



TESIS - BM185407

**PENGARUH SUHU TERHADAP KLASIFIKASI
KEMURNIAN DAGING SAPI BERBASIS ELECTRONIC
NOSE**

SINARRING AZI LAGA
09211750054001

Dosen Pembimbing:
Prof. Drs. Ec. Ir. RIYANARTO SARNO, M.Sc, PhD

**Departemen Manajemen Teknologi
Fakultas Bisnis Dan Manajemen Teknologi
Institut Teknologi Sepuluh Nopember
2019**

LEMBAR PENGESAHAN TESIS

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar

Magister Manajemen Teknologi (M.MT)

di

Institut Teknologi Sepuluh Nopember

Oleh:

Sinarring Azi Laga

NRP: 09211750054001

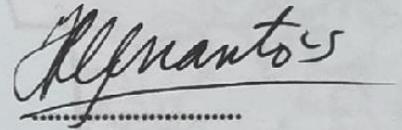
Tanggal Ujian: 15 Juli 2019

Periode Wisuda: September 2019

Disetujui oleh:

Pembimbing:

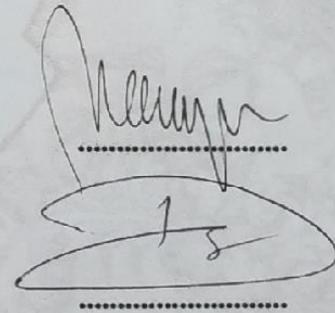
1. Prof. Drs. Ec. Ir. Riyanarto Sarno, MSc, PhD
NIP: 195908031986011001



.....

Penguji:

1. Dr. Techn. Ir. R. V. Hari Ginardi, M.Sc.
NIP: 196505181992031003



.....

2. Daniel O. Siahaan, S.Kom. M.Sc, PD.Eng
NIP: 197411232006041001

Kepala Departemen Manajemen Teknologi
Fakultas Bisnis dan Manajemen Teknologi



Prof. Ir. J. Nyoman Pujawan, M.Eng, Ph.D, CSCP
NIP: 196912311994121076

PENGARUH SUHU TERHADAP KLASIFIKASI KEMURNIAN DAGING SAPI BERBASIS *ELECTRONIC NOSE*

Nama : Sinarring Azi Laga
NRP : 09211750054001
Pembimbing : Prof. Drs.Ec. Ir. Riyanarto Sarno, M.Sc, Ph.D

ABSTRAK

Tingginya tingkat permintaan serta tingginya harga bahan baku pangan seperti daging banyak dimanfaatkan oleh penjual daging pada pasar tradisional untuk melakukan kecurangan demi meraup keuntungan yang lebih. Kecurangan tersebut berupa praktik pencampuran daging sapi dengan daging babi atau yang lebih dikenal dengan daging oplosan. Hal ini menjadi isu keamanan pangan yang marak terjadi beberapa tahun belakangan di Indonesia. Keamanan dan penjaminan pangan terdiri dari beberapa aspek mencakup kesehatan, higienitas dan labelitas halal. Isu-isu keamanan pangan tersebut dapat diantisipasi melalui bau nya.

Dengan menggunakan *electronic nose* yang telah dirakit menggunakan sensor elektrokimia dan sensor udara baik temperatur, tekanan, maupun kelembapan udara dapat mendeteksi kemurnian daging sapi atau daging sapi campuran. Sensor tersebut dapat mendeteksi gas sesuai selektivitasnya dengan menghasilkan arus kecil yang diakibatkan dari reaksi kimia antara oksigen di dalam sensor dengan gas.

Pada penelitian ini menerapkan metode klasifikasi k-NN, SVM, Naïve Bayes, dan Random Forest pada 5 variasi daging dengan perbandingan 0%, 10%, 50%, 90%, dan 100% dengan 3 suhu yang berbeda untuk mengetahui suhu yang paling optimal untuk klasifikasi kemurnian daging sapi. Adapun suhu yang digunakan pada penelitian ini yaitu -22°C , suhu ruang, dan 55°C .

Hasil penelitian menunjukkan adanya pengaruh suhu terhadap peningkatan akurasi, yaitu pada suhu -22°C . Karena semakin rendah suhu maka semakin stabil pula nilai yang didapatkan oleh *electronic nose*. Pada suhu -22°C , metode yang menghasilkan akurasi tertinggi adalah metode *Random Forest*.

Kata kunci: *Electronic nose*, Klasifikasi, k-NN, SVM, Naïve Bayes, Random Forest

[Halaman ini sengaja dikosongkan]

TEMPERATURE EFFECT ON BEEF PURITY CLASSIFICATION USING ELECTRONIC NOSE

Name : Sinarring Azi Laga
NRP : 09211750054001
Supervisor : Prof. Drs.Ec. Ir. Riyanarto Sarno, M.Sc, Ph.D

ABSTRACT

The higher level of demand and price of groceries, especially beef make many sellers at traditional market cheating to making a profit. The fraud is mixing beef with pork. In this time become a food security issue that has been rampant in recent years in Indonesia. Food guaranty and food security consist of several aspect including health, hygiene, and halal label. These food safety issues can be distinguished by their smell.

By using electronic nose that has been assembled using electrochemical sensors and air sensors for temperature, pressure, and humidity can detect the purity of beef or beef mixture. That sensors can detect the selectivity gas by producing a small current which result from a chemical reaction between oxygen in sensors with gas.

In this study, applying k-NN, SVM, Naïve Bayes, and Random Forest as methods for classification with 5 variations of meat with ratio 0%, 10%, 50%, 90%, 100% and 3 different temperatures to find out the most optimal temperature for the classification of purity of beef. The temperature used in this study is : -22°C, Room temperature, and 55°C.

The results showed the influence of temperature on increasing accuracy, ie at -22°C. Because the lower the temperature the more stable the value obtained by electronic nose is. At -22°C temperature the best method to calculate accuracy is random forest.

Key words: Electronic nose, Classification, k-NN, SVM, Naïve Bayes, Random Forest

[Halaman ini sengaja dikosongkan]

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Tesis yang berjudul **“PENGARUH SUHU TERHADAP KLASIFIKASI KEMURNIAN DAGING SAPI BERBASIS *ELECTRONIC NOSE*”**. Tesis ini diajukan untuk memenuhi prasyarat untuk menyelesaikan studi magister di Program Studi Magister Manajemen Teknologi, Konsentrasi Manajemen Teknologi Informasi, Institut Teknologi Sepuluh Nopember Surabaya.

Dalam penyelesaian Tesis ini, penulis telah mendapatkan banyak dukungan moral maupun material dari banyak pihak. Atas bantuan yang telah diberikan penulis ingin menyampaikan penghargaan dan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Ir. I Nyoman Pujawan., M.Eng., Ph.D., CSCP selaku Kepala Departemen Magister Manajemen Teknologi.
2. Bapak Dr. Tech, Ir. R. V. Hari Ginardi, M.Sc. selaku Kepala Program Studi Departemen Magister Manajemen Teknologi dan Dosen Wali yang telah meluangkan waktu, tenaga dan pikiran dalam memberikan bimbingan, pengarahan, dan ilmu pengetahuan.
3. Prof. Drs.Ec. Ir. Riyanarto Sarno, M.Sc, Ph.D selaku Pembimbing Tesis yang telah meluangkan waktu, tenaga dan pikiran dalam memberikan bimbingan, masukan, pengarahan, dan ilmu pengetahuan.
4. Seluruh dosen pengajar yang telah memberikan pengajaran dan ilmu yang begitu banyak. Serta seluruh karyawan MMT-ITS yang telah banyak membantu dalam berbagai hal selama masa perkuliahan. Terima kasih atas ilmu yang telah diajarkan kepada penulis.
5. Ayah Umar dan Ibu Choireem yang selalu memberikan dukungan baik melalui doa ataupun material untuk kesuksesan dan kelancaran penelitian ini.
6. Andhika Dwi Aristyawan, S.Farm., Apt yang selalu memberikan obat semangat ketika sedang membutuhkan .

7. Teman-teman MTI angkatan 2018 yang selalu memotivasi, mengingatkan, memberi masukan, dan selalu memberi suntikan semangat kepada penulis dalam penyusunan Tesis ini.
8. Semua pihak yang tidak dapat disebutkan satu persatu, yang telah banyak memberikan berbagai macam bantuan dalam penyusunan Tesis ini.

Akhir kata, penulis berharap Tesis ini dapat memberikan manfaat kepada pembaca mengenai pengaruh suhu terhadap klasifikasi kemurnian daging sapi. Penulis menyadari bahwa tesis ini masih jauh dari kesempurnaan dan memiliki banyak kekurangan. Oleh karena itu, dengan kerendahan hati penulis mengharapkan masukan dan saran yang membangun dari pembaca untuk perbaikan ke depan.

Surabaya, Juni 2019

Sinarring Azi Laga

DAFTAR ISI

LEMBAR PENGESAHAN	ii
ABSTRAK	iii
ABSTRACT	v
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
1 BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	3
1.5 Kontribusi Penelitian	3
1.6 Batasan Masalah	3
1.7 Sistematika Penulisan	4
2 BAB II KAJIAN PUSTAKA	5
2.1 Daging	5
2.2 Suhu	6
2.3 <i>Electronic Nose</i>	6
2.4 Sensor Gas Elektrokimia	6
2.5 Klasifikasi	8
2.6 K-Nearest Neighbor (k-NN)	10
2.7 Support Vector Machine (SVM)	11
2.7.1 Fungsi Kernel	14

2.8	Naïve Bayes	16
4.8.1.	Bayesian Classifier	16
2.8.2	Naïve Bayes untuk Klasifikasi	17
2.9	Random Forest	17
2.10	Kelebihan dan Kekurangan Masing-Masing Metode	19
2.11	K-Fold Cross Validation	21
2.12	Confussion Matrix	22
2.13	Kurva ROC	24
2.14	WEKA	24
3	BAB III METODOLOGI PENELITIAN.....	27
3.1	Diagram Alur Metode Penelitian	27
3.2	Perumusan Masalah	28
3.3	Pengumpulan Data	28
3.4	<i>Preprocessing</i> Data	29
3.5	Ekstraksi Fitur	30
3.5.1	Kurtosis	30
3.5.2	Skewness	31
3.5.3	Mean.....	31
3.6	Klasifikasi Data.....	32
3.6.1	K Nearest Neighbor (k-NN).....	32
3.6.2	<i>Support Vector Machine</i> (SVM)	32
3.6.3	<i>Naïve Bayes</i>	33
3.6.4	<i>Random Forest</i>	33
3.7	Penarikan Kesimpulan	33
3.8	Jadwal Penelitian	34
4	BAB IV HASIL PENELITIAN DAN PEMBAHASAN	35

4.1	Pengujian 5 variasi komposisi daging terhadap suhu ruang.....	35
4.2	Skenario Pengujian.....	36
4.3	Pengujian Skenario 1 (Metode k-NN).....	37
4.3.1	Pengujian S0B100.....	37
4.3.2	Pengujian S10B90.....	39
4.3.3	Pengujian S50B50.....	40
4.3.4	Pengujian S90B10.....	41
4.3.5	Pengujian S100B0.....	43
4.4	Pengujian Skenario 2 (<i>Metode Support Vector Machine</i>).....	44
4.4.1	Pengujian S0B100.....	45
4.4.1	Pengujian S10B90.....	46
4.4.2	Pengujian S50B50.....	48
4.4.1	Pengujian S90B10.....	49
4.4.1	Pengujian S100B0.....	51
4.5	Pengujian Skenario 3 (Metode Naïve Bayes).....	52
4.5.1	Pengujian S0B100.....	53
4.5.2	Pengujian S10B90.....	54
4.5.3	Pengujian S50B50.....	55
4.5.4	Pengujian S90B10.....	57
4.5.5	Pengujian S100B0.....	58
4.6	Pengujian Skenario 4 (Metode <i>Random Forest</i>).....	60
4.6.1	Pengujian S0B100.....	60
4.6.2	Pengujian S10B90.....	62
4.6.3	Pengujian S50B50.....	63
4.6.4	Pengujian S90B10.....	65
4.6.5	Pengujian S100B0.....	66

4.7	Rekapitulasi Nilai ROC Area dari masing-masing variasi daging	68
4.8	Nilai ROC berdasarkan suhu dan metode.....	72
5	BAB V KESIMPULAN DAN SARAN.....	77
5.1	Kesimpulan	77
5.2	Saran	78
	DAFTAR PUSTAKA.....	79
	LAMPIRAN	83
	BIODATA PENULIS.....	835

DAFTAR GAMBAR

Gambar 2.1 Skema <i>Electronic Nose</i>	7
Gambar 2.2 Skema dasar proses klasifikasi	8
Gambar 2.3 Ilustrasi Klasifikasi Dua Kelas	9
Gambar 2.4 Ilustrasi Klasifikasi Multi Kelas	10
Gambar 2.5 Penentuan hyperplane terbaik	12
Gambar 2.6 Metode Pembagian Data dengan <i>K-Fold Cross-Validation</i>	22
Gambar 2.7 Tampilan awal GUI WEKA	25
Gambar 3.1 Diagram Alir Penelitian	27
Gambar 3.2 <i>output</i> serial monitor yang dihasilkan <i>electronic nose</i>	29
Gambar 4.1 Perbandingan Akurasi Berdasarkan Variasi Daging	36
Gambar 4.2 Pengaruh suhu terhadap akurasi S0B100	38
Gambar 4.3 Pengaruh Suhu Terhadap Akurasi S10B90	39
Gambar 4.4 Pengaruh Suhu Terhadap Akurasi S50B50	41
Gambar 4.5 Pengaruh Suhu Terhadap Akurasi S90B10	42
Gambar 4.6 Pengaruh Suhu Terhadap Akurasi S100B0	44
Gambar 4.7 Pengaruh Suhu Terhadap Akurasi S0B100	46
Gambar 4.8 Pengaruh Suhu Terhadap Akurasi S10B90	47
Gambar 4.9 Pengaruh Suhu Terhadap Akurasi S50B50	49
Gambar 4.10 Pengaruh Suhu Terhadap Akurasi S90B10	50
Gambar 4.11 Pengaruh Suhu Terhadap Akurasi S100B0	52
Gambar 4.12 Pengaruh Suhu Terhadap Akurasi S0B100	53
Gambar 4.13 Pengaruh Suhu Terhadap Akurasi S10B90	55
Gambar 4.14 Pengaruh Suhu Terhadap Akurasi S50B50	56
Gambar 4.15 Pengaruh Suhu Terhadap Akurasi S90B10	58
Gambar 4.16 Pengaruh Suhu Terhadap Akurasi S100B0	59
Gambar 4.17 Pengaruh Suhu Terhadap Akurasi S0B100	61
Gambar 4.18 Pengaruh Suhu Terhadap Akurasi S10B90	63
Gambar 4.19 Pengaruh Suhu Terhadap Akurasi S50B50	64
Gambar 4.20 Pengaruh Suhu Terhadap Akurasi S90B10	66
Gambar 4.21 Pengaruh Suhu Terhadap Akurasi S100B0	67

Gambar 4.22 Grafik Perbandingan Metode terhadap Suhu.....	75
Gambar 4.23 Sinyal Gas yang dihasilkan pada masing-masing variasi suhu	76

DAFTAR TABEL

Table 2.1 Kelebihan dan Kekurangan Metode.....	20
Table 2.2 Confussion Matrix untuk 2 Kelas	22
Table 3.1 Distribusi normal pada masing-masing sensor	30
Table 3.2 Jadwal Pelaksanaan Penelitian.....	34
Table 4.1 <i>Confussion Matrix</i> variasi komposisi daging.....	35
Table 4.2 Detail Akurasi dari masing-masing variasi daging.....	35
Table 4.3 Skenario Pengujian	37
Table 4.4 Pengaruh suhu terhadap <i>true positive</i> pada S0B100.....	37
Table 4.5 Pengaruh suhu terhadap <i>true positive</i> pada S10B90.....	39
Table 4.6 Pengaruh suhu terhadap <i>true positive</i> pada S50B50.....	40
Table 4.7 Pengaruh suhu terhadap <i>true positive</i> pada S90B10.....	42
Table 4.8 Pengaruh suhu terhadap <i>true positive</i> pada S100B0.....	43
Table 4.9 Pengaruh suhu terhadap <i>true positive</i> pada S0B100.....	45
Table 4.10 Pengaruh suhu terhadap <i>true positive</i> pada S10B90.....	47
Table 4.11 Pengaruh suhu terhadap <i>true positive</i> pada S50B50.....	48
Table 4.12 Pengaruh suhu terhadap <i>true positive</i> pada S90B10.....	50
Table 4.13 Pengaruh suhu terhadap <i>true positive</i> pada S100B0.....	51
Table 4.14 Pengaruh suhu terhadap <i>true positive</i> pada S0B100.....	53
Table 4.15 Pengaruh suhu terhadap <i>true positive</i> pada S10B90.....	54
Table 4.16 Pengaruh suhu terhadap <i>true positive</i> pada S50B50.....	56
Table 4.17 Pengaruh suhu terhadap <i>true positive</i> pada S90B10.....	57
Table 4.18 Pengaruh suhu terhadap <i>true positive</i> pada S100B0.....	59
Table 4.19 Pengaruh suhu terhadap <i>true positive</i> pada S100B0.....	60
Table 4.20 Pengaruh suhu terhadap <i>true positive</i> pada S10B90.....	62
Table 4.21 Pengaruh suhu terhadap <i>true positive</i> pada S50B50.....	64
Table 4.22 Pengaruh suhu terhadap <i>true positive</i> pada S90B10.....	65
Table 4.23 Pengaruh suhu terhadap <i>true positive</i> pada S100B0.....	67
Table 4.24 Rekapitulasi nilai ROC Area	68
Table 4.25 Rekapitulasi nilai ROC Area <i>sort by</i> nilai ROC tertinggi.....	70
Table 4.26 Nilai ROC berdasarkan suhu -22°C	72

Table 4.27 Nilai ROC Berdasarkan Suhu Ruang	73
Table 4.28 Nilai ROC Berdasarkan Suhu 55°C.....	74
Table 4.29 Perbandingan metode terhadap suhu	75

BAB I

PENDAHULUAN

Pada bab ini akan dijelaskan tentang latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika penulisan laporan penelitian dengan judul “Pengaruh Suhu Terhadap Klasifikasi Daging Sapi Berbasis *Electronic Nose*”.

1.1 Latar Belakang

Tingginya tingkat permintaan serta tingginya harga bahan baku pangan seperti daging banyak dimanfaatkan oleh penjual daging pada pasar tradisional untuk melakukan kecurangan demi meraup keuntungan yang lebih. Kecurangan tersebut berupa praktik pencampuran daging sapi dengan daging babi atau yang lebih dikenal dengan daging oplosan. Hal ini menjadi isu keamanan pangan yang marak terjadi beberapa tahun belakangan di Indonesia. Keamanan dan penjaminan pangan terdiri dari beberapa aspek mencakup kesehatan, higienitas dan lebelitas halal (Erwanto, Abidin, Sugiyono, & Rohman, 2014). Pencampuran daging sapi dengan daging babi sering ditemukan dalam keadaan segar. Hal ini terjadi karena daging babi merupakan sumber protein hewani yang harganya murah dan mudah diperoleh di pasaran, sehingga para produsen nakal mendapatkan keuntungan yang lebih dari dari pemalsuan daging sapi. Kasus pencampuran daging sapi ini sangat menimbulkan permasalahan yang besar, karena mengingat Indonesia merupakan negara mayoritas muslim terbesar di dunia. Selain itu, beberapa golongan masyarakat juga memiliki hipersensitivitas atau intoleran terhadap daging babi (Ong, et al., 2007). Isu-isu keamanan pangan tersebut dapat dibedakan melalui bau nya.

Dengan menggunakan *electronic nose* yang telah dirakit menggunakan sensor elektrokimia dan sensor udara baik temperatur, tekanan, maupun kelembapan udara dapat mendeteksi kemurnian daging sapi atau daging sapi campuran. Sensor tersebut dapat mendeteksi gas sesuai selektivitasnya dengan

menghasilkan arus kecil yang diakibatkan dari reaksi kimia antara oksigen di dalam sensor dengan gas.

Penelitian sebelumnya telah dilakukan oleh Daiva, A. F. (2017) mengenai klasifikasi kemurnian daging sapi menggunakan *electronic nose* dengan 11 variasi sampel daging dengan perbandingan 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, dan 100%. Namun pada penelitian ini belum menggunakan variable suhu untuk menghasilkan tingkat akurasi yang tinggi .

Pada penelitian ini diterapkan analisis lebih lanjut menggunakan algoritma *machine learning* dengan 3 perbedaan suhu pada masing-masing 5 variasi sampel data daging untuk menghasilkan hasil klasifikasi yang paling optimal. Algoritma *machine learning* yang digunakan adalah *k-Nearest Neighbor* (k-NN), *Support Vector Machine* (SVM), *Naïve Bayes*, dan *Random Forest* .

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disebutkan maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana proses klasifikasi daging sapi murni dan daging sapi campuran menggunakan metode k-NN, SVM, Naïve Bayes, dan *Random Forest* ?
2. Bagaimana pengaruh suhu terhadap klasifikasi daging sapi murni?
3. Bagaimana pengaruh metode terhadap masing-masing suhu ?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

1. Untuk mengetahui pengaruh suhu pada klasifikasi kemurnian daging sapi dan daging sapi campuran yang paling optimal dengan menggunakan metode k-NN, SVM, Naïve Bayes, dan *Random Forest*.
2. Untuk mengetahui suhu terbaik yang digunakan pada klasifikasi kemurnian daging sapi dan daging sapi campuran.
3. Untuk mengetahui metode terbaik pada masing- masing suhu

1.4 Manfaat Penelitian

Manfaat penelitian yang ingin dicapai adalah menghasilkan tingkat akurasi yang tinggi terhadap klasifikasi kemurnian daging sapi berdasarkan 3 variasi suhu dengan 5 variasi daging yang berbeda dengan perbandingan 0%, 10%, 50%, 90%, dan 100%.

1.5 Kontribusi Penelitian

Kontribusi dari penelitian ini adalah memberikan informasi bahwa adanya pengaruh suhu terhadap klasifikasi kemurnian daging sapi berbasis *electronic nose*

1.6 Batasan Masalah

Batasan masalah yang ditentukan dalam penelitian ini adalah sebagai berikut:

1. Alat yang digunakan untuk pengumpulan data adalah *electronic nose* yang telah dirakit menggunakan sensor gas elektrokimia yang tersambung pada Arduino
2. Pada penelitian ini tidak membahas mengenai perakitan *electronic nose*
3. Jenis sampel daging sapi yang digunakan adalah jenis sapi dewasa
4. Terdapat 5 variasi sampel daging yang diuji, yaitu:
 - Sapi 0% Babi 100%,
 - Sapi 10% Babi 90%,
 - Sapi 50% Babi 50%,
 - Sapi 100% Babi 0%,
 - Sapi 90% Babi 10%,
5. Pengumpulan data dilakukan terhadap 3 variasi suhu, yaitu : -22°C, Suhu ruang, dan 55°C
6. Metode klasifikasi yang digunakan adalah *K-Nearest Neighbor*(k-NN) , *Support Vector Machine* (SVM), Naïve Bayes, dan *Random Forest* .

1.7 Sistematika Penulisan

Berikut adalah sistematika penulisan yang digunakan pada penelitian ini:

Bab I Pendahuluan

Bab ini menyajikan mengenai latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, metodologi, dan sistematika penulisan.

