratama Pekalongan

E-mail: binabars@gmail.com, ivandarialkaromi@gmail.com

## RINGKASAN

Kanker merupakan salah satu penyakit yang banyak menyebabkan kematian di seluruh dunia. Dari data International Agency for Research of Cancer (IARC) pada tahun 2012 tercatat lebih dari 8,2 juta jiwa meninggal dunia akibat penyakit kanker. Sedangkan keseluruhan kasus yang tercatat pada tahun tersebut mencapai 14.067.894 jiwa. Dalam 5 tahun terakhir penderita kanker payudara merupakan yang terbanyak yaitu 19,2% dari keseluruhan kasus. Selama ini pencatatan perihal pasien penyakit kanker sudah banyak dilakukan. Pencatatan tersebut nantinya dapat digunakan untuk menganalisa serta mendeteksi pasien lain yang memiliki ciri yang sama. Data mining merupakan ilmu yang menggunakan data lampau untuk menghasilkan informasi dan pengetahuan baru. Salah satu fungsi utama data mining adalah klasifikasi. Beberapa teknik klasifikasi digunakan untuk menghasilkan sebuah model atau tingkat akurasi. Salah satu algoritma klasifikasi terbaik adalah KNN. Dari hasil penelitian menunjukkan bahwa KNN dengan nilai k 13 memiliki tingkat akurasi terbaik yaitu 97,28% dengan nilai error sebesar 1,5% dan nilai mikro sebesar 97,28%.

**Kata Kunci :** Breast cancer, KNN, Optimasi parameter

## 1. PENDAHULUAN

### 1.1 Latar Belakang Masalah

Penyakit kanker merupakan salah satu penyakit yang banyak mengakibatkan kematian di seluruh dunia. Lebih dari 8,2 juta jiwa meninggal dunia akibat penyakit kanker (Iarc. 2012). Berdasarkan data yang didapatkan dari GLOBOCAN, International Agency for Research of Cancer (IARC) pada tahun 2012 setidaknya ada 14.067.894 penderita kanker baru dan menyebabkan kematian atas 8.201.575 jiwa. Kanker payudara merupakan jenis kanker yang terbanyak diderita di seluruh dunia dalam 5 tahun terakhir.

Gambar 1 merupakan grafik presentase penderita penyakit kanker dalam 5 tahun terakhir (Iarc. 2012). Pada grafik tersebut Breast cancer (kanker payudara) merupakan jenis kanker dengan presentase terbesar yaitu sebanyak 19,2% dari keseluruhan penderita kanker. Data secara lebih rinci dapat dilihat pada tabel 1.

Banyak penderita penyakit kanker yang mengetahui diagnosa setelah mengalami beberapa komplikasi. Ketika penyakit kanker terdeteksi ternyata penyakit tersebut sudah mengalami perkembangan di dalam tubuh. Secara medis penyakit dengan taraf stadium yang lebih tinggi akan lebih sulit ditangani

