

Klasifikasi Mutu Susu Sapi menggunakan Metode *Modified K-Nearest Neighbor* (MKNN)

Reyvaldo Aditya Pradana¹, Imam Cholissodin², Nurul Hidayat³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹reyvaldo.shikamaru@gmail.com, ²imamcs@ub.ac.id, ³ntayadih@ub.ac.id

Abstrak

Susu merupakan salah satu bahan makanan yang paling kompleks, karena di dalam kandungan susu ada banyak kandungan yang dibutuhkan bagi tubuh manusia. Oleh karena itu, pentingnya pemantauan untuk menguji kualitas mutu susu sapi sangat diperlukan. Sehingga dapat menghasilkan produk olahan susu berkualitas tinggi. Perkembangan teknologi pangan semakin tahun semakin pesat terutama bagi teknologi pangan produk olahan susu sapi yang banyak digemari oleh masyarakat. Diketahui UPT Laboratorium Kesehatan Hewan Malang membuat sistem aplikasi yang dapat melakukan klasifikasi terhadap mutu susu sapi. Sistem aplikasi mutu susu sapi ini menggunakan masukan berupa komposisi kimiawi dimana komposisi ini diambil dengan alat Milkscope Julie C2. Komposisi kimiawi susu terdiri dari lemak, bahan kering tanpa lemak, kekentalan, laktosa dan protein. Ada beberapa metode yang digunakan. Penelitian ini menggunakan metode Modified K-Nearest Neighbor dan dataset yang akan digunakan sebanyak 269 data mutu susu sapi dengan 5 parameter dan 2 kelas hasil. Berdasarkan dari beberapa penelitian, metode Modified K-Nearest Neighbor dapat digunakan dalam proses pengklasifikasian dan mendapatkan tingkat akurasi yang cukup tinggi. Berdasarkan hasil pengujian, diperoleh rata – rata nilai akurasi pengujian nilai K sebesar 91.1%, kemudian rata – rata nilai akurasi pengujian pengaruh jumlah data latih sebesar 88.4%, dan rata – rata nilai akurasi pengujian kelas seimbang dan tidak seimbang sebesar 86.12%. Maka dapat disimpulkan bahwa metode Modified K-Nearest Neighbor dapat diimplementasikan dan diuji kedalam sistem klasifikasi mutu susu sapi.

Kata Kunci: Modified K-Nearest Neighbor, Klasifikasi, Data Mining, Mutu Susu Sapi

Abstract

Milk is a very complex food ingredient, because milk has many ingredients needed by the human body. It is necessary to control the quality of cow's milk in order to produce high quality dairy products. With the development of food technology for cow's milk products, UPT Laboratory of Animal Health Malang by creating an application system that can classify the quality of cow's milk. This cow's milk quality application system uses input in the form of chemical composition where this composition is taken with the Julie C2 Milkscope tool. The chemical composition of milk consists of fat, lean dry matter, viscosity, lactose and protein. There are various methods used. This study uses the Modified K-Nearest Neighbor method and the dataset used is 269 cow's milk quality data with 5 parameters and 2 yield classes. Based on several studies, the Modified K-Nearest Neighbor method can be used in the classification process and obtain a fairly high level of accuracy. Based on the test results, the average test accuracy value of K value is 91.1%, then the average value of the accuracy of testing the effect of the amount of training data is 88.4%, and the average value of balanced and unbalanced class testing accuracy is 86.12%. It can be concluded that the Modified K-Nearest Neighbor method can be implemented and tested into the cow's milk quality classification system.

Keyword: Modified K-Nearest Neighbor, Classification, Data Mining, Cow's Milk

1. PENDAHULUAN

Susu sapi merupakan salah satu bahan pangan yang rentan terhadap kontaminasi dan pembusukan bakteri akibat pertumbuhan bakteri yang terkandung dalam susu. Oleh karena itu,

untuk mendapatkan susu yang berkualitas tinggi, susu harus diolah dengan baik. Sehingga, untuk mendapatkan produk susu yang berkualitas perlu dilakukan pengendalian kualitas susu.