Bab II Tinjauan Pustaka

Bab ini menyajikan tentang kajian literatur mengenai teori dan penelitian sebelumnya yang telah dilakukan yang mendasari penelitian.

Bab III Metode Penelitian

Bab ini menyajikan metode dan langkah-langkah yang dilakukan oleh peneliti dalam melakukan penelitian ini.

Bab IV Hasil Penelitian dan Pembahasan

Bab ini menyajikan hasil penelitian sesuai dengan scenario yang ada serta pembahasan mengenai hasil penelitian tersebut.

Bab V Kesimpulan dan Saran

Bab ini menyajikan kesimpulan dari penelitian yang telah dilakukan dan saran untuk peneliti selanjutnya.

BAB II

KAJIAN PUSTAKA

Bab ini menjelaskan tentang informasi dari tempat studi kasus serta teori-teori penunjang yang digunakan sebagai acuan dalam pembuatan laporan penelitian dengan judul “Pengaruh Suhu Terhadap Klasifikasi Daging Sapi Berbasis *Electronic Nose*”.

2.1 Daging

Definisi daging berdasarkan Codex Alimentarius adalah seluruh bagian dari hewan yang aman digunakan dan sesuai untuk konsumsi manusia. Daging terdiri dari air, protein, asam amino, mineral, lemak, asam lemak, vitamin dan komponen bioaktif lainnya serta sedikit karbohidrat. Daging merupakan sumber protein yang sangat baik, mengandung asam amino penting dan tinggi kandungan vitamin dan mineral. Kelebihan pangan yang berasal dari daging dengan pangan vegetarian adalah tingginya kandungan vitamin B12 dan zat besi ([FAO], 2007). Daging terdiri dari air, lemak, protein, mineral dan sedikit karbohidrat. Adapun jenis daging yang akan dibahas adalah jenis daging yang digunakan sebagai sampel data untuk penelitian ini antara lain :

a. Daging Sapi

Ciri daging sapi dewasa antara lain berwarna merah cerah, memiliki tekstur yang kasar, konsistensi daging keras, jumlah marbling banyak dan warna lemak putih kekuningan. Menurut (Potter, 1993) daging sapi memiliki warna cerah, bau dan rasa aromatis, berserat halus dengan sedikit lemak, konsistensi liat/kenyal, permukaan mengkilat, dan bersih tidak ada darah

b. Daging Babi

Daging babi memiliki warna merah muda, tekstur daging halus, konsistensi daging empuk, jumlah marbling banyak, dan warna lemak putih. Menurut (Naruki & Kononi, 1992) warna daging babi pucat sehingga merah muda, serabut halus dengan konsistensi padat dan berbau spesifik, dan pada umur tua daging berwarna lebih tua, sedikit lemak dan serabut kasar.

2.2 Suhu

Suhu merupakan ukuran energi kinetik rata-rata dari pergerakan molekul-molekul. Suhu suatu benda ialah keadaan yang menentukan kemampuan benda tersebut, untuk memindahkan (transfer) panas ke benda – benda lain atau menerima panas dari benda – benda lain tersebut.

Pada suhu panas, daging akan lebih cepat mengeluarkan gas-gas yang terkandung pada daging itu sendiri seperti gas NH_3 , Methane, dan CO_2 . Suhu panas ini juga dapat mengakibatkan tidak konsistennya asap yang keluar pada daging saat dipanaskan, sehingga data yang terbaca pada *electronic nose* tidak stabil. Berbeda dengan suhu ruang dan -22°C , daging tidak mengeluarkan asap, sehingga data pada *electronic nose* dapat terbaca stabil.

2.3 *Electronic Nose*

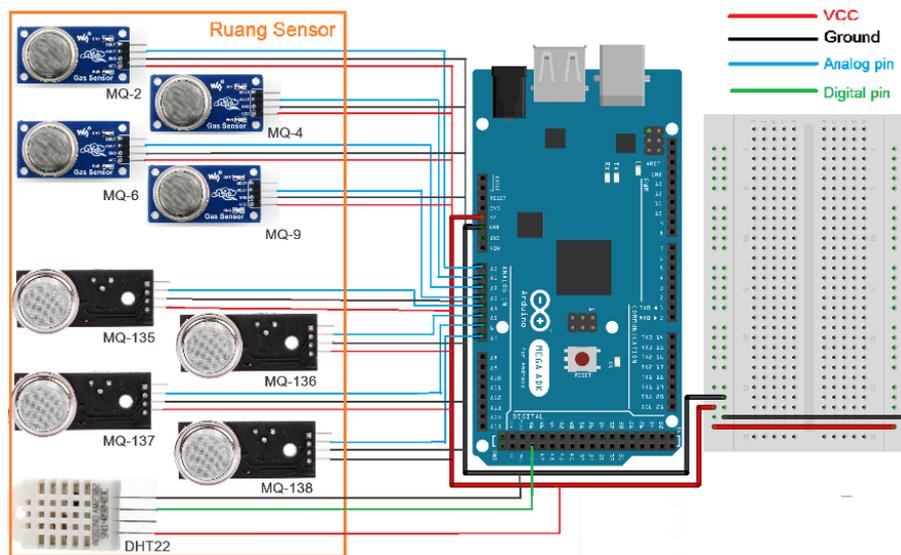
Electronic nose adalah sebuah instrumen yang mereplikasi indra penciuman manusia untuk mendeteksi bau (Nurjuliana, Che Man, & Mat Hashim, 2011). Data mengenai bau didapatkan melalui sensor kimia yang terdapat dalam *electronic nose*. Sensor-sensor yang terinstalasi pada *electronic nose* dalam penelitian ini adalah sensor-sensor yang memiliki selektivitas terhadap hidrogen (H_2), metana (CH_4), alkohol, karbon monoksida (CO), karbon dioksida (CO_2), hidrogen sulfat (H_2S), dan keton; digunakan untuk membedakan beberapa tingkat kemurnian daging sapi berdasarkan persentase pencampuran dengan daging babi. (Daiva, 2017)

2.4 Sensor Gas Elektrokimia

Sensor merupakan komponen yang berfungsi untuk mendeteksi atau mengukur dari sebuah objek yang di teliti, yang di gunakan di berbagai alat. Saat ini sensor dibuat dengan ukuran yang sangat kecil agar memudahkan dalam pemakaian dan menghemat dalam segi energi. Sensor merupakan bagian dari transduser yang berfungsi untuk “merasakan dan menangkap” adanya perubahan dari suatu energi eksternal yang masuk ke dalam bagian input bagian converter dari transduser untuk diubah menjadi energi listrik (Rusmadi, 2001).

Informasi mengenai gas yang dikeluarkan dari bau daging didapat dengan menggunakan sensor elektrokimia. Penelitian ini menggunakan delapan sensor

elektrokimia dan 1 sensor suhu dan kelembapan. Kedelapan sensor elektrokimia tersebut memiliki kepekaan terhadap gas-gas yang sebagaimana dipaparkan pada tabel 2.1. Kombinasi dari delapan sensor gas tersebut diperlukan untuk mengidentifikasi pola kandungan gas dari daging yang diuji coba. Sensor suhu dan kelembapan diperlukan untuk meyakinkan bahwa pengukuran gas dilakukan pada suhu dan kelembapan yang tidak bervariasi. Dengan demikian, pengukuran kandungan gas pada daging tidak terganggu oleh perbedaan lingkungan uji coba.



Gambar 2.1 Skema *Electronic Nose*

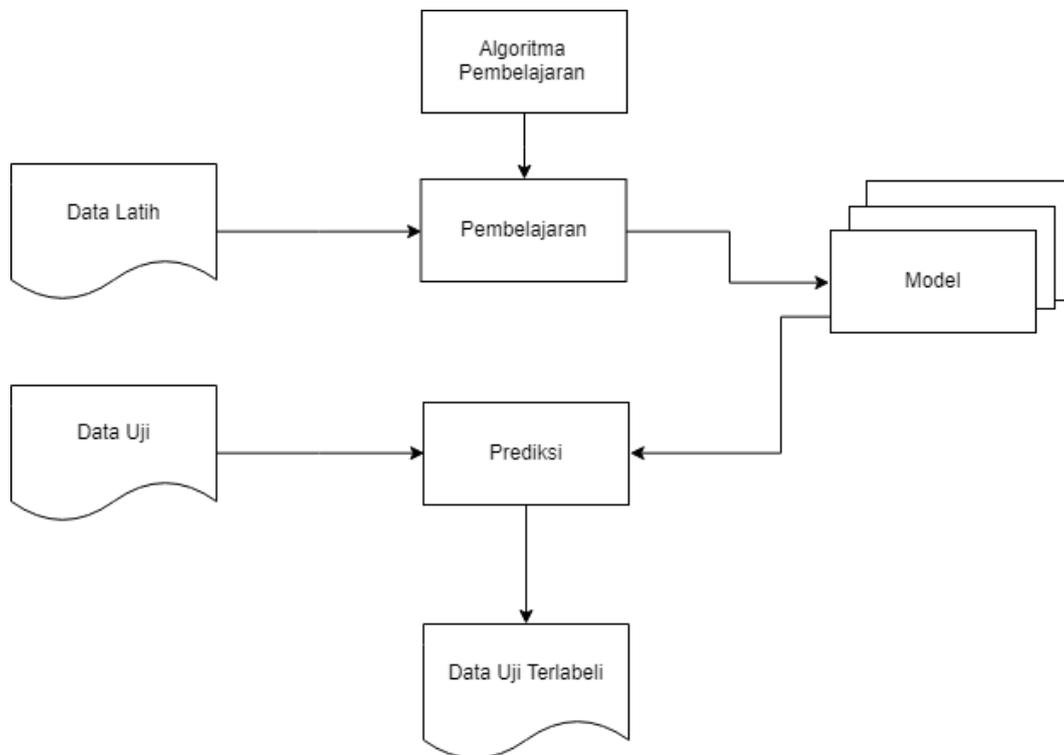
Tabel 2.1 Daftar Sensor Gas dan Fungsinya

Sensor	Fungsi
MQ2	LPG, i-butane, propane, methane, alcohol, H ₂ , smoke
MQ4	Methane (CH ₄) Natural gas
MQ6	LPG, iso-butane, propane
MQ9	Methane, Propane and CO
MQ135	Carbon Dioxide (CO ₂)
MQ136	Hydrogen Sulfide (H ₂ S)
MQ137	Amonia (NH ₃)
MQ138	Ketone
DHT22	Temperature-Humidity

2.5 Klasifikasi

Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu : Pembangunan model sebagai prototype untuk disimpan sebagai memori dan Penggunaan model tersebut untuk melakukan pengenalan, klasifikasi, prediksi pada suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang mudah disimpan .

Berbagai metode klasifikasi menggunakan pendekatan yang mirip untuk menentukan kelas dari suatu data. Setiap teknik klasifikasi menggunakan sebuah algoritme pembelajaran untuk membentuk suatu model yang menjelaskan relasi antara variable dan kelas dari data yang telah diketahui kelasnya. Model yang dihasilkan seharusnya mampu mengenali pula kelas dari suatu data baru yang belum diketahui kelasnya. Dengan demikian sebuah algoritme klasifikasi adalah membentuk model yang memiliki generalisasi yang baik.

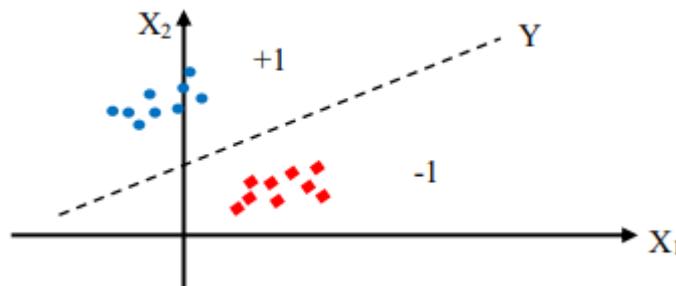


Gambar 2.2 Skema dasar proses klasifikasi

Gambar 2.2 menjelaskan proses dasar dari permasalahan klasifikasi. Sekumpulan data yang telah diketahui kelasnya digunakan sebagai data latih

(*training set*). Data latih digunakan oleh suatu algoritme klasifikasi untuk proses pembelajaran. Proses pembelajaran akan mengenali relasi antara kelas dan variable-variabel pada data latih. Pembelajaran yang menggunakan contoh data yang telah diketahui kelasnya disebut dengan . Relasi yang telah dikenali disebut dengan model. Model dapat dimanfaatkan pada proses prediksi untuk menentukan kelas dari data uji (*training set*), yaitu sekumpulan data baru yang belum diketahui kelasnya.

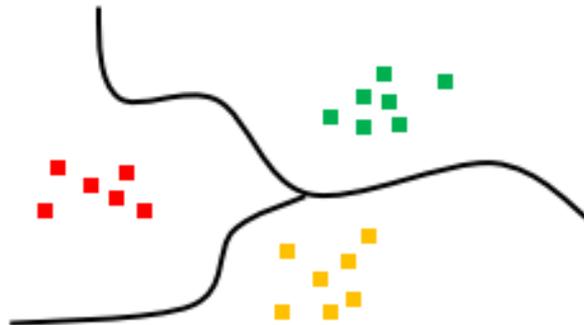
Menurut (Santosa B. , 2007), klasifikasi dibagi menjadi dua macam yakni klasifikasi dua kelas dan multi kelas. Yang pertama, klasifikasi dua kelas dapat dijelaskan sebagai berikut. Misalkan kita memiliki *set data training* $(X_i, Y_i), i = 1, \dots, l$ dengan data input $X = \{x_1, x_2, \dots, x_l\} \subseteq \mathbb{R}^N$ dan output yang bersangkutan $Y = \{y_1, \dots, y_l\} \subseteq \{\pm\}^l$. Tujuan dari klasifikasi dua kelas adalah menemukan suatu fungsi keputusan (*decision fuction*) $F(x)$ yang secara akurat memprediksi kelas dari data test (x, y) yang berasal dari fungsi distribusi yang sama dengan data untuk *training* seperti Gambar 2.3 . *Set data* $(X_i, Y_i), i = 1, \dots, l$ biasa dinamakan *set training*, dimana X_i berkaitan dengan parameter input dan Y_i menunjukkan parameter *output*



Gambar 2.3 Ilustrasi Klasifikasi Dua Kelas

Untuk klasifikasi Multi Kelas, memiliki *set data* untuk *training* $(X_i, Y_i), i = 1, \dots, l$ dengan data input $X = \{x_1, x_2, \dots, x_l\} \subseteq \mathbb{R}^N$ dan output yang bersangkutan = $\{y_1, \dots, y_l\} \subseteq \{1, 2, \dots, k\}^l$. Terlihat bahwa output Y tidak lagi terbatas ± 1 seperti dalam kasus dua kelas. Output dari data kita bisa 1,2,3,4 atau bahkan 10. Ilustrasi ditampakkan dalam Gambar 2.4. Dalam hal ini harus mengelompokkan obyek yang di pelajari ke dalam lebih dari dua kelas atau ke dalam k kelas dimana nilai k lebih

dari 2. Dalam kasus demikian tugas klasifikasi menjadi lebih rumit dan perlu teknik khusus untuk mengatasinya.



Gambar 2.4 Ilustrasi Klasifikasi Multi Kelas

2.6 K-Nearest Neighbor (k-NN)

Klasifikasi dengan k-NN adalah mengelompokkan data baru, dalam penelitian ini adalah data uji berdasarkan jarak ke beberapa data k tetangga terdekat terhadap data latih (Hanselman, 1998). Prinsip kerja k-NN adalah menghitung jarak menggunakan jarak *Euclidean*. Jarak *Euclidean* digunakan untuk menghitung jarak antara dua vektor yang berfungsi menguji ukuran yang bisa digunakan sebagai interpretasi kedekatan jarak antara dua obyek (Mcandrew, 2004); yang direpresentasikan dalam persamaan (1).

$$d(X, Y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \dots\dots\dots (2.1)$$

Keterangan :

$d(X, Y)$: Jarak Euclidean antara vector x dan vector y

X_i : Fitur ke i dari vector x

Y_i : Fitur ke i dari vector y

n : Jumlah Fitur pada vektor x dan y

Klasifikasi k-NN dilakukan dengan mencari tetangga terdekat dari data uji dan memilih kelas dengan anggota terbanyak. Adapun langkah - langkah klasifikasi k-NN adalah sebagai berikut:

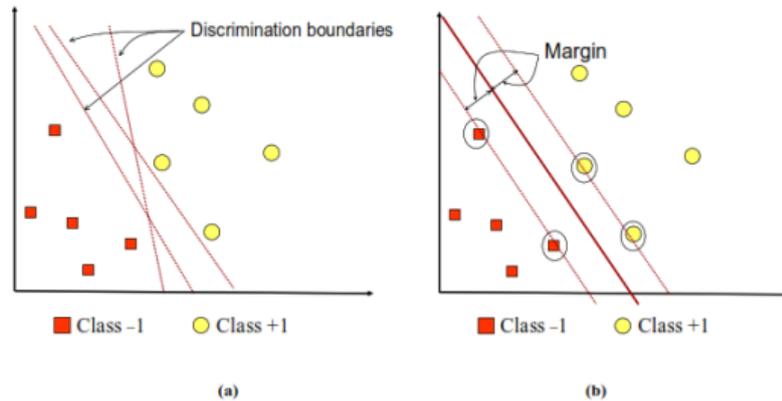
1. Mementukan parameter k (jumlah tetangga paling dekat).

2. Menghitung kuadrat jarak *Euclid* (*query instance*) masing-masing objek terhadap data sampel yang diberikan.
3. Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *Euclid* terkecil.
4. Mengumpulkan kategori y (klasifikasi *Nearest Neighbor*).
5. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksi nilai *query instance* yang telah dihitung.

k-NN memiliki beberapa kelebihan yaitu ketangguhan terhadap *training data* yang memiliki banyak *noise* dan efektif apabila *training data*-nya besar. Sedangkan, kelemahan k-NN adalah k-NN perlu menentukan nilai dari parameter k (jumlah dari tetangga terdekat), *training* berdasarkan jarak tidak jelas mengenai jenis jarak apa yang harus digunakan dan atribut mana yang harus digunakan untuk mendapatkan hasil terbaik, dan biaya komputasi cukup tinggi karena diperlukan perhitungan jarak dari tiap sampel uji pada keseluruhan sampel latih.

2.7 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan salah satu teknik klasifikasi data dengan proses pelatihan (*supervised learning*). Salah satu ciri dari metode klasifikasi SVM adalah menemukan garis pemisah (*hyperplane*) terbaik sehingga diperoleh ukuran margin yang maksimal. Titik yang terdekat dengan *hyperplane* disebut *support vector*. Margin adalah dua kali jarak antara *hyperplane* dengan *support vector*. Ilustrasi SVM untuk data yang terpisahkan secara linier dapat dilihat pada Gambar 2. 5.



Gambar 2.5 Penentuan hyperplane terbaik

Gambar 2.5a memperlihatkan beberapa *pattern* yang merupakan anggota dari dua buah *class* : +1 dan -1. *Pattern* yang tergabung pada *class* - 1 disimbolkan dengan warna merah (kotak), sedangkan *pattern* pada *class* +1, disimbolkan dengan warna kuning(lingkaran). Permasalahan klasifikasi dapat diterjemahkan dengan usaha menemukan garis (*hyperplane*) yang memisahkan antara kedua kelompok tersebut. Berbagai alternatif garis pemisah (*discrimination boundaries*) ditunjukkan pada gambar 2.2a.

Hyperplane pemisah terbaik antara kedua *class* dapat ditemukan dengan mengukur margin *hyperplane* dan mencari titik optimum *hyperplane* tersebut. Margin adalah jarak antara *hyperplane* dengan *pattern* terdekat dari masing-masing *class*. *Pattern* yang paling dekat disebut sebagai *support vector*. Garis solid pada gambar 2.1b menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua *class*, sedangkan titik merah dan kuning yang berada dalam lingkaran hitam adalah *support vector*. Usaha untuk mencari lokasi *hyperplane* ini merupakan inti dari proses pembelajaran pada SVM.

Misal data yang tersedia dinotasikan sebagai $\vec{x}_i \in \mathfrak{R}^d$ sedangkan label masing-masing dinotasikan $y_i A \in \{-1,+1\}$ untuk $i= 1,2, \dots , l$ dimana l adalah banyaknya data. Diasumsikan kedua kelas -1 dan +1 dapat terpisah secara sempurna oleh *hyperplane* berdimensi d , yang didefinisikan:

$$\vec{w} \cdot \vec{x} + b = 0 \dots\dots\dots (2.2)$$

Pattern \vec{x} yang termasuk class -1 (sampel negatif) dapat dirumuskan sebagai pattern yang memenuhi pertidaksamaan

$$\vec{w} \cdot \vec{x} + b \leq 0 \dots\dots\dots (2.3)$$

Sedangkan pattern \vec{x} yang termasuk class +1 (sampel positif)

$$\vec{w} \cdot \vec{x} + b \geq 0 \dots\dots\dots (2.4)$$

Margin terbesar dapat ditemukan dengan mengoptimalkan nilai jarak antara hyperplane dan titik terdekatnya, yaitu $1/\|\vec{w}\|$. Hal tersebut dapat dirumuskan sebagai Quadratic Programming (QP) problem, yaitu mencari titik minimal persamaan 5, dengan memperhatikan constraint persamaan 6.

$$\min_w \tau(w) = \frac{1}{2} \|\vec{w}\|^2 \dots\dots\dots (2.5)$$

Constraint

$$y_i(\vec{w} \cdot \vec{x}_i + b) - 1 \geq 0, \quad \forall_i \dots\dots\dots (2.6)$$

Data input dinotasikan x_i , adalah keluaran dari data x_i , w , b adalah parameter-parameter yang di cari nilainya. Formulasi di atas, ingin meminimalkan fungsi tujuan (obyektif function) $\frac{1}{2} \|\vec{w}\|^2$ atau memaksimalkan kuantitas $\|\vec{w}\|^2$ dengan memperhatikan pembatas sebagaimana persamaan 3 dan 4. Bila output data $y_i = +1$, maka pembatas menjadi $\vec{w} \cdot \vec{x} + b \geq +1$. Sebaliknya bila $y_i = -1$, pembatas menjadi $\vec{w} \cdot \vec{x} + b \leq -1$.