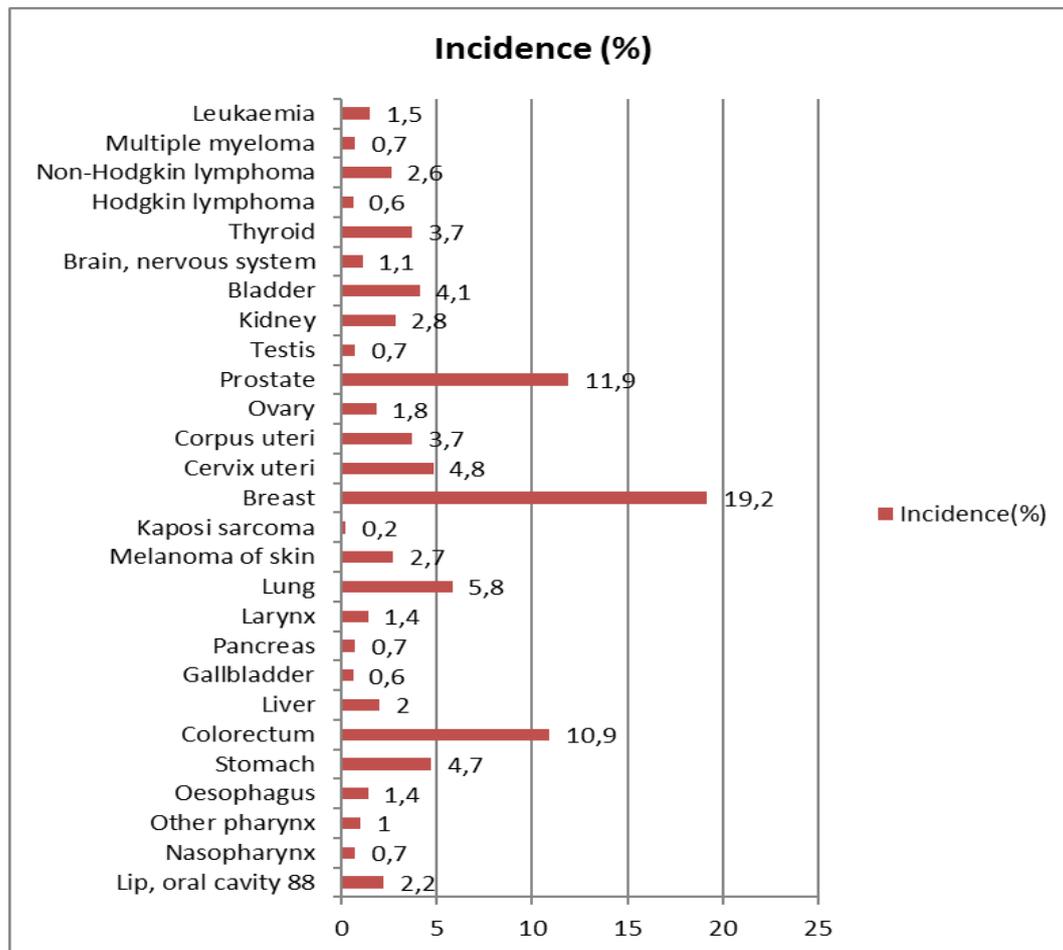
dibandingkan dengan penyakit yang terdeteksi lebih dini.

Data mining merupakan satu bidang ilmu yang memanfaatkan data yang sebelumnya kurang terpakai untuk mendapatkan suatu informasi atau pengetahuan baru. Teknik data mining dapat digunakan untuk proses klasifikasi dengan memanfaatkan data lampau. Tipe data sangat mempengaruhi performa dan akurasi suatu algoritma (Amancio et al. 2013). Algoritma terbaik untuk sebuah tipe data belum tentu baik untuk tipe data yang lain (Patel, Vala, and Pandya 2014). Bahkan dimungkinkan suatu algoritma yang baik akan menjadi sangat buruk untuk tipe data yang lain (Ragab et al. 2014) (Ashari, Paryudi, and Tjoa 2013). Atribut dari sebuah dataset sangat mempengaruhi performa suatu algoritma (Han and Kamber 2006) (Maimoon 2010) (Alpaydin 2010). Salah satu algoritma klasifikasi data mining yang terbaik adalah KNN (Wu et al. 2007).

KNN merupakan algoritma yang banyak dipakai untuk proses klasifikasi. Algoritma ini menghitung jarak antara keseluruhan data *training* dengan data *testing*. Kemudian data terdekat atau data dengan kemiripan yang paling tinggi nantinya akan diambil untuk klasifikasi. Salah satu kelemahan KNN adalah ketika parameter K yang digunakan adalah satu maka hasil klasifikasi akan terlihat kaku. Sedangkan

bila parameter K yang digunakan terlalu besar maka hasilnya akan menjadi samar. Dalam penelitian ini optimalisasi parameter k dilakukan

dengan menggunakan nilai K yaitu ganjil dari k=1 sampai dengan k=49.