Pengujian kualitas susu sapi yang dilakukan

di UPT Laboratorium Kesehatan Hewan Malang adalah pengujian kimia menggunakan alat Milkscope Julie C2 yang memiliki 5 parameter yaitu lemak (*Fat*), bahan kering tanpa lemak (*Solid NonFat*), kekentalan (*Density*), protein dan laktosa, komposisi yang dijadikan parameter untuk melakukan pengklasifikasian. Tetapi pada komposisi tersebut ternyata hanya 3 komposisi yaitu protein, *Solid NonFat*, dan lemak yang digunakan untuk pertimbangan Standar Nasional Indonesia yang dikeluarkan oleh badan standarisasi Indonesia dalam menentukan mutu susu sapi. Sedangkan 2 komposisi yaitu kekentalan dan laktosa tidak digunakan dalam pertimbangan mutu susu sapi. Oleh karena itu pengklasifikasian mutu susu sapi kurang maksimal. Kemudian pihak UPT Laboratorium Kesehatan Hewan Malang ingin 2 komposisi yang tidak dipertimbangkan ikut dalam pertimbangan mutu susu sapi untuk membantu memaksimalkan pengklasifikasian mutu susu sapi.

Klasifikasi mutu susu sapi dirancang dengan mengadaptasi bidang ilmu ke bidang ilmu komputer, yaitu *data mining*. *Data mining* adalah proses menemukan hubungan atau pola yang bermakna dengan memeriksa kumpulan data besar yang disimpan dalam arsip menggunakan teknik pengenalan pola seperti metode statistik dan matematika.

Dalam penelitian ini menggunakan metode *Modified K-Nearest Neighbor* (MKNN) yang sebelumnya ada pada penelitian yang telah dilakukan oleh Parvin, MKNN merupakan metode modifikasi dari algoritma *K-Nearest Neighbor* (KNN) yang diberi tambahan beberapa proses diantaranya proses validasi data latih dan *weight voting*. Keunggulan dari metode algoritma *Modified K-Nearest Neighbor* adalah metode ini digunakan secara sederhana dan mudah dipelajari, pelatihan sangat cepat, efektif jika data pelatihan besar.

2. DASAR TEORI

2.1. Susu

Susu adalah bahan makanan yang kompleks dan dapat diperoleh dengan memeras susu dengan benar. Kandungan alami susu tidak ditambahkan atau dikurangi tanpa perawatan apa pun. Jadi susu mengandung zat-zat yang dibutuhkan tubuh manusia seperti protein, lemak, karbohidrat, vitamin dan mineral. Zat-zat ini dicerna dengan baik oleh tubuh manusia.

Susu umumnya didistribusikan di kalangan masyarakat umum dan dianggap sebagai minuman yang sehat karena mengandung nutrisi. Namun, tentu perlu memastikan bahwa tes untuk memeriksa mutu susu sudah sesuai prosedur.

2.2. Atribut Susu Sapi

Data diperoleh dari UPT Laboratorium Kesehatan Hewan, Dataset susu berupa susu. Susu adalah cairan yang diperoleh dari pemerahan yang benar dari sapi yang sehat tanpa ditambah maupun dikurangi kecuali proses pendinginan untuk penyimpanan.

Dataset pada susu memiliki 269 data, 5 atribut dan mempunyai 2 kelas, yaitu kelas 0 (Tidak Sesuai SNI) dan kelas 1 (Sesuai SNI). Data komposisi kimiawi pada susu berupa:

1. Lemak (*Fat*)

Lemak merupakan salah satu komposisi susu. Dalam satu gram lemak dapat menghasilkan 9 kalori. Lemak susu termasuk asam butirat, asam kaporat, asam miristat, asam laurat, asam oleat dan asam stearate. Menurut SNI, kandungan lemaknya minimal 3.0%.