Permasalahan tersebut dapat dipecahkan dengan berbagai teknik komputasi, di antaranya Lagrange Multiplier.

$$L(w, b, \alpha) = \frac{1}{2} \|\vec{w}\|^2 - \sum_{i=1}^l \alpha_i (y_i(\vec{x}_i \cdot \vec{w} + b) - 1) \dots\dots\dots (2.7)$$

$$L(w, b, \alpha) = \frac{1}{2} \|\vec{w}\|^2 - \sum_{i=1}^l \alpha_i y_i(\vec{x}_i \cdot \vec{w} - b) \sum_{i=1}^l \alpha_i y_i + \sum_{i=1}^l \alpha_i \dots\dots\dots (2.8)$$

Tambahan konstrain, $\alpha_1 \geq 0$ (nilai dari koefisien lagrange). Meminimumkan L terhadap w dan b.

$$\frac{\partial}{\partial b} L(w, b, \alpha) = 0 \dots\dots\dots (2.9)$$

$$\frac{\partial}{\partial w} L(w, b, \alpha) = 0 \dots\dots\dots(2.10)$$

Dari persamaan 9 dan persamaan 10 diperoleh persamaan berikut:

$$\sum_{i=1}^l \alpha_i y_i = 0 \dots\dots\dots(2.11)$$

$$\vec{w} = \sum_{i=1}^l \alpha_i y_i x_i \dots\dots\dots(2.12)$$

Nilai \vec{w} sering kali besar, tetapi nilai $\alpha_i \geq 0$ terhingga. Untuk itu, formula *lagrangian* L_p (*primal problem*) diubah kedalam *dual problem*. Dengan mensubsitusikan persamaan 12 ke LP diperoleh *dual problem* L_d dengan konstrain berbeda.

$$L_d(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \vec{x}_i \vec{x}_j \dots\dots\dots(2.13)$$

Dimana α_i adalah *Lagrange Multiplier* $\alpha_i \geq 0$. Nilai optimal dari persamaan 13 dapat dihitung dengan meminimalkan L terhadap \vec{w} dan b, dan memaksimalkan L terhadap α_i , dengan memperhatikan sifat bahwa pada titik optimal *gradient* $L = 0$ persamaan 13 dapat dimodifikasi sebagai maksimalisasi problem yang hanya mengandung α_i , sebagaimana terlihat pada persamaan 14 dan 15 dibawah ini.

$$\max_{\alpha} L_d = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \vec{x}_i \vec{x}_j \dots\dots\dots(2.14)$$

$$\alpha_i \geq 0 \quad (i = 1, 2, \dots, l) \quad \sum_{i=1}^l \alpha_i y_i = 0 \dots\dots\dots(2.15)$$

Berdasarkan persamaan diatas, maka akan diperoleh α_i yang kebanyakan bernilai positif yang disebut sebagai *support vector* (Vapnik, 1995).

2.7.1. Fungsi Kernel

Pada mulanya klasifikasi dikembangkan dengan asumsi kelinieran. Sehingga algoritma yang dihasilkan terbatas untuk kasus-kasus yang linear saja. Akan tetapi untuk menghadapi kasus yang tidak linier maka dapat menggunakan bantuan berbagai macam fungsi Kernel. Kernel trick memberikan berbagai kemudahan, karena dalam proses pembelajaran SVM. Untuk menentukan support vector, maka cukup dengan mengetahui fungsi Kernel yang dipakai, dan tidak perlu mengetahui wujud dari fungsi non-linear.

Menurut (Karatzoglou, Smola, Hornik, & Zeieis, 2004) ada beberapa fungsi Kernel yang sering digunakan dalam literature SVM antara lain sebagai berikut :

- a) Kernel linear adalah kernel yang paling sederhana dari semua fungsi kernel. Kernel ini biasa digunakan dalam kasus klasifikasi teks.

$$K(x_i, x_j) = x_i^T x_j \dots\dots\dots(2.16)$$

- b) Kernel Polynominal adalah kernel yang sering digunakan untuk klasifikasi gambar.

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^p, \gamma > 0 \dots\dots\dots(2.17)$$

- c) Kernel Radial Basis Gaussian adalah kernel yang umum digunakan untuk data yang sudah valid (available) dan merupakan default dalam tools SVM.

$$K(x_i, x_j) = \exp(-\gamma ||x_i - x_j||^2) , \gamma > 0 \dots\dots\dots(2.18)$$

- d) Kernel Tangent Hyperbolic adalah kernel yang sering digunakan untuk neural networks.

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_{j+r}) \dots\dots\dots(2.19)$$

Pemilihan fungsi Kernel yang tepat merupakan hal yang sangat penting karena akan menentukan feature space, dimana fungsi classifier akan dicari. Sepanjang fungsi kernelnya sesuai/cocok, SVM akan beroperasi secara benar meskipun tidak tahu pemetaan yang digunakan (Santosa B. , 2007). Menurut (Scholkopf & Smola, 2002), fungsi Kernel Gaussian RBF memiliki kelebihan yaitu secara otomatis menentukan nilai rentang tak terhingga. Gaussian RBF juga efektif menghindari overfitting dengan memilih nilai yang tepat untuk parameter C dan γ dan RBF yang baik digunakan ketika tidak ada pengetahuan terdahulu. Fungsi kernel yang direkomendasikan adalah fungsi kernel RBF karena dapat memetakan hubungan tidak linier, RBF lebih robust terhadap outlier karena fungsi kernel RBF berada pada selang $(-\infty, \infty)$ sedangkan fungsi

Kernel yang lain memiliki rentang antara (-1 sampai dengan 1) (Hsu, Chang, & Lin, 2003).

2.8 Naïve Bayes

4.8.1. Bayesian Classifier

Metode Bayesian Classifier menggunakan pendekatan teori peluang untuk melakukan klasifikasi. Metode yang menjadi acuan pada Bayesian Classifier adalah Teorema Bayes. Teorema Bayes menjelaskan peluang suatu kejadian berdasarkan kombinasi antara pengetahuan sebelumnya tentang kejadian tersebut dan bukti-bukti baru yang dikumpulkan dari data. Secara matematis, Teorema Bayes dapat diekspresikan sebagai berikut :

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \dots\dots\dots (2.20)$$

Keterangan :

$P(A|B)$ = Peluang terjadinya kejadian A dengan syarat kejadian B telah terjadi.

$P(B|A)$ = Peluang terjadinya kejadian B dengan syarat kejadian A telah terjadi.

$P(A)$ = Peluang terjadinya kejadian A, tanpa pengaruh kejadian yang lain.

$P(B)$ = Peluang terjadinya kejadian B, tanpa pengaruh kejadian yang lain.

Persamaan Teorema Bayes tersebut dapat dimanfaatkan untuk melakukan klasifikasi. Pada permasalahan klasifikasi, variable B pada persamaan Teorema Bayes menyatakan sekumpulan variable-variabel pada data, sedangkan A menyatakan kelas pada data. Jika kelas pada suatu data memiliki hubungan non deterministic dengan variable-variabelnya maka A dan B dapat dianggap sebagai variable acak dan relasinya dapat dinyatakan sebagai $P(A|B)$. Nilai $P(A|B)$ bisa disebut peluang posterior dan $P(A)$ disebut sebagai peluang prior.

Selama proses training, probabilitas posterior $P(A|B)$ ditelaah dari setiap kombinasi A dan B dari data latih. Berdasarkan nilai probabilitas posterpr, sebuah data uji B' dapat diklasifikasikan ke dalam kelas A' yang memiliki nilai probabilitas posterior $P(A'|B')$ tertinggi.

2.8.2 Naïve Bayes untuk Klasifikasi

Naïve Bayes classifier menggunakan dasar Teorema Bayes untuk memperkirakan kejadian bersyarat munculnya suatu kelas dengan asumsi bahwa variable-variabel yang mempengaruhinya bersifat independen. Menggunakan asumsi independensi antar variabsle, peluang bersyarat munculnya suatu kelas, $P(A|B)$ dapat dihitung menggunakan persamaan sebagai berikut :

$$P(A|B) = \frac{P(A) \prod_{i=1}^d P(B_i|A)}{P(B)} \dots\dots\dots(2.21)$$

Keterangan :

$P(B_i|A)$ = Peluang terjadinya setiap variable B_i dengan syarat A

Nilai $P(B)$ bersifat konstan untuk setiap kejadian A, sehingga nilai $P(B)$ dapat dihilangkan dari persamaan. Dengan demikian, penentuan kelas suatu data dapat dilakukan mencari nilai yang maksimal dari $P(A_k|B)$ untuk setiap kelas yang terdapat pada data. Dalam bentuk matematika, pemilihan kelas dilakukan dengan menggunakan persamaan berikut :

$$C = \underset{k \in \{1, \dots, K\}}{\operatorname{argmax}} P(A_k) \prod_{i=1}^d P(B_i|A) \dots\dots\dots(2.22)$$

Keterangan :

K = Himpunan kelas yang terdapat pada data

2.9 Random Forest

Metode *Random Forest* merupakan salah satu metode dalam *Decision Tree*. *Decision Tree* atau pohon pengambil keputusan adalah sebuah diagram alir yang berbentuk seperti pohon yang memiliki sebuah *root node* yang digunakan untuk mengumpulkan data, Sebuah *inner node* yang berada pada *root node* yang berisi tentang pertanyaan tentang data dan sebuah *leaf node* yang digunakan untuk memecahkan masalah serta membuat keputusan. *Decision tree* mengklasifikasikan suatu sampel data yang belum diketahui kelasnya kedalam kelas – kelas yang ada. Penggunaan *decision tree* bertujuan agar dapat menghindari *overfitting* pada sebuah set data saat mencapai akurasi yang maksimum. *Random forest* adalah kombinasi dari masing – masing *tree* yang baik kemudian dikombinasikan ke dalam satu

model. Algoritma ini dapat menyediakan variabel dependen pada kelas sejumlah tree yang dibentuk. Skema ini pertama kali dicetuskan oleh Leo Breiman pada tahun 2000 untuk membangun prediktor dengan sekumpulan decision tree yang berkembang secara acak pada subruang data. Kelas yang dihasilkan dari proses klasifikasi dipilih dari kelas yang paling banyak yang dihasilkan oleh pohon keputusan yang ada (Biau, 2012). *Random Forest* bergantung pada sebuah nilai vector random dengan distribusi yang sama pada semua pohon yang masing masing *decision tree* memiliki kedalaman yang maksimal. *Random forest* adalah *classifier* yang terdiri dari *classifieryang* berbentuk pohon $\{h(\mathbf{x}, \theta_k), k = 1, \dots\}$ dimana θ_k adalah random vector yang didistribusikan secara independen dan masing masing tree pada sebuah unit kan memilih class yang paling populer pada input x . Berikut ini karakteristik akurasi pada *random forest*.

- Memusatkan *random forest*

Terdapat classifier $h_1(x), h_2(x), \dots, h_k(x)$ dan dengan training set dari distribusi random vector Y, X . Berikut fungsi yang terbentuk:

$$mg(X, Y) = au_k I(h_k(X) = Y) - \max_{j \neq Y} au_k I(h_k(X) = j) \dots\dots\dots(2.23)$$

Fungsi eror yang digunakan:

$$PE^* = P_{X,Y}(mg(X, Y) < 0) \dots\dots\dots(2.24)$$

Hasil dari penggabungan fungsi :

$$P_{X,Y}(P_{\Theta}(h(X, \Theta) = Y) - \max_{j \neq Y} P_{\Theta}(h(X, \Theta) = j) < 0) \dots\dots\dots(2.25)$$

- Kekuatan dan korelasi

Untuk *random forest*, batas atas dapat diturunkan untuk kesalahan generalisasi dari dua parameter yang mengukur seberapa akurat masing-masing klasifikasi dan ketergantungannya. Interaksi antara keduanya memberikan landasan bagi memahami cara kerja *random forest*. Fungsi yang dihasilkan adalah :

$$PE^* \leq \sum_i var (P_{\Theta}(h(X, \Theta) = Y) - P_{\Theta}(h(X, \Theta) = j)) s_j^2 \dots\dots\dots(2.26)$$

Pada fungsi tersebut, kekuatan tidak bergantung pada *forest*.

- *Random forest* menggunakan seleksi input yang *random*

Bagging digunakan untuk pemilihan fitur secara *random*. Masing – masing *training set* diambil dengan penggantian dari *training set* asli. Kemudian sebuah *tree* di tanam pada sebuah *training set* menggunakan seleksi fitur *random*. Ada dua alasan penggunaan *bagging* yaitu yang pertama penggunaan *bagging* untuk meningkatkan akurasi ketika fitur *random* digunakan. Yang kedua *bagging* digunakan untuk memberikan perkiraan dari kesalahan generalisasi (PE*) dari gabungan *tree*, untuk memperkirakan kekuatan dan korelasi. *Random Forest* yang paling sederhana dengan fitur *random* dibentuk dengan seleksi secara *random*, pada masing – masing node, sebuah grup kecil dari input variable yang terbagi. Membentuk *tree* menggunakan metodologi CART ke ukuran yang maksimum.

- *Random forest* menggunakan kombinasi input yang linear

Misalkan terdapat beberapa input, M, F mengambil fraksi pada M yang akan memimpin dalam meningkatkan kekuatan tetapi pada korelasi yang tinggi. Pendekatan yang lain terbentuk dengan mendefinisikan lebih banyak fitur dengan mengambil kombinasi *random linear* dari sejumlah variable input. Fitur tersebut variabel L yaitu jumlah dari variable yang dikombinasikan. Variabel L secara *random* diseleksi dan ditambahkan bersama dengan koefisien yang memiliki nomor *random* [-1,1]. Kombinasi linear F dihasilkan. Prosedur ini di sebut *Forest-RC*.

2.10 Kelebihan dan Kekurangan Masing-Masing Metode

Pada subbab ini membahas tentang kelebihan dan kekurangan dari 3 metode yang digunakan, yaitu metode *k-Nearest Neighbor*, *Support Vector Machine* dan *Naïve Bayes* yang di visualisasi pada table 2.1 .

Table 2.1 Kelebihan dan Kekurangan Metode

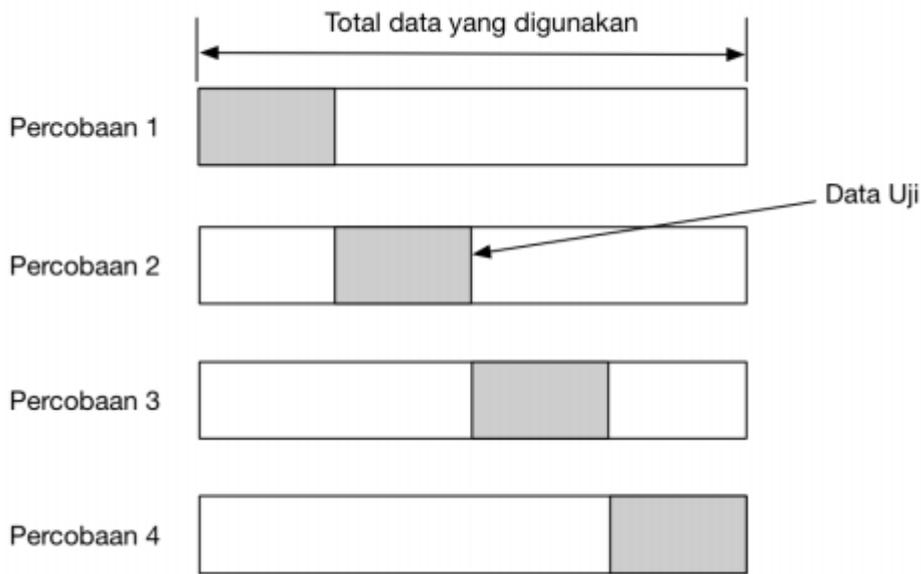
Metode	Kelebihan	Kekurangan
k-Nearest Neighbor	Robust terhadap data yang noisy	Perlu menunjukkan parameter K (jumlah tetangga terdekat)
	Efektif jika training data berjumlah banyak	Berdasarkan perhitungan nilai jarak (Distance based learning), tidak jelas perhitungan jarak mana yang sebaiknya digunakan dan atribut mana yang memberikan hasil yang baik.
Support Vector Machine	Generalisasi : kemampuan suatu metode untuk mengklasifikasikan suatu pattern, yang tidak termasuk data yang dipakai dalam fase pembelajaran metode itu	Sulit dipakai problem berskala besar. Dalam hal ini dimaksudkan dengan jumlah sampel yang diolah.
	Curse of dimensionality : masalah yang dihadapi suatu metode pattern recognition dalam mengestimasi parameter dikarenakan jumlah sampel data yang relatif lebih sedikit dibandingkan dengan dimensional ruang vektor data tersebut.	SVM secara teoritik dikembangkan untuk problem klasifikasi dengan dua kelas. Dewasa ini SVM telah dimodifikasi agar dapat menyelesaikan masalah dengan lebih dari dua kelas
Naïve Bayes	Lebih cepat dalam proses perhitungan	Tidak berlaku jika probabilitas kondisionalnya adalah nol, apabila nol maka probabilitas prediksi akan bernilai nol juga.

	Menangani kuantitatif dan data diskrit	Mengasumsikan variabel bebas.
Random Forest	Random Forest merupakan salah satu algoritma yang cukup efektif dalam melakukan klasifikasi. Hal ini terlihat dari grafik persentase error yang membandingkan dengan algoritma klasifikasi lain. Random Forest memperlihatkan persentase error yang bersaing dengan algoritma lain, dengan di beberapa data memiliki persentase error terkecil.	Random Forest memiliki kelemahan dalam hal kestabilan akurasi yang dihasilkan. Dengan parameter input dan data yang sama, untuk eksekusi lebih dari satu kali yang dilakukan berurutan akan menghasilkan akurasi yang berbeda-beda

2.11 K-Fold Cross Validation

Validation adalah proses untuk mengevaluasi keakuratan prediksi dari model. Validasi digunakan untuk memperoleh prediksi menggunakan model yang ada dan kemudian membandingkan hasil tersebut dengan hasil yang sudah diketahui, ini mewakili langkah paling penting dalam proses membangun sebuah model.

Cross validation adalah metode statistik yang mengevaluasi dan membandingkan algoritma pembelajaran dengan membagi data menjadi dua yaitu data training dan data testing. Bentuk dari cross validation adalah k-fold cross validation. Prinsip dari k-fold cross validation adalah membagi tiap kelompok data menjadi k bagian kelompok data yang selanjutnya, data tersebut secara bergantian akan digunakan untuk training dan testing sejumlah k pengujian. Gambar 2.4 merupakan ilustrasi prosedur dari K-Fold *cross validation*



Gambar 2.6 Metode Pembagian Data dengan *K-Fold Cross-Validation*

2.12 Confussion Matrix

Confusion matrix adalah tabel yang merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas. Tabel ini memperlihatkan prediksi mana yang benar dan prediksi mana yang salah dari aktualnya. Dengan melihat tabel ini maka dapat diketahui dengan jelas seberapa sering suatu kelas mendapatkan label yang salah. Hal ini berguna untuk menguji performa algoritma klasifikasi.

Table 2.2 Confussion Matrix untuk 2 Kelas

<i>Classification</i>	<i>Predicted Class</i>		
		Class = Yes	Class = No
<i>Observed Class</i>	<i>Class = Yes</i>	A (<i>True Positive</i> –TP)	B (<i>False Negative</i> – FN)
	<i>Class = No</i>	C (<i>False Positive</i> - FP)	D (<i>True Negative</i> – TN)

Keterangan :

- a) True Positive (TP) = proporsi positif dalam data set yang diklasifikasikan positif.
- b) True Negative (TN) = proporsi negative dalam data set yang diklasifikasikan negative.
- c) False Positive (FP) = proporsi negatif dalam data set yang diklasifikasikan positif.
- d) False Negative (FN) = proporsi negative dalam data set yang diklasifikasikan negative

Pada Tabel 2.2 nilai TP (*true positive*) dan TN (*true negative*) menunjukkan tingkat ketepatan klasifikasi. Umumnya semakin tinggi nilai TP dan TN semakin baik pula tingkat klasifikasi dari akurasi, presisi, dan recall. Jika label prediksi keluaran bernilai benar (*true*) dan nilai sebenarnya bernilai salah (*false*) disebut sebagai *false positive* (FP). Sedangkan jika prediksi label keluaran bernilai salah (*false*) dan nilai sebenarnya bernilai benar (*true*) maka hal ini disebut sebagai *false negative* (FN). Berdasarkan table 2.2 perhitungan evaluasi performansi metode klasifikasi dapat dilakukan dengan formula sebagai berikut :

$$TP\ Rate = \frac{TP}{TP+FN} \dots\dots\dots(2.27)$$

$$FP\ Rate = \frac{FP}{FP+TN} \dots\dots\dots(2.28)$$

$$Akurasi = \frac{TN+TP}{TN+TP+FN=FP} \times 100\% \dots\dots\dots(2.29)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \dots\dots\dots(2.30)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \dots\dots\dots(2.31)$$

$$F - Measure = \frac{2 \times Precision \times Recall}{(Precision+Recall)} \dots\dots\dots(2.33)$$

2.13 Kurva ROC

Menurut (Gorunescu, 2011) , kurva ROC banyak digunakan untuk menilai hasil prediksi, kurva ROC adalah teknik untuk memvisualisasikan, mengatur, dan memilih pengklasifikasian berdasarkan kinerja mereka . Kurva ROC adalah *tool* dua dimensi yang digunakan untuk menilai kinerja klasifikasi yang menggunakan dua *class* keputusan, masing-masing objek dipetakan ke salah satu elemen dari himpunan pasangan, positif atau negatif. Pada kurva ROC, TP *rate* diplot pada sumbu Y dan FP *rate* diplot pada sumbu X.

Untuk klasifikasi *data mining* menurut Gorunescu, nilai AUC atau Area di bawah kurva yang sering digunakan sebagai jumlah dari kurva ROC dan sebagai ukuran kinerja pengklasifikasi dapat dibagi menjadi beberapa kelompok :

- 0,90 – 1.00= *Excellent Classification*
- 0,80 – 0,90= *Good Classification*
- 0,70 – 0,80= *Fair Classification*
- 0,60 – 0,70= *Poor Classification*
- 0,50 – 0,60= *Failure*

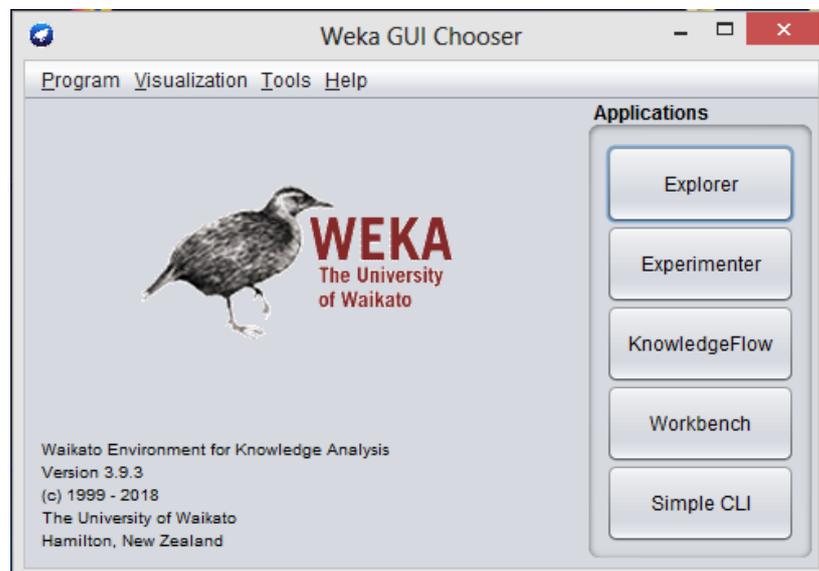
2.14 WEKA

WEKA merupakan API Java yang menyediakan API untuk pengolahan dalam data mining yang berbasis open source (GPL) dan berengine JAVA. WEKA dikembangkan oleh Universitas Waikato di Selandia Baru dan juga merupakan perangkat lunak gratis yang tersedia dibawah GNU (General Public License). WEKA menyediakan penggunaan teknik klasifikasi menggunakan pohon keputusan dengan algoritma J48. Teknik yang digunakan WEKA adalah classifier.

Menurut situs resmi WEKA (Waikato Environment for Knowledge Analysis), “WEKA is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own Java code. Weka contains tools for data pre-processing, classification,

regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.”