Gambar 1. Presentase Penderita Jenis Kanker (IARC, 2012)

Tabel 1. Data Penderita Kanker (Iarc. 2012)

Cancer	Incidence			Mortality			5-year prevalence		
	Number	(%)	ASR (W)	Number	(%)	ASR (W)	Number	(%)	Prop
Lip, oral cavity	300373	2.1	4	145353	1.8	1.9	702149	2.2	13.5
Nasopharynx	86691	0.6	1.2	50831	0.6	0.7	228698	0.7	4.4
Other pharynx	142387	1	1.9	96105	1.2	1.3	309991	1	6
Oesophagus	455784	3.2	5.9	400169	4.9	5	464063	1.4	8.9
Stomach	951594	6.8	12.1	723073	8.8	8.9	1538127	4.7	29.6
Colorectum	1360602	9.7	17.2	693933	8.5	8.4	3543582	11	68.2
Liver	782451	5.6	10.1	745533	9.1	9.5	633170	2	12.2
Gallbladder	178101	1.3	2.2	142823	1.7	1.7	205646	0.6	4
Pancreas	337872	2.4	4.2	330391	4	4.1	211544	0.7	4.1
Larynx	156877	1.1	2.1	83376	1	1.1	441675	1.4	8.5
Lung	1824701	13	23.1	1589925	19	19.7	1893078	5.8	36.5

Melanoma of skin	232130	1.7	3	55488	0.7	0.7	869754	2.7	16.8
Kaposi sarcoma	44247	0.3	0.6	26974	0.3	0.3	80395	0.2	1.5
<b>Breast</b>	<b>1671149</b>	<b>12</b>	<b>43.1</b>	<b>521907</b>	<b>6.4</b>	<b>12.9</b>	<b>6232108</b>	<b>19</b>	<b>240</b>
Cervix uteri	527624	3.8	14	265672	3.2	6.8	1547161	4.8	59.6
Corpus uteri	319605	2.3	8.3	76160	0.9	1.8	1216504	3.7	46.8
Ovary	238719	1.7	6.1	151917	1.9	3.8	586624	1.8	22.6
Prostate	1094916	7.8	30.7	307481	3.7	7.8	3857500	12	149
Testis	55266	0.4	1.5	10351	0.1	0.3	214666	0.7	8.3
Kidney	337860	2.4	4.4	143406	1.7	1.8	906746	2.8	17.5
Bladder	429793	3.1	5.3	165084	2	1.9	1319749	4.1	25.4
Brain, nervous system	256213	1.8	3.4	189382	2.3	2.5	342914	1.1	6.6
Thyroid	298102	2.1	4	39771	0.5	0.5	1206075	3.7	23.2
Hodgkin lymphoma	65950	0.5	0.9	25469	0.3	0.3	188538	0.6	3.6
Non-Hodgkin lymphoma	385741	2.7	5.1	199670	2.4	2.5	832843	2.6	16
Multiple myeloma	114251	0.8	1.5	80019	1	1	229468	0.7	4.4
Leukaemia	351965	2.5	4.7	265471	3.2	3.4	500934	1.5	9.6
All cancers excl. non-melanoma skin cancer	14067894	100	182	8201575	100	102	32455179	100	625

## 2. METODE PENELITIAN

### 2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data *public* yaitu data *breast cancer wisconsin* yang didapatkan dari *UCI repository*. Dataset tersebut memiliki 699 *record*, 1 atribut id, 9 atribut informasi dan 1 atribut kelas. Dalam dataset tersebut terindikasi ada 458 terjangkit kanker jinak dan 241 lainnya terjangkit kanker ganas. Atribut informasi yang digunakan antara lain: *Clump Thickness*, *Uniformity of Cell Size*, *Uniformity of Cell Shape*, *Marginal Adhesion*, *Single Epithelial Cell Size*, *Bare Nuclei*, *Bland Chromatin*, *Normal Nucleoli*, serta *Mitoses*. Tabel 2 merupakan metadata dari dataset yang didapatkan.