2. Bahan Kering Tanpa Lemak (SNF / *Solid NonFat*)

Solid NonFat merupakan padatan yang dihilangkan unsur lemaknya. Jadi *Solid NonFat* tersusun dari vitamin, protein, hidrat arang dan beberapa mineral. Menurut SNI kadar *Solid NonFat* minimum 8.0%.

3. Kekentalan

Kekentalan adalah ukuran kepadatan susu yang dinyatakan banyaknya berat massa pada susu dalam satuan volume. Menurut SNI kadar kekentalan minimum 21.0%.

4. Laktosa

Laktosa merupakan karbohidrat utama dalam susu. Ini terdiri dari dua komponen laktosa, glukosa dan galaktosa. Namun, laktosa tidak semanis gula dan memiliki kelarutan 20% pada suhu antara 100 dan 130 °C. Menurut SNI, tidak ada informasi batas minimum, tetapi UPT Kesehatan Hewan Malang menetapkan batas kandungan laktosa adalah 4,6% dalam keadaan larut dalam air.

5. Protein

Protein tersusun dari enzim, kasein,

lactoglobulin, laktalbumin dan imonoglobulin. Kandungan protein terbanyak terdapat pada kasein. Pada protein memberikan 4,1 kalori dalam satu gram. Menurut SNI kadar protein minimal 2.7%.

2.3. Data Mining

Data mining adalah bidang ilmiah yang menggabungkan pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk memecahkan masalah penggalian informasi dari database besar.

Dalam data mining terdapat beberapa tahapan – tahapan berikut adalah rincian langkah - langkah yang diperlukan untuk menemukan pola dalam data:

1. Data Selection

Pemilihan data dari dataset aktif harus dilakukan sebelum langkah ekstraksi informasi. Data yang dihasilkan dipilih untuk digunakan dalam penambahan data disimpan dalam file terpisah dari database produksi.

2. Pre-Processing

Sebelum dapat melakukan data mining, perlu dilakukan preprocessing pada data tersebut. Preprocessing termasuk menghapus data duplikat, memeriksa ketidakkonsistenan data, dan memperbaiki kesalahan data.

3. Transformation Data

Transformation data merupakan sebuah teknik untuk mengubah data ke dalam format yang sesuai untuk digunakan dalam data mining. Transformasi dan pemilihan data ini juga menentukan kualitas hasil data mining, karena karakteristik tertentu dari teknologi data mining.

4. Data Mining

Data mining merupakan proses menemukan pola atau informasi yang menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma penambahan data sangat berbeda. Memilih metode atau algoritme yang tepat sangat bergantung pada tujuan dan proses Anda secara keseluruhan.

5. Evaluasi / Interpretasi

Model informasi yang dihasilkan oleh proses data mining harus ditampilkan dalam format

yang dapat dipahami oleh para pemangku kepentingan. Langkah - langkah tersebut termasuk memverifikasi bahwa pola atau informasi yang ditemukan tidak sesuai dengan fakta dan asumsi yang ada.

2.4. Klasifikasi

Klasifikasi adalah proses untuk menemukan pola atau fungsi yang menggambarkan atau membedakan suatu konsep atau kelas data untuk tujuan memperkirakan kelas fitur yang tidak diketahui

Dalam pengklasifikasian data terdapat dua proses yang dilakukan, yaitu:

1. Proses Training

Proses pelatihan menggunakan data pelatihan yang ditentukan dan label yang diketahui untuk menghasilkan model atau fungsi mentah.

2. Proses Testing

Selama pengujian, model yang dilatih terdiri dari data uji bersama dengan data lain untuk menentukan seberapa akurat model tersebut. Jika tingkat akurasi cukup, model ini dapat digunakan untuk memprediksi kelas data yang tidak diketahui.

Di dalam proses klasifikasi terdapat beberapa metode algoritma, seperti *Bayesian*, *Decision Tree*, *Neural Network*, *Fuzzy*, *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN).