The Waikato Environment for Knowledge Analysis (Weka) adalah rangkaian lengkap perpustakaan kelas Java yang mengimplementasikan banyak state-of-the-art pembelajaran mesin dan algoritma data mining. Weka tersedia secara bebas di World Wide Web dan menyertai teks baru pada dokumen data mining dan sepenuhnya menjelaskan semua algoritma yang dikandungnya. Aplikasi yang ditulis menggunakan library class pada Weka yang dapat dijalankan pada komputer manapun dengan kemampuan browsing Web, ini memungkinkan pengguna untuk menerapkan teknik pembelajaran mesin untuk data mereka sendiri terlepas dari platform komputer (Witten, Frank, & Hall, 2011).



Gambar 2.7 Tampilan awal GUI WEKA

WEKA mulai dikembangkan sejak tahun 1994 dan telah menjadi software data mining open source yang paling populer. WEKA mempunyai kelebihan seperti mempunyai banyak algoritma data mining dan machine learning, kemudahan dalam penggunaannya, selalu up-to-date dengan algoritma-algoritma yang baru. Software WEKA tidak hanya digunakan untuk akademik saja namun cukup banyak dipakai oleh perusahaan untuk meramalkan bisnis dari suatu perusahaan. Ian H. Witten merupakan latar belakang dibalik kesuksesan WEKA. Beliau merupakan profesor di Universitas of Waikato, New Zealand, yang menekuni Digital Library, Text

Mining, Machine Learning dan Information Retrieval. Pada Weka ada beberapa metode pemilihan variable dari suatu dataset, diantaranya BestFirst, ExhaustiveSearch, FCBFSearch, GeneticSearch, GreedyStepwise, RaceSearch, RandomSearch, Rankerdan, RankerSearch. Metode atau Teknik yang digunakan Weka adalah Predictive dan Descriptive karena Weka mendukung teknik-teknik data preprocessing, clustering, classification, regression, visualization, dan feature Reduction. Semua teknik Weka adalah didasarkan pada asumsi bahwa data tersedia sebagai flat file tunggal atau hubungan, dimana setiap titik data digambarkan oleh sejumlah tetap atribut (biasanya, atribut numeric atau nominal, tetapi beberapa jenis atribut lain juga didukung). Metode evaluasi yang didukung oleh WEKA adalah :

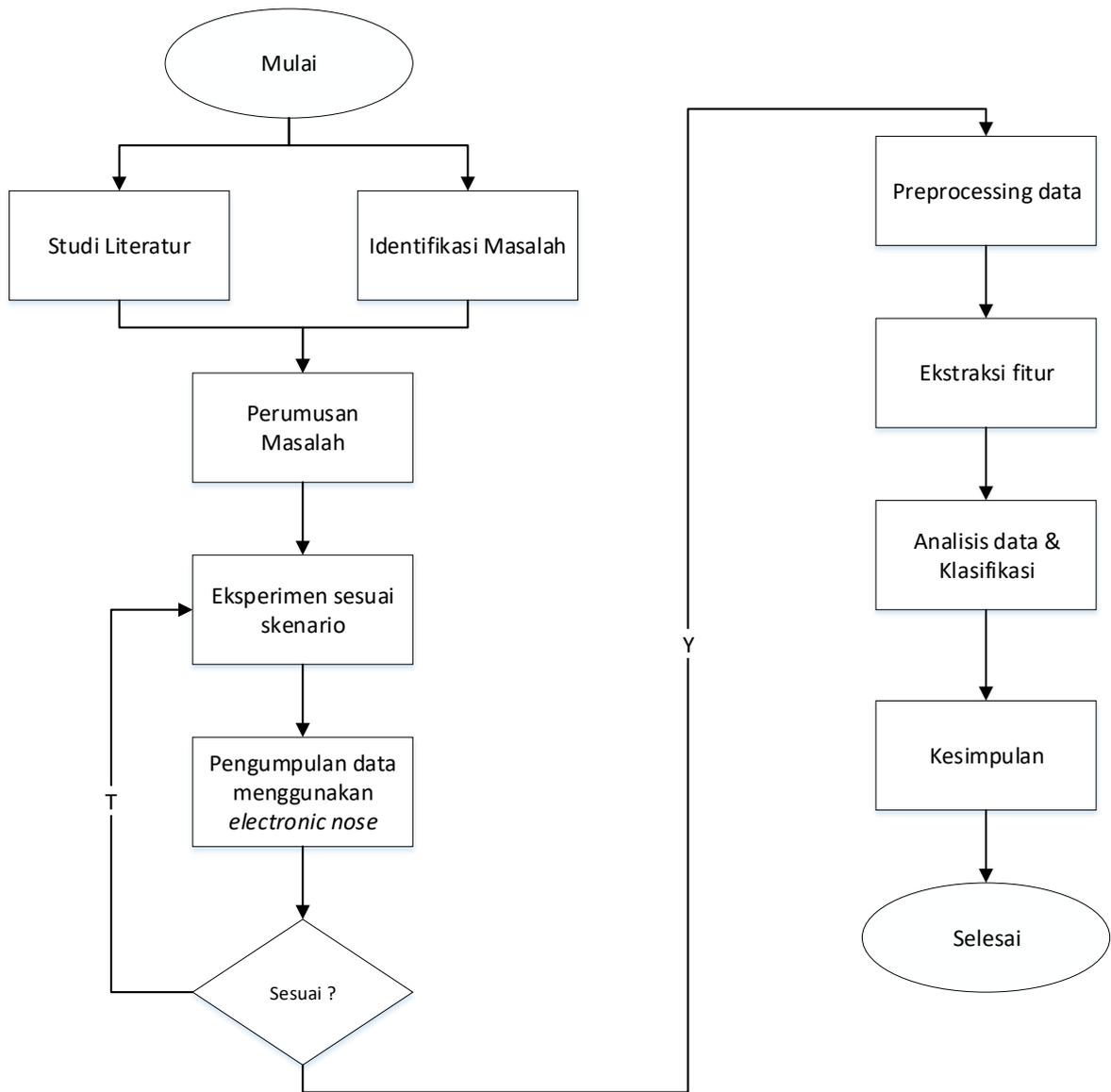
- Use training set . Pada metode ini Weka akan menggunakan data training yang diinputkan sebelumnya sebagai data testing. Dengan kata lain, proses *training* dan *testing* menggunakan data yang sama
- Supplied test set. Metode ini mengharuskan pengguna untuk memiliki data *training* dan *testing* terpisah
- Cross-validation. Menggunakan metode cross validation, dimana data latih akan dibagi menjadi k bagian secara acak. Selanjutnya k-1 bagian digunakan sebagai data latih dan 1 bagian digunakan sebagai data uji. Proses tersebut diulang sehingga setiap bagian berkesempatan menjadi data uji.
- Percentage split. Data yang diinputkan pada langkah sebelumnya akan dibagi menjadi data latih dan data uji berdasarkan presentase tertentu.

BAB III METODOLOGI PENELITIAN

Bab ini akan menggambarkan metodologi penelitian yang berisi tahapan penelitian dan rencana waktu pengerjaannya. Metodologi penelitian ini akan menjadi panduan dalam mengerjakan penelitian agar dapat diselesaikan secara sistematis, terarah dan jelas.

3.1 Diagram Alur Metode Penelitian

Tahapan penelitian yang akan dilakukan oleh penulis digambarkan dalam diagram alur metodologi yang tersaji pada Gambar 3.1 di bawah ini.



Gambar 3.1 Diagram Alir Penelitian

3.2 Perumusan Masalah

Sebelum melakukan penelitian, peneliti melakukan studi kasus terhadap permasalahan yang sering dialami khususnya pada pasar Indonesia. Maraknya praktek kecurangan daging sapi campuran atau sering disebut daging oplosan merupakan hal yang harus ditindaklanjuti karena mengingat 87,18% dari penduduk Indonesia adalah pemeluk agama islam (Badan Pusat Statistik, Sensus Penduduk 2010). Berdasarkan permasalahan tersebut maka peneliti mengangkat topik pengaruh suhu terhadap klasifikasi daging sapi berbasis *electronic nose*.

3.3 Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan menggunakan *electronic nose* yang telah dirakit untuk mengetahui perubahan konsentrasi gas, suhu, dan kelembapan pada sampel daging. Sampel daging yang digunakan sebagai ground truth adalah daging sapi dan daging babi yang telah dicincang halus lalu ditimbang dengan total berat 10 gram setiap variasi sampel daging. Adapun rincian dari sampel daging yang digunakan yaitu :

- S0B100 : 100% daging babi (daging babi murni)
- S10B90 : 10% daging sapi, 90% daging babi
- S50B50 : 50% daging sapi, 50% daging babi
- S90B10 : 90% daging sapi, 10% daging babi
- S100B0 : 100% daging sapi (daging sapi murni)

5 sampel data daging yang ada, kemudian diambil datanya menggunakan *electronic nose* dengan 3 variasi suhu yang masing-masing suhunya mewakili jenis-jenis suhu untuk mengetahui tingkat akurasi tertinggi. Adapun variasi suhu yang digunakan pada penelitian ini yaitu:

- -22°C
- Suhu ruang \approx 24°C
- 55°C

Total data yang digunakan pada penelitian ini adalah 750 data, dengan rincian:

- S0B100 = 150 data
- S10B90 = 150 data
- S50B50 = 150 data
- S90B10 = 150 data
- S100B0 = 150 data

Data yang dihasilkan oleh electronic nose berupa angka-angka seperti pada gambar 3.2 yang kemudian diolah menggunakan Microsoft Excel untuk diambil data statistiknya berupa nilai minimum, maksimum, kurtosis, skewness, mean, dan standar deviasi



Gambar 3.2 *output* serial monitor yang dihasilkan *electronic nose*

3.4 Preprocessing Data

Pada tahap *preprocessing* data, peneliti menerapkan data *cleaning*. *Data cleaning* merupakan proses membuang duplikasi data, memeriksa data yang tidak

konsisten, dan memperbaiki kesalahan pada data, seperti kesalahan penulisan. Pada umumnya data yang diperoleh baik dari hasil eksperimen, memiliki isi yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekadar salah ketik. Selain itu, ada juga atribut atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data cleaning juga akan memengaruhi hasil informasi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

3.5 Ekstraksi Fitur

Untuk memastikan apakah analisis yang dibuat benar atau tidak maka pertama kali yang harus dipastikan adalah apakah data tersebut terdistribusi normal atau tidak. Nilai parameter statistik yang dimanfaatkan pada tahap ekstraksi fitur adalah nilai minimum, maksimum, kurtosis, skewness, mean dan stdev yang dihitung menggunakan alat bantu Microsoft excel. Tabel 3.1 merupakan salah satu contoh perhitungan distribusi normal menggunakan Microsoft Excell.

Table 3.1 Distribusi normal pada masing-masing sensor

Fitur Sensor	Kurtosis	Skewness	Mean	Stdev
MQ2	0,702943	0,26631332	343,5313	3,044578
MQ4	9,523292	2,41261746	554,7875	6,609872
MQ6	10,92701	-0,1685022	709,0375	1,491643
MQ9	2,9748	-1,2291393	859,4	1,705338
MQ135	-1,42625	-0,282968	574,35	5,154469
MQ136	2,961708	0,86558918	428,0625	5,712931
MQ137	-1,00748	-0,2300753	549,4813	12,39845
MQ138	-0,68318	-0,2717227	465,1125	8,178866

3.5.1 Kurtosis

Kurtosis adalah ukuran seberapa datar bagian atas distribusi simetris bila dibandingkan dengan distribusi normal dari varian yang sama. Kurtosis sebenarnya lebih dipengaruhi oleh skor di ekor distribusi daripada skor di pusat distribusi.

Biasanya hal ini dibandingkan dengan distribusi normal yang mempunyai koefisien kurtosis sama dengan 3, yang dinamakan mesokurtik. Distribusi simetrik yang mempunyai koefisien kurang dari 3 berpuncak tajam, dinamakan leptokurtik, sedangkan distribusi yang ditandai dengan puncak datar, yang mempunyai koefisien kurtosis lebih dari 3, dinamakan platikurtik. Kurtosis adalah derajat dari kurva frekuensi distribusi data

$$Kurtosis (K) = \frac{1}{T\sigma^4} \sum_{t=1}^T (T_t - \mu)^4 \dots\dots\dots(3.1)$$

3.5.2 Skewness

Skewness adalah kemiringan kurva frekuensi suatu data distribusi (Ankarali, Yazici, & Ankarali, 2009). Ada dua jenis kecondongan positif dan negatif. Kemiringan positif, kemiringan lereng kurva frekuensi suatu distribusi data cenderung banyak di sebelah kiri median (nilai di bawah median). Jadi ekor kemiringan kurva frekuensi lebih menuju kanan. Demikian juga, dalam kemiringan negatif, ekor kemiringan kurva frekuensi lebih ke arah kiri. Persamaan 2 adalah rumus skewness, di mana r_t berada nilai dari setiap data yang diamati, σ adalah standar deviasi dari data sampel, μ adalah rata-rata dari data sampel, dan T adalah jumlah observasi.

$$Skewness (S) = \frac{1}{T\sigma^3} \sum_{t=1}^T (r_t - \mu)^3 \dots\dots\dots(3.2)$$

3.5.3 Mean

Mean adalah proses dimana frekuensi sinyal akan dipusatkan dengan menjumlahkan semua data dari kelompok sampel, kemudian dibagi dengan jumlah sampel data. Persamaan 3 adalah rumus skewness, di mana r_2 adalah nilainya dari setiap data yang diamati dan T adalah jumlah sampel pengamatan.

$$Mean (\mu) = \frac{1}{T} \sum_{t=1}^T r_2 \dots\dots\dots(3.3)$$

3.5.4 Standar Deviasi

Nilai tinggi σ menunjukkan bahwa nilai data menyebar dari nilai tengahnya μ . Jika σ rendah, maka nilainya didapat dikelompokkan pada nilai rata-rata. Persamaan 4 adalah rumus dari skewness, di mana r_2 adalah nilai dari setiap data

yang diamati, μ adalah rata-rata dari data sampel, dan T adalah jumlah observasi sampel

$$\text{Variansi } (\sigma^2) = \frac{1}{T} \sum_{t=1}^T (r_t - \mu)^2 \dots\dots\dots(3.4)$$

3.6 Klasifikasi Data

Pada proses klasifikasi dibagi menjadi dua fase: learning & test. Pada fase learning, sebagian data yang telah diketahui kelas datanya diumpangkan untuk membentuk model perkiraan. Kemudian pada fase test model yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tersebut. Bila akurasinya mencukupi model ini dapat dipakai untuk prediksi kelas data yang belum diketahui. Pada penelitian ini terdapat 4 skenario untuk klasifikasi data, yaitu:

- Skenario 1 (Metode *k-Nearest Neighbor*)
- Skenario 2 (Metode *Support Vector Machine*)
- Skenario 3 (Metode *Naïve Bayes*)
- Skenario 4 (Metode *Random Forest*)

3.6.1 K Nearest Neighbor (k-NN)

Dalam penelitian ini menggunakan metode k-NN untuk mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga (neighbor) terdekatnya dalam data *training*. Data training diproyeksikan ke ruang berdimensi banyak, yang mana masing-masing dimensi menjelaskan fitur dari data (Whidhiasih, Wahanani, & Supriyanto, 2013). Selain itu k-NN memiliki keunggulan yaitu lebih efektif untuk mengelola data yang besar serta dapat menghasilkan data yang lebih akurat.

3.6.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi (Santosa B. , 2007). Dalam penelitian ini juga menggunakan metode SVM karena dinilai memiliki prinsip dasar linier classifier atau kasus klasifikasi yang secara linier dapat dipisahkan, akan tetapi

SVM telah dikembangkan agar dapat bekerja pada problem non-linier dengan memasukkan konsep kernel pada ruang kerja berdimensi tinggi.

3.6.3 *Naïve Bayes*

Selain menggunakan metode k-NN dan SVM, pada penelitian ini juga menggunakan Naïve Bayes untuk memprediksi probabilitas isi dari suatu class. Bayesian classification didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan neural netrok dan decision tree. Bayesian classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar. Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan

3.6.4 *Random Forest*

Random Forest merupakan sebuah metode bisa yang digunakan untuk klasifikasi, regresi, ataupun tujuan lainnya. Random Forest bekerja dengan cara membangun lebih dari 1 Decision Tree secara random saat training. Hasil yang diberikan oleh Random Forest untuk klasifikasi adalah modus dari decision tree-decision tree nya. Dengan membuat banyak Decision Tree secara random, maka sebenarnya banyak dari pohon-pohon yang dibuat oleh metode Random Forest menjadi kurang berguna. Namun Random Forest mampu menjadi sebuah metode klasifikasi yang cukup baik, karena beberapa Decision Tree yang ikut dibuat saat konstruksi, ternyata memiliki kemampuan prediksi yang baik. Saat dilakukan pemilihan untuk menentukan klasifikasi secara keseluruhan, pohon-pohon yang buruk akan membuat prediksi yang acak dan saling bertentangan, sehingga jawaban dari beberapa decision tree yang merupakan prediktor yang baik akan muncul sebagai jawaban.

3.7 Penarikan Kesimpulan

Setelah dilakukan uji klasifikasi terhadap 5 variasi sampel data daging dengan 3 variasi suhu yang becrbeda serta dengan scenario yang ada maka dapat ditarik

kesimpulan berupa hasil akurasi tertinggi pada suhu-suhu tertentu dan metode terbaik yang dapat digunakan pada suhu tersebut.

3.8 Jadwal Penelitian

Berikut ini adalah jadwal rencana pengerjaan dan kegiatan penelitian mulai dari penyusunan proposal sampai sidang. Jadwal rencana dan pengerjaan kegiatan penelitian ditunjukkan pada Tabel 3.2.

Table 3.2 Jadwal Pelaksanaan Penelitian

Aktivitas	Waktu Pelaksanaan																			
	Februari				Maret				April				Mei				Juni			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Perumusan Masalah	■	■																		
Penyusunan Proposal Tesis		■	■	■	■															
Seminar proposal						■	■													
Klasifikasi Data							■	■	■	■	■	■								
Pengujian dan Evaluasi									■	■	■	■	■	■						
Penyusunan Buku Tesis													■	■	■	■	■	■	■	■

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

Pada bab ini dibahas mengenai pengujian klasifikasi pada masing-masing data yang telah didapat dari *electronic nose* serta evaluasi hasil

4.1 Pengujian 5 variasi komposisi daging terhadap suhu ruang

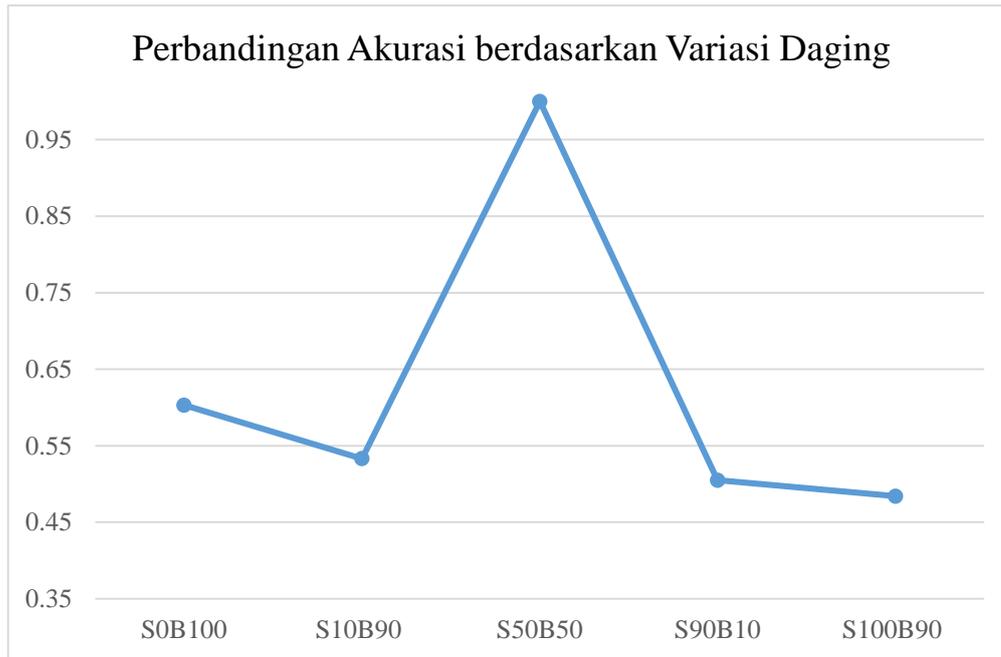
Untuk mencari komposisi daging terbaik, maka perlu dilakukan uji coba 5 variasi komposisi daging terhadap suhu ruang. Pengujian pada tabel 4.1 dilakukan dengan menggunakan metode naïve bayes k-fold cross-validation, dengan k=10.

Table 4.1 *Confussion Matrix* variasi komposisi daging

Variasi		PREDIKSI				
		Daging	S0B100	S10B90	S50B50	S90B10
AKTUAL	S0B100	9	28	0	8	5
	S10B90	11	21	0	10	7
	S50B50	2	0	48	0	0
	S90B10	7	33	0	2	16
	S100B0	7	36	0	6	7

Table 4.2 Detail Akurasi dari masing-masing variasi daging

Variasi Daging	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
S0B100	0.180	0.136	0.250	0.180	0.209	0.050	0.603	0.253
S10B90	0.429	0.485	0.178	0.429	0.251	-0.045	0.533	0.197
S50B50	0.960	0.000	1.000	0.960	0.980	0.975	1.000	0.999
S90B10	0.040	0.121	0.077	0.040	0.053	-0.106	0.505	0.189
S100B0	0.020	0.101	0.048	0.020	0.028	-0.116	0.484	0.184
AVG	0.325	0.167	0.311	0.325	0.304	0.153	0.625	0.365



Gambar 4.1 Perbandingan Akurasi Berdasarkan Variasi Daging

Pada tabel 4.2 dan gambar 4.1 menunjukkan nilai ROC tertinggi yaitu pada variasi daging S50B50 dengan predikat *excellent classification*, sedangkan nilai ROC tertinggi kedua yaitu variasi daging S0B100 dengan predikat *poor classification*. Pada variasi komposisi daging S10B90, S90B10, dan S100B0 memiliki nilai ROC dibawah 0,50 dimana komposisi daging tersebut memiliki predikat *failure*. Untuk mengatasi permasalahan tersebut, maka peneliti menambahkan variable suhu untuk meningkatkan akurasi pada variasi komposisi daging yang ada .

4.2 Skenario Pengujian

Skenario pengujian yang dilakukan adalah uji klasifikasi kemurnian daging sapi dengan menggunakan metode kNN, *Support Vector Machine*, *naïve bayes*, dan *Random Forest*. Uji klasifikasi dilakukan dengan menggunakan *software* WEKA 3.8. Sampel dibagi menjadi 5 kelas dengan 3 variasi suhu.

Table 4.3 Skenario Pengujian

Variasi Daging	Daging Sapi	Daging Babi	Suhu		
			-22°C	Suhu Ruang	55°C
S0B100	0%	100%	✓	✓	✓
S10B90	10%	90%	✓	✓	✓
S50B50	50%	50%	✓	✓	✓
S90B10	90%	10%	✓	✓	✓
S100B90	100%	90%	✓	✓	✓

Selanjutnya divisualisasikan table confusion matrix dan dilakukan proses perhitungan nilai *TP Rate*, *FP Rate*, *Precision*, *Recall*, *F-Measure*, *MCC*, dan *ROC Area*.