### 2.2 Desain Eksperimen dan Pengujian

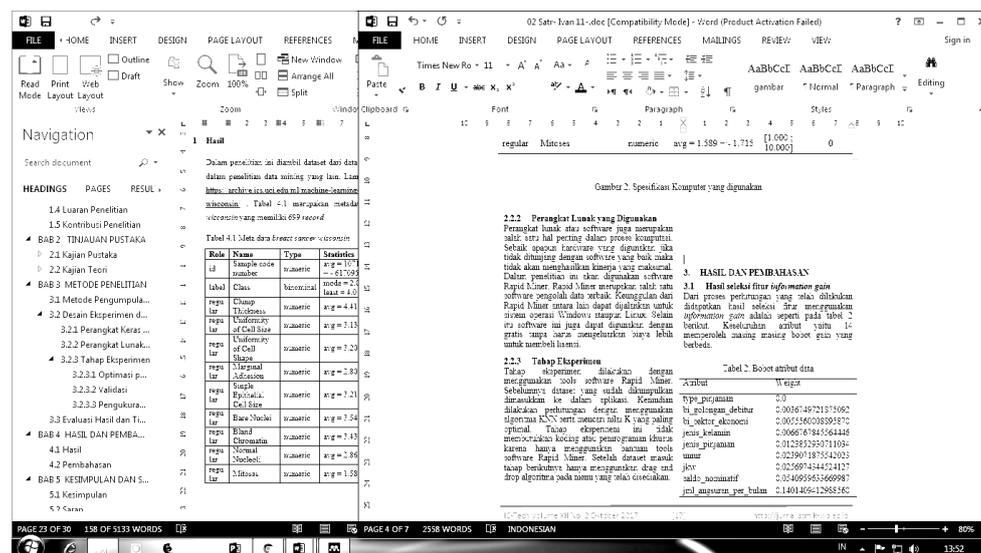
#### 2.2.1 Perangkat Keras yang Digunakan

Perangkat keras atau hardware merupakan salah satu penentu berjalannya sebuah proses komputasi. Ketika sebuah proses berjalan diatas perangkat keras dengan spesifikasi yang minim maka proses tersebut tidak akan berjalan dengan baik. Sebaliknya, dengan menggunakan perangkat keras yang memenuhi spesifikasi minimal atau lebih akan menjamin kelancaran sebuah proses komputasi. Penelitian ini menggunakan perangkat keras yang baik agar mempercepat proses komputasi. Processor yang digunakan adalah processor intel core i7 generasi ke 7 dengan *clock speed* mencapai 3,4 GHz. Memory yang digunakan adalah jenis DDR3L dengan kapasitas 16 GB. Kapasitas penyimpanan SSD 512GB. Adapun gambar 2 merupakan spesifikasi komputer yang digunakan.

Tabel 2. Metadata *breast cancer wisconsin*

Role	Name	Type	Statistics	Range	Missings
id	Sample code number	numeric	avg = 1071704.099 +/- 617095.730	[61634 ; 13454352]	0
label	Class	binominal	mode = 2.0 (458), least = 4.0 (241)	2.0 (458), 4.0 (241)	0

regular	Clump Thickness	numeric	avg = 4.418 +/- 2.816	[1.000 ; 10.000]	0
regular	Uniformity of Cell Size	numeric	avg = 3.134 +/- 3.051	[1.000 ; 10.000]	0
regular	Uniformity of Cell Shape	numeric	avg = 3.207 +/- 2.972	[1.000 ; 10.000]	0
regular	Marginal Adhesion	numeric	avg = 2.807 +/- 2.855	[1.000 ; 10.000]	0
regular	Single Epithelial Cell Size	numeric	avg = 3.216 +/- 2.214	[1.000 ; 10.000]	0
regular	Bare Nuclei	numeric	avg = 3.545 +/- 3.644	[1.000 ; 10.000]	16
regular	Bland Chromatin	numeric	avg = 3.438 +/- 2.438	[1.000 ; 10.000]	0
regular	Normal Nucleoli	numeric	avg = 2.867 +/- 3.054	[1.000 ; 10.000]	0
regular	Mitoses	numeric	avg = 1.589 +/- 1.715	[1.000 ; 10.000]	0



Gambar 2. Spesifikasi Komputer yang digunakan

## 2.2.2 Perangkat Lunak yang Digunakan

Perangkat lunak atau software juga merupakan salah satu hal penting dalam proses komputasi. Sebaik apapun hardware yang digunakan jika tidak ditunjang dengan software yang baik maka tidak akan menghasilkan kinerja yang maksimal. Dalam penelitian ini akan digunakan software Rapid Miner. Rapid Miner merupakan salah satu software pengolah data terbaik. Keunggulan dari Rapid Miner antara lain dapat dijalankan untuk sistem operasi Windows ataupun Linux. Selain itu software ini juga dapat digunakan dengan gratis tanpa harus mengeluarkan biaya lebih untuk membeli lisensi.