2.5. K-Nearest Neighbor (KNN)

Algoritma K-Nearest Neighbor (KNN) merupakan metode yang digunakan untuk mengurutkan objek berdasarkan data latih yang memiliki jarak terdekat dengan objek tersebut. KNN adalah algoritma pembelajaran terawasi di mana hasil versi kueri hanya diklasifikasikan menurut sebagian besar jenis algoritma KNN. Jika kelas muncul lebih dari satu kali, itu adalah hasil dari klasifikasi KNN.

Prinsip yang digunakan pada algoritma *K-Nearest Neighbor* adalah bandingan data uji yang diberikan dengan data latih untuk menemukan jarak terdekat antara dua data berperingkat K dari tetangga terdekat dalam data latih. Skala dekat atau jauh dapat dihitung berdasarkan variabel. Perhitungan jarak yang umum digunakan adalah *Euclidean*. Jarak

Euclidean dihitung menggunakan Persamaan (2.1).

$$D(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.1)$$

Dimana:

$D(x, y)$ = Jarak *Euclidean* antara titik data latih x dengan data uji y

x_i = Sampel data latih

y_i = Data uji

n = Dimensi atribut

Dimana x sebagai variabel data latih dan y sebagai variabel data uji. Kemudian data kontinu harus dinormalisasi sebelum klasifikasi. Normalisasi itu sendiri dimaksudkan untuk mencegah atribut memiliki rentang 0 hingga 1 agar menghindari atribut bernilai rendah dan bernilai terlalu besar. Perhitungan *min - max* normalisasi dapat digunakan untuk mengubah nilai atribut dapat jatuh pada range tertentu. *min - max* dapat dihitung dengan Persamaan (2.2).

$$ndata = \frac{(v - min) * (nmax - nmin)}{max - min} + nmin \quad (2.2)$$

Dimana:

$ndata$ = Data hasil normalisasi

v = Data yang akan dinormalisasi

min = Nilai minimum dari data

max = Nilai maksimum dari data

$nmax$ = Skala maksimum yang kita berikan

$nmin$ = Skala minimum yang kita berikan

Tahapan yang digunakan dalam menghitung metode algoritma *K-Nearest Neighbor* (KNN) mulailah dengan menentukan parameter K (tetangga terdekat). Kemudian menghitung jarak *Euclidean* untuk setiap data pelatihan yang diberikan. Hasil perhitungan jarak *Euclidean* diurutkan dari minimum hingga maksimum dan menentukan nilai K yang digunakan.

2.6. Modified K-Nearest Neighbor (MKNN)

Algoritma *Modified K-Nearest Neighbor* (MKNN) merupakan evolusi dari algoritma *K-Nearest Neighbor* dengan berbagai langkah yaitu perhitungan nilai validitas dan perhitungan bobot.

2.6.1 Perhitungan Validitas Data

Pada algoritma MKNN, semua data dalam data pelatihan harus divalidasi terlebih dahulu. Validitas setiap data tergantung pada masing-masing data yang berdekatan. Proses ini harus dilakukan pada semua data dalam data latih. Hasil perhitungan validitas untuk setiap data digunakan sebagai informasi tambahan tentang data tersebut. Persamaan (2.3) digunakan untuk menghitung nilai validitas.

$$Validity(x) = \frac{1}{H} \sum_{i=0}^n S(label(x), label(N_i(x))) \quad (2.3)$$

Dimana:

H = Jumlah K titik terdekat.

$label(x)$ = Kelas data x .

$label(N_i(x))$ = Label kelas titik ke - i terdekat data x .

Fungsi S digunakan ketika menghitung kesamaan antara titik x dan data ke - i dari tetangga terdekat. Persamaan (2.4) digunakan untuk menghitung fungsi S .

$$S(a, b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \quad (2.4)$$

Dimana:

a = Kelas a pada data latih.

b = Kelas lain selain a pada data latih.