4.3 Pengujian Skenario 1 (Metode k-NN)

Pada skenario pengujian ini dilakukan uji coba klasifikasi daging dengan menggunakan metode k-Nearest Neighbor dengan 5 macam variasi daging dengan 3 variasi suhu berbeda . Pada pengujian scenario 1 ini dilakukan dengan memecah data fitur ekstraksi menjadi *data training* dan *data testing* dengan proporsi 70%:30%, dan k=3.

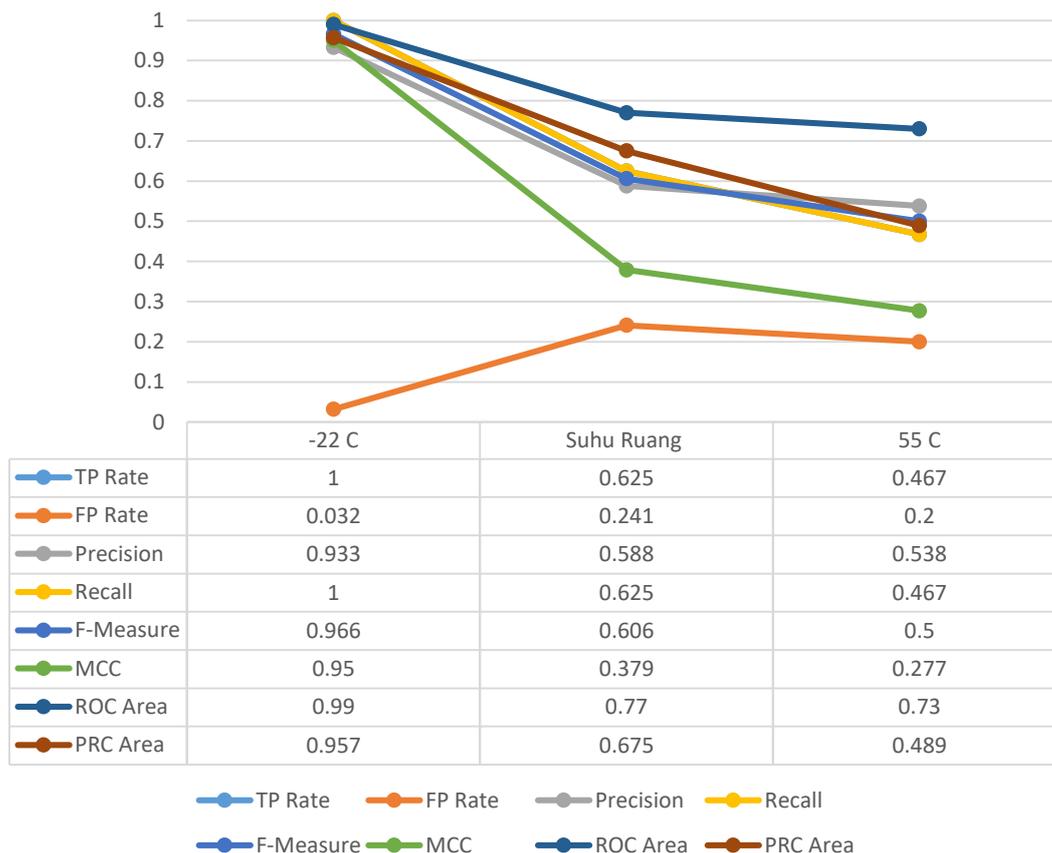
4.3.1 Pengujian S0B100

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 0% dan daging babi 100% . Tabel 4.4 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S0B100. Jumlah prediksi benar S0B100 pada suhu -22°C adalah 14, sedangkan jumlah prediksi benar S0B100 pada suhu ruang adalah 10 dan jumlah prediksi benar S0B100 pada suhu 55°C adalah 7.

Table 4.4 Pengaruh suhu terhadap *true positive* pada S0B100

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	14	17	1	0
Suhu Ruang	10	21	7	6
55°C	7	24	6	8

Pengaruh Suhu Terhadap Akurasi S0B100



Gambar 4.2 Pengaruh suhu terhadap akurasi S0B100

Pada Gambar 4.2 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada tabel 4.5 menunjukkan nilai yang sangat baik pada suhu -22 yaitu dengan nilai 1. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0.032. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0,990 dengan predikat *excellent classification*

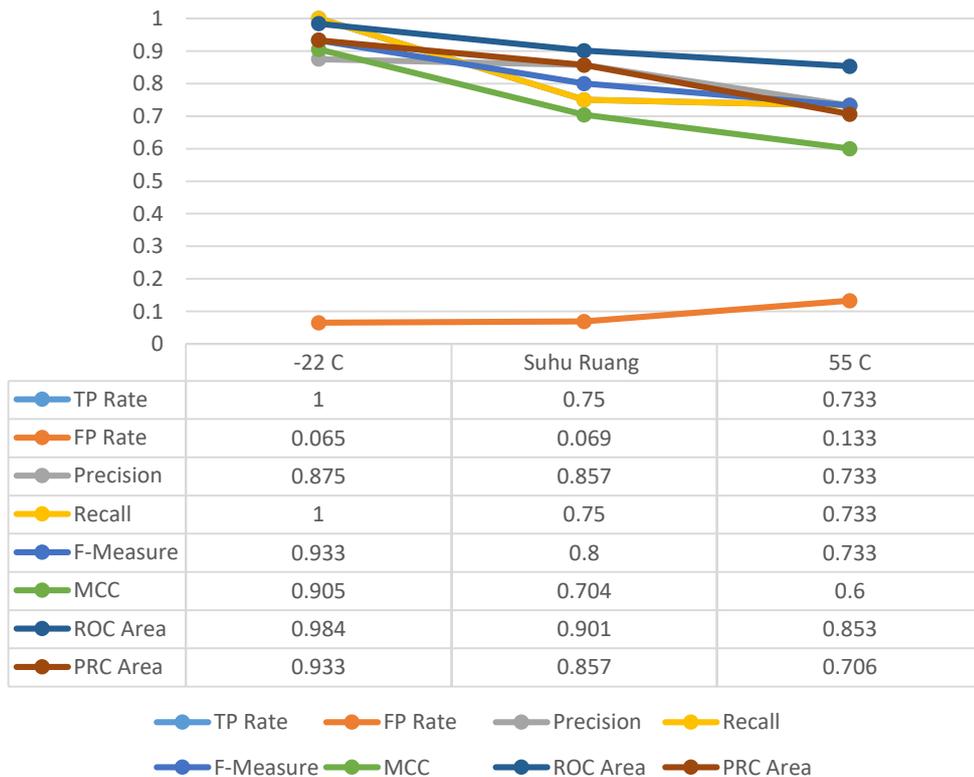
4.3.2 Pengujian S10B90

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 10% dan daging babi 90% . Tabel 4.5 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S10B90. Jumlah prediksi benar S10B90 pada suhu -22°C adalah 14, sedangkan jumlah prediksi benar S10B90 pada suhu ruang adalah 12 dan jumlah prediksi benar S10B90 pada suhu 55°C adalah 11.

Table 4.5 Pengaruh suhu terhadap *true positive* pada S10B90

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	14	23	2	0
Suhu Ruang	12	25	2	4
55°C	11	26	4	4

Pengaruh Suhu Terhadap Akurasi S10B90



Gambar 4.3 Pengaruh Suhu Terhadap Akurasi S10B90

Pada Gambar 4.3 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.3 menunjukkan nilai yang sangat baik pada suhu -22 yaitu dengan nilai 1. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0.065. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0,984 dengan predikat *excellent classification*.

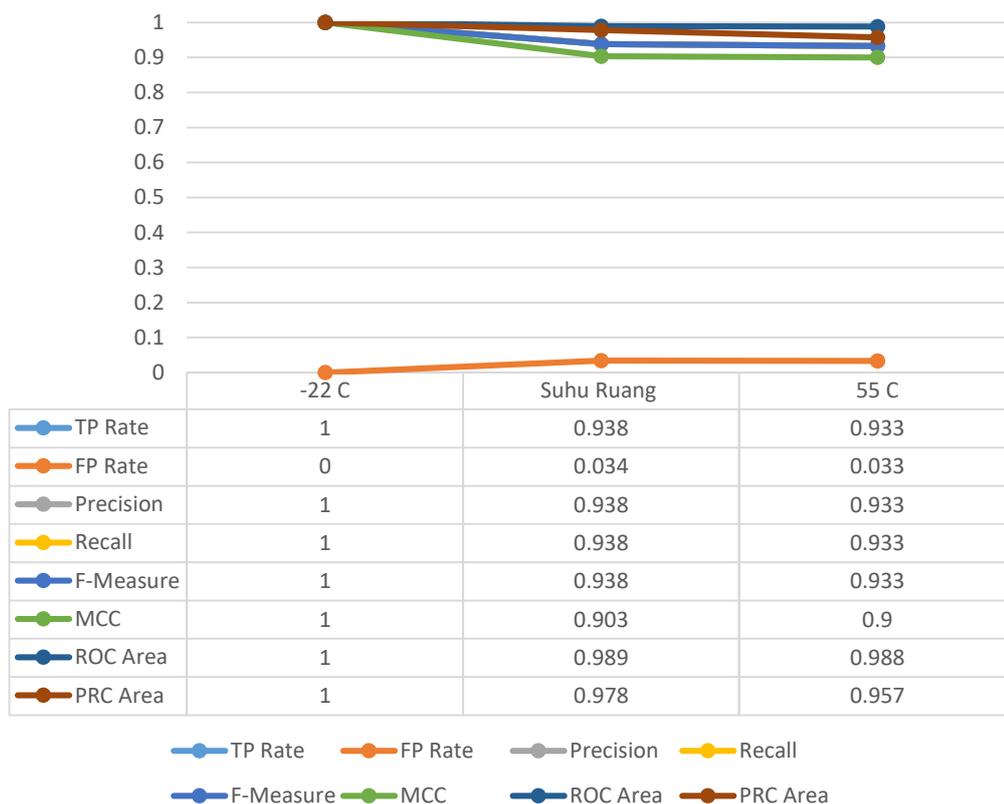
4.3.3 Pengujian S50B50

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 50% dan daging babi 50% . Tabel 4.6 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S50B50. Jumlah prediksi benar S50B50 pada suhu -22°C adalah 14, sedangkan jumlah prediksi benar S50B50 pada suhu ruang adalah 15 dan jumlah prediksi benar S50B50 pada suhu 55°C adalah 14.

Table 4.6 Pengaruh suhu terhadap *true positive* pada S50B50

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	14	29	0	0
Suhu Ruang	15	28	1	1
55°C	14	29	1	1

Pengaruh Suhu Terhadap Akurasi S50B50



Gambar 4.4 Pengaruh Suhu Terhadap Akurasi S50B50

Pada Gambar 4.4 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.4 menunjukkan nilai yang sangat baik pada suhu -22 yaitu dengan nilai 1. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 1 dengan predikat *excellent classification*.

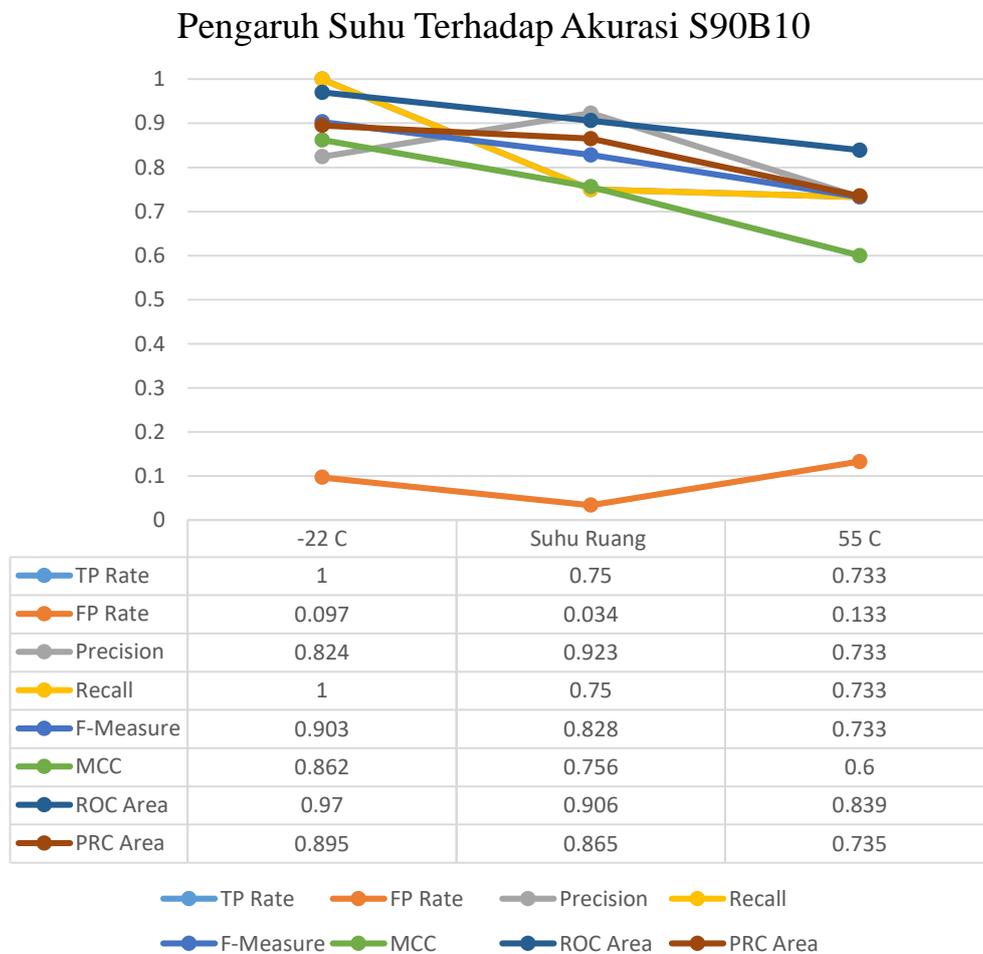
4.3.4 Pengujian S90B10

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 90% dan daging babi 10% . Tabel 4.7 merepresentasikan secara visual jumlah aktual dan

prediksi dari masing-masing kelas terhadap S90B10. Jumlah prediksi benar S90B10 pada suhu -22°C adalah 14, sedangkan jumlah prediksi benar S90B10 pada suhu ruang adalah 12 dan jumlah prediksi benar S90B10 pada suhu 55°C adalah 11.

Table 4.7 Pengaruh suhu terhadap true positive pada S90B10

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	14	23	3	0
Suhu Ruang	12	25	1	4
55°C	11	26	4	4



Gambar 4.5 Pengaruh Suhu Terhadap Akurasi S90B10

Pada Gambar 4.5 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.5 menunjukkan nilai yang sangat baik pada suhu -22 yaitu dengan nilai 1. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu ruang dengan nilai 0.034. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu ruang. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.97 dengan predikat *excellent classification*.

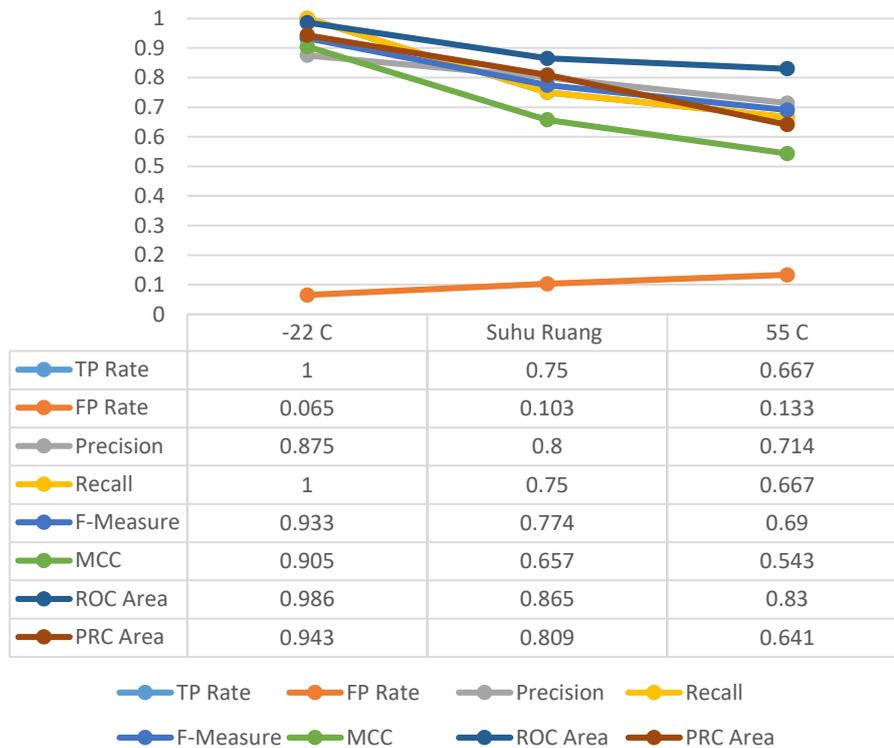
4.3.5 Pengujian S100B0

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 100% dan daging babi 0%. Tabel 4.8 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S100B0. Jumlah prediksi benar S100B0 pada suhu -22°C adalah 14, sedangkan jumlah prediksi benar S100B0 pada suhu ruang adalah 12 dan jumlah prediksi benar S100B0 pada suhu 55°C adalah 10.

Table 4.8 Pengaruh suhu terhadap true positive pada S100B0

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	14	22	2	0
Suhu Ruang	12	24	3	4
55°C	10	26	4	5

Pengaruh Suhu Terhadap Akurasi S100B0



Gambar 4.6 Pengaruh Suhu Terhadap Akurasi S100B0

Pada Gambar 4.6 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.6 menunjukkan nilai yang sangat baik pada suhu -22 yaitu dengan nilai 1. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0.065 . Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.986 dengan predikat *excellent classification*.

4.4 Pengujian Skenario 2 (*Metode Support Vector Machine*)

Pada skenario pengujian ini dilakukan uji coba klasifikasi daging dengan menggunakan metode *Support Vector Machine* dengan 5 macam variasi daging dengan 3 variasi suhu berbeda . Pada pengujian scenario 2 ini dilakukan dengan

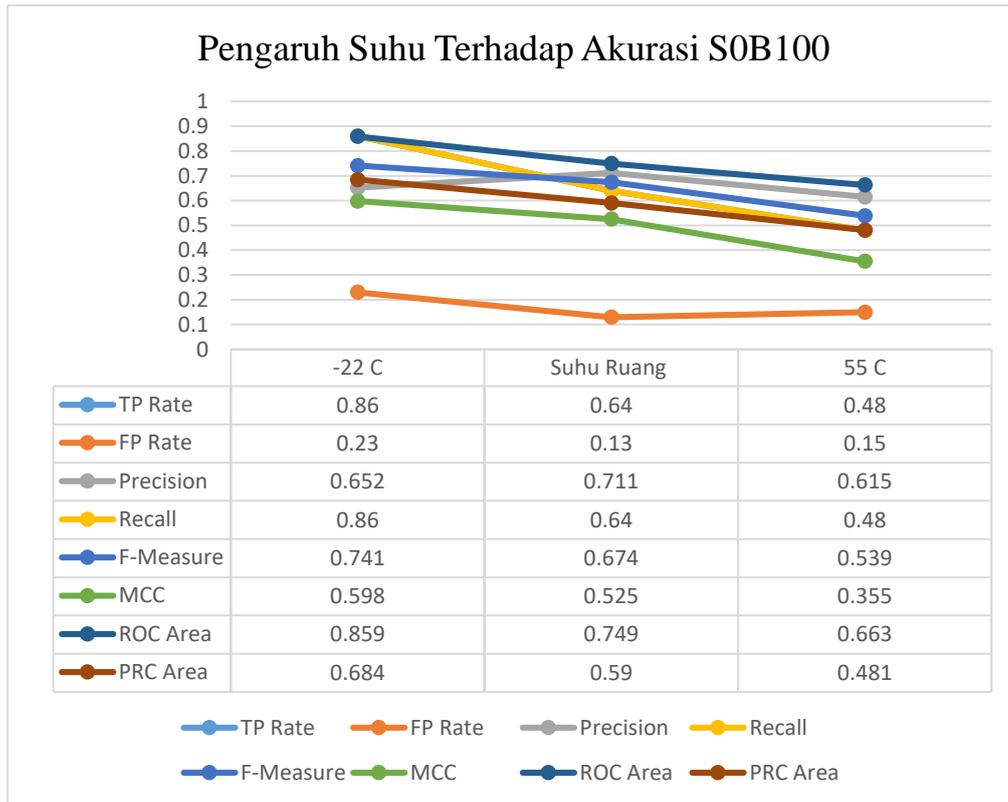
menggunakan *k-fold cross-validation*, dengan $k=10$ untuk kernel RBF. Tujuan dilakukan pengujian menggunakan *k-fold cross-validation* adalah untuk memilih parameter suhu terbaik berdasarkan akurasi paling tinggi sehingga dapat meningkatkan akurasi SVM dalam mengklasifikasi kemurnian daging sapi.

4.4.1 Pengujian S0B100

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 0% dan daging babi 100% . Tabel 4.9 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S0B100. Jumlah prediksi benar S0B100 pada suhu -22°C adalah 43, sedangkan jumlah prediksi benar S0B100 pada suhu ruang adalah 32 dan jumlah prediksi benar S0B100 pada suhu 55°C adalah 24.

Table 4.9 Pengaruh suhu terhadap true positive pada S0B100

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	43	56	23	7
Suhu Ruang	32	67	13	18
55°C	24	75	15	26



Gambar 4.7 Pengaruh Suhu Terhadap Akurasi S0B100

Pada Gambar 4.7 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.7 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.86. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu ruang dengan nilai 0.13 . Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu ruang. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.859 dengan predikat *good classification*.

4.4.1 Pengujian S10B90

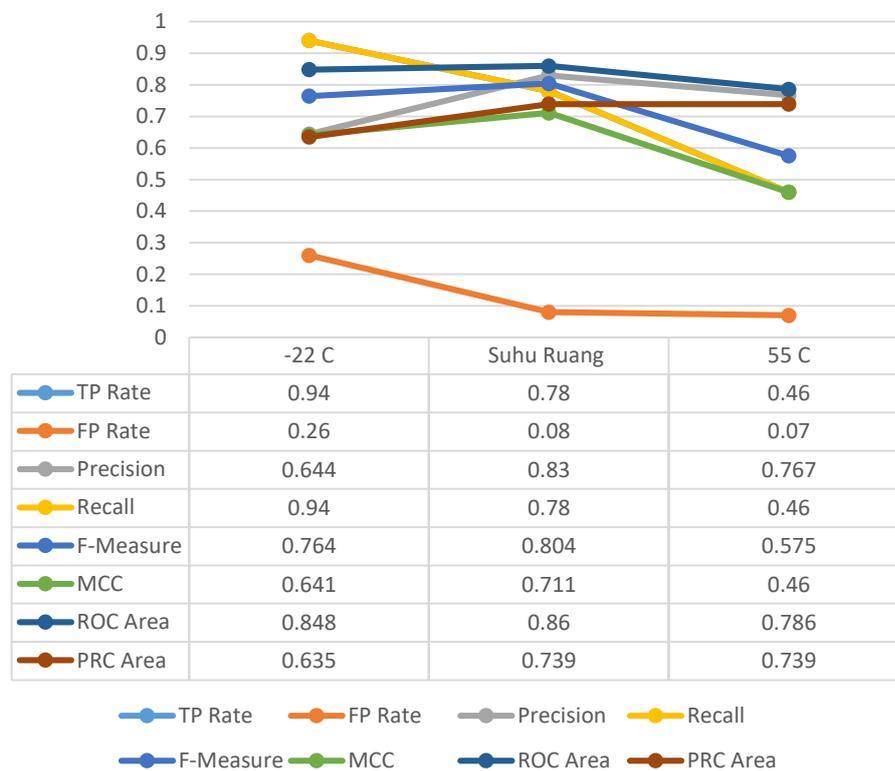
Pada subab ini, dilakukan pengujian dengan variasi daging sapi 10% dan daging babi 90% . Tabel 4.10 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S10B90. Jumlah prediksi benar S10B90 pada suhu -22°C adalah 47, sedangkan jumlah prediksi benar S10B90

pada suhu ruang adalah 39 dan jumlah prediksi benar S10B90 pada suhu 55°C adalah 23.

