DETEKSI KELAYAKAN FISIK AIR UNTUK KONSUMSI MENGGUNAKAN *NAIVE BAYES CLASSIFIER*

SKRIPSI

Oleh:
SAIFULLAH
NIM. 13650071

JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2019

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

DETEKSI KELAYAKAN FISIK AIR UNTUK KONSUMSI MENGGUNAKAN *NAIVE BAYES CLASSIFIER*

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh:
SAIFULLAH
NIM. 13650071

JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2019

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

LEMBAR PERSETUJUAN

DETEKSI KELAYAKAN FISIK AIR UNTUK KONSUMSI MENGGUNAKAN *NAIVE BAYES CLASSIFIER*

SKRIPSI

Oleh :
SAIFULLAH
NIM. 13650071

Telah Diperiksa dan Disetujui untuk Diuji

Tanggal : 28 November 2019

Dosen Pembimbing I	Dosen Pembimbing II
Irwan Budi Santoso, M.Kom	Dr. M. Amin Hariyadi, M.T
NIP. 19770103 201101 1 004	NIP. 19670118 200501 1 001

Mengetahui,
Ketua Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

LEMBAR PENGESAHAN

DETEKSI KELAYAKAN FISIK AIR UNTUK KONSUMSI MENGGUNAKAN *NAIVE BAYES CLASSIFIER*

SKRIPSI

Oleh :
SAIFULLAH
NIM. 13650071

Telah Dipertahankan di Depan Dewan Penguji
dan Diyantakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Pada Tanggal : 06 Desember 2019

Susunan Dewan Penguji			Tanda Tangan
1. Penguji Utama	:	Dr. Cahyo Crysdian NIP. 19740424 200901 1 008	( )
2. Ketua Penguji	:	A'la Syauqi, M.Kom NIP. 19771201 200801 1 007	( )
3. Sekretaris Penguji	:	Irwan Budi Santoso, M.Kom NIP. 19770103 201101 1 004	( )
4. Anggota Penguji	:	Dr. M. Amin Hariyadi, M.T NIP. 19670118 200501 1 001	( )

Mengetahui,
Ketua Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama	: Saifullah
NIM	: 13650071
Fakultas/Jurusan	: Sains dan Teknologi/Teknik Infomatika
Judul Skripsi	: **DETEKSI KELAYAKAN FISIK AIR UNTUK KONSUMSI MENGGUNAKAN *NAIVE BAYES CLASSIFIER***

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan Skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 06 Desember 2019

Hormat saya,

Saifullah

NIM. 13650071

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

“Menyerah Hanya Untuk Orang-Orang Yang Kurang Bersyukur”

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

HALAMAN PERSEMBAHAN

Ku Persembahkan Karya Kecil Ini Untuk Almarhum Ayahanda Tercinta Muttaqin, Dan Ibunda Tercinta Tuliyatin. Terimakasih Telah Merawat, Mendidik, Dan Membesarkanku Sampai Saat Ini. Terimakasih Juga Atas Dukungan Moril Dan Materil Sehingga Putra Jenengan Ini Bisa Sampai Di Titik Ini.

Dan Tak Lupa Ku Persembahkan Karya Ini Untuk Mbak Muzdalifah, Mbak Illiyyun, Dan Mas Amrullah.

Terimakasih Atas Dukunganya, Semangatnya Dan Nasihat-Nasihatnya Disaat Adiknya Ini Sudah Mulai Patah Semangat Untuk Menyelesaikan Pendidikan Ini.

Terimakasih Banyak.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Segala puji bagi Allah SWT, karena atas segala rahmat dan karunia-Nya, sehingga penulis mampu menyelesaikan studi di Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang sekaligus menyelesaikan Skripsi dengan baik dan lancar. Shalawat serta salam selalu tercurah kepada tauladan terbaik Nabi Muhammad SAW yang telah membimbing umatnya dari zaman *jahiliyah* menuju islam yang *rahmatan lil alamin*.

Dalam proses menyelesaikan skripsi ini, penulis telah mendapatkan banyak bantuan dari berbagai pihak, baik secara moril, nasihat dan semangat maupun materiil. Oleh karena itu, ucapan terima kasih penulis sampaikan kepada:

1. Bapak Irwan Budi Santoso, M.Kom, selaku dosen pembimbing I yang telah bersedia meluangkan waktu untuk membimbing, memotivasi, dan mengarahkan selama proses pengerjaan skripsi ini.
2. Bapak Dr. M. Amin Hariyadi, M.T, selaku dosen pembimbing II yang telah senantiasa memberikan masukan dan nasihat serta petunjuk dalam penyusunan skripsi ini.
3. Segenap dosen dan staf Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang yang telah memberikan bimbingan dan ilmu sehingga dapat menjadi bekal untuk penulis dalam menyelesaikan studi dan skripsi selama masa studi.
4. Ayah dan Ibu tercinta serta keluarga besar penulis yang selalu memberikan doa, motivasi, nasihat, pengalaman berharga, serta membiayai penulis dalam menuntut ilmu dan menyelesaikan skripsi ini.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

5. Teman-teman Jurusan Teknik Informatika, khususnya angkatan 2013, yang telah memberikan motivasi, informasi, dan masukannya pada penulis sehingga dapat menyelesaikan skripsi ini dengan baik.

Berbagai kekurangan dan kesalahan mungkin dapat ditemukan dalam penulisan skripsi ini, oleh karena itu penulis menerima segala kritik dan saran akan penulis terima dengan senang hati dan akan menjadi bahan pertimbangan bagi penulis selanjutnya untuk menyempurnakan skripsi ini. Semoga karya ini senantiasa dapat memberi manfaat. Amin.

Wassalamualaikum Wr. Wb.

Malang, Desember 2019

Penulis

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

DAFTAR ISI

HALAMAN JUDUL ..... **i**
HALAMAN PENGAJUAN ..... **ii**
LEMBAR PERSETUJUAN ..... **iii**
LEMBAR PENGESAHAN ..... **iv**
PERNYATAAN KEASLIAN TULISAN ..... **v**
MOTTO ..... **v**
HALAMAN PERSEMBAHAN ..... **vii**
KATA PENGANTAR ..... **viii**
DAFTAR ISI ..... **x**
DAFTAR GAMBAR ..... **xiii**
DAFTAR TABEL ..... **xiv**
ABSTRAK ..... **xv**
ABSTRACT ..... **xvi**
الملخص ..... **xvii**
BAB I PENDAHULUAN **1**
1.1. Latar Belakang Masalah 1
1.2. Identifikasi Masalah 3
1.3. Tujuan Masalah 3
1.4. Batasan Masalah 4
1.5. Manfaat Penelitian 4
1.6. Sistematika Penelitian 4
BAB II TINJAUAN PUSTAKA **6**
2.1 Penelitian Terkait 6
2.2 Landasan Teori 7
2.2.1 Pengertian Air 7
2.2.2 Kelayakan Air Minum 8
2.2.3 Parameter Fisik Air minum 9
2.3 Naive Bayes Classifier 11

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

2.3.1 Teorema *Bayes* 11
2.3.2 *Naive Bayes* Untuk Klasifikasi 12
2.4 *Confusion Matrix* 14
BAB III ANALISA DAN PERANCANGAN SISTEM 17
3.1 Prosedur Penelitian 17
3.2 Alat dan Bahan 18
3.3 Deskripsi Data 20
3.3.1 Data Sintetik 21
3.3.2 Data *Training* 22
3.3.3 Data *Testing* 22
3.4 Desain Sistem 23
3.4.1 Akuisisi Citra 25
3.4.2 *Cropping* 25
3.4.3 *Resize* 26
3.4.4 Ekstraksi Fitur 27
3.4.5 Estimasi Parameter 30
3.4.6 Naive Bayes Classifier 34
3.4.7 Pengukuran Kinerja Klasifikasi 40
3.5 Desain *Interface* 42
3.5.1 Form *Training* 42
3.5.2 Form Testing 43
3.6 Implementasi Sistem 44
3.6.1 Implementasi Sistem 45
BAB IV UJI COBA DAN PEMBAHASAN 46
4.1. Prosedur Pengujian 46
4.1.1 Persiapan Data 47
4.1.2 Preproses 47

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

4.1.3 Proses Klasifikasi *Naive Bayes Classifier* 48
4.2. Hasil Uji Coba 48
4.2.1. Hasil Proses *Training* 49
4.2.2. Hasil Proses *Testing* 56
4.3. Pembahasan 61
BAB V PENUTUP 65
5.1 Kesimpulan 65
5.2 Saran 66
DAFTAR PUSTAKA 67
LAMPIRAN

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

DAFTAR GAMBAR

Gambar 3.1 Alat pengambilan image air 18

Gambar 3.2 Termometer 19

Gambar 3.3 Ph Test Paper 19

Gambar 3.4 Tahapan Pengumpulan Data *Sintetik* 21

Gambar 3.5 Desain Sistem Parameter Fisik Air 23

Gambar 3.6 Sebuah *Image* Air di *Cropping* 26

Gambar 3.7 Ilustrasi Metode *Interpolasi Nearest Neighbour* 26

Gambar 3.8 Fitur Image Air di Ekstraksi Fitur 28

Gambar 3.9 Pseudocode resize dan ekstraksi fitur data training 28

Gambar 3.10 Pseudocode untuk mengelompokkan data menjadi perkelas 29

Gambar 3.11 *Flowchart* menghitung *mean* 30

Gambar 3.12 Flowchart menghitung varian 31

Gambar 3.13 *Pseudocode* estimasi parameter 34

Gambar 3.14 *Flowchart* kaidah *Naive Bayes Classifier* 35

Gambar 3.15 *Flowchart Naive Bayes Classifier* Data Kualitatif dan Kontinue 36

Gambar 3.16 *Pseudocode Naive Bayes Classifier* 39

Gambar 3.17 *Pseudocode Confusion Matrix* 41

Gambar 3.18 *Interface form training* 43

Gambar 3.19 *Interface form testing* 44

Gambar 4.1 Hasil *Training* Data 50

Gambar 4.2 Hasil *testing* sistem 57

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

DAFTAR TABEL

Tabel 3.1 Contoh fitur data *training* 32
Tabel 3.2 Hasil Perhitungan *Mean* dan *Varian* 33
Tabel 3.3 Hasil probabilitas kemunculan nilai bau 33
Tabel 3.4 Hasil probabilitas kemunculan nilai rasa 34
Tabel 3.5 Contoh fitur data *testing* 37
Tabel 4.1 Hasil uji coba data *training* tanpa fitur bau dan rasa 51
Tabel 4.2 Confusion matrix hasil data training tanpa fitur bau dan rasa 52
Tabel 4.3 Hasil uji coba data training dengan fitur bau dan rasa 53
Tabel 4.4 Confusion matrix hasil data training tanpa fitur bau dan rasa 55
Tabel 4.5 Hasil akurasi data *training* 55
Tabel 4.6 Hasil uji coba data *testing* tanpa fitur bau dan rasa 58
Tabel 4.7 *Confusion matrix* hasil data *testing* tanpa fitur bau dan rasa 58
Tabel 4.8 Hasil uji coba data *testing* dengan fitur bau dan rasa 59
Tabel 4.9 *Confusion matrix* hasil data *testing* dengan fitur bau dan rasa 60
Tabel 4.10 Hasil akurasi data *testing* 61

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

ABSTRAK

Saifullah. 2019. **Deteksi Kelayakan Fisik Air Untuk Konsumsi Menggunakan *Naive Bayes Classifier.*** Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing (I) Irwan Budi Santoso, M.Kom. (II) Dr. M. Amin Hariyadi, M.T

Kata Kunci : Air Sumur, Fisik, Konsumsi, *Naive Bayes classifier*.