2.6.2 Perhitungan Weight Voting

Dalam metode MKNN, bobot setiap tetangga awalnya dihitung menggunakan $1 / (de + 0.5)$. Kemudian dilakukan penimbangan setiap bobot data pelatihan berdasarkan jarak *Euclidean*. Persamaan (2.5) digunakan untuk menghitung *Weight voting*.

$$W(i) = Validity(i) * \frac{1}{de + 0.5} \quad (2.5)$$

Dimana:

$W(i)$ = Perhitungan bobot *voting*.

$Validity(i)$ = Nilai Validitas.

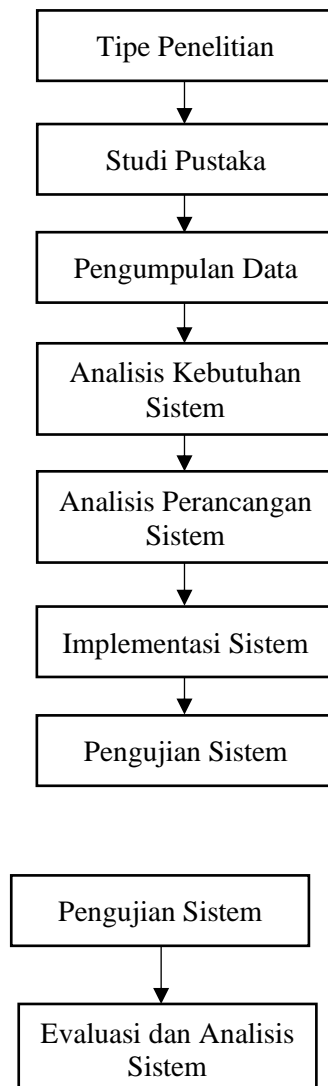
de = Jarak *Euclidean* antara data latih dengan data uji.

Teknik perhitungan bobot ini mempunyai pengaruh penting dalam data tersebut memiliki nilai validitas tertinggi dan paling dekat dengan data tersebut. Selain itu, juga mengatasi semua kekurangan data di kejauhan bobot, yang memiliki banyak masalah dibagian luar. Maka dari itu algoritma MKNN diusulkan secara

signifikan lebih kuat daripada metode KNN tradisional yang hanya didasarkan pada jarak (Parvin, 2008)

3. METODOLOGI

Metodologi menggambarkan langkah-langkah yang dilakukan peneliti untuk memecahkan masalah penelitian menentukan klasifikasi mutu susu. Tahapan-tahapan yang dilakukan dapat dilihat di Gambar 3.1.



Gambar 3.1 Diagram Alir

4. PENGUJIAN DAN ANALISIS

4.1. Pengujian Pengaruh Nilai K

Pengaruh nilai K ini diuji dengan memasukkan nilai K yang berbeda ke dalam sistem untuk mengetahui pengaruh akurasi terhadap nilai yang dihasilkan. Nilai K yang digunakan adalah dari K = 1 sampai K = 5. Data yang digunakan dalam pengujian ini yang pertama data uji sebesar 25% dan data latih

sebesar 75%, yang kedua data uji sebesar 50% dan data latih sebesar 50%, yang ketiga data uji sebesar 75% dan data latih sebesar 25% dapat dilihat di Tabel 4.1 Hasil Pengujian Pengaruh Nilai K.

Tabel 4.1 Hasil Pengujian Pengaruh Nilai K

Nilai K	Uji Coba	Akurasi (%)		
		Data Uji 25%	Data Uji 50%	Data Uji 75%
1	1	75.0	74.6	91.5
	2	77.9	85.8	90.0
	3	86.7	80.5	94.5
	4	86.7	87.3	92.5
	5	85.2	83.5	90.5
2	1	80.8	79.8	86.5
	2	85.2	85.8	88.0
	3	88.2	84.3	91.5
	4	88.2	86.5	90.5
	5	85.2	84.3	90.0
3	1	88.2	93.2	77.1
	2	92.6	92.5	80.0
	3	91.1	91.7	91.0
	4	92.6	94.0	89.5
	5	91.1	91.0	87.5
4	1	73.5	86.5	92.5
	2	76.4	80.5	93.5
	3	88.2	85.8	93.5
	4	89.7	92.5	93.5
	5	88.2	85.8	93.5
5	1	72.0	88.9	89.0
	2	73.5	77.6	89.5
	3	89.7	88.8	90.0
	4	89.7	90.2	89.5
	5	88.2	85.0	89.5