Table 4.10 Pengaruh suhu terhadap true positive pada S10B90

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	47	62	26	3
Suhu Ruang	39	70	8	11
55°C	23	86	7	27

Pengaruh Suhu Terhadap Akurasi S10B90



Gambar 4.8 Pengaruh Suhu Terhadap Akurasi S10B90

Pada Gambar 4.8 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.8 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.94. Berbeda dengan TP rate,

nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu 55°C dengan nilai 0.07. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu ruang. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu ruang sebesar 0.86 dengan predikat *good classification*.

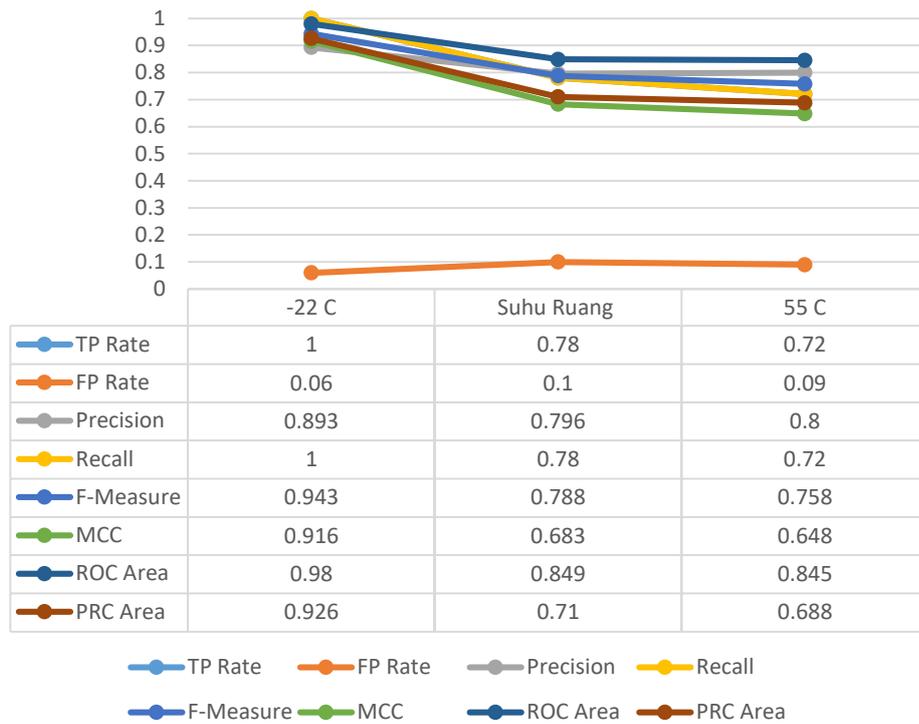
4.4.2 Pengujian S50B50

Pada subbab ini, dilakukan pengujian dengan variasi daging sapi 50% dan daging babi 50%. Tabel 4.11 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S50B50. Jumlah prediksi benar S50B50 pada suhu -22°C adalah 50, sedangkan jumlah prediksi benar S50B50 pada suhu ruang adalah 39 dan jumlah prediksi benar S50B50 pada suhu 55°C adalah 36.

Table 4.11 Pengaruh suhu terhadap true positive pada S50B50

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	50	75	6	0
Suhu Ruang	39	86	10	11
55°C	36	89	9	14

Pengaruh Suhu Terhadap Akurasi S50B50



Gambar 4.9 Pengaruh Suhu Terhadap Akurasi S50B50

Pada Gambar 4.9 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.9 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 1. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22 °C dengan nilai 0.06 . Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.98 dengan predikat *excellent classification*.

4.4.1 Pengujian S90B10

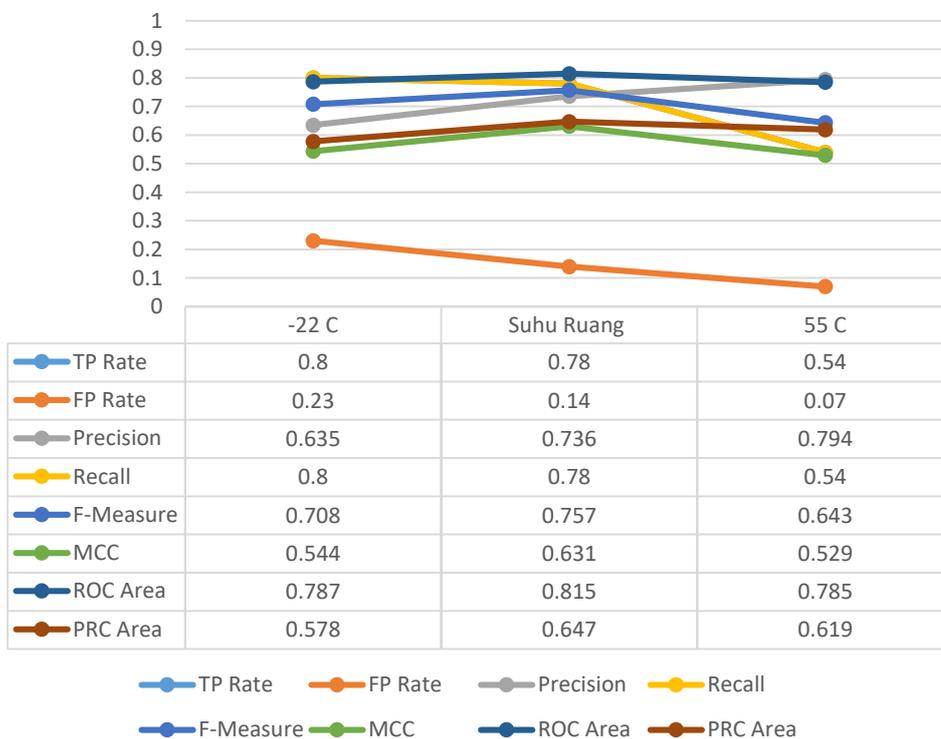
Pada subab ini, dilakukan pengujian dengan variasi daging sapi 90% dan daging babi 10% . Tabel 4.12 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S90B10. Jumlah prediksi benar

S90B10 pada suhu -22°C adalah 40, sedangkan jumlah prediksi benar S90B10 pada suhu ruang adalah 39 dan jumlah prediksi benar S90B10 pada suhu 55°C adalah 27.

Table 4.12 Pengaruh suhu terhadap true positive pada S90B10

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	40	66	23	10
Suhu Ruang	39	67	14	11
55°C	27	79	7	23

Pengaruh Suhu Terhadap Akurasi S90B10



Gambar 4.10 Pengaruh Suhu Terhadap Akurasi S90B10

Pada Gambar 4.10 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.10 menunjukkan nilai yang

sangat baik pada suhu -22°C yaitu dengan nilai 0.8. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu 55°C dengan nilai 0.07. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu 55°C . Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu ruang sebesar 0.815 dengan predikat *good classification*.

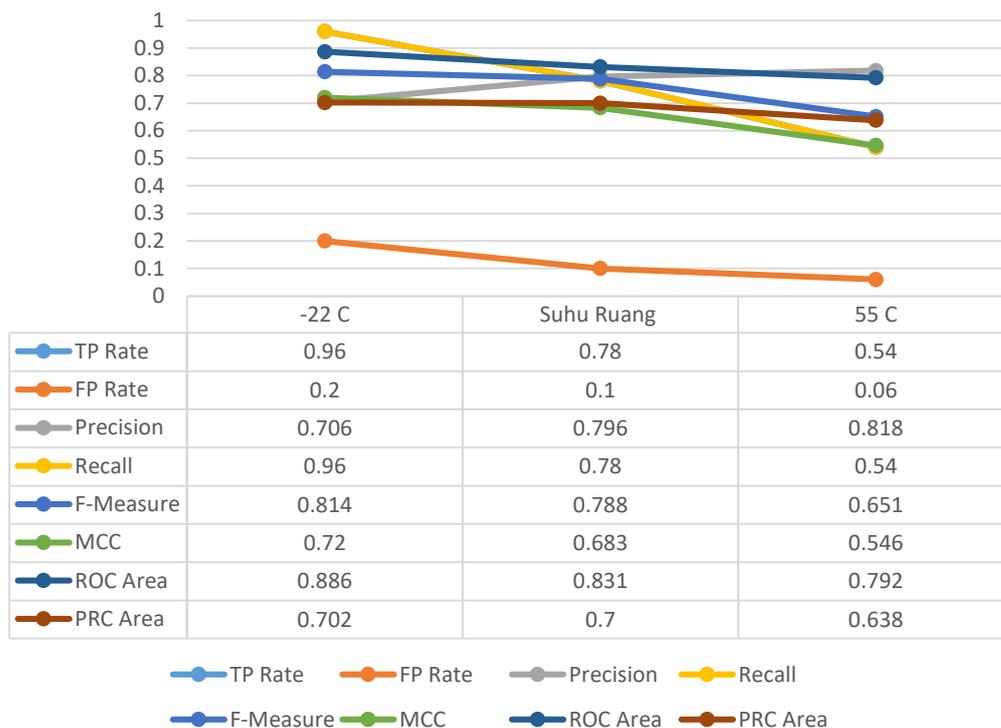
4.4.1 Pengujian S100B0

Pada subbab ini, dilakukan pengujian dengan variasi daging sapi 100% dan daging babi 0%. Tabel 4.13 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S100B0. Jumlah prediksi benar S100B0 pada suhu -22°C adalah 48, sedangkan jumlah prediksi benar S100B0 pada suhu ruang adalah 39 dan jumlah prediksi benar S100B0 pada suhu 55°C adalah 27.

Table 4.13 Pengaruh suhu terhadap true positive pada S100B0

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	48	66	20	2
Suhu Ruang	39	75	10	11
55°C	27	87	6	23

Pengaruh Suhu Terhadap Akurasi S100B0



Gambar 4.11 Pengaruh Suhu Terhadap Akurasi S100B0

Pada Gambar 4.11 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.11 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.96. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu 55°C dengan nilai 0.06. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu 55°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.886 dengan predikat *good classification*.

4.5 Pengujian Skenario 3 (Metode Naïve Bayes)

Pada skenario pengujian ini dilakukan uji coba klasifikasi daging dengan menggunakan metode Naïve Bayes dengan 5 macam variasi daging dengan 3 variasi suhu berbeda. Pada pengujian scenario 3 ini menggunakan *k-fold cross-validation*, dengan k=10.

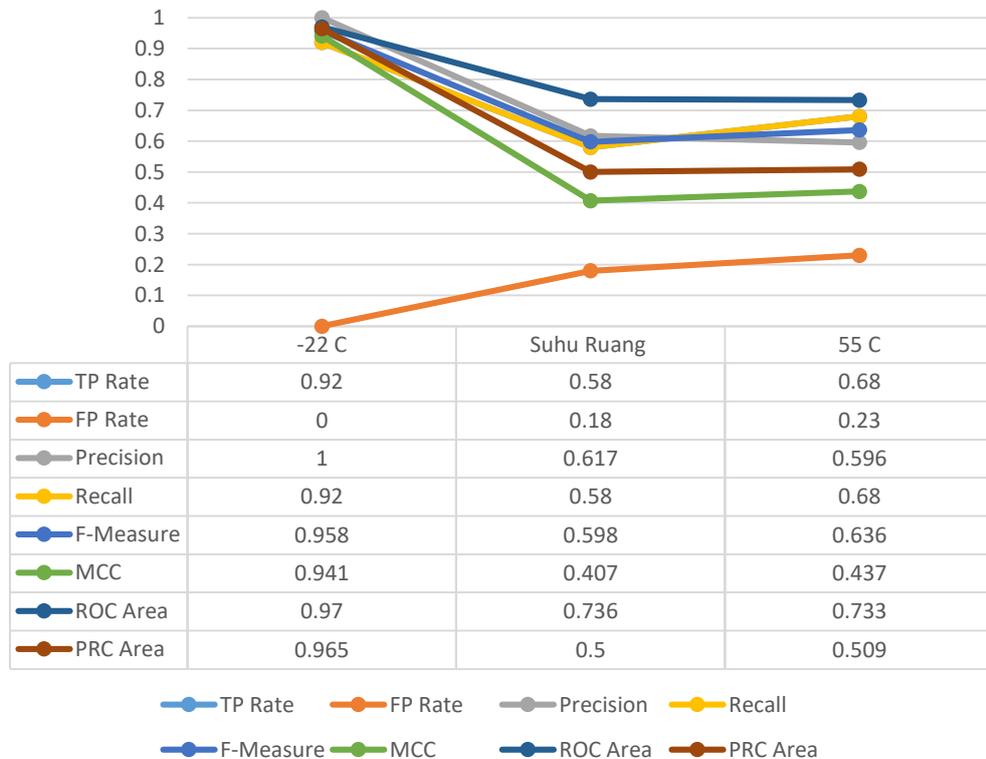
4.5.1 Pengujian S0B100

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 0% dan daging babi 100% . Tabel 4.14 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S0B100. Jumlah prediksi benar S0B100 pada suhu -22°C adalah 46, sedangkan jumlah prediksi benar S0B100 pada suhu ruang adalah 29 dan jumlah prediksi benar S0B100 pada suhu 55°C adalah 34.

Table 4.14 Pengaruh suhu terhadap *true positive* pada S0B100

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	46	63	0	4
Suhu Ruang	29	80	18	21
55°C	34	75	23	16

Pengaruh Suhu Terhadap Akurasi S0B100



Gambar 4.12 Pengaruh Suhu Terhadap Akurasi S0B100

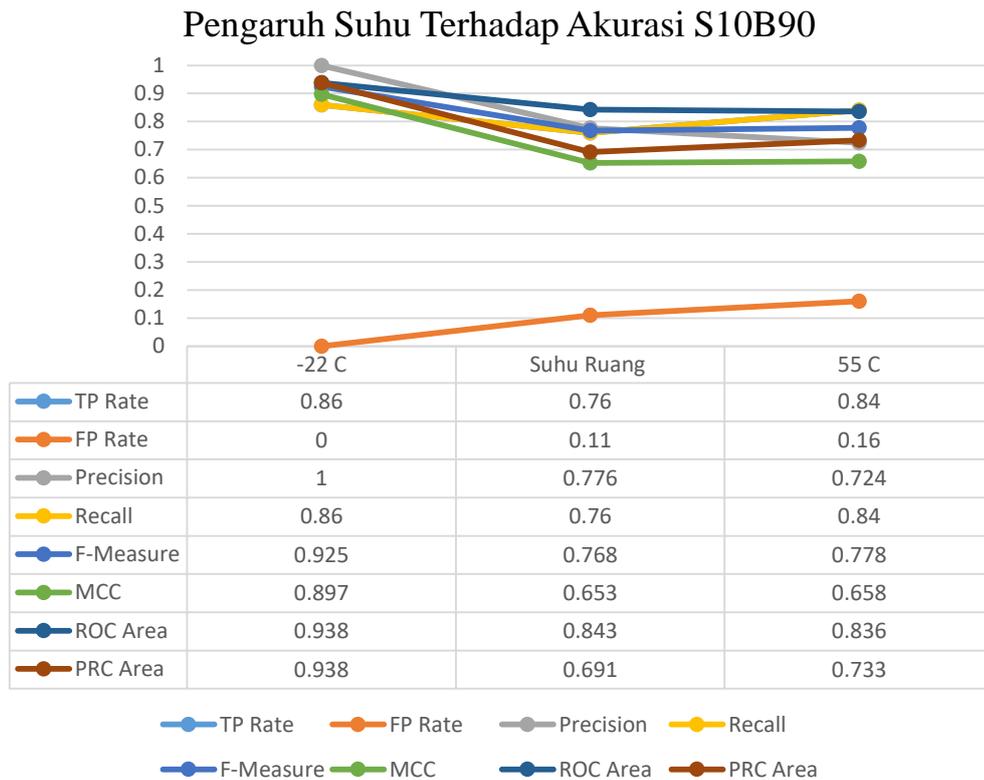
Pada Gambar 4.12 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.12 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.92. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.97 dengan predikat *excellent classification*.

4.5.2 Pengujian S10B90

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 10% dan daging babi 90% . Tabel 4.15 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S10B90. Jumlah prediksi benar S10B90 pada suhu -22°C adalah 43, sedangkan jumlah prediksi benar S10B90 pada suhu ruang adalah 38 dan jumlah prediksi benar S10B90 pada suhu 55°C adalah 42.

Table 4.15 Pengaruh suhu terhadap *true positive* pada S10B90

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	43	80	0	7
Suhu Ruang	38	85	11	12
55°C	42	81	16	8



Gambar 4.13 Pengaruh Suhu Terhadap Akurasi S10B90

Pada Gambar 4.13 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.13 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.86. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.938 dengan predikat *excellent classification*.

4.5.3 Pengujian S50B50

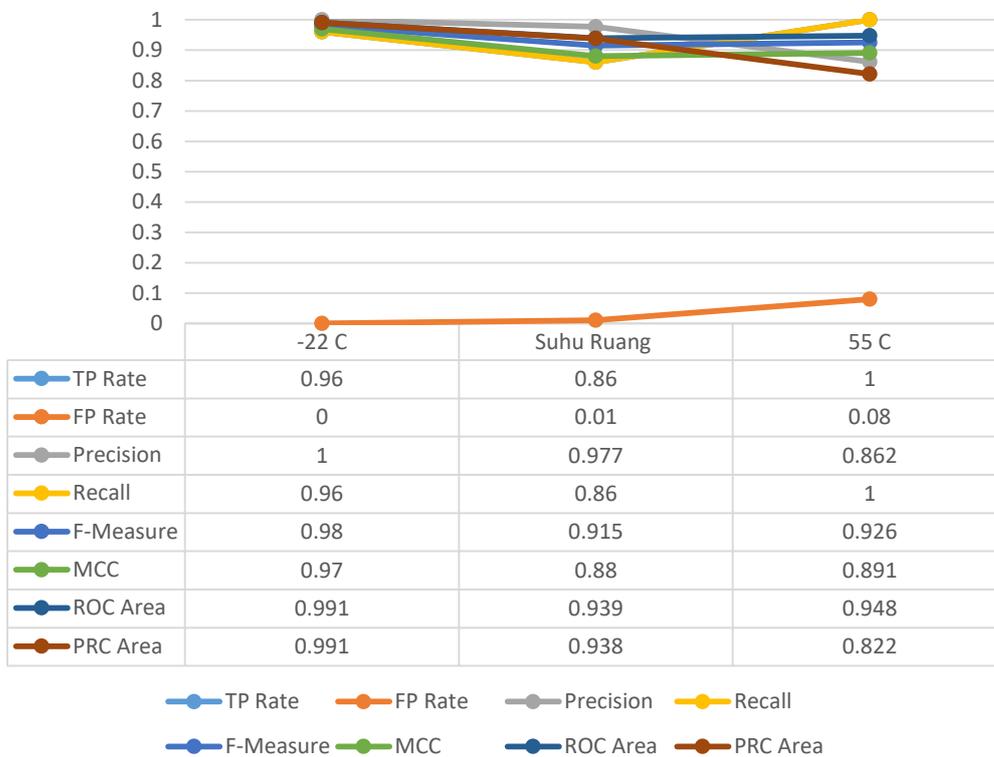
Pada subab ini, dilakukan pengujian dengan variasi daging sapi 50% dan daging babi 50% . Tabel 4.16 merepresentasikan secara visual jumlah aktual dan

prediksi dari masing-masing kelas terhadap S50B50. Jumlah prediksi benar S50B50 pada suhu -22°C adalah 48, sedangkan jumlah prediksi benar S50B50 pada suhu ruang adalah 43 dan jumlah prediksi benar S50B50 pada suhu 55°C adalah 50.

Table 4.16 Pengaruh suhu terhadap true positive pada S50B50

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	48	93	0	2
Suhu Ruang	43	98	1	7
55°C	50	91	8	0

Pengaruh Suhu Terhadap Akurasi S50B50



Gambar 4.14 Pengaruh Suhu Terhadap Akurasi S50B50

Pada Gambar 4.14 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode

klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.14 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.96. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C . Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.991 dengan predikat *excellent classification*.

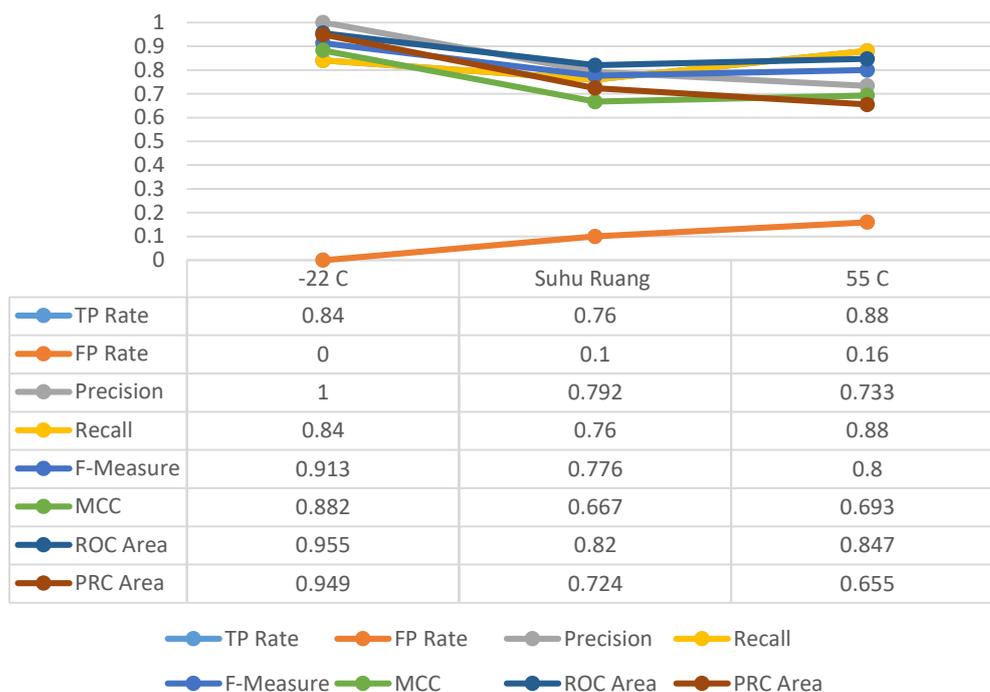
4.5.4 Pengujian S90B10

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 90% dan daging babi 10% . Tabel 4.17 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S90B10. Jumlah prediksi benar S90B10 pada suhu -22°C adalah 42, sedangkan jumlah prediksi benar S90B10 pada suhu ruang adalah 38 dan jumlah prediksi benar S90B10 pada suhu 55°C adalah 44.