## 2.2.3 Tahap Eksperimen

Tahap eksperimen dilakukan dengan menggunakan tools software Rapid Miner. Sebelumnya dataset yang sudah dikumpulkan dimasukkan ke dalam aplikasi. Kemudian dilakukan perhitungan dengan menggunakan algoritma KNN serta mencari nilai K yang paling optimal. Tahap eksperimen ini tidak membutuhkan coding atau pemrograman khusus karena hanya menggunakan bantuan tools software Rapid Miner. Setelah dataset masuk tahap berikutnya hanya menggunakan drag and drop algoritma pada menu yang telah disediakan.

### 2.2.3.1 Optimasi parameter k pada KNN

Dalam tahap ini akan dilakukan beberapa kali percobaan klasifikasi deteksi penyakit kanker payudara dengan menggunakan algoritma KNN. Dataset yang sudah ada sebelumnya akan dibagi menjadi dua bagian. Satu bagian menjadi data *training* dan bagian yang lain akan menjadi data *testing*. Pemilihan nilai parameter k akan ditentukan dengan membandingkan akurasi dari beberapa percobaan yang dilakukan. Nilai parameter k paling optimal didapatkan dengan mengambil hasil akurasi yang paling baik.

### 2.2.3.2 Validasi

Dalam proses validasi penelitian ini akan digunakan *10 folds cross validation*. Proses ini banyak digunakan oleh peneliti karena sudah terbukti baik dan menghasilkan tingkat akurasi yang stabil. Secara tori *10 folds cross validation* sudah dijelaskan secara lebih terinci dalam bab sebelumnya (Witten, Frank, and Hall 2011).

### 2.2.3.3 Pengukuran akurasi algoritma

Pengukuran dari suatu algoritma merupakan suatu pembuktian yang banyak dilakukan peneliti (Amancio et al. 2013). Dalam prosesnya banyak cara dapat digunakan untuk mengetahui performa suatu algoritma. Salah satu yang paling banyak digunakan adalah dengan menghitung akurasi algoritma. Perhitungan akurasi adalah presentase dari jumlah data *testing* dengan klasifikasi yang sesuai dengan aslinya dibagi keseluruhan data. Cara lain adalah dengan menghitung *Error rate*. *Error rate* adalah kebalikan dari tingkat akurasi, yaitu presentase kesalahan klasifikasi dibagi dengan keseluruhan dataset.

Dalam penelitian ini digunakan *confusion matrix* sebagai alat ukur performa algoritma klasifikasi. *Confusion matrix* merupakan salah satu alat untuk menghitung nilai akurasi suatu algoritma. Dalam matrix ini dapat dilihat keseluruhan data *testing* yang sesuai dengan klasifikasi sebenarnya serta yang tidak sesuai. Perhitungan secara lebih terperinci dapat dilihat dalam bab sebelumnya.

## 2.3 Evaluasi Hasil dan Tingkat Akurasi

Selanjutnya setelah dilakukan perhitungan tingkat akurasi dari masing masing pemilihan parameter k maka berikutnya dilakukan pencatatan tingkat akurasi. Nilai k paling optimal dalam klasifikasi akan digunakan untuk mengembangkan aplikasi deteksi penyakit kanker payudara.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Hasil akurasi algoritma KNN

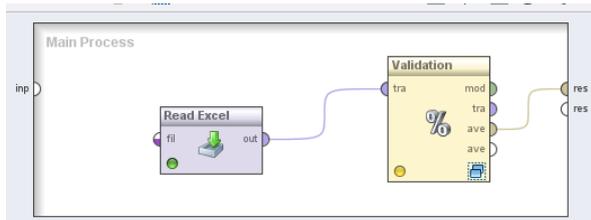
Dari beberapa percobaan komputasi dengan menggunakan rapid miner didapatkan hasil sebagaimana tabel 3 berikut. Dalam tabel 3 ini terdapat beberapa penggunaan variabel K dalam perhitungan algoritma KNN.