Setelah mendapatkan hasil yang benar dari lima kali pengujian dengan 25% data uji, 75% data latih, 50% data uji 50% data latih, 75% data uji 25% data latih dengan nilai K yang berbeda rata – rata akurasi pengujian pengaruh nilai K dapat dilihat di Tabel 6.2.

Tabel 4.2 Rata – Rata Hasil Pengujian Pengaruh Nilai K

Nilai K	Rata – Rata Akurasi (%)		
	Data Uji 25%	Data Uji 50%	Data Uji 75%
1	77.9	84.6	87.5
2	81.1	84.4	88.2
3	89.1	86.2	91.0
4	89.4	90.1	91.1
5	87.6	85.9	90.2

Berdasarkan uji coba diatas rata – rata hasil pengujian pengaruh nilai K di Tabel 6.2, mendapatkan hasil rata – rata akurasi tertinggi sebesar 91.1% pada nilai K = 4 menggunakan data uji 75% dan hasil rata – rata akurasi terendah sebesar 77.9% pada nilai K = 1 menggunakan data uji 25%. Disimpulkan bahwa nilai K yang digunakan mempengaruhi hasil dan akurasi yang dicapai oleh sistem.

4.2. Pengujian Pengaruh Jumlah Data Latih

Untuk pengujian pengaruh jumlah data latih yang diperoleh dengan data latih hingga 40 data latih, hingga 50 data latih, dan hingga 60 data latih. Untuk mengetahui pengaruh jumlah data latih yang digunakan untuk data uji, nilai data uji adalah 100. Nilai K yang digunakan dalam pengujian ini adalah K = 4. Dapat dilihat di Tabel 4.3 Hasil Pengujian Pengaruh Jumlah Data Latih dibawah ini.

Tabel 4.3 Hasil Pengujian Pengaruh Jumlah Data Latih

Data Latih	Data Uji	Uji Coba	Akurasi (%)
40	100	1	90
		2	86
		3	85
		4	89
		5	85
50	100	1	89
		2	87
		3	87
		4	90
		5	86
60	100	1	90
		2	88
		3	88
		4	89
		5	87

Berdasarkan hasil akurasi di Tabel 4.3 Hasil Pengujian Pengaruh Jumlah Data Latih diatas, Tabel tersebut mencakup hasil rata-rata untuk akurasi, termasuk hingga 40 data pelatihan, 50 data pelatihan, dan 60 data pelatihan. Rata-rata akurasi yang diperoleh dari hasil pengujian rata-rata untuk pengaruh jumlah data latih ditunjukkan di Tabel 4.4.

Tabel 4.4 Rata – Rata Hasil Pengujian Pengaruh Jumlah Data Latih

Data Latih	Data Uji	Rata – Rata Akurasi (%)
------------	----------	-------------------------

40	100	87
50		87.8
60		88.4

Berdasarkan hasil uji coba yang ditunjukkan di Tabel 4.4, dapat menyimpulkan bahwa jumlah data pelatihan yang berbeda dapat mempengaruhi hasil akurasi sistem. Pada hasil pengujian, akurasi minimum adalah 87% pada data latih 40, nilai 87.8% saat data latih 50, dan akurasi maksimal saat data latih 60 88.4 %.