Table 4.17 Pengaruh suhu terhadap true positive pada S90B10

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	42	82	0	8
Suhu Ruang	38	84	10	12
55°C	44	80	16	6

Pengaruh Suhu Terhadap Akurasi S90B10



Gambar 4.15 Pengaruh Suhu Terhadap Akurasi S90B10

Pada Gambar 4.15 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.14 menunjukkan nilai yang sangat baik pada suhu 55°C yaitu dengan nilai 0.88. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.955 dengan predikat *excellent classification*.

4.5.5 Pengujian S100B0

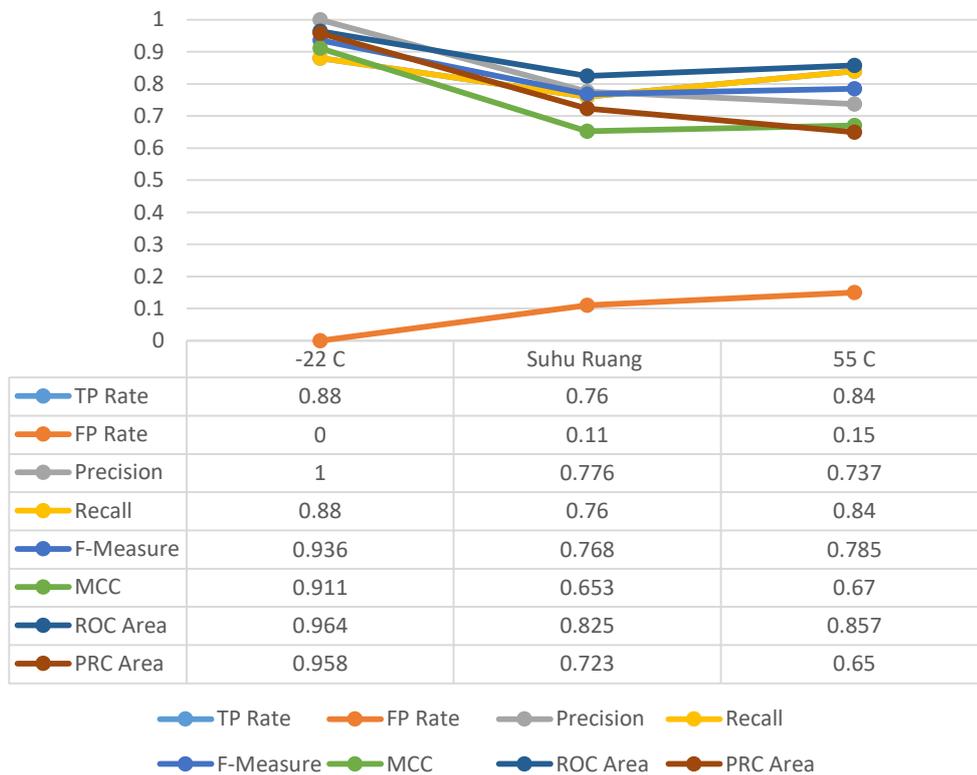
Pada subab ini, dilakukan pengujian dengan variasi daging sapi 100% dan daging babi 0% . Tabel 4.18 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S100B0. Jumlah prediksi benar S100B0 pada suhu -22°C adalah 44, sedangkan jumlah prediksi benar S100B0

pada suhu ruang adalah 38 dan jumlah prediksi benar S100B0 pada suhu 55°C adalah 42.

Table 4.18 Pengaruh suhu terhadap *true positive* pada S100B0

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	44	80	0	6
Suhu Ruang	38	86	11	12
55°C	42	82	15	8

Pengaruh Suhu Terhadap Akurasi S100B0



Gambar 4.16 Pengaruh Suhu Terhadap Akurasi S100B0

Pada Gambar 4.16 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.16 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.88. Berbeda dengan TP rate,

nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C . Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.964 dengan predikat *excellent classification*.

4.6 Pengujian Skenario 4 (Metode *Random Forest*)

Pada skenario pengujian ini dilakukan uji coba klasifikasi daging dengan menggunakan metode *random forest* dengan 5 macam variasi daging dengan 3 variasi suhu berbeda. Pada pengujian skenario ini menggunakan *k-fold cross-validation*, dengan $k=10$.

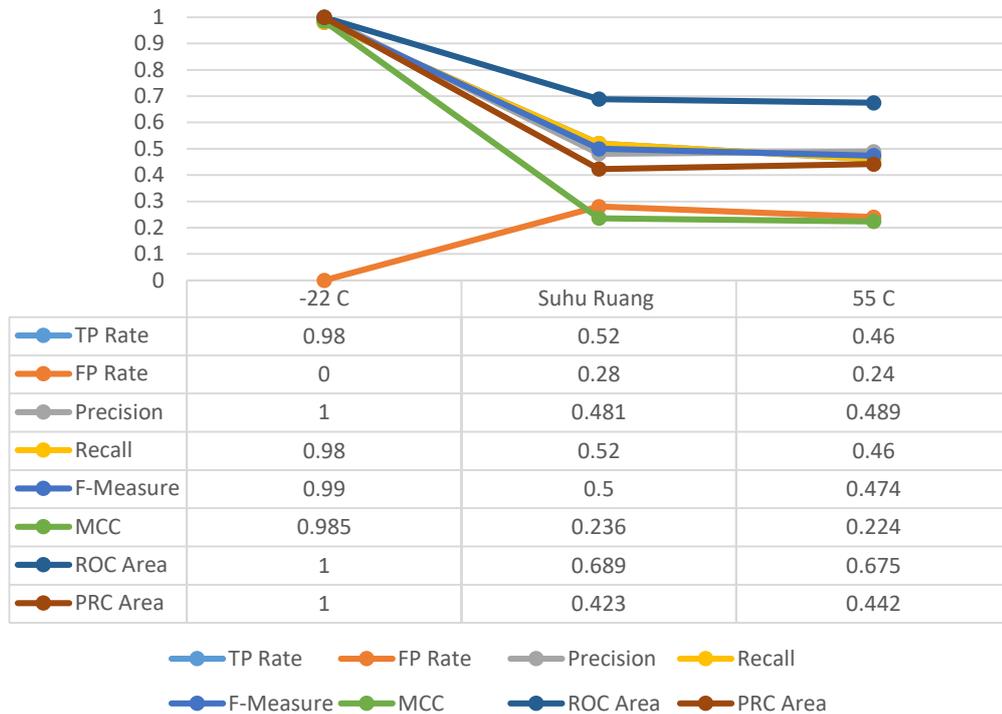
4.6.1 Pengujian S0B100

Pada subbab ini, dilakukan pengujian dengan variasi daging sapi 0% dan daging babi 100%. Tabel 4.19 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S0B100. Jumlah prediksi benar S0B100 pada suhu -22°C adalah 49, sedangkan jumlah prediksi benar S0B100 pada suhu ruang adalah 26 dan jumlah prediksi benar S0B100 pada suhu 55°C adalah 23.

Table 4.19 Pengaruh suhu terhadap *true positive* pada S100B0

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	49	49	0	1
Suhu Ruang	26	72	28	24
55°C	23	75	24	27

Pengaruh Suhu Terhadap Akurasi S0B100



Gambar 4.17 Pengaruh Suhu Terhadap Akurasi S0B100

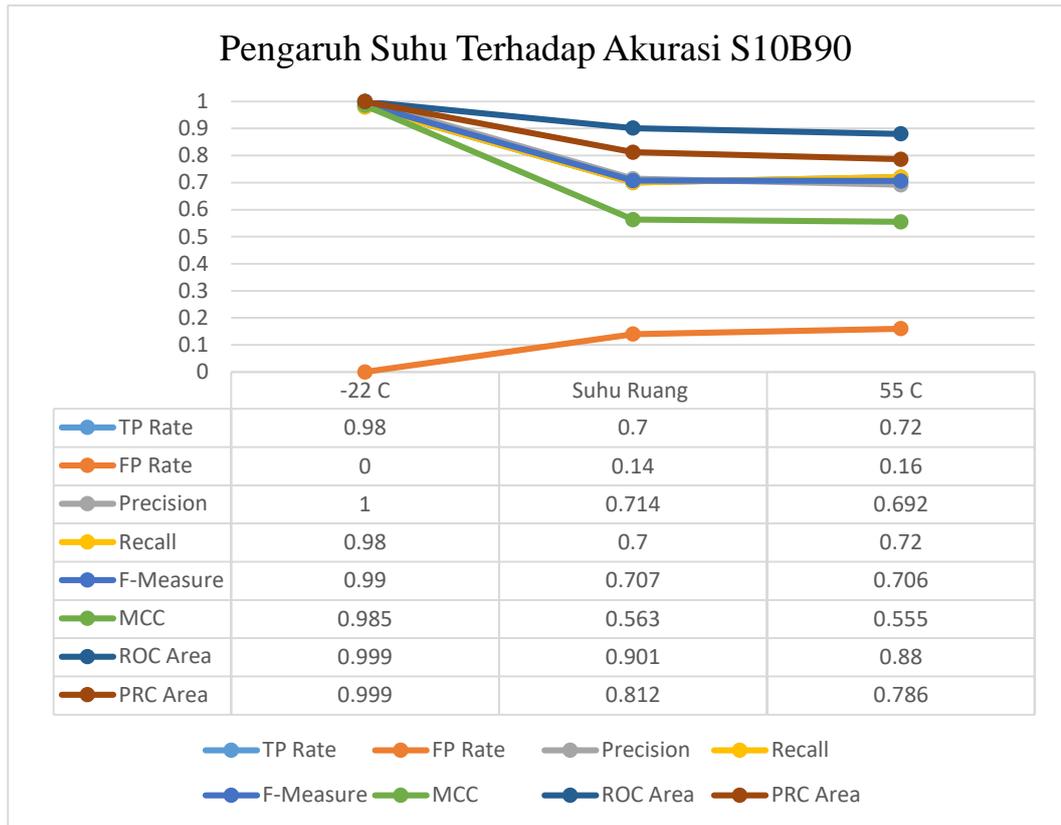
Pada Gambar 4.17 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.16 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.98. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C . Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 1 dengan predikat *excellent classification*.

4.6.2 Pengujian S10B90

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 10% dan daging babi 90% . Tabel 4.20 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S10B90. Jumlah prediksi benar S10B90 pada suhu -22°C adalah 49, sedangkan jumlah prediksi benar S10B90 pada suhu ruang adalah 35 dan jumlah prediksi benar S10B90 pada suhu 55°C adalah 36.

Table 4.20 Pengaruh suhu terhadap true positive pada S10B90

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	49	71	0	1
Suhu Ruang	35	85	14	15
55°C	36	84	16	14



Gambar 4.18 Pengaruh Suhu Terhadap Akurasi S10B90

Pada Gambar 4.18 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.18 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.98. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.999 dengan predikat *excellent classification*.

4.6.3 Pengujian S50B50

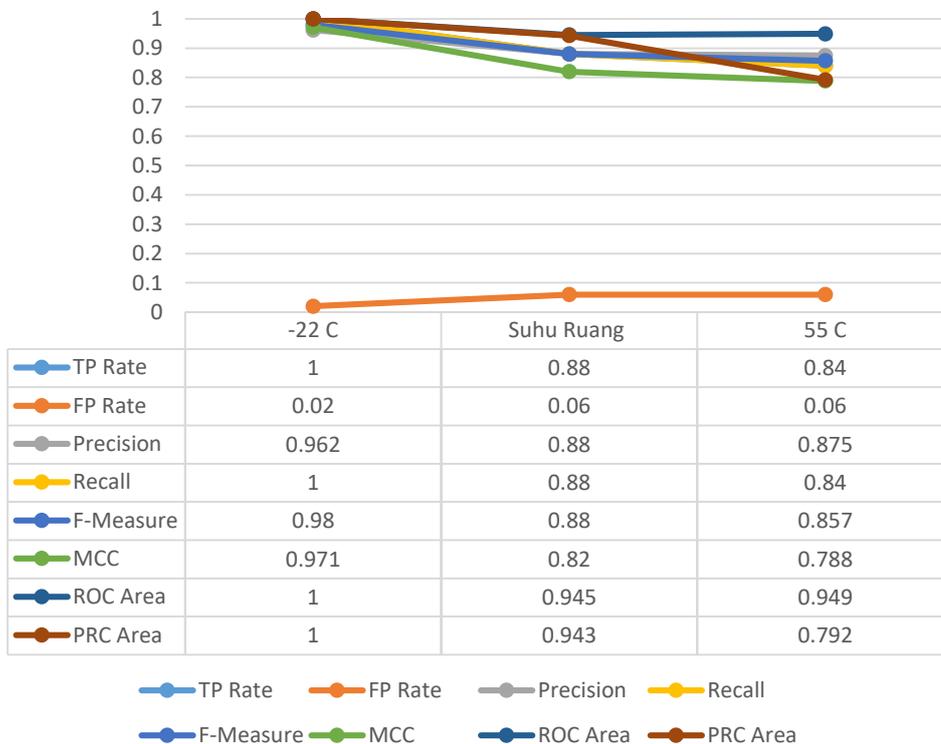
Pada subab ini, dilakukan pengujian dengan variasi daging sapi 50% dan daging babi 50% . Tabel 4.21 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S50B50. Jumlah prediksi benar

S50B50 pada suhu -22°C adalah 50, sedangkan jumlah prediksi benar S50B50 pada suhu ruang adalah 44 dan jumlah prediksi benar S50B50 suhu 55°C adalah 42.

Table 4.21 Pengaruh suhu terhadap *true positive* pada S50B50

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	50	86	0	0
Suhu Ruang	44	92	6	6
55°C	42	94	6	6

Pengaruh Suhu Terhadap Akurasi S50B50



Gambar 4.19 Pengaruh Suhu Terhadap Akurasi S50B50

Pada Gambar 4.19 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC

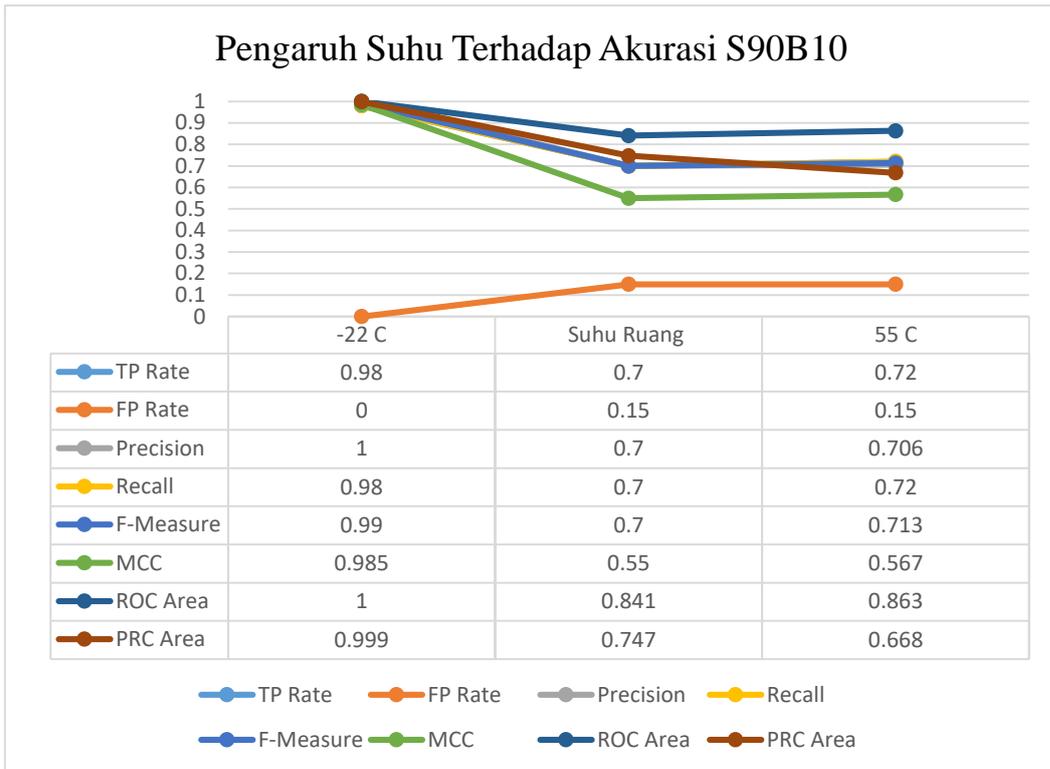
Area dan PRC Area. TP rate atau Recall pada gambar 4.19 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 1. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0.02. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C . Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.971 dengan predikat *excellent classification*.

4.6.4 Pengujian S90B10

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 90% dan daging babi 10% . Tabel 4.22 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S90B10. Jumlah prediksi benar S90B10 pada suhu -22°C adalah 49, sedangkan jumlah prediksi benar S90B10 pada suhu ruang adalah 35 dan jumlah prediksi benar S90B10 suhu 55°C adalah 36.

Table 4.22 Pengaruh suhu terhadap *true positive* pada S90B10

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	49	71	0	1
Suhu Ruang	35	85	15	15
55°C	36	84	15	14



Gambar 4.20 Pengaruh Suhu Terhadap Akurasi S90B10

Pada Gambar 4.20 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.20 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 1. Berbeda dengan TP rate, nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0.02. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 0.971 dengan predikat *excellent classification*.

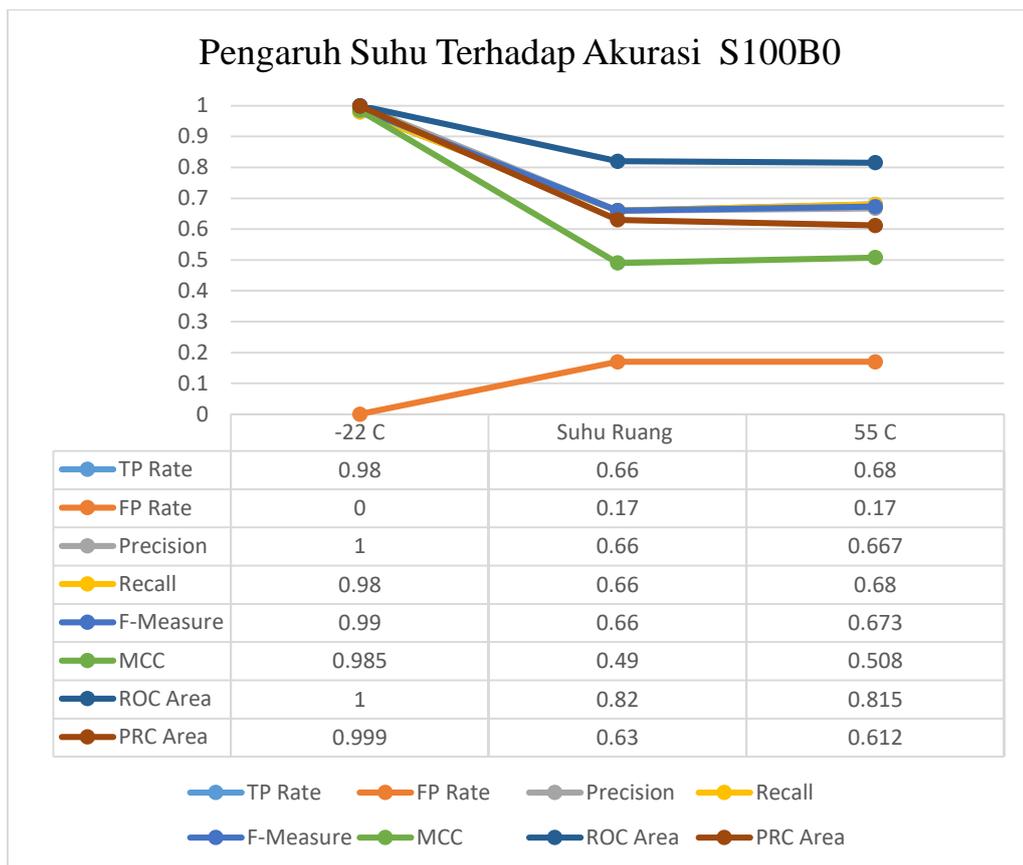
4.6.5 Pengujian S100B0

Pada subab ini, dilakukan pengujian dengan variasi daging sapi 100% dan daging babi 0% . Tabel 4.23 merepresentasikan secara visual jumlah aktual dan prediksi dari masing-masing kelas terhadap S100B0. Jumlah prediksi benar S100B0 pada suhu -22°C adalah 49, sedangkan jumlah prediksi benar S100B0

pada suhu ruang adalah 33 dan jumlah prediksi benar S100B0 suhu 55°C adalah 34.

Table 4.23 Pengaruh suhu terhadap *true positive* pada S100B0

Suhu	True Positive	True Negative	False Positive	False Negative
-22°C	49	67	0	2
Suhu Ruang	33	83	17	17
55°C	34	82	17	16



Gambar 4.21 Pengaruh Suhu Terhadap Akurasi S100B0

Pada Gambar 4.21 merepresentasikan Pengaruh Suhu Terhadap Akurasi dari masing-masing *class*. Terdapat 8 variabel evaluasi performansi metode klasifikasi yaitu TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area dan PRC Area. TP rate atau Recall pada gambar 4.21 menunjukkan nilai yang sangat baik pada suhu -22°C yaitu dengan nilai 0.98. Berbeda dengan TP rate,

nilai FP rate semakin baik bila mendekati nilai 0. Dalam hal ini FP rate yang paling baik adalah pada suhu -22°C dengan nilai 0. Sama halnya dengan TP Rate, nilai precision yang baik bila mendekati nilai 1 yaitu pada suhu -22°C. Sedangkan nilai ROC Area yang paling unggul yaitu pada suhu -22°C sebesar 1 dengan predikat *excellent classification*.

4.7 Rekapitulasi Nilai ROC Area dari masing-masing variasi daging

Tabel 4.44 merupakan tabel rekapitulasi nilai ROC Area dari semua variasi daging berdasarkan metode. Sedangkan tabel 4.45 merupakan tabel rekapitulasi nilai ROC dengan urutan nilai ROC tertinggi.