Penurunan kualitas air akibat pembuangan senyawa kimia yang makin meningkat, membuat sumber air di daerah-daerah tercemar sehingga memungkinkan masyarakat telah mengkonsumsi air sumur yang tidak layak untuk di konsumsi dan bisa membahayakan bagi tubuh mereka. Ada satu cara sebagai pencegahan untuk mengantisipasi itu semua yaitu dengan uji parameter fisik air. Ini akan membawa keuntungan karena dengan parameter fisik bisa mengidentifikasi kelayakan parameter lainya. Sehingga menurunkan resiko masyarakat dalam mengkonsumsi air sumur yang tidak layak konsumsi. Metode naive bayes classifier dirancang untuk mengidentifikasi kelayakan fisik air sumur. dengan menggunakan ph meter untuk fitur ph, termometer untuk fitur suhu, organoleptik untuk fitur bau dan rasa, dan fitur warna dengan meletakkan air sumur ke dalam prototype yang telah dibuat, kemudian mengambil gambar dengan kualitas tinggi. Dari gambar kamera yang dihasilkan, kemudian diekstraksi fitur warna. Jumlah total data yang digunakan dalam penelitian ini 80 air sumur yang bersumber dari survei di lapangan. Data pelatihan menggunakan 60 data, 30 data air sumur tidak layak secara fisik dan 30 data air sumur layak secara fisik. Sedangkan untuk pengujian menggunakan 20 data, yang terbagi 10 data air sumur layak secara fisik dan 10 data air sumur tidak layak secara fisik. Berdasarkan hasil uji coba pada sistem dari data training diperoleh akurasi terbaik yaitu sebesar 90%, dan untuk hasil uji coba data testing diperoleh akurasi terbaik sebesar 85%.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

ABSTRACT

Saifullah. 2019. **Detection of Physical Feasibility of Water for Consumption with Naive Bayes Classifier.** Thesis. Department of Informatics, Faculty of Science and Technology, State Islamic University of Maulana Malik Ibrahim Malang.
Supervisor (I) Irwan Budi Santoso, M.Kom. (II) Dr. M. Amin Hariyadi, M.T

Keywords: Well Water, Physical, Consumption, Naive Bayes classifier

The decrease of water quality due to the disposal of chemical compounds is increasing, which create water sources in polluted areas, therefore the people can consume well water which is unfit for consumption and can be harmful to their bodies. There is one way to prevent this from anticipating all, namely by testing the physical parameters of water. This will bring the benefits because the physical parameters can identify the feasibility of the other parameters. That can reduce the risk of the community for consume the well water which is not suitable for consumption. The naive bayes classifier method is designed to identify the physical feasibility of the well water. This using a ph meter for ph features, thermometers for temperature features, organoleptics for odor and taste features, and color features by putting well water into the prototype which has been made, then taking high quality pictures. Therefore, from the camera images produced, color features are extracted. The total amount of data which is used in this study was 80 the well water sourced from field surveys. The training data uses 60 data, which are divided into 30 physically unfit the well water data and 30 physically the well water data. As for testing used 20 data, 10 the physical well water data and 10 the well water data are divided. Based on the results of trials on the system of training data are obtained the best accuracy that is equal to 90%, and for the results of testing data are obtained the best accuracy of 85%.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

الملخص

سيف الله. 2019. الكشف عن الجدوى المادية للمياه للاستهلاك باستخدام تصنيف السذاجة بايز. البحث الجامعي. قسم المعلومة التقنيقة. كلية العلوم والتكنولوجيا. جامعة مولانا مالك إبراهيم مالانج الإسلامية الحكومية.
المشرف (الأول) إروان بودي سانتوسو،الماجيستر (الثاني) الدكتور محمد أمين هريادي الماجيستر

كلمات البحث : مياه الآبار، الفيزيائية، الاستهلاك، تصنيف السذاجة بايز

يتزايد انخفاض جودة المياه بسبب التخلص من المركبات الكيميائية، يجعل مصادر المياه في المناطق الملوثة حتى يتمكن الناس من استهلاك مياه الآبار غير الصالحة للاستهلاك ويمكن أن تكون ضارة بأجسامهم. هناك طريقة واحدة لمنع ذلك كله أي باختبار المعلمات الفيزيائية للمياه. سيؤدي ذلك إلى تحقيق فوائد لأن المعلمات المادية يمكنها لتحديد جدوى المعلمات الأخرى. حيثما تقليل مخاطر المجتمع في استهلاك مياه الآبار غير المناسبة للاستهلاك. تصميم طريقة تصنيف السذاجة بايز لتحديد الجدوى المادية لمياه الآبار. باستخدام مقياس فتاه لميزات فتاه، ومقاييس الحرارة لميزات درجة الحرارة، ومثبطات الرائحة للرائحة وميزات الذوق، وميزات اللون عن طريق وضع مياه البئر في النموذج الأولي المنشاؤة، ثم التقاط صور بجودة عالية. من صور ألة التصوير المنتجة، ثم استخراج ميزات اللون. المبلغ الإجمالي كمية البيانات المستخدمة في هذا البحث 80 مياه بئر مصدرها المسوحات الميدانية. تستخدم بيانات التدريب 60 بيانة، التي تنقسم إلى 30 بيانة مياه الآبار غير الصالحة جسديًا و 30 بيانة مياه الآبار الجسدية. بالنسبة للاختبار باستخدام 20 بيانة، يتم تقسيم 10 بيانات عن مياه الآبار الفيزيائية و 10 بيانات عن مياه الآبار. بناءً على نتائج التجارب على نظام بيانات التدريب المحصولة عليها، حصلت على أفضل دقة تساوي 90٪، ولبيانات اختبار النتائج المحصولة عليها، فقد حصلت على أفضل دقة بنسبة 85٪.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Air merupakan sumber kehidupan bagi seluruh makhluk hidup yang ada di bumi. Tumbuhan dan hewan sangat membutuhkan air untuk bertahan hidup, terlebih lagi manusia. Ketergantungan manusia terhadap air sangatlah besar. Selain untuk diminum, air juga digunakan pada berbagai kegiatan lainnya seperti mencuci, mandi, memasak, dan lain-lain.

Di Indonesia sendiri ketersediaan air sangat melimpah. hampir disetiap daerah ditemukan sumber mata air, Namun tidak semua sumber air itu dapat digunakan untuk kebutuhan sehari-hari terutama untuk minun. Karena kualitas air disetiap daerah berbeda-beda dan harus memenuhi persyaratan fisika, mikrobiologis, dan radioaktif sebagai kualitas air yang bisa untuk dikonsumsi (Permenkes No. 492, 2010).

Masalah utama yang dihadapi berkaitan dengan sumber daya air adalah penurunan kualitas air akibat pembuangan senyawa kimia tertentu yang makin meningkat, terutama akibat kegiatan industri dan transportasi dan akibat penggunaan bahan-bahan berbahaya oleh manusia. Kegiatan industri, transportasi, dan aktivitas manusia yang banyak mengeluarkan bahan pencemaran seringkali merusak keadaan lingkungan, salah satunya adalah air (Achmad, 2004). Sehingga memungkinkan tercemarnya air sumur yang ada. Dan mengakibatkan masyarakat mengkonsumsi air sumur yang sebenarnya tidak layak untuk dikonsumsi. Allah SWT. berfirman dalam Al-Qur'an surat Al-Baqarah ayat 168 yang berbunyi :

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

يَٰٓأَيُّهَا ٱلنَّاسُ كُلُواْ مِمَّا فِي ٱلۡأَرۡضِ حَلَٰلٗا طَيِّبٗا وَلَا تَتَّبِعُواْ خُطُوَٰتِ ٱلشَّيۡطَٰنِۚ إِنَّهُۥ لَكُمۡ عَدُوّٞ مُّبِينٌ ١٦٨

Artinya : "Hai sekalian manusia, makanlah yang halal lagi baik dari apa yang terdapat di bumi, dan janganlah kamu mengikuti langkah-langkah syaitan; karena sesungguhnya syaitan itu adalah musuh yang nyata bagimu(QS.Al-Baqarah[2]:168)"(Al-Kalam, 2009).

Dalam tafsir ibnu katsir dijelaskan bahwa ayat ini memperbolehkan manusia makan dari apa yang ada di Bumi, yaitu yang dihalakan bagi mereka lagi baik dan tidak membahayakan tubuh serta akal mereka.

Dari tafsir tersebut dijelaskan bahwa kita sebagai manusia di minta untuk memakan ataupun mengkonsumsi makanan yang bersih, dan sehat bagi tubuh. Sehingga kita bisa terhindar dari segala macam penyakit yang ada. Air sendiri menjadi salah satu cairan yang sangat sering dikonsumsi oleh manusia. Sehingga dalam penggunaannya, apabila air yang digunakan terkontaminasi oleh bakteri ataupun zat kimia lainnya, maka akan menimbulkan penyakit bagi manusia (Candra, 2014).

Dalam rangka meminimalisir penggunaan masyarakat terhadap air sumur yang tidak layak konsumsi akibat pencemaran tersebut diperlukan sebuah tindakan agar masyarakat bisa mengetahui lebih dini apakah air yang digunakan untuk kebutuhan minum sehari-hari tersebut memang layak atau tidak sebagai air minum. Dan dari keempat parameter air tersebut ada satu parameter yang paling mudah untuk diketahui tingkat kelayakannya, yaitu parameter fisik air. parameter tersebut merupakan hal yang harus diperhatikan pertama kali ketika ingin menggunakan air sebagai air minum. Karena jika parameter fisik air ini sudah tidak memiliki

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

kelayakan, maka air tersebut memang sudah tidak layak untuk digunakan sebagai air minun.

Berdasarkan permasalah tersebut diperlukan adanya suatu sistem deteksi kelayakan fisik air berbasis *Artificial Intelligence* yang membantu dalam menentukan kelayakan parameter fisik air.

Oleh karena itu, dalam penelitian ini diusulkan penggunaan metode *naive bayes classifier* untuk memberikan keputusan apakah parameter fisik tersebut masuk ke klasifikasi air layak minum atau tidak dengan menggunaan data2 yang terdahulu.

Diharapkan dengan adanya penelitian ini. Dapat meminimalisir penggunaan air sumur yang tidak layak untuk konsumsi. Sehingga dapat mengurangi hal-hal yang dapat mengganggu kesehatan masyarakat dari air minum yang dikonsumsi tersebut. Untuk mengidentifikasi kualitas fisik air layak minum ini menggunakan sampel berupa citra air yang akan di identifikasi dan beberapa parameter fisik air.

1.2. Identifikasi Masalah

Dari penjelasan latar belakang tersebut, maka bisa disimpulkan bahwa permasalahan yang diangkat dalam penelitian kali ini ialah sebagai berikut :

Seberapa akurat penggunaan algoritma *Naive Bayes* untuk mendeteksi kualitas fisik air untuk konsumsi?

1.3. Tujuan Masalah

Mengukur akurasi penggunaan algoritma *Naive Bayes* untuk mendeteksi kualitas fisik air untuk konsumsi.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

1.4. Batasan Masalah

Untuk mencegah penelitian ini tidak keluar dari apa yang telah dijabarkan di atas, maka perlu adanya batasan-batasan dalam pengerjaan ini. Berikut batasan masalah dalam penelitian ini antara lain :

1. Deteksi air layak untuk konsumsi hanya secara fisik, yaitu : ph, bau, rasa, warna, dan suhu.
2. Air yang dideteksi merupakan air sumur

1.5. Manfaat Penelitian

Manfaat penelitian yang diharapkan dengan dilakukannya penelitian ini adalah untuk mempermudah masyarakat menentukan air layak minum dengan parameter fisiknya agar bisa diproses lebih lanjut kedepanya.

1.6. Sistematika Penelitian

Berikut adalah sistematika penelitian yang terdiri dari 5 bab :

BAB I PENDAHULUAN

Pendahuluan membahas tentang latar belakang masalah , identifikasi masalah , tujuan penelitian , batasan masalah , manfaat penelitian serta sistematika penelitian.

BAB II TINJAUAN PUSTAKA

Tinjauan Pustakan berisikan tentang teori-teori dan hasil-hasil penelitian yang relevan.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

BAB III ANALISA DAN PERANCANGAN SISTEM

Menganalisa kebutuhan sistem yang berguna dalam pembuatan aplikasi serta berisikan perancangan serta langkah pembuatan aplikasi deteksi air minum layak minum.

BAB IV UJI COBA DAN PEMBAHASAN

Dalam bab ini membahas tentang implementasi metode kedalam sistem dan melakukan pengujian sistem dalam melakukan perhitungan metode.

BAB V KESIMPULAN DAN SARAN

Bab ini merupakan penjelasan dari kesimpulan yang didapat pada hasil penelitian yang telah dilakukan serta saran terhadap penelitian

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Penelitian terkait merupakan penelitian yang sudah ada sebelumnya yang akan dijadikan sebagai acuan penelitian. Oleh karena itu penelitian ini juga mengambil beberapa teori-teori dari penelitian sebelumnya.

Beberapa penelitian yang terkait yaitu :

Acuan penelitian yang pertama yaitu oleh Ronald Misigo. yang meneliti cara membedakan varietas buah apel. Aplikasi ini menggunakan sampel citra apel sebagai bahan yang diuji. Tujuan utama penelitian ini adalah untuk menyelidiki penerapan dan kinerja algoritma *naive bayes* dalam klasifikasi varietas buah apel. Hasil dari penelitian ini menenjukkan bahwa hasil dari klasifikasi naive bayes lebih tinggi dari pada analisis komponen utama logika *fuzzy* dan *MPLNeural* (Misigo, 2016).

Penelitian selanjutnya yaitu penelitian dari Rahman. Penelitian ini untuk menentukan kualitas air minum secara fisik, kimia dan biologis pada air minum kemasan yang beredar di Kota Makassar. Dalam penelitian ini diambil sampel sebanyak 17 merek air minum kemasan yang ada di Kota Makassar dengan teknik *purposive sampling* dan dengan metode organoleptik, spektrofotometer SPADNS, sistem multiple tukes. Hasil yang diperoleh dari penelitian ini menunjukkan bahwa dari ke-17 sampel tersebut kualitas parameter fisik menunjukkan kelayakan, dari bau, warna, dan rasa tidak ada sama sekali. Namun dari parameter kimia 14 sampel memenuhi syarat dan 3 sampel tidak memenuhi syarat. Dan untuk parameter

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

biologi dari ke-17 sampel hanya satu yang tidak layak secara kualitas (Rahman, 2017).

Penelitian yang dilakukan Park Dongchul. Penelitian yang bertujuan membandingkan klasifikasi konvensional dan metode *Naive Bayes*. Dalam penelitian ini *Naïve Bayes Classifier* dapat menghasilkan hasil klasifikasi yang sangat akurat dengan waktu pelatihan minimum bila dibandingkan dengan algoritma pembelajaran konvensional yang diawasi atau tidak diawasi. Hasilnya menunjukkan bahwa Naïve Bayes Classifier yang diusulkan melebihi pengklasifikasi konvensional dalam hal kecepatan pelatihan dan akurasi klasifikasi (Dongchul, 2016).

2.2 Landasan Teori

2.2.1 Pengertian Air

Air sangat erat hubungannya dengan kehidupan manusia, air merupakan suatu sarana utama untuk meningkatkan derajat kesehatan masyarakat, karena dalam penggunaannya, apabila air yang digunakan terkontaminasi oleh bakteri ataupun zat kimia lainnya, maka akan menimbulkan penyakit bagi manusia (Candra, 2014).

Peningkatan kuantitas air adalah syarat kedua setelah kualitas, karena semakin maju tingkat hidup seseorang, maka akan semakin tinggi pula tingkat kebutuhan air dari masyarakat tersebut. Untuk keperluan minum maka dibutuhkan air rata-rata sebesar 5 liter/hari, sedangkan secara keseluruhan kebutuhan air di suatu rumah tangga untuk masyarakat Indonesia diperlukan sekitar 60 liter/hari.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Jadi untuk negara-negara yang sudah maju kebutuhan air pasti lebih besar dari kebutuhan untuk negara-negara yang sedang berkembang.

Dengan pertumbuhan penduduk yang pesat, sumber-sumber air telah menjadi salah satu kekayaan yang sangat penting. Air tidak hanya menjadi hal pokok bagi konsumsi dan sanitasi umat manusia, tapi juga untuk produksi barang industri. Air tersebar tidak merata di atas bumi, sehingga ketersediaannya disuatu tempat akan sangat bervariasi.

Perencanaan yang didasarkan keahlian serta pengolahan yang seksama merupakan hal yang penting untuk mencapai tingkat efisiensi pemanfaatan air yang akan dibutuhkan di masa yang akan datang. Walaupun demikian, usaha-usaha ini haruslah mempunyai lingkup yang lebih luas. Investasi dalam pengembangan sumber daya air dipengaruhi oleh pertimbangan. pertimbangan ekonomi, sosial, dan politis serta kenyataan-kenyataan teknik dasar.

Air merupakan salah satu sumber daya alam yang memiliki fungsi yang sangat vital bagi kehidupan makhluk hidup yang ada di muka buni. Untuk itu air perlu dilindungi agar dapat tetap bermanfaat bagi kehidupan manusia serta mahluk hidup lainnya. Pengertian tersebut menunjukkan bahwa air memiliki peran yang sangat strategis dan harus tetap tersedia dan lestari, sehingga mampu mendukung kehidupan dan pelaksanaan pembangunan di masa kini maupun di masa mendatang. Tanpa adanya air maka kehidupan tidak dapat berjalan normal.

2.2.2 Kelayakan Air Minum

Air harus tidak berasa dan tidak berbau untuk dikonsumsi oleh konsumen. Dalam menilai mutu air minum, pada dasarnya para konsumen dapat mengandalkan

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

mereka. Konstituen mikroba, kimiawi, dan fisik air dapat mempengaruhi tampilan, bau atau rasa air; dan para konsumen akan menilai mutu dan kelayakan air berdasarkan kelayakan tersebut. Walaupun substansi tersebut tidak memiliki efek langsung pada kesehatan. Air yang sangat keruh, sangat berwarna, atau memiliki rasa atau bau yang sangat tidak mengenakkan, akan dianggap tidak aman oleh para konsumen dan mungkin tidak akan dikonsumsi. Pada kondisi ekstrem, konsumen mungkin akan menghindari air dari sumber yang aman, tetapi tampilan estetinya tidak menyenangkan, dan mereka lebih memilih air yang sumbernya tidak aman, tetapi tampilan estetis yang menyenangkan. Oleh karena itu, persepsi konsumen harus diwaspadai secara bijak serta pedoman kesehatan dan keriteria estetis harus diperhitungkan ketika memeriksa suplai air minum dan saat menyusun standar serta regulasi.

Perubahan penampilan, rasa, atau bau normal pada suplai air minum menandakan adanya penurunan mutu sumber air mentah atau peroses pengolahan air tidak memadai dan harus diinvestigasi.

2.2.3 Parameter Fisik Air minum

Air yang memenuhi persyaratan fisik adalah air yang tidak berbau, tidak berasa, tidak berwarna, tidak keruh atau jernih, dan dengan suhu sebaiknya di bawah suhu udara sedemikian rupa sehingga menimbulkan rasa nyaman, dan jumlah zat padat terlarut (TDS) yang rendah. air yang tidak berbau dan berwarna merupakan air yang baik, sebaliknya air yang mempunyai warna tertentu pasti mengandung bahan kimia. Demikian pula dengan bau, bila air berbau biasanya mengandung bahan-bahan organik (Suyanta, 2002).

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

- Bau

Air yang berbau selain tidak estetis juga tidak akan disukai oleh masyarakat. Bau air dapat memberi petunjuk akan kualitas air.

- Rasa

Air yang bersih biasanya tidak memberi rasa/tawar. Air yang tidak tawar dapat menunjukkan kehadiran berbagai zat yang dapat membahayakan kesehatan.

- Warna

Air sebaiknya tidak berwarna untuk alasan estetis dan untuk mencegah keracunan dari berbagai zat kimia maupun mikroorganisme yang berwarna. Warna dapat disebabkan adanya tannin dan asam humat yang terdapat secara alamiah di air rawa, berwarna kuning muda, menyerupai urin, oleh karenanya orang tidak mau menggunakannya. Selain itu, zat organik ini bila terkena khlor dapat membentuk senyawa-senyawa khloroform yang beracun. Warna pun dapat berasal dari buangan industri.

- Kekeruhan

Kekeruhan air disebabkan oleh zat padat yang tersuspensi, baik yang bersifat anorganik maupun yang organik. Zat anorganik, biasanya berasal dari lapukan batuan dan logam, sedangkan yang organik dapat berasal dari lapukan tanaman atau hewan. Buangan industri dapat juga merupakan sumber kekeruhan.

- Suhu

Suhu air sebaiknya sejuk atau tidak panas terutama agar tidak terjadi pelarutan zat kimia yang ada pada saluran/pipa yang dapat membahayakan kesehatan, menghambat reaksi-reaksi biokimia didalam saluran/pipa, mikroorganisme

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

patogen tidak mudah berkembang biak, dan bila diminum air dapat menghilangkan dahaga.

- Jumlah Zat Padat Terlarut

Jumlah zat padat terlarut (TDS) biasanya terdiri atas zat organik, garam anorganik, dan gas terlarut. Bila TDS bertambah maka kesadahan akan naik pula. Selanjutnya, efek TDS ataupun kesadahan terhadap kesehatan tergantung pada spesies kimia penyebab masalah tersebut.

2.3 Naive Bayes Classifier

2.3.1 Teorema *Bayes*

Teorema bayes digunakan untuk menghitung peluang atau probabilitas terjadinya suatu peristiwa berdasarkan pengaruh yang di dapat dari hasil observasi. Metode *Bayesian* yang dipilih ialah metode *Naïve Bayesian*. Model statistic dalam metode *Naïve Bayesian* merupakan salah satu model yang sangat andal sebagai sistem pendukung keputusan. Algoritma ini memanfaatkan metode probabilitas dan statistic sederhana dengan asumsi bahwa antar satu kelas dengan kelas yang lain tidak saling tergantung (Nugroho, 2013) . Metode ini berfungsi untuk mencari nilai probabilitas tiap faktor sehingga nantinya hasilnya akan dijadikan perhitungan untuk menentukan apakah resikonya kecil, sedang atau besar.

Prediksi *bayes* didasarkan pada teorema bayes dengan formula umum pada persamaan 2.1 berikut :

$$P(H|E) = \frac{P(E|H)\ x\ P(H)}{P(E)} \qquad (2.1)$$

Penjelasan dari persamaan 2.1 adalah sebagai berikut :

Parameter Keterangan

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

P(H|E) Probabilitas akhir bersyarat *(conditional probability)* suatu hipotesis H terjadi jika diberikan bukti *(evidence)* E terjadi.

P(E|H) Probabilitas sebuah bukti E terjadi akan memengaruhi hipotesis H.

P(H) Probabilitas awal (priori) hipotesis H terjadi terjadi tanpa memandang bukti apapun.

P(E) Probabilitas awal (priori) bukti E terjadi tanpa memandang hipotesis/bukti yang lain

Ide dasar dari aturan *bayes* adalah bahwa hasil dari hipotesis atau peristiwa (H) dapat diperkirakan berdasarkan pada beberapa bukti (E) yang diamati. Ada beberapa hal penting dari aturan *bayes* tesebut, yaitu :

1. Sebuah probabilitas awal/prior H atau P(H) adalah probabilitas dari suatu hipotesis sebelum bukti diamati.
2. Sebuah probabilitas akhir H atau P(H|E) adalah probabilitas dari suatu hipotesis setelah bukti diamati.

2.3.2 *Naive Bayes* Untuk Klasifikasi

Kaitan antara *Naive Bayes* dengan klasifikasi, korelasi hipotesis dan bukti klasifikasi adalah bahwa hipotesis dalam teorema *bayes* merupakan label kelas yang menjadi target pemetaan dalam klasifikasi, sedangkan bukti merupakan fitur-fitur yang menjadikan masukan dalam model klasifikasi. Jika X adalah vektor masukan yang berisi fitur dan Y adalah label kelas, *naive bayes* dituliskan dengan (P(X|Y). Notasi tersebut berarti probabilitas label kelas Y didapatkan setelah fitur-fitur X diamati. Notasi ini disebut juga probabilitas akhir *(posterior probability)* untuk Y, sedangkan P(Y) disebut probabilitas awal *(prior probability)* Y.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Selama proses pelatihan harus dilakukan pembelajaran probabilitas akhir P(Y|X) pada model untuk setiap kombinasi X dan Y berdasarkan informasi yang didapat dari data latih. Dengan membangun model tersebut, suatu data uji X' dapat diklasifikasikan dengan mencari nilai Y' dengan memaksimalkan nilai P(X'|Y') yang didapat.

Formulasi Naive Bayes untuk klasifikasi dapat dilihat pada persamaan 2.2 berikut :

$$P(Y|X) = \frac{P(Y)\prod_{I=1}^{q} P(X_i|Y)}{P(X)} \quad (2.2)$$

P(Y|X) adalah probabilitas data dengan vektor X pada kelas Y. P(Y) adalah probabilitas awal kelas Y. $\prod_{I=1}^{q} P(X_i|Y)$ adalah probabilitas independen kelas Y dari semua fitur dalam vektor X. Nilai P(X) selalu tetap sehingga dalam perhitungan prediksi nantinya kita tinggal menghitung bagian $P(Y)\prod_{I=1}^{q} P(X_i|Y)$ dengan memilih yang terbesar sebagai kelas yang dipilih sebagai hasil prediksi. Sementara probabilitas independen $\prod_{I=1}^{q} P(X_i|Y)$tersebut merupakan pengaruh semua fitur dari data terhadap setiap kelas Y, yang dinotasikan dengan persamaan 2.3 sebagai berikut :

$$P(X|Y = y) = \prod_{I=1}^{q} P(X_i|Y = y) \quad (2.3)$$

Setiap set fitur $X = \{X_1, X_{2,} X_{3,\ldots\ldots,} X_q\}$ terdiri atas q atribut (q dimensi).

Umumnya *bayes* mudah dihitung untuk fitur bertipe kategoris seperti pada kasus klasifikasi hewan dengan fitur "jenis kelamin" dengan nilai {pria, wanita}. Namun fitur dengan tipe numerik (kontinu) ada perlakuan khusus sebelum dimasukkan *naive bayes*. Caranya adalah :

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

1. Melakukan diskretiasi pada setiap fitur kontinu dan mengganti nilai fitur kontinu tersebut dengan nilai interval diskret. Pendekatan ini dilakukan dengan mentransformasikan fitur kontinu kedalam fitur ordinal
2. Mengasumsikan bentuk tertentu dari distribusi probabilitas untuk fitur kontinu dan memperkirakan parameter distribusi dengan data pelatihan. Distribusi *Gaussian* biasanya dipilih untuk merepresentasikan probabilitas bersyarat dari fitur kontinu pada sebuah kelas $P(X_i|Y)$, sedangkan distribusi *Gaussian* dikarakteristikkan dengan dua parameter *mean* dan *varian*. Untuk setiap kelas *yj*, probabilitas bersyarat kelas *yj* untuk fitur X_i dapat dilihat pada persamaan 2.4 sebagai berikut :

$$P(X_i = x_i|Y_j = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} exp^{\frac{(X_i-\mu_{ij})^2}{2\sigma_{ij}^2}} \quad (2.4)$$

Parameter μ_{ij} bisa didapat dari *mean* sampel X_i dari semua data latih yang menjadi milik kelas *yj*, sedangkan σ_{ij}^2 dapat diperkirakan dari *varian* sampel (S^2) dari data latih.

2.4 *Confusion Matrix*

Matriks konfusi merupakan tabel pencatat hasil kerja klasifikasi. Matriks konfusi berisi informasi tentang klasifikasi yang dapat diprediksi dengan sistem klasifikasi. Kinerja sistem umumnya dievaluasi menggunakan data dalam bentuk matriks (Santra dan Christy, 2012). Melalui *Confusion Matrix*, keakuratan, tingkat kesalahan, ketepatan dan nilai penarikan dapat diketahui. Kuantitas matriks konfusi dapat diringkus menjadi dua nilai, yaitu akurasi dan laju *error*. Dengan mengetahui jumlah data yang diklasifikasin dengan benar, kita dapat mengetahui akurasi hasil

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

prediksi dan dengan mengetahui jumlah data yang diklasifikasin secara salah, kita dapat mengetahi laju *error* dari prediksi yang dilakukan.

Pada pengukuran kinerja menggunakan *confusion matrix*, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah *True Positive* (TP), *True Negative* (TN), *False Positive*(FP) dan *False Negative* (FN). Nilai *True Negative* (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan *False Positive* (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, *True Positive* (TP) merupakan data positif yang terdeteksi benar. *False Negative* (FN) merupakan kebalikan dari *True Positive*, sehingga data posifit, namun terdeteksi sebagai data negatif.

Berdasarkan nilai *True Negative* (TN), *False Positive* (FP), *False Negative* (FN), dan *True Positive* (TP) dapat diperoleh nilai akurasi, presisi, *error* dan *recall*. Namun untuk penelitian ini hanya kan membahas tentang akurasi dan *error* pada sistem. Akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar. Dengan kata lain, nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Nilai akurasi dapat diperoleh dengan Persamaan 2.5

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \quad (2.5)$$

Sedangkan *error* merupakan seberapa tidak akuratnya sistem dalam mengklasifikasikan data secara benar. Nilai *error* diperoleh dengan Persamaan 2.6

$$Error = \frac{FP+FN}{TP+TN+FP+FN} * 100\% \quad (2.6)$$

dimana:

- *TP* adalah *True Positive,* yaitu jumlah data positif yang terklasifikasi dengan benar oleh sistem untuk kelas ke-i.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

- *TN* adalah *True Negative*, yaitu jumlah data negatif yang terklasifikasi dengan benar oleh sistem untuk kelas ke-i.
- *FN* adalah *False Negative*, yaitu jumlah data negatif namun terklasifikasi salah oleh sistem untuk kelas ke-i.
- *FP* adalah *False Positive*, yaitu jumlah data positif namun terklasifikasi salah oleh sistem.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

BAB III

ANALISA DAN PERANCANGAN SISTEM

Pada bab ini akan menjelaskan tentang analisa dan perancangan sistem dari Aplikasi deteksi kualitas fisik air menggunakan metode *Naive Bayes* yang meliputi prosedur penelitian, alat dan bahan untuk pengambilan data, desain sistem, desain *interface*, hingga implementasinya. Analisa dan perancangan sistem dilakukan berdasarkan kebutuhan dengan tujuan untuk mempermudah proses pembuatan aplikasi.

3.1 Prosedur Penelitian

Prosedur penelitian adalah langkah-langkah yang digunakan sebagai alat untuk mengumpulkan data dalam penelitian. dengan pra penelitian kemudian studi literatur untuk mengumpulkan penelitian yang terkait sebagai acuan awal penelitian. Setelah itu dilakukan identifikasi masalah dan dilanjutkan dengan pengumpulan data yang akan digunakan untuk penelitian sehingga penelitian bisa dilakukan. Selanjutnya data yang sudah diperoleh diseleksi sesuai dengan kebutuhan penelitian sehingga data yang sudah dikumulkan benar-benar data yang valid.

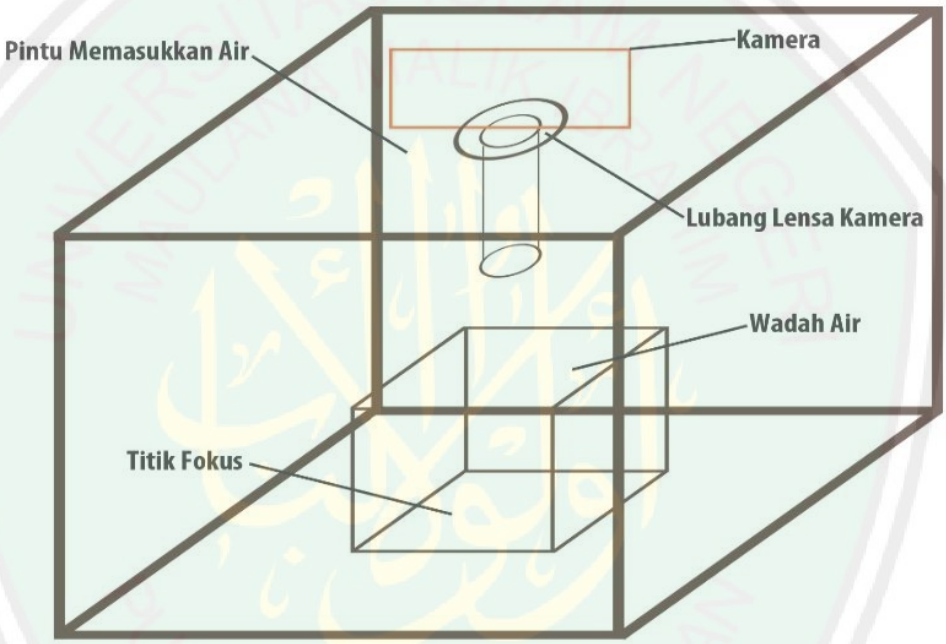
Ketika data sudah terkumpul dan tervalidasi maka tahap selanjutnya adalah pangaplikasian metode *naive bayes*. dalam metode *naive bayes* ini bertujuan untuk pengklasifikasian data yang telah diperoleh. Selanjutnya adalah tahap pengujian dan evaluasi, dalam tahapan kali ini yaitu pengujian aplikasi yang telah dibuat dan mengevaluasi apa saja yang kurang pada aplikasi tersebut. Langkah yang terakhir adalah menganalisa hasil akhir dan memberikan kesimpulan.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.2 Alat dan Bahan

Alat yang digunakan dalam peneltian ini terdiri dari beberapa alat. Mulai dari alat yang dibuat sendiri sampai yang sudah ada. Karena memang untuk datanya sendiri tidak hanya terdiri dari *image* air, ada beberapa data yang harus diperoleh selain image air seperti, rasa, bau, ph, dan suhu. Berikut adalah alat yang digunakan untuk pengambilan data :

1. Alat pengambilan *image* air



Gambar 3.1 Alat pengambilan image air

Gambar 3.1 di atas merupakan rancangan alat untuk pngambilan data *image* air. Alat tersebut berbentuk kubus dan terbuat dari akrilik yang bersifat kedap cahaya, di atasnya juga telah tersedia sebuah lubang yang akan digunakan untuk menempatkan lensa kamera *handphone*. Di dalamnya juga terdapat beberapa lampu led yang dipasang sebagai pencahayaan agar gambar yang dihasilkan nantinya lebih jelas, terang dan sesuai dengan yang diinginkan.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

2. Termometer

Gambar 3.2 *Termometer*

Gambar 3.2 di atas merupakan alat ukur yang digunakan untuk mengukur suhu air. Alat ini juga sudah tersedia dibanyak daerah atau sudah diperjual belikan. Untuk penggunaanya sendiri cukup memasukkan ujung termometer kedalam air.

3. Ph *test paper*

Gambar 3.3 *Ph* Test Paper

Gambar 3.3 di atas merupakan kertas Ph pengukur keasaman atau kebasaan air dengan satuan pH.Meter ideal untuk mengukur kualitas air pada air minum, food processing, kolam renang, rumah sakit, aquarium, rumah, Lab, industri, dan water testing secara umum. PH adalah tingkat keasaman atau kebasaan yang diukur dengan menggunakan skala pH antara 0 sampai 14.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Adapun alat dan bahan yang perlu disiapkan untuk mengambil *image* air adalah sebagai berikut:

- Akrilik 50 cm x 20 cm
- Lem akrilik
- Lakban hitam 1 buah
- *Cutter*
- Lampu led 4 biji
- Adaptor 1
- Steaker 1
- Gergaji
- Solder
- Sakelar
- Amplas
- Kabel
- Penggaris
- Gunnting
- Baterai
- Bor

3.3 Deskripsi Data

Data yang akan digunakan dalam penelitian bersumber dari data riset bersama laboratorium *artifical intelligence* dan *computer vision*. Jumlah keseluruhan data yang akan digunakan dalam penelitian ini sebanyak 80 data air yang meliputi yang diambil langsung dari sumur warga Desa Kepuharjo Kecamatan Karangploso Kabupaten Malang.

Untuk pengambilan data parameter bau dan rasa supaya hasil yang diperoleh dapat digunakan sebagai acuan maka dalam hal ini dilakukan dengan teknik *binary voting* yaitu dengan menunjuk beberapa orang untuk mendeskripsikan atau memberikan penilaian tentang rasa dan bau air tersebut sebagai acuan. Untuk jumlah pemilihnya akan melibatkan 3 orang. Setelah itu untuk menentukan hasilnya dipilih atau diambil dari suara terbanyak dari pemilih tersebut sebagai hasil.

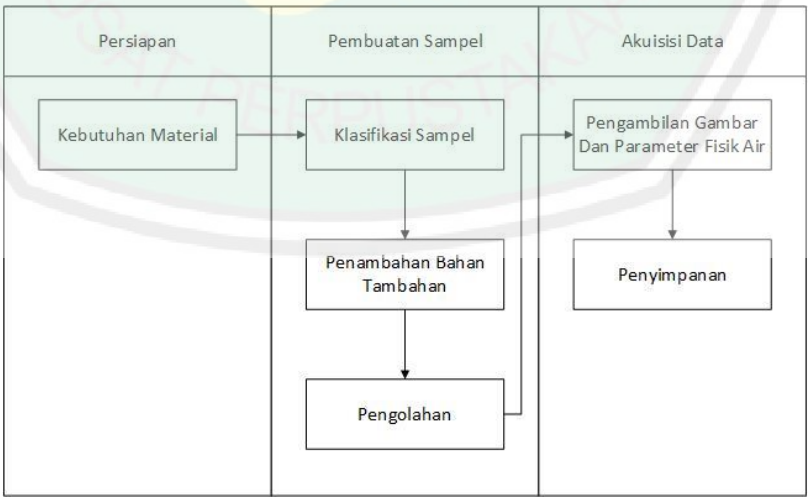
CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Selanjutnya untuk mengetahui kelayakan kualitas fisik air untuk minum harus memenuhi beberapa *standart* sebagai berikut :

1. Tidak memiliki warna, jernih atau tidak keruh
2. Tidak memiliki bau
3. Tidak memiliki rasa/tawar
4. Ph antara 6,5 – 8,5
5. Temperaturnya normal

3.3.1 Data Sintetik

Data *sintetik* merupakan data sampel air yang diluar data utama, yaitu data yang dibuat sendiri oleh penulis dengan menggunakan air layak minum sebagai bahan utama. Pembuatanya dilakukan dengan cara mencampurkan bahan seperti *detergen* dan bahan-bahan lain ke air yang benar-benar memiliki kualitas fisik layak untuk konsumsi agar dapat merubah setiap parameter-parameter air yang akan diuji. Untuk pembuatanya sendiri melalui beberapa tahapan seperti yang terlihat dalam Gambar 3.4 berikut :



Gambar 3.4 Tahapan Pengumpulan Data *Sintetik*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Secara umum, terdapat 3 tahapan yang dilakukan untuk mengumpulkan data buatan yaitu persiapan yang meliputi pemenuhan kebutuhan material pada saat proses pembuatan sampel yang diawali dengan mengumpulkan data air yang benar-benar layak untuk dikonsumsi sesuai dengan peraturan kementerian kesehatan Indonesia. Setelah itu baru kemudian mulai mencampurkkan bahan yang dapat merubah nilai dari masing-masing parameter fisik air tersebut mulai dari warna, suhu, bau, ph dan rasa

3.3.2 Data *Training*

Data *training* diperoleh dari hasil pengambilan data *image* air dan parameter fisik dengan menggunakan alat yang dibuat oleh peneliti maupun yang sudah disediakan peneliti. Banyak data yang akan digunakan untuk proses *training* adalah sejumlah 60 air yang diambil dari data keseluruhan.

Jadi dari 60 data tersebut nantinya akan digunakan untuk proses training dengan rincian 30 data kualitas fisik air layak minum dan 30 data kualitas fisik air yang tidak layak untuk diminum.

3.3.3 Data *Testing*

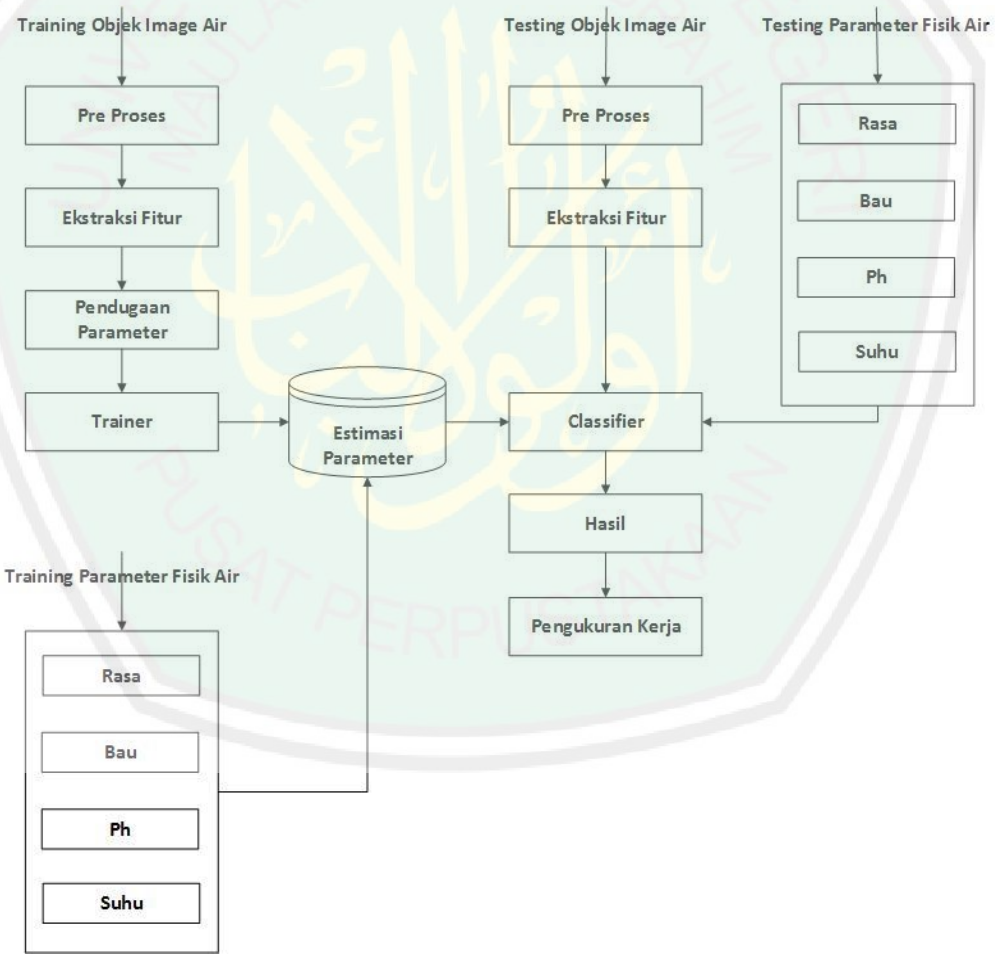
Data *testing* diperoleh dari tahapan yang hampir sama dengan proses *training* data, yaitu dari data pengambilan *image* air dan parameter fisik menggunakan alat yang telah dibuat oleh peneliti dan alat yang disediakan oleh peneliti. Banyak data yang digunakan untuk proses *testing* kali ini adalah sejumlah 20 buah data yang terdiri dari dua kategori, yaitu 10 data dengan kualitas fisik air

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

layak minum dan 10 data dengan kualitas fisik air tidak layak minum. 20 data ini diperoleh dari hasil survey dilapangan.

3.4 Desain Sistem

Tahap ini dibentuk untuk memenuhi langkah suatu sistem dalam menyelesaikan masalah yang menjadi kajian pada objek penelitian ini. Tahap ini bertujuan untuk memberikan gambaran apa yang akan dikerjakan dan bagaimana aplikasi deteksi air layak minum ini bekerja. Berikut gambaran desain sistem yang telah dibuat dapat dilihat pada Gambar 3.5 sebagai berikut :



Gambar 3.5 Desain Sistem Parameter Fisik Air

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Gambar 3.5 merupakan desain sistem dari aplikasi deteksi air minum layak minum, terdapat dua runtutan proses yang terdapat dalam desain sistem yaitu proses *training* dan *testing.* Sistem *training* merupakan bagian awal dari aplikasi untuk mengidentifikasi kualitas air layak minum pada sumur. Sistem ini berfungsi untuk menyiapkan segala informasi sebagai bahan pelatihan dan pengenalan yang akan digunakan sebagai bahan acuan dalam testing sistem.

Ada beberapa tahapan yang akan dilakukan dalam fase ini, diantaranya yaitu mengambil data-data yang sudah diolah kemudian dipisah kedalam data trraining. Kemudian dilakukan proses pengolahan citra mengambil nilai warna citra dengan menggunakan hue saturation value untuk mendapatkan nilai warna dari citra air tersebut serta melakukan ekstraksi fitur terhadap data latih, dilanjutan dengan mengestimasi parameter data dimana output pada proses ini adalah nilai parameter yang memiliki *mean* dan *varian* tiap fitur tiap kelas. Selanjutnya data nilai *mean* dan *varian* yang telah diperoleh akan disimpan untuk digunakan dalam proses testing sistem.

Sedangkan dalam perancangan sistem testing langkah harus dilakukan tidak jauh berbeda dengan langkah sistem *training.* Pada tahap ini metode naive bayes akan diimplementasikan dengan mengacu kepada nilai *mean* dan *varian* tiap fitur data populasi tiap kelas yang telah diperoleh pada proses pelatihan sistem. Selanjutnya hasil dari tahapan ini akan dilakukan perhitungan tingkat akurasi dengan membandingkan hasil idenifikasi sistem dengan data aktual.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.4.1 Akuisisi Citra

Sampel data yang akan digunakan dalam penelitian ini yaitu air sumur dan air sungai yang baru saja diambil dari sumur dan sungai, kemudian air sumur tersebut diambil data cira dengan menggunakan *prototype candling* yang kedap cahaya dari luar, sehingga pengambilan data citranya menjadi lebih sempura karena tidak ada gangguan cahaya dari luar. Citra air disimpan dengan format jpg dengan ukuran 200 x 200 dimensi citra air tersebut. Sehingga semua citra yang akan dijadikan training dan testing memiliki ukuran yang sama.

3.4.2 *Cropping*

Cropping merupakan sebuah proses pemotongan citra pada koordinat tertentu. Proses ini digunakan untuk memotong bagian citra melalui dua koordinat, yaitu koordinat awal yang merupakan titik dimulainya bagian citra yang akan diambil atau dibuang dan koordinat akhir yang merupakan titik akhir pengambilan gambar hasil *cropping*. Setelah hasil *cropping* dilakukan maka akan membentuk sebuah bangun persegi empat baru yang tiap-tiap pixel pada area koordinat yang dipilih akan disimpan dalam citra yang baru. Proses *Cropping* ini digunakan untuk membuang dan mengambil bagian citra yang ingin digunakan dan tidak digunakan. Bagian citra yang akan dipilih yaitu citra air yang seikit memantulkan cahaya lampu pada citra tersebut. Berikut citra sebelum dan sesudah mengalami proses *cropping* dapat dilihat pada Gambar 3.6.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

(a) *image* sebelum *cropping* (b) *image* sesudah *cropping*

Gambar 3.6 Sebuah *Image* Air di *Cropping*

3.4.3 *Resize*

Resize merupakan sebuah fungsi yang digunakan untuk mengubah ukuran panjang dan lebar sebuah citra *(image size)* menjadi lebih besar atau kecil dari ukuran aslinya. Proses *resize* ini sendiri nantinya akan digunakan untuk mengubah ukuran *image* air. Agar semua image air untuk data training dan data testing memiliki ukuran yang dimensi yang sama. Pada metode kali ini menggunakan metode *Interpolasi Nearest Neighbour.* Cara kerja dari *Interpolasi Neirest Neighbour* ini yaitu nilai *pixel* diambil dari *pixel* asal yang paling dekat dengan koordinasi hasil perhitungan transformasi spasial. Berikut ini contoh memperbesar *size* gambar menggunakan metode *Interpolasi Nearest Neighbour*. untuk mengubah *pixel* tersebut menjadi 3x5 dapat dilihat ilustrasi pada Gambar 3.7

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 4 & 0 & 5 & 4 \\ 0 & 7 & 0 & 8 & 7 \end{bmatrix}$$

(a) *Pixel asli* (b) *Pixel Sesudah Interpolasi*

Gambar 3.7 Ilustrasi Metode *Interpolasi Nearest Neighbour*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Yang harus dilakukan pertama kali adalah menambah zero matrik pada kolom pertama dan kolom kedua, kemudian kolom selanjutnya berisi kolom kedua dari matrik asal, kemudian kolom selanjutnya adalah kolom ketiga dari matrik asal,untuk kolom selanjutnya diisi dengan kolom ketiga dari kiri dan kolom keempat dari kiri sampai akhirnya kembali pada kolom terakhir dari matriks asal. Kemudian dari baris matrik tersebut diambil matriks kolom yang ganjil.

3.4.4 Ekstraksi Fitur

Ekstraksi fitur merupakan suatu pengambilan ciri / *feature* dari seuatu bentuk yang nantinya nilai yang didapatkan akan dianalisis untuk proses selanjutnya. Ada tiga macam yang bisa digunakan untuk ekstraksi fitur yaitu ekstraksi warna, bentuk dan tekstur. Dalam penelitian kali ini akan menggunakan ekstraksi warna yang akan digunakan pada citra air. Karena faktor pembeda dari kualitas fisik air yang layak konsumsi dan yang tidak layak konsumsi yaitu dari segi warna nya. Untuk pengambilan warna dari citra air tersebut akan mengambil nilai RGB *(Red, Green, Blue).* Nilai RGB inilah yang akan digunakan sebagai pembeda dari objek yang lain. Jadi setiap data citra air akan diambil warna R (*Red*/Merah), G (*Green*/Hijau), dan B (*Blue*/Biru). Setelah itu nilai masing-masing RGB *(Red, Green, Blue)* tersebut harus dijadikan ke dalam bentuk *array* dengan dimensi 1xN yang biasa disebut dengan vektor.

Warna air yang layak konsumsi dan tidak memiliki beberapa karakteristik yang dapat dibedakan secara kasat mata, yaitu warna air yang tidak layak konsumsi cenderung keruh dan memiliki warna lain sedangkan air yang layak untuk konsumsi memiliki warna yang jernih atau bisa dikatakan tidak berwarna. Warna yang ada

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

dalam air menandakan bahwa ada material lain yang tercampur dalam air tersebut sehingga membuat air tersebut menjadi tidak layak konsumsi. Gambar 3.8 menunjukkan sebuah perbedaan fitur *image* air sebelum dan sesudah dilakukan proses ekstraksi fitur.

(a) Fitur setelah di *resize* (b) Fitur *resize* setelah di ekstraksi fitur

Gambar 3.8 Fitur *Image* Air di Ekstraksi Fitur

Source code untuk proses *resize* dan ekstraksi fitur data *training* dapat dilihat pada gambar Gambar 3.9

```
function [DataR, DataG, DataB, Data_c] =
ekstraksi_fitur(direktorigambarcitra,DataPerKelas,kelas, baris,
kolom)
for sampel = 1:10
    FimageAir =
strcat(direktorigambarcitra,'\',int2str(sampel),'.jpg');
    x  = imread(FimageAir);
    R=x(:,:,1);
    G=x(:,:,2);
    B=x(:,:,3);
    AirR = imresize(R,[kolom,baris]);
    AirG = imresize(G,[kolom,baris]);
    AirB = imresize(B,[kolom,baris]);
    DataR(sampel,1:kolom*baris) = reshape(AirR,1,kolom*baris);
    DataG(sampel,1:kolom*baris) = reshape(AirG,1, kolom*baris);
    DataB(sampel,1:kolom*baris) = reshape(AirB,1, kolom*baris);
end
sum = 0;
for i = 1:kelas
    for j = 1:DataPerKelas
        sum = sum+1;
        Data_c(sum)=i;
    end
end
DataR = double(DataR);
DataG = double(DataG);
DataB = double(DataB);
Data_c = double(Data_c);
save('DataR.mat','DataR');
save('DataG.mat','DataG');
save('DataB.mat','DataB');
save('Data_c.mat','Data_c');
save('DataPerKelas.mat','DataPerKelas');
end
```

Gambar 3.9 *Pseudocode resize* dan ekstraksi fitur data *training*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Dari hasil ekstraksi fitur tersebut kemudian datata yang diperoleh dikelompokan berdasarkan kelasnya. Kelas kualitas fisik air layak konsumsi dimulai dari data 1 sampai 30 sedangkan kelas kualitas fisik air tidak layak konsumsi dimulai dari 31 sampai 60. *Source code* untuk pengelompokkan data sesuai dengan kelasnya dapat dilihat pada Gambar 3.10

```
load Data_c;
load meanRGB;
[N n]=size(meanRGB);
BanyakSampelKelas=[];
BanyakKelas=max(Data_c);
kode=[];
count=0;
for k=1:BanyakKelas
    for j=1:N
        if (Data_c(j)==k)
            count=count+1;
        end
    end
    BanyakSampelKelas(k)=count;
    count=0;
end
SampelAwal=0;
for k=1:BanyakKelas
    if(k==1)
        for baris=1:BanyakSampelKelas(1)
            for kolom=1:n
 DataKelas(baris, kolom,k)= meanRGB(baris,kolom);
            end
        end
    else
        SampelAwal=SampelAwal+BanyakSampelKelas(k-1);
        SampelAkhir =SampelAwal+BanyakSampelKelas(k);
        for baris = (SampelAwal+1) : SampelAkhir
            for kolom=1:n
                DataKelas((baris-
SampelAwal),kolom,k)=meanRGB(baris,kolom);

            end
        end
    end
end
save BanyakKelas;
save DataKelas;
```

Gambar 3.10 *Pseudocode* untuk mengelompokkan data menjadi perkelas

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.4.5 Estimasi Parameter

Estimasi parameter merupakan metode statistika yang berfungsi untuk memperkirakan nilai karakteristik dari populasi atau parameter populasi berdasarkan nilai karakteristik sampel atau statistik sampel. *Estimator* adalah setiap statistik (rata-rata sampel, presentase sampel, variansi sampel, dan lain-lain) yang digunakan untuk mengestimasi sebuah parameter. Jika berhadapan dengan data kontinu, asumsi khas yang digunakan adalah distribusi Gaussian. Namun jika dengan parameter model dari *mean* dan *varian*. Terdapat beberapa tahapan yang perlu dilakukan untuk nilai *mean*. Berikut rumus yang digunakan untuk menghitung *mean :*

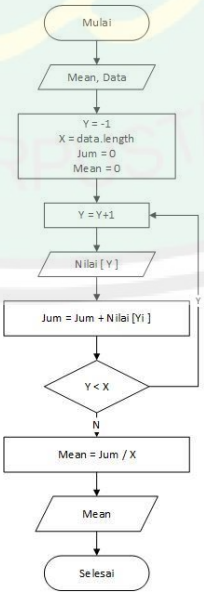
$$\mu_{ic=\frac{1}{n}} \sum_{j}^{n} F_{ij} \quad (3.5)$$

Keterangan :

μ_{ic} = *mean* fitur ke-i pada kelas c

n = banyaknya citra pada kelas c

F_{ij} = fitur ke-*i* pada kelas c



Gambar 3.11 *Flowchart* menghitung *mean*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Flowchart pada Gambar 3.11. merupakan *flowchart* untuk menghitung *mean*, proses yang pertama kali dilakukan yaitu input fitur dan jumlah fitur (n), jika terdapat lebih dari satu fitur maka akan ada proses looping dalam proses pencarian *mean* untuk menambah fitur pertama sampai dengan fitur ke-n, kemudian hasil dari penambahan fitur satu sampai ke-n dibagi dengan jumlah fitur (n).

Sedangkan rumus yang digunakan untuk menghitung *varian* yaitu :

$$\sigma_{ic}^2 = \frac{1}{n-1}\sum_{i=j}^{n}(F_{ij} - \mu_{ic})^2 \quad (3.6)$$

Keterangan:

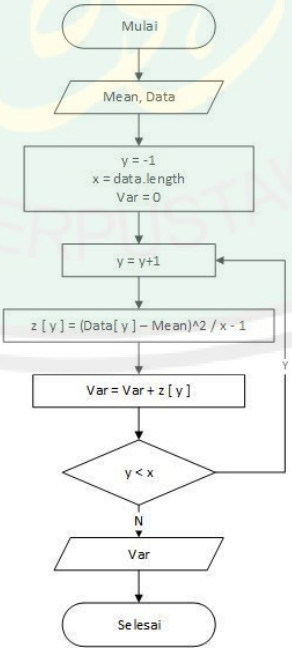
σ_{ic}^2 = *varian* dari fitur ke i pada kelas c

n = banyaknya data pada kelas c

F_{ij} = fitur ke-i

μ_{ic} = *mean* fitur ke-i pada kelas c

Sedangkan tahapan-tahaan untuk menghitung nilai *varian* adalah seperti yang terdapat pada *flowchart* yang tampak pada Gambar 3.11 sebagai berikut.



Gambar 3.12 *Flowchart* menghitung *varian*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Flowchart pada Gambar 3.12 merupakan *flowchart* untuk menghitung *varian*, proses pertama yaitu input firur, *mean* dan jumlah fitur (n), jika jumlah fitur yang ada lebih dari satu maka akan terdapat perulangan atau *looping* untuk memproses semua fitur yang ada dalam sekali perintah. Untuk proses pencarian nilai *varian* dengan cara yaitu fitur dikurangi *mean* lalu dikuadratka sampai fitur ke-n kemudian kuadratkan $(F_{ij} - \mu_{ic})^2$ kemudian hasil dari penambahan fitur satu sampai ke-n tersebut dibagi dengan jumlah fitur-1 (n-1).

Sedangkan data yang bersifat diskrit seperti yang ada pada fitur bau dan rasa maka tidak menggunakan *mean* dan *varian*. Jadi untuk dua fitur ini akan menggunakan probabilitas kemunculan setiap nilai

Misalkan data *training* sebanyak 10 data dan diantara 10 data tersebut terdiri dari dua kelas yaitu 5 data kelas air layak minum dan 5 data kelas air tidak layak minum dengan fitur RGB dan parameter fisik(bau, rasa, ph, suhu). Maka akan memperoleh tabel seperti Tabel 3.1 seperti berikut :

Tabel 3.1 Contoh fitur data *training*

data ke	Red	Green	Blue	Suhu	Ph	Bau	Rasa	Kelas
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	
1	7	5	3	2	1	1	1	Layak Minum
2	1	3	6	2	9	1	1	Layak Minum
3	8	7	1	7	9	1	1	Layak Minum
4	3	2	1	9	1	1	1	Layak Minum
5	2	8	3	5	9	0	0	Layak Minum
6	7	6	2	1	7	1	0	Tidak Layak Minum
7	1	1	4	6	9	0	1	Tidak Layak Minum
8	4	4	4	3	5	0	0	Tidak Layak Minum
9	6	2	6	5	3	0	0	Tidak Layak Minum
10	2	7	1	8	1	1	0	Tidak Layak Minum

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Keterangan :

- Nilai 1 pada fitur bau menunjukkan bahwa nilai atribut fitur yaitu tidak berbau sedangkan nilai 0 menunjukkan nilai atribut berbau atau memiliki bau
- Nilai 1 pada fitur rasa menunjukkan bahwa nilai atribut fitur tidak berasa sedangkan 0 menunjukan nilai atribut berasa atau memiliki rasa

Dari fitur X_1 setiap kelas dihitung nilai *mean* dan *varian* begitu juga X_3, X_4, sampai X_5. Sedangkan untuk fitur X_6 dan X_7 yaitu bau dan rasa menggunakan probabilitas kemunculan setiap nilai. Hasil dari perhitungan *mean* dan *varian* dapat dilihat pada Tabel 3.2.

Tabel 3.2 Hasil Perhitungan *Mean* dan *Varian*

Fitur	Mean	Std	Kelas
X_1	4,2	9,7	Layak Minum
X_2	5	6,5	Layak Minum
X_3	2,8	4,2	Layak Minum
X_4	5	9,5	Layak Minum
X_5	5,8	19	Layak Minum
X_1	4	6,5	Tidak Layak Minum
X_2	4	6,5	Tidak Layak Minum
X_3	3,4	3,8	Tidak Layak Minum
X_4	4,6	7,3	Tidak Layak Minum
X_5	5	10	Tidak Layak Minum

Sedangkan hasil probabilitas kemunculan dari fitur bau dan rasa dapat dilihat dari Tabel 3.3 dan 3.4.

Tabel 3.3 Hasil probabilitas kemunculan nilai bau

Bau	jumlah kejadian		Probabilitas	
	layak	tidak layak	layak	tidak layak
Tidak Berbau	4	3	4/5	3/5
berbau	1	2	1/5	2/5
jumlah	5	5	1	1

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Tabel 3.4 Hasil probabilitas kemunculan nilai rasa

Rasa	jumlah kejadian		Probabilitas	
	layak	tidak layak	layak	tidak layak
Tidak Berasa	4	2	4/5	2/5
berbau	1	3	1/5	3/5
jumlah	5	5	1	1

Menghitung estimasi parameter dalam pemrograman matlab cukup dengan menggunakan *function std* untuk menghitung nilai *varian* dan untuk menghitung nilai *mean* cukup menggunakan *function mean. Source code* untuk proses estimasi parameter dapat dilihat pada Gambar 3.13.

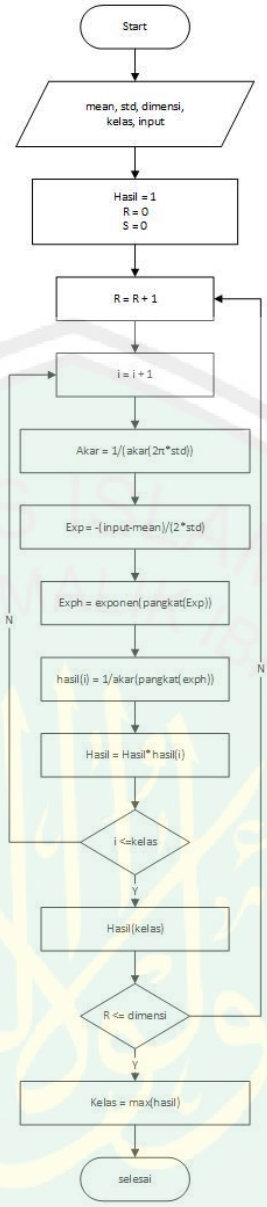
```
function estimasiparameter()
save BanyakKelas;
save DataKelas;
for kelas=1:BanyakKelas
    StandartD=std(DataKelas(:,:,kelas));
    hasilVarian = StandartD;
    hasilMean(:,:,kelas)= mean(DataKelas(:,:,kelas));
end
save ('hasillMean.mat','hasillMean');
save ('hasillVarian.mat','hasillVarian');
end
```

Gambar 3.13 *Pseudocode* estimasi parameter

3.4.6 Naive Bayes Classifier

Cara kerja metode *Naive Bayes Classifier* adalah menghitung probabilitas setiap kelas untuk dibandingkan dan melakukan pengklasifikasian dengan cara menghitung peluang terbesar dan menghitung peluang fitur perkelas melalui inputan *mean* dan *varian* dari data training, fitur yang akan di *testing*, jumlah data setiap kelas dan jumlah data keseluruhan. Cara menghitung *naive bayes* menggunakan persamaan 2.6. *Flowchart* klasifikasi dengan menggunakan *naive bayes* dapat dilihat pada Gambar 3.14.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG



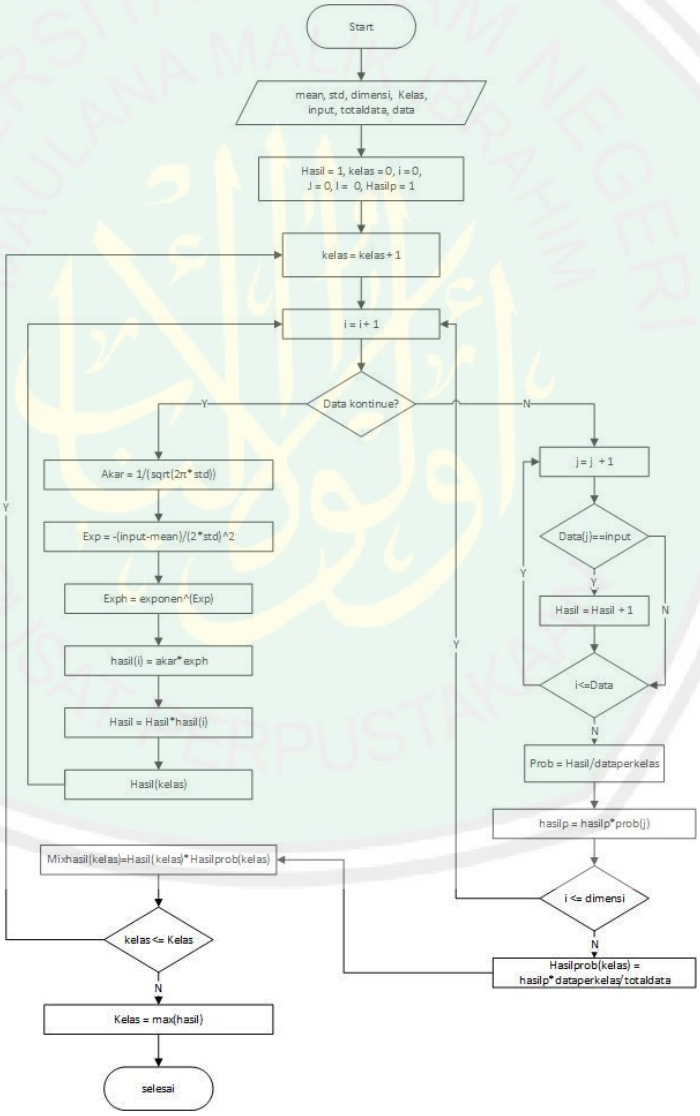
Gambar 3.14 *Flowchart* kaidah *Naive Bayes Classifier*

Dari Gambar 3.14 digambarkan cara pertama kali yang harus dilakukan untuk klasifikasi menggunakan metode *naive bayes classifier* adalah menginputkan fitur yang akan diuji coba kemudian *mean* dan *varian* serta data perkelas dan total data. Kemudian menghitung akar dari 2phi dikali dengan *varian,* sehingga diperoleh nilai dari $1/\sqrt{2\pi\sigma}$, kemudian menghitung nilai pangkat yang ada pada exponen dengan cara input dikurangi *mean* akar 2 dibagi dengan 2 dikali *varian*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

akar 2 ($\frac{-(input-\mu)^2}{2\sigma^2}$), setelah diperoleh nilai dari pangkat exponen langkah selanjutnya yaitu menghitung nilai exponen sesuai dengan akarnya kemudian menghitung nilai $1/\sqrt{2\pi\sigma}$ dikali dengan hasil dari exponen. Setelah diperoleh hasil dari perhitungan tersebut maka langkah yang terakhir yaitu menentukan *maximal argument* dari setiap kelas dengan menggunakan persamaan 2.4.

Sedangkan *Flowchart* perhitungan *naive bayes* dengan model data kualitatif dan kontinue dapat dilihat pada Gambar 3.15.



Gambar 3.15 *Flowchart Naive Bayes Classifier* Data Diskrit dan Kontinu

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Gambar 3.15 merupakan algoritma gabungan dari rumus naive bayes pada persamaan 2.4 dan probabilitas kemunculan nilai pada setiap fitur. Dari gambar tersebut cara perhitungan menggunakan dua cara yaitu data kontinue dan data diskrit. Ketika fitur data pertama merupakan data kontinue maka perhitunganya menggunakan perhitungan *gaussian*. Pertama variabel Akar, Exp, dan Exph merupakan pelabelan pecahan perhitungan yang ada dalam persamamaan 2.4 untuk memudahkan dalam perhitungan, kemudian variabel Hasil merupakan perkalian antara hasil perhitungan setiap fitur, setelah data kontinue sudah habis maka beralih ke perhitungan data diskrit. Variabel data(j) merupakan jumlah data dari data training, kemudian variabel Hasilp merupakan jumlah data yang sesuai dengan inputan tiap fitur. Setelah itu variabel prob yang merupakan hasil data yang sesuai dengan inputan fitur dibagi dengan data perkelas, sehingga diperoleh hasil probabilitas data fitur kelas ke 1. Kemudian setiap nilai fitur kelas saling dikalikan, disimpan dalam variabel hasilp. Jika sudah pada tahap ini, maka variabel hasilp dikalikan data perkelas *training* dan dibagi dengan total data *training*. Setelah itu perkalian antara hasil data kontinue dengan hasil data probabilitas saling dikalikan. Sehingga diperoleh nilai hasil perhitungan naive bayes data diskrit dan data kontinu kelas ke 1. Untuk perhitungan kelas ke dua caranya sama seperti di atas. Dan untuk yang terakhir setelah data kelas 1 dan kelas 2 telah diperoleh maka dicari nilai tertinggi dari kedua kelas.

Misalkan data *testing* memiliki fitur seperti Tabel 3.5 sebagai berikut:

Tabel 3.5 Contoh fitur data *testing*

Red	Green	Blue	Ph	Suhu	Bau	Rasa
3	1	4	7	2	1	1

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Maka hasil perhitungan *Naive Bayes Classifier* setiap kelas dengan menggunakan data training dari tabel 3.1 adalah:

Kelas 1 :

- fitur Red = $\frac{1}{\sqrt{2\pi.1,58}} 2,7183^{\frac{-(3-4,2)^2}{2.9,7}} = 0,1191$
- fitur Green = $\frac{1}{\sqrt{2\pi.2,54}} 2,7183^{\frac{-(1-5)^2}{2.6,5}} = 0,0455$
- fitur Blue = $\frac{1}{\sqrt{2\pi.2,04}} 2,7183^{\frac{-(4-2,8)^2}{2.4,2}} = 0,1645$
- fitur Ph = $\frac{1}{\sqrt{2\pi.3,08}} 2,7183^{\frac{-(7-5)^2}{2.9,5}} = 0,1049$
- fitur Suhu = $\frac{1}{\sqrt{2\pi.4,38}} 2,7183^{\frac{-(2-5,8)^2}{2.19}} = 0,0625$
- fitur Bau = 4/5
- fitur Rasa = 4/5

Hasil kelas 1 = 0,1191x0,0455x0,1645x0,1049x0,0625x4/5x4/5x5/10

= 1.87022795505e-06

Kelas 2 =

- fitur Red = $\frac{1}{\sqrt{2\pi.2,54}} 2,7183^{\frac{-(3-4)^2}{2.6,5}} = 0,1454$
- fitur Green = $\frac{1}{\sqrt{2\pi.2.54}} 2,7183^{\frac{-(1-4)^2}{2.6,5}} = 0,0782$
- fitur Blue = $\frac{1}{\sqrt{2\pi.1,94}} 2,7183^{\frac{-(4-3,4)^2}{2.3,8}} = 0,1960$
- fitur Ph = $\frac{1}{\sqrt{2\pi.2,70}} 2,7183^{\frac{-(7-4,6)^2}{2.7,3}} = 0,0995$
- fitur Suhu = $\frac{1}{\sqrt{2\pi.3,16}} 2,7183^{\frac{-(2-5)^2}{2.10}} = 0,0804$
- fitur Bau = 3/5

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

- fitur Rasa = 2/5

  Hasil Kelas 2 = 0,1454x0,0782x0,1960x0,0995x0,0804x3/5x2/5x5/10

  = 2.13937839900288e-06

Dari perhitungan diatas fitur RGB, ph, dan suhu menggunakan persamaan 2.4 untuk mendapat nilai dari setiap fitur menggunakan mean, varian dari dari tabel 3.2, sedangkan fitur bau dan rasa menggunakan nilai probabilitas kemunculan dari tabel 3.3 dan tabel 3.4. Setelah di peroleh semua nilai fitur maka semua nilai fitur tersebut akan saling dikalikan untuk memperoleh nilai dari masing-masing kelas. bisa dilihat bahwa kelas 1 memiliki nilai yang lebih besar dibandingkan dengan kelas ke 2, sehingga dari hasil yang diperoleh menunjukkan bahwa data *testing* tersebut masuk ke klasifikasi kualitas fisik air yang layak konsumsi.

Source code untuk klasifikasi menggunakan metode *Naive Bayes Classifier* dapat dilihat pada Gambar 3.16.

```
function naivebayesclassification()
    load hasilMean;
    load hasilStd;
    load DataPerKelas;
     for k = 1:2
        perhitungan1=sqrt(2*pi);
        perhitungan2(:,:,k)=1./(perhitungan1*stdD(:,:,k));
         perhitungan3(:,:,k)=-(dataRRGGBB(:,:,:)-
                              hasilMean(:,:,k)).^2;
        perhitungan4(:,:,k)= 2*stdD(:,:,k).^2;

perhitungan5(:,:,k)=perhitungan3(:,:,k)./perhitungan4(:,:,k);
        perhitungan6(:,:,k)= exp(1).^perhitungan5(:,:,k);
  perhitungan7(:,:,k)=perhitungan2(:,:,k).*perhitungan6(:,:,k);
        hasil1(:,:,k)= perhitungan7(1,1,k)*perhitungan7(1,2,k)*
                       perhitungan7(1,3,k)*perhitungan7(1,4,k)*
                       perhitungan7(1,5,k)*perkalianbauu(:,:,k)*
                       perkalianrasaa(:,:,k);
     end
    [argmax(:,:), keles]=max(hasil1);
    hasilee = keles;

end
```

Gambar 3. 16 *Pseudocode Naive Bayes Classifier*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.4.7 Pengukuran Kinerja Klasifikasi

Pengukuran keakurasian program merupakan sebuah cara mengetahui seberapa akurat aplikasi dalam menentukan hasil klasifikasi. Sebuah sistem yang melakukan klasifikasi diharapkan dapat melakukan klasifikasi semua set data dengan benar, tetapi tidak dapat dipungkiri bahwa knerja suatu sistem tidak bisa 100% benar sehingga sebuah sistem klasifikasi harus diukur kinerjannya. Untuk mengetahi keakurasian program maka diperlukan metode atau suatu cara untuk mengukur seberapa akurat aplikasi tersebut dalam melakukan klasifikasi/deteksi. Dalam penelitian ini untuk mengetahui seberapa akurat aplikasi dalam menentukan sebuah deteksi kelayakan fisik air minum maka akan digunakan metode *confusion matrix*. Dalam *confusion matrix* dapat diperoleh nilai akurasi, presisi, *error* dan *recall*. Namun untuk penelitian ini hanya kan membahas tentang akurasi dan *error* pada sistem. Akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar. Sedangkan *error* merupakan seberapa tidak akuratnya sistem dalam mengklasifikasikan data secara benar. tahapan-tahaan untuk menghitung nilai akurasi mencari nilai *true positive, true negative, false positive, dan false negative* kemudian menggunakan persamamaan 2.5 untuk menghitung nilai akurasi, sedangkan untuk menghitung nilai error menggunakan Persamaan 2.6. untuk *pseudecode* confusion matrikx bisa dilihat pada Gambar 3.17 sebagai berikut :

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

```
function [ b, OveAcc, Kappa, error, UserAcc, Akurasi ] =
AccuracyEvaluationByConfusionMatrix( datahasil, datasebenarnya,
jmlkelas )
  [m,n] = size(datasebenarnya);
  b = a(jmlkelas,jmlkelas);
  for i=1:m
      for j=1:n
          if(datasebenarnya(i,j)==0)
              continue;
          end
          t=datahasil(i,j);
          k=datasebenarnya(i,j);
          b(k,t)=b(k,t)+1
      end

  end
 [m,n] = size(b);
  r=m;
  rowsum=a(1,m);
  columnsum=a(1,n);
  N=0;
  Diagsum=0;
  error=a(m,1);
  UserAcc=a(m,1);
  Akurasi=a(2,1);
  for i=1:m
      for j=1:n
          rowsum(i)=rowsum(i)+b(i,j);
          columnsum(j)=columnsum(j)+b(i,j);
          N=N+b(i,j);
          if(i==j)
              Diagsum=Diagsum+b(i,i);
          end
      end
  end

  Pc=0;
  for i=1:r
      Pc=Pc+rowsum(i)*columnsum(i);
  end

  Kappa=(N*Diagsum-Pc)/(N*N-Pc);
  OveAcc=Diagsum/N;
 for i=1:m
     error(i)=b(i,i)/rowsum(i);
     UserAcc(i)=b(i,i)/columnsum(i);
 end
 errror(1)=sum(error)/jmlkelas*100
 Akurasi(2)=sum(UserAcc)/jmlkelas*100
end
```

Gambar 3.17 *Pseudocode Confusion Matrix*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.5 Desain *Interface*

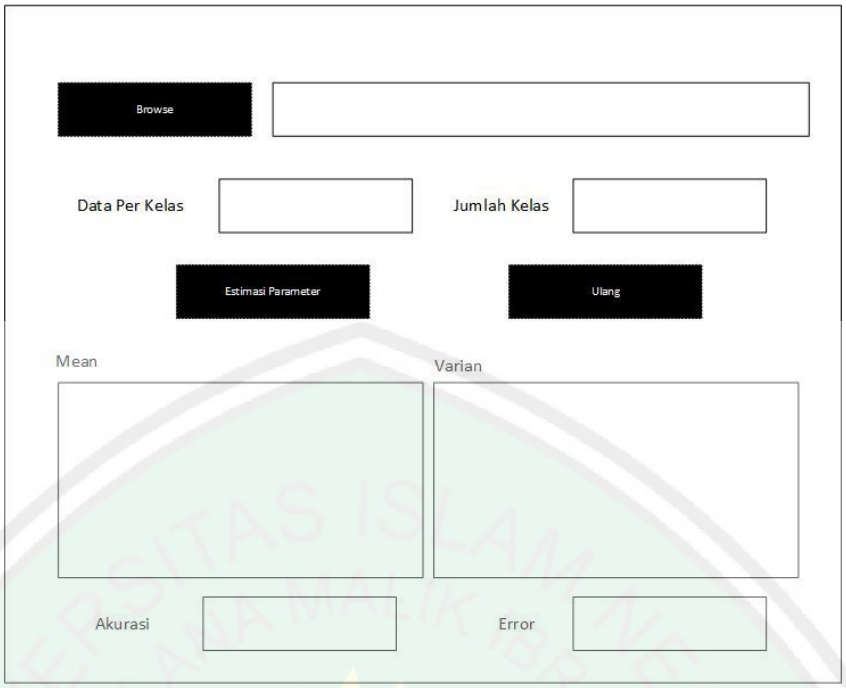
Perancangan sistem deteksi air sumur layak minum ini akan diterapkan dengan tampilan yang berbasis Matlab. Unruk sistem *testing* dan sistem *training* akan menggunakan platform *matlab*. Berikut adalah desain *interface* yang akan dibangun.

3.5.1 Form *Training*

Form *training* merupakan form yang digunakan untuk melatih image air dan parameter fisik air(ph, suhu, bau, dan rasa) sumur dengan output estimasi parameter meliputi *mean* dan *varian* tiap fitur yang direpresentasikan dalam bentuk tabel. Dalam form ini juga terdapat presentasi tingkat kesalahan dalam proses training.

Sedangkan untuk inputnya terdiri dari data mengenai jumlah sampel perkelas, dan jumlah kelas. fitur pertama yaitu mengambil gambar/citra air sumur yang ada di media penyimpanan untuk dilakukan training. Dan fitur utama selanjutnya adalah training sistem, yang berguna untuk memproses citra air sumur untuk diperoleh nilai dari warna tersebut beserta proses perhitungan nilai-nilai parameter air. Adapun desain interface Form *Training* adalah seperti yang tampak pada Gambar 3.18.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG



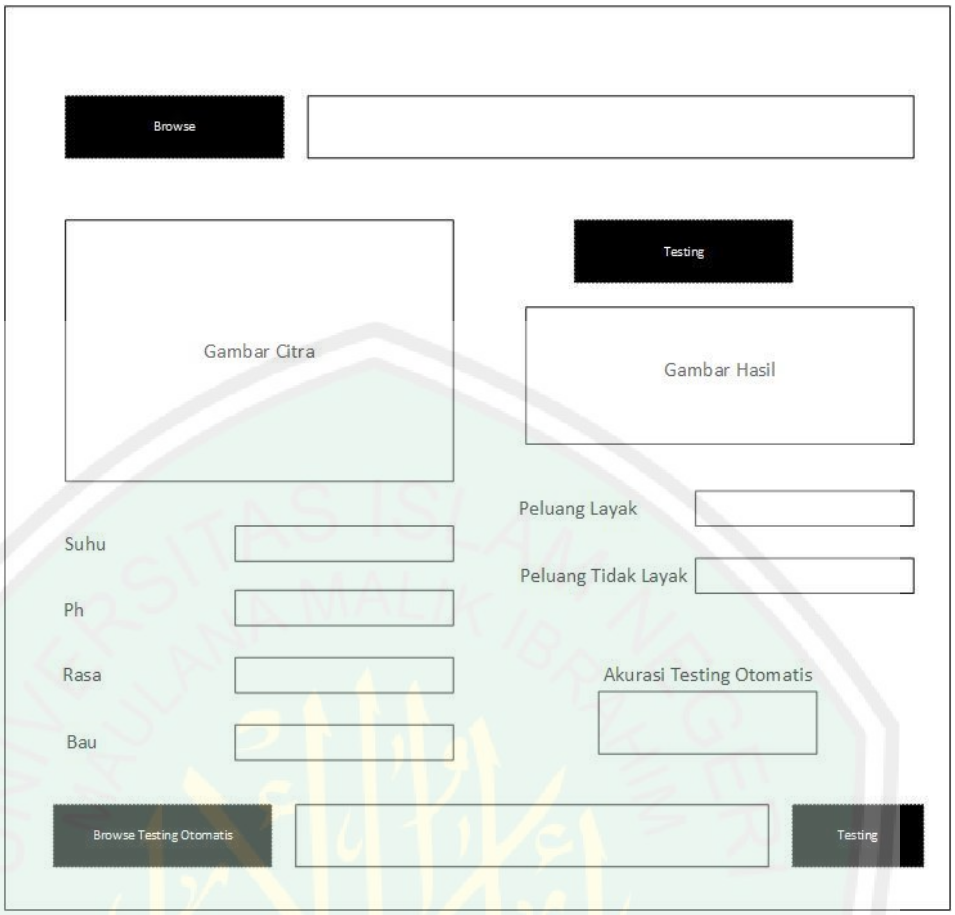
Gambar 3.18 ***Interface form training***

Bisa dilihat dari Gambar 3.18 bahwa untuk inputan hanya dengan image air saja sedangkan untuk parameter fisik lain seperti ph, suhu, bau, dan rasa langsung mengambil dari media penyimpanan secara otomatis. Sehingga tidak perlu memasukkan satu persatu data tersebut kedalam interface.

3.5.2 Form Testing

Form *testing* merupakan *form* yang digunakan untuk menguji data testing. Bisa dilihat pada Gambar 3.19.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG



Gambar 3.19 *Interface form testing*

Dari Gambar 3.19 bisa dilihat bahwa *desain interface form testing* memiliki dua fitur utama yaitu testing dengan data satu persatu dan fitur yang kedua yaitu testing dengan data otomatis. Untuk testing data satu persatu untuk inputannya sendiri membutuhkan gambar, inputan hasil suhu, ph, rasa, dan bau. Semuanya dimasukan secara manual pada sistem. Sedangkan untuk testing data otomatis cukup untuk memasukan direktori gambar saja. Dan untuk parameter suhu, ph, rasa, dan bau akan otomatis diambil dari media penyimpana oleh sitem.

3.6 Implementasi Sistem

Dalam proses pengaplikasianya sistem ini membutuhkan beberapa komponen. Apabila semua komponen pendukung aplikasi untuk mengidentifikasi

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