4.3 Pengujian Kelas Seimbang dan Tidak Seimbang

Pengujian kelas seimbang dan tidak seimbang ini digunakan untuk mengetahui pengaruh jumlah data yang terdistribusi antara kelas seimbang dan tidak seimbang terhadap nilai akurasi yang dihasilkan oleh sistem. Nilai K yang digunakan pada pengujian ini adalah K = 4. Hal ini dikarenakan nilai paling akurat yang diperoleh pada pengujian sebelumnya adalah K = 4. Data yang akan digunakan pada pengujian ini adalah data latih hingga 40 data latih, hingga 50 data latih dan hingga 60 data latih. Dapat dilihat di Tabel 4.5 Hasil Pengujian Kelas Seimbang dan Tidak Seimbang dibawah ini.

Tabel 4.5 Hasil Pengujian Kelas Seimbang dan Tidak Seimbang

Jenis Data	Data Latih	Data Uji	Uji Coba	Akurasi (%)
Seimbang	40	110	1	84.5
			2	85.4
			3	84.5
			4	83.6
			5	83.6
	50	110	1	85.4
			2	83.6
			3	85.4
			4	86.3
			5	84.5
	60	110	1	86.3
			2	86.3
			3	87.2
			4	85.4
			5	85.4
Tidak Seimbang	40	110	1	80.9
			2	81.8
			3	82.7
			4	80.9
			5	80.9
	50	110	1	80.0

			2	80.9
			3	80.9
			4	80.0
			5	81.8
	60	110	1	81.8
			2	81.8
			3	80.9
			4	82.7
			5	80.9

Berdasarkan hasil akurasi pada Tabel 4.5 Hasil Pengujian Kelas Seimbang dan Tidak Seimbang tabel di atas mencakup hasil akurasi rata - rata dan data pelatihan 40 data latih, 50 data latih dan 60 data latih yang seimbang dan tidak seimbang. Rata – rata akurasi yang dihasilkan dari rata – rata hasil pengujian kelas seimbang dan tidak seimbang dapat dilihat pada Tabel 4.6.

Tabel 4.6 Rata – Rata Hasil Pengujian Kelas Seimbang dan Tidak Seimbang

Jenis Data	Data Latih	Data Uji	Rata – Rata Akurasi(%)
Seimbang	40	110	84.32
	50		85.04
	60		86.12
Tidak Seimbang	40	110	81.44
	50		80.72
	60		81.62

Berdasarkan hasil pengujian yang ditunjukkan di Tabel 4.6, rata-rata akurasi tertinggi diperoleh sebesar 86.12% dengan data latih seimbang sebanyak 60 data latih dan akurasi rerata terendah sebesar 80.72% diperoleh dengan data latih tidak seimbang sebesar 50 data latih tidak seimbang.

5. PENUTUP

Berdasarkan hasil pengujian yang dilakukan pada klasifikasi mutu susu menggunakan metode Modified K-Nearest Neighbor, dapat disimpulkan bahwa:

1. Metode *Modified K-Nearest Neighbor* hal ini dapat dilakukan dalam klasifikasi mutu susu menggunakan 5 parameter dan 2 status kategori hasil.
2. Sistem klasifikasi mutu susu sapi menggunakan metode *Modified K-Nearest Neighbor* dapat menguji dan mendapatkan hasil akurat berikut:

A. Berdasarkan dari hasil pengujian pengaruh nilai K, nilai akurasi

tertinggi sebesar 91.1%. Didapatkan hasil bahwa nilai K yang paling bagus digunakan adalah K = 4, ketika K = 5 akurasi akan cenderung turun. Dikarenakan tetangga terdekat yang digunakan sebagai pertimbangan pengambilan keputusan akan semakin banyak. Namun ketika K = 5 tetangga terdekat akan mengambil hasil yang kelanya mendominasi sehingga berdampak pada nilai akurasi yang dihasilkan cenderung turun dikarenakan error rate yang terjadi ketika hasil prediksi sistem semua berstatus Tidak Sesuai SNI maka peluang terpilihnya ketika pembobotan dengan status Tidak Sesuai SNI semakin tinggi begitu pula sebaliknya.