Table 4.24 Rekapitulasi nilai ROC Area

Metode	Variasi Daging	Suhu	Nilai ROC Area
k-NN	S0B100	-22 °C	0.99
	S0B100	Suhu Ruang	0.77
	S0B100	55 °C	0.73
	S10B90	-22 °C	0.984
	S10B90	Suhu Ruang	0.901
	S10B90	55 °C	0.853
	S50B50	-22 °C	1.000
	S50B50	Suhu Ruang	0.989
	S50B50	55 °C	0.988
	S90B10	-22 °C	0.97
	S90B10	Suhu Ruang	0.906
	S90B10	55 °C	0.839
	S100B0	-22 °C	0.986
	S100B0	Suhu Ruang	0.865
	S100B0	55 °C	0.83
SVM	S0B100	-22 °C	0.859
	S0B100	Suhu Ruang	0.749
	S0B100	55 °C	0.663
	S10B90	-22 °C	0.848
	S10B90	Suhu Ruang	0.86
	S10B90	55 °C	0.786
	S50B50	-22 °C	0.98
	S50B50	Suhu Ruang	0.849
	S50B50	55 °C	0.845
	S90B10	-22 °C	0.787

	S90B10	Suhu Ruang	0.815
	S90B10	55 °C	0.785
	S100B0	-22 °C	0.886
	S100B0	Suhu Ruang	0.831
	S100B0	55 °C	0.792
Naïve bayes	S0B100	-22 °C	0.97
	S0B100	Suhu Ruang	0.736
	S0B100	55 °C	0.733
	S10B90	-22 °C	0.938
	S10B90	Suhu Ruang	0.843
	S10B90	55 °C	0.836
	S50B50	-22 °C	0.991
	S50B50	Suhu Ruang	0.939
	S50B50	55 °C	0.948
	S90B10	-22 °C	0.991
	S90B10	Suhu Ruang	0.939
	S90B10	55 °C	0.948
	S100B0	-22 °C	0.964
	S100B0	Suhu Ruang	0.825
	S100B0	55 °C	0.857
	Random Forest	S0B100	-22 °C
S0B100		Suhu Ruang	0.689
S0B100		55 °C	0.675
S10B90		-22 °C	0.999
S10B90		Suhu Ruang	0.901
S10B90		55 °C	0.880
S50B50		-22 °C	1.000
S50B50		Suhu Ruang	0.945
S50B50		55 °C	0.949
S90B10		-22 °C	1.000
S90B10		Suhu Ruang	0.841
S90B10		55 °C	0.863
S100B0		-22 °C	1.000
S100B0		Suhu Ruang	0.820
S100B0		55 °C	0.815

Setelah mendapatkan nilai rekapitulasi ROC Area berdasarkan metode dan variasi daging, maka langkah selanjutnya adalah mengelompokan nilai rekapitulasi tersebut berdasarkan nilai ROC tertinggi. Hal tersebut bertujuan untuk melihat suhu yang memiliki akurasi tertinggi .

Table 4.25 Rekapitulasi nilai ROC Area *sort by* nilai ROC tertinggi

Metode	Variasi Daging	Suhu	Nilai ROC Area
k-NN	S50B50	-22 °C	1.000
<i>Random Forest</i>	S0B100	-22 °C	1.000
<i>Random Forest</i>	S50B50	-22 °C	1.000
<i>Random Forest</i>	S90B10	-22 °C	1.000
<i>Random Forest</i>	S100B0	-22 °C	1.000
<i>Random Forest</i>	S10B90	-22 °C	0.999
Naïve bayes	S50B50	-22 °C	0.991
Naïve bayes	S90B10	-22 °C	0.991
k-NN	S0B100	-22 °C	0.99
k-NN	S100B0	-22 °C	0.986
k-NN	S10B90	-22 °C	0.984
SVM	S50B50	-22 °C	0.98
k-NN	S90B10	-22 °C	0.97
Naïve bayes	S0B100	-22 °C	0.97
Naïve bayes	S100B0	-22 °C	0.964
Naïve bayes	S10B90	-22 °C	0.938
SVM	S100B0	-22 °C	0.886
SVM	S0B100	-22 °C	0.859
SVM	S10B90	-22 °C	0.848
SVM	S90B10	-22 °C	0.787
k-NN	S50B50	Suhu Ruang	0.989
k-NN	S50B50	55 °C	0.988
<i>Random Forest</i>	S50B50	55 °C	0.949
Naïve bayes	S50B50	55 °C	0.948
Naïve bayes	S90B10	55 °C	0.948
<i>Random Forest</i>	S50B50	Suhu Ruang	0.945
Naïve bayes	S50B50	Suhu Ruang	0.939
Naïve bayes	S90B10	Suhu Ruang	0.939
k-NN	S90B10	Suhu Ruang	0.906
k-NN	S10B90	Suhu Ruang	0.901
<i>Random Forest</i>	S10B90	Suhu Ruang	0.901

<i>Random Forest</i>	S10B90	55 °C	0.88
k-NN	S100B0	Suhu Ruang	0.865
<i>Random Forest</i>	S90B10	55 °C	0.863
SVM	S10B90	Suhu Ruang	0.86
Naïve bayes	S100B0	55 °C	0.857
k-NN	S10B90	55 °C	0.853
SVM	S50B50	Suhu Ruang	0.849
SVM	S50B50	55 °C	0.845
Naïve bayes	S10B90	Suhu Ruang	0.843
<i>Random Forest</i>	S90B10	Suhu Ruang	0.841
k-NN	S90B10	55 °C	0.839
Naïve bayes	S10B90	55 °C	0.836
SVM	S100B0	Suhu Ruang	0.831
k-NN	S100B0	55 °C	0.83
Naïve bayes	S100B0	Suhu Ruang	0.825
<i>Random Forest</i>	S100B0	Suhu Ruang	0.82
SVM	S90B10	Suhu Ruang	0.815
<i>Random Forest</i>	S100B0	55 °C	0.815
SVM	S100B0	55 °C	0.792
SVM	S10B90	55 °C	0.786
SVM	S90B10	55 °C	0.785
k-NN	S0B100	Suhu Ruang	0.77
SVM	S0B100	Suhu Ruang	0.749
Naïve bayes	S0B100	Suhu Ruang	0.736
Naïve bayes	S0B100	55 °C	0.733
k-NN	S0B100	55 °C	0.73
<i>Random Forest</i>	S0B100	Suhu Ruang	0.689
<i>Random Forest</i>	S0B100	55 °C	0.675
SVM	S0B100	55 °C	0.663

Tabel 4.43 membuktikan bahwasanya suhu berpengaruh terhadap nilai akurasi. Semakin rendah suhu maka semakin stabil pula nilai yang didapatkan oleh *electronic nose*. Hal ini dapat di buktikan dengan 20 peringkat teratas pada nilai ROC adalah pada suhu -22°C. Sedangkan pada suhu 55°C masih terdapat ketidakstabilan data yang disebabkan oleh asap yang keluar pada daging tidak konsisten.

4.8 Nilai ROC berdasarkan suhu dan metode

Setelah mendapatkan nilai rekap ROC Area dari semua variasi daging berdasarkan metode, langkah selanjutnya adalah mengelompokkan nilai ROC tersebut berdasarkan Suhu dan metodenya untuk mengetahui metode yang terbaik berdasarkan suhu.

Table 4.26 Nilai ROC berdasarkan suhu -22°C

Suhu	Metode	Variasi Daging	Nilai ROC
-22 °C	k-NN	S0B100	0.990
		S10B90	0.984
		S50B50	1.000
		S90B10	0.970
		S100B0	0.986
	Rata-rata Nilai ROC		0.986
	SVM	S0B100	0.859
		S10B90	0.848
		S50B50	0.980
		S90B10	0.787
		S100B0	0.886
	Rata-rata Nilai ROC		0.872
	Naïve bayes	S0B100	0.970
		S10B90	0.938
		S50B50	0.991
		S90B10	0.991
		S100B0	0.964
	Rata-rata Nilai ROC		0.971
	Random Forest	S0B100	1.000
		S10B90	0.999
		S50B50	1.000
S90B10		1.000	
S100B0		1.000	
Rata-rata Nilai ROC		1.000	

Tabel 4.46 menjelaskan bahwa pada suhu -22°C metode yang memiliki akurasi paling tinggi adalah metode *random forest* dengan nilai rata-rata ROC 1,000. Sedangkan metode yang memiliki akurasi tinggi ke 2 yaitu metode *k-Nearest Neighbor* dengan nilai rata-rata ROC 0,986 . Metode naïve bayes termasuk peringkat ke 3 dengan nilai rata-rata ROC 0.971 dan metode SVM merupakan metode yang memiliki akurasi paling rendah dengan nilai rata-rata ROC 0,872.

Table 4.27 Nilai ROC Berdasarkan Suhu Ruang

Suhu	Metode	Variasi Daging	Nilai ROC
Suhu Ruang	k-NN	S0B100	0.770
		S10B90	0.901
		S50B50	0.989
		S90B10	0.906
		S100B0	0.865
	Rata-rata Nilai ROC		0.886
	SVM	S0B100	0.749
		S10B90	0.86
		S50B50	0.849
		S90B10	0.815
		S100B0	0.831
	Rata-rata Nilai ROC		0.820
	Naïve bayes	S0B100	0.736
		S10B90	0.843
		S50B50	0.939
		S90B10	0.939
		S100B0	0.825
	Rata-rata Nilai ROC		0.856
	Random Forest	S0B100	0.689
		S10B90	0.901
		S50B50	0.945
S90B10		0.841	
S100B0		0.820	
Rata-rata Nilai ROC		0.839	

Tabel 4.47 menjelaskan bahwa pada suhu ruang metode yang memiliki akurasi paling tinggi adalah metode *k-Nearest Neighbor* dengan nilai rata-rata ROC 0,886. Sedangkan metode yang memiliki akurasi tinggi ke 2 yaitu metode Naïve bayes dengan nilai ROC rata-rata 0,856 . Metode *Random Forest* termasuk peringkat ke 3 dengan nilai rata-rata ROC 0.839 dan metode SVM merupakan metode yang memiliki akurasi paling rendah dengan nilai rata-rata ROC 0,820.

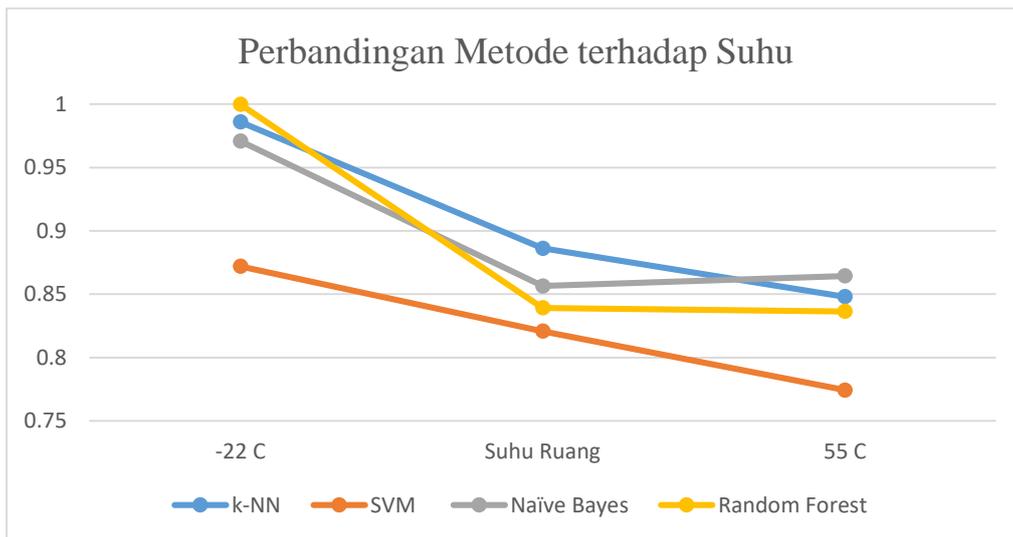
Table 4.28 Nilai ROC Berdasarkan Suhu 55°C

Suhu	Metode	Variasi Daging	Nilai ROC
55°C	k-NN	S0B100	0.730
		S10B90	0.853
		S50B50	0.988
		S90B10	0.839
		S100B0	0.830
	Rata-rata Nilai ROC		0.848
	SVM	S0B100	0.663
		S10B90	0.786
		S50B50	0.845
		S90B10	0.785
		S100B0	0.792
	Rata-rata Nilai ROC		0.774
	Naïve bayes	S0B100	0.733
		S10B90	0.836
		S50B50	0.948
		S90B10	0.948
		S100B0	0.857
	Rata-rata Nilai ROC		0.864
	Random Forest	S0B100	0.675
		S10B90	0.880
		S50B50	0.949
		S90B10	0.863
		S100B0	0.815
	Rata-rata Nilai ROC		0.836

Tabel 4.48 menjelaskan bahwa pada suhu ruang metode yang memiliki akurasi paling tinggi adalah metode Naïve Bayes dengan nilai rata-rata ROC 0,857. Sedangkan metode yang memiliki akurasi tinggi ke 2 yaitu metode *K-Nearest Neighbor* dengan nilai rata-rata ROC 0,848 . Metode *Random Forest* termasuk peringkat ke 3 dengan nilai rata-rata ROC 0.836 dan metode SVM merupakan metode yang memiliki akurasi paling rendah dengan nilai rata-rata ROC 0,774.

Table 4.29 Perbandingan metode terhadap suhu

	k-NN	SVM	Naïve Bayes	Random Forest
-22 C	0.986	0.872	0.971	1.000
Suhu Ruang	0.886	0.821	0.856	0.839
55 C	0.848	0.774	0.864	0.836

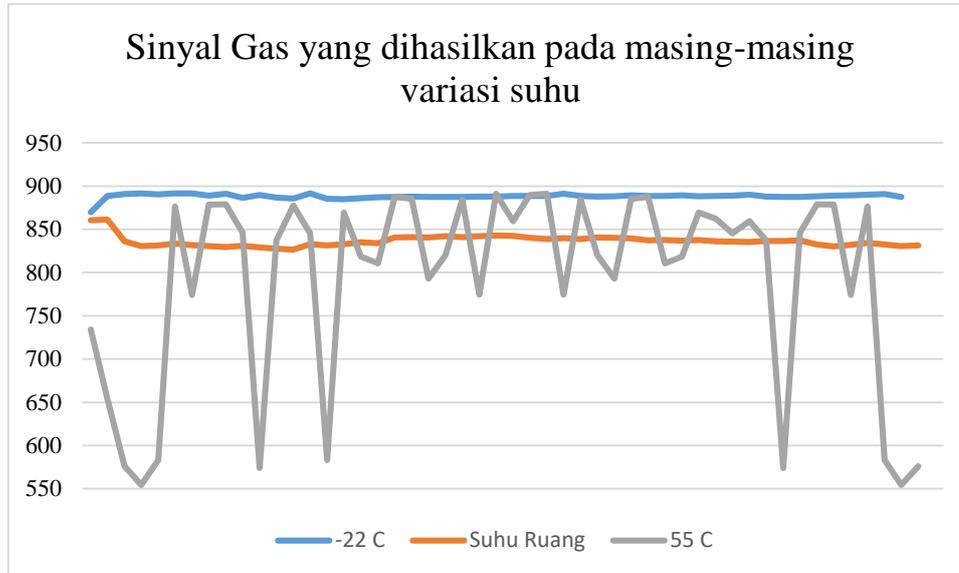


Gambar 4.22 Grafik Perbandingan Metode terhadap Suhu

Tabel 4.49 merupakan rangkuman nilai rata-rata ROC pada masing-masing metode terhadap suhu. Pada gambar 4.22 terlihat adanya kenaikan akurasi pada suhu -22°C dari masing-masing metode. Pada suhu -22°C metode yang memiliki akurasi tinggi yaitu metode *Random Forest* dan metode yang memiliki akurasi paling rendah yaitu *Support Vector Machine*. Pada suhu ruang metode yang memiliki akurasi tinggi yaitu metode *K-Nearest Neighbor* dan metode yang memiliki akurasi paling rendah yaitu *Support Vector Machine*. Pada suhu 55°C metode yang memiliki akurasi tinggi yaitu metode *Naïve Bayes* dan metode yang memiliki akurasi paling rendah yaitu *Support Vector Machine*.

Berdasarkan gambar 4.22, Nilai ROC tertinggi adalah pada suhu -22°C , hal tersebut dikarenakan pada suhu -22°C daging menghasilkan gas yang lebih stabil.

Sedangkan pada suhu 55°C sinyal gas yang dihasilkan tidak stabil, karena asap yang keluar pada daging tidak konsisten pada saat proses pemanasan



Gambar 4.23 Sinyal Gas yang dihasilkan pada masing-masing variasi suhu

Pada gambar 4.23 terlihat bahwasnya pada suhu 55°C gas yang dihasilkan oleh daging tidak stabil dimana hal tersebut dapat menyebabkan turunnya akurasi pada proses klasifikasi. Sedangkan pada suhu -22°C dan suhu ruang mengeluarkan gas yang cukup konstan. Maka dari itu suhu memiliki pengaruh terhadap proses klasifikasi. Semakin rendah suhu, maka semakin konstan pula gas yang dikeluarkan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan oleh penulis diperoleh kesimpulan sebagai berikut :

1. Peneliti membagi percobaan menjadi 4 skenario dengan masing-masing komposisi 5 variasi daging (Sapi0% Babi100%, Sapi10% Babi90%, Sapi50% Babi50%, Sapi90% Babi10% dan Sapi100 Babi0%) dengan 3 variasi suhu (-22°C, Suhu Ruang, dan 55°C) yaitu :
 - a. Metode *k-Nearest Neighbor*
 - b. Metode *Support Vector Machine*
 - c. Metode *Naïve Bayer*
 - d. Metode *Random Forest*
2. Terdapat pengaruh suhu terhadap peningkatan akurasi, yaitu pada suhu -22°C. Karena semakin rendah suhu maka semakin stabil pula nilai yang didapatkan oleh *electronic nose*.
3. Berikut merupakan metode yang memiliki akurasi tinggi berdasarkan suhu:
 - a) Pada suhu -22oC, urutan metode yang memiliki akurasi paling tinggi yaitu Random Forest dengan nilai rata-rata ROC 1,00 ; K-Nearest Neighbor dengan nilai rata-rata ROC 0,986 ; Naïve Bayes dengan nilai rata-rata ROC 0,971 dan Support Vector Machine dengan nilai rata-rata ROC 0,872.
 - b) Pada suhu ruang, urutan metode yang memiliki akurasi paling tinggi yaitu K-Nearest Neighbor dengan nilai rata-rata ROC 0,886; Naïve Bayes dengan nilai rata-rata ROC 0,856; Random Forest dengan nilai rata-rata ROC 0,839 dan Support Vector Machine dengan nilai rata-rata ROC 0,821.
 - c) Pada suhu 55°C, urutan metode yang memiliki akurasi paling tinggi yaitu Naïve Bayes dengan nilai rata-rata ROC 0,864 ; K-Nearest Neighbor

dengan nilai rata-rata ROC 0,848; Random Forest dengan nilai rata-rata ROC 0,836 dan Support Vector Machine dengan nilai rata-rata ROC 0,774.

5.2 Saran

Dari hasil penelitian ini, dapat diberikan saran yang lebih banyak tertuju pada pengembangan untuk penelitian selanjutnya demi kesempurnaan dan manfaat penelitian, yakni :

1. Kajian lebih lanjut untuk pemilihan jenis-jenis daging sapi dan jenis-jenis daging babi yang lebih bervariasi
2. Kajian lebih lanjut untuk makanan olah / makanan cepat saji yang menggunakan campuran daging sapi dengan daging babi .

DAFTAR PUSTAKA

- [FAO]. (2007, Maret 18). Retrieved from Food and Agriculture Organization:
http://www.fao.org/ag/againfo/themes/en/meat/backgr_composition.html
- Ankarali, H., Yazici, A. C., & Ankarali, S. (2009). A bootstrap confidence interval for skewness and kurtosis and properties of t-test in small samples from normal distribution. *Balkan Medical Journal*.
- Biau, G. (2012). Analysis of a Random Forests Model. *Journal of Machine Learning Research*, 1063-1095.
- Daiva, A. F. (2017). *Klasifikasi Kemurnian Daging Sapi Berbasis Electronic Nose*. Departemen Informatika : Fakultas Teknologi Informasi dan Komunikasi Institut Teknologi Sepuluh Nopember Surabaya.
- Erwanto, Y., Abidin, M. Z., Sugiyono, E. P., & Rohman, A. (2014). Identification of Pork Contamination in Meatballs of Indonesia Local Market Using Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) Analysis. *Asian-Australas J Anim Sci*.
doi:10.5713/ajas.2014.14014
- Gorunescu, F. (2011). *Intelligent Systems Reference Library* (Vol. 12). Berlin: Springer.
- Hanselman, D. L. (1998). *Mastering Matlab 5, A Comprehensive Tutorial and Reference*. Prentice-Hall Inc.
- Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). *A Practical Guide to Support Vector Classification*.
- Indonesia, B. (2010). *Sensus Penduduk*.
- Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. (2004). kernlab – An S4 Package for Kernel Methods in R. *Journal of Statistical Software*, 11(9), 1-20.

- McAndrew, A. (2004). *An Introduction to Digital Image Processing with Matlab*. Australia: Thomson.
- Naruki, S., & Kononi, S. (1992). *Kimia Teknologi Pengolahan Hasil Hewan*. Universitas Gadjah Mada.
- Nurjuliana, M., Che Man, Y. B., & Mat Hashim, D. (2011). Analysis of lard's aroma by an electronic nose for rapid Halal authentication. *JAOCS, J.Am. Oil Chem. Soc*, 88, 75-82.
- Ong, S., Zuraini, M., Jurin, M., Cheah, Y., Tunung, R., Chai, L., . . . Son, R. (2007). Meat molecular detection : sensitivity of polymerase chain reaction-restriction fragment length polymorphism in species differentiation of meat from animal origin. *ASEAN FOOD JOURNAL*, 51-59.
- Potter, M. D. (1993). *Principle of Meat Science 2th* . Iowa: Publishing Co.
- Rusmadi, D. (2001). *Mengenal Elektronika*. Bandung: Pionir Jaya.
- Santosa, B. (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Santosa, B. (2007). Klasifikasi Tingkat Keganasan Breast Cancer dengan menggunakan Regresi Logistik Ordinal dan SupportVector Machine (SVM). *Jurnal Sains dan Seni ITS, Vol.1, No.1*.
- Scholkopf, B., & Smola, A. (2002). *Learning with Kernels*. Cambridge . Massachusetts: The MIT Press.
- Soeparno. (1994). *Ilmu dan Teknologi Daging cetakan II*. Yogyakarta: Fakultas Peternakan Universitas Gadjah Mada.
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag.
- Whidhiasih, R. N., Wahanani, N. A., & Supriyanto. (2013). Klasifikasi buah belimbing berdasarkan citra red-green-blue menggunakan kkn dan ida. *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic1 (1)*, 29-35.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining Practical Machine Learning Tools and Techniques (3rd ed)*. USA: Elsevier.

[Halaman ini sengaja dikosongkan]

LAMPIRAN I
FOTO PERCOBAAN



Gambar 5.1 Proses penimbangan daging sapi dan daging babi



Gambar 5.2 Proses pengambilan data daging menggunakan electronic nose dengan suhu ruang dan suhu -22°C



Gambar 5.3 Proses pengambilan data daging menggunakan electronic nose dengan suhu 55°C

BIODATA PENULIS



Nama : Sinarring Azi Laga
Tempat Tanggal Lahir : Surabaya, 1 Juni 1995
Email : sinarringazil@gmail.com

Penulis merupakan mahasiswa yang berasal dari Kota Surabaya, Jawa Timur. Penulis merupakan anak kedua dari dua bersaudara. Penulis menamatkan pendidikan dasar di SD Muhammadiyah 4 Surabaya pada tahun 2007, menamatkan pendidikan Sekolah Menengah Pertama di SMP Negeri 29 Surabaya pada tahun 2010 dan menamatkan pendidikan Sekolah Menengah Atas di SMA Negeri 6 Surabaya pada tahun 2013. Setelah mengenyam pendidikan bangku sekolah, penulis melanjutkan studi ke jenjang S1 Teknik Informatika di Universitas Pembangunan Nasional “Veteran” Jawa Timur dan lulus pada tahun 2017.

Setelah menamatkan jenjang pendidikan S1, penulis melanjutkan pendidikan ke jenjang Magister (S2) pada tahun 2017 di Departemen Manajemen Teknologi, Fakultas Bisnis dan Manajemen Teknologi Institut Teknologi Sepuluh Nopember Surabaya (MMT-ITS) dengan mengambil bidang konsentrasi/keahlian Manajemen Teknologi Informasi. Penulis memiliki ketertarikan pada bidang *data mining*.