Tabel 3. Tingkat akurasi algoritma KNN

<b>K</b>	<b>Akurasi</b>	<b>+/-</b>	<b>mikro</b>
1	95.13 %	2.67	95.14 %
3	96.28 %	2.25	96.28 %
5	96.99 %	1.35	97 %
7	96.99 %	1.5	97 %
9	96.42 %	1.83	96.42 %
11	96.71 %	2.64	96.71 %
<b>13</b>	<b>97.28 %</b>	<b>1.5</b>	<b>97.28 %</b>
15	96.28 %	1.95	96.28 %
17	96.85 %	1.54	96.85 %
19	96.57 %	1.94	96.57 %
21	96.42 %	2.95	96.42 %
23	96.42 %	1.95	96.42 %
25	96.28 %	2.04	96.28 %
27	96.28 %	1.46	96.28 %
29	96.14 %	2.12	96.14 %
31	96.13 %	2.94	96.14 %
33	96.42 %	2.15	96.42 %
35	96.14 %	1.7	96.14 %
37	96.14 %	1.57	96.14 %
39	96.14 %	1.81	96.14 %
41	95.99 %	2.85	95.99 %
43	95.99 %	2.46	95.99 %
45	95.99 %	1.54	95.99 %
47	95.99 %	1.67	95.99 %
49	96.99 %	1.79	96.99 %

### 3.2 Pembahasan

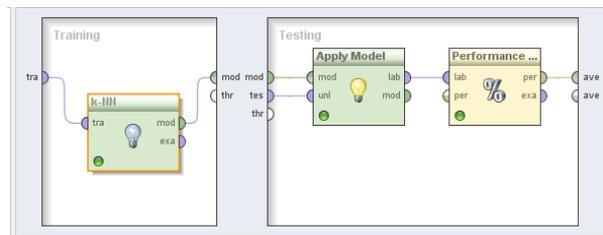
Dalam penelitian ini digunakan *10 folds cross validation* untuk proses validasi. Yang artinya dataset akan dibagi menjadi 10 bagian dengan 1 bagian untuk data testing dan 9 bagian lainnya digunakan sebagai data training. Proses ini diulang sampai 10 kali sehingga keseluruhan data mendapat porsi satu kali untuk menjadi data testing. Gambar 3 merupakan proses validasi yang dilakukan dalam aplikasi rapid miner



Gambar 3 Proses validasi dalam aplikasi

Dalam gambar 3 terdapat *read excell* yang fungsinya sebagai pembaca dari dataset yang ada. Digunakan *read excell* karena dataset yang ada

dalam bentuk excell. Sedangkan output dari *read excell* tersebut masuk ke dalam proses validasi. Di dalam proses validasi terdapat proses yang dapat dilihat pada gambar 4 berikut.



Gambar 4 Proses evaluasi *confusion matrix*

Dari gambar 4 terlihat bahwa algoritma K-NN digunakan untuk *training* dan *apply model* serta *performance* yaitu menggunakan *confusion matrix* digunakan dalam *testing*. Hasil dari *confusion matrix* ini nantinya akan ditampilkan dan diambil nilai rerata dari 10 kali percobaan dalam validasi. Gambar 5 dan gambar 6 merupakan hasil keluaran dari program dengan menggunakan nilai k sebesar 1 dan k sebesar 13.

accuracy: 95.13% +/- 2.67% (mikro: 95.14%)			
	true 2.0	true 4.0	class precision
pred. 2.0	445	21	95.49%
pred. 4.0	13	220	94.42%
class recall	97.16%	91.29%	

Gambar 5 *Confusion matrix* untuk k=1

accuracy: 97.28% +/- 1.50% (mikro: 97.28%)			
	true 2.0	true 4.0	class precision
pred. 2.0	448	9	98.03%
pred. 4.0	10	232	95.87%
class recall	97.82%	96.27%	

Gambar 6 *Confusion matrix* untuk k=13

## 4. KESIMPULAN DAN SARAN

### 4.1 Kesimpulan

Dalam proses penelitian terdapat 16 *missing value* dalam dataset yaitu dalam atribut *bare nuclei*. Hal ini dapat menurunkan tingkat akurasi pada umumnya. Hasil dari penelitian menunjukkan bahwa algoritma KNN baik digunakan untuk klasifikasi dataset *breast cancer wisconsin*. Tingkat akurasi dari keseluruhan percobaan nilai k mulai dari k=1 sampai dengan k=49 menunjukkan performa yang baik. Performa ini terbukti dengan tingkat akurasi dari

keseluruhan yang lebih dari 95% dengan nilai mikro minimal sama dengan tingkat akurasi. Tingkat akurasi tertinggi didapatkan ketika klasifikasi menggunakan nilai k sebesar 13. Tingkat akurasi yang muncul mencapai 97,28% dengan nilai error +/- sebesar 1,5 % dan nilai mikro 97,28%.

### 4.2 Saran

Dalam penelitian ini masih menggunakan penetapan nilai k secara manual. Diharapkan dalam penelitian berikutnya dapat digunakan algoritma lain guna menentukan nilai variabel k

yang paling baik dalam klasifikasi. Pengujian dalam penelitian ini juga hanya menggunakan aplikasi rapid miner, pengujian lain semisal menggunakan pemrograman web dapat digunakan untuk pengembangan aplikasi pada tahap selanjutnya.

## 5. DAFTAR PUSTAKA

- Alkaromi, M Adib. 2014. "Information Gain Untuk Pemilihan Fitur Pada Klasifikasi Heregistrasi Calon Mahasiswa Dengan Menggunakan K-NN."
- Alpaydin, Ethem. 2010. *Introduction to Machine Learning Second Edition*. London: The MIT Press.
- Amancio, D. R., C. H. Comin, D. Casanova, G. Travieso, O. M. Bruno, F. a. Rodrigues, and L. Da F. Costa. 2013. "A Systematic Comparison of Supervised Classifiers," October. <http://arxiv.org/abs/1311.0202v1>.
- Ashari, Ahmad, Iman Paryudi, and A Min Tjoa. 2013. "Performance Comparison between Naïve Bayes , Decision Tree and K-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool" 4 (11): 33–39.
- Azhagusundari, B, and Antony Selvadoss Thanamani. 2013. "Feature Selection Based on Information Gain," no. 2: 18–21.
- Azwany, Faraby. 2010. "Sistem Pendukung Keputusan Pemberian Kredit Usaha Rakyat Pada Bank Syariah Mandiri Cabang Medan Menggunakan Metode AHP." *Program Studi Ilmu Komputer Departemen Ilmu Komputer Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Sumatera Utara Medan*.
- Dwi Cahyani, Bangun. 2013. "Sistem Pendukung Keputusan Persetujuan Permohonan Kredit Pinjaman Pada Bank Muamalat Indonesia Cabang Yogyakarta." *Sekolah Tinggi Manajemen Informatika Dan Komputer AMIKOM Yogyakarta*.
- Ginanjari Mabur, Angga, and Riani Lubis. 2012. "Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit." *Jurnal Komputer Dan Informatika ( KOMPUTA )* 1.
- Han, Jiawei, and Micheline Kamber. 2006. "Data Mining: Concepts and Techniques Second Edition" 40 (6). Elsevier: 9823. doi:10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C.
- Iarc., International Agency for Research on Cancer. World Health Organization. 2012. "GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012." *Globocan*. doi:10.1002/ijc.27711.
- Indrayanti, Devi Sugianti, and M Adib Al Karomi. 2017. "Optimasi Parameter K Pada Algoritma K-Nearest Neighbour Untuk Klasifikasi Penyakit Diabetes Mellitus." *Prosiding SNATIF Buku 3 (2017)*: 823–29.
- Ivandari, Tria Titiani Chasanah, Satriedi Wahyu Binabar, and M Adib Al Karomi. 2017. "Data Attribute Selection with Information Gain to Improve Credit Approval Classification Performance Using K-Nearest Neighbor Algorithm." *IJIBEC I*: 15–24.
- Karegowda, Asha Gowda, A S Manjunath, and M. A Jayaram. 2010. "Comparative Study of Attribute Selection Using Gain Ratio and Correlation Based Feature Selection." *International Journal of Information Technology and Knowledge Management 2 (2)*: 271–77.
- Karomi, M Adib Al. 2015. "Optimasi Parameter K Pada Algoritma KNN Untuk Klasifikasi Heregistrasi Mahasiswa Program Studi Teknik Informatika STMIK Widya Pratama Jl . Patriot 25 Pekalongan Email : Adib.comp@gmail.com." *IC-TECH X (285)*: 5.
- Koprinska, Irena. 2010. "Feature Selection for Brain-Computer Interfaces," 100–111.
- Maimoon, Oded, and Lior Rokach. 2010. *Data Mining and Knowledge Discovery Handbook*. Vol. 40. Springer. doi:10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C.
- Maulana, Much. Rifqi, and M Adib Al Karomi. 2016. "Sistem Pendukung Keputusan Persetujuan Kredit Menggunakan Algoritma C4.5." *Jurnal IC-Tech Vol. XI No (1)*: 29–38. [IC-Tech Volume XII No. 2 Oktober 2017](http://jurnal.stmik-</a></p></div><div data-bbox=)

wp.ac.id/gdl.php?mod=browse&op=read&i  
d=ictech--muchrifqim-80.

Menggunakan Metode Bayes.” *Sekolah  
Tinggi Manajemen Informatika Dan  
Komputer AMIKOM Yogyakarta.*

- Melissa, Ira, and Raymond S Oetama. 2013. “Analisis Data Pembayaran Kredit Nasabah Bank Menggunakan Metode Data Mining” IV (1): 18–27.
- Patel, Kanu, Jay Vala, and Jaymit Pandya. 2014. “Comparison of Various Classification Algorithms on Iris Datasets Using WEKA” 1 (1): 1–7.
- Ragab, Abdul Hamid M., Amin Y. Noaman, Abdullah S. Al-Ghamdi, and Ayman I. Madbouly. 2014. “A Comparative Analysis of Classification Algorithms for Students College Enrollment Approval Using Data Mining.” *Proceedings of the 2014 Workshop on Interaction Design in Educational Environments - IDEE '14*. New York, New York, USA: ACM Press, 106–13. doi:10.1145/2643604.2643631.
- Suhari, Yohanes, Muji Sukur, and Sri Eniyati. 2009. “SISTEM PENDUKUNG KEPUTUSAN PEMBERIAN KREDIT PADA PT . BPR ARTAMANUNGGAL ABADI MRANGGEN.” *Dinamika Informatika Fakultas Teknologi Informasi Universitas Stikubank Semarang* I (1).
- Wahyu Oktaputra, Alif, and Edi Noersasongko. 2014. “Sistem Pendukung Keputusan Kelayakan Pemberian Kredit Motor Menggunakan Metode Simple Additive Weighting Pada Perusahaan Leasing HD Finance.” *Program Studi Sistem Informasi - SI, Fakultas Ilmu Komputer Universitas Dian Nuswantoro, Semarang*, 1–9.
- Wasana, Emil. 2010. “Sistem Pendukung Keputusan Kelayakan Pengajuan Kredit Motor Menggunakan Metode Scoring System.” *Jurusan Siste Informasi, Sekolah Tinggi Manajemen Informatika & Teknik Komputer Surabaya*, 1–10.
- Wu, Xindong, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan, et al. 2007. *Top 10 Algorithms in Data Mining. Knowledge and Information Systems*. Vol. 14. doi:10.1007/s10115-007-0114-2.
- Zahid, Ahmad. 2013. “Sistem Pendukung Keputusan Persetujuan Penerimaan Pinjaman Di PD.BPR BKK Lasem Dengan