- B. Berdasarkan hasil pengujian pengaruh jumlah data latih, akurasi tertinggi adalah 88.4%. Hasilnya menunjukkan bahwa hasil akurasi secara bertahap meningkat seiring dengan meningkatnya jumlah data latih yang digunakan. Hal ini dikarenakan semakin banyak data latih yang digunakan maka akan semakin banyak pula data yang dibandingkan pada saat proses klasifikasi.
- C. Berdasarkan hasil pengujian kelas seimbang dan tidak seimbang, nilai akurasi tertinggi adalah 86.12%. Hasilnya menunjukkan bahwa dengan meningkatnya keseimbangan data pelatihan, akurasi data pelatihan seimbang meningkat. Sementara data pelatihan tidak seimbang, akurasi yang dihasilkan meningkat dan akurasi yang dihasilkan menurun. Hal ini disebabkan tidak proporsionalnya data pelatihan kelas yang mendominasi proses klasifikasi, karena hasil yang diperoleh tidak tepat dan cenderung dipilih oleh kelas terbanyak yang lebih tinggi.

Berdasarkan hasil penelitian tersebut, maka dimungkinkan untuk membuat rekomendasi yang penulis buat untuk penelitian selanjutnya Klasifikasi Mutu Susu Sapi Menggunakan

Metode *Modified K-Nearest Neighbor* dengan menambahkan wawancara dengan pakar terkait case mutu susu sapi agar informasi atau data yang dikumpulkan lebih lengkap, akurat, seimbang dan objektif dan diharapkan adanya kombinasi dari metode *Modified K-Nearest Neighbor* dengan metode lainnya sehingga nilai akurasi yang dihasilkan dapat lebih optimal dan error yang dihasilkan oleh sistem dapat dikurangi.

6. DAFTAR PUSTAKA

- Fayyad, U., 1996. *Discovering in Knowledge Discovery and Data Mining*. s.l., MIT Press.
- Ganidar, F. R., 2015. Pengklasifikasian Mutu Susu Sapi Menggunakan Metode Learning Vector Quantization (LVQ). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.
- Hadiwiyoto, S., 1994. *Teori dan Prosedur Pengujian Mutu Susu dan Hasil Olahannya*. Yogyakarta: Liberty.
- Kusnawi, 2007. *Pengantar Solusi Data Mining. Seminar Nasional Teknologi*. Yogyakarta, STMIK AMIKOM Yogyakarta.
- Larose, D. T., 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. s.l.:John Willey & Sons, Inc.
- Parvin, H., 2008. *MKNN: Modified K-Nearest Neighbor. Proceedings of the World Congress on Engineering and Computer Science*. San Francisco, USA.
- Rasepta, K. M., 2016. Klasifikasi Status Gizi Balita Menggunakan Metode Modified K-Nearest Neighbor. <http://repository.uin-suska.ac.id/id/eprint/3012>.
- Ressang AA, & N. A., 1998. Ilmu Kesehatan Susu (Milk Hygiene). Bogor: Institut Pertanian Bogor.
- Sari, P., 2017. Klasifikasi Kualitas Susu Sapi Menggunakan Metode Support Vector Machine (SVM). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.
- Sholeh, A., 2013. Pengembangan Sistem Pengenalan Wajah 2D dengan Implementasi Algoritma Eigenface dan Manhattan Distance. *repository.upi.edu*.
- Simanjuntak, T. H., 2017. Implementasi Modified K-Nearest Neighbor Dengan Otomatisasi Nilai K Pada Pengklasifikasian Penyakit Tanaman Kedelai. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.
- Suryadi, 2017. Perbandingan Metode Naive Bayes dan K-Nearest Neighbor Untuk Klasifikasi Mutu Susu Sapi. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.
- Turban, E., 2005. *Decision Support Systems and Intelligent Systems Edisi Bahasa Indonesia Jilid 1*. Yogyakarta: Andi.
- Wafiyah, F., 2017. Implementasi Algoritma Modified K-Nearest Neighbor (MKNN) Untuk Klasifikasi Penyakit Demam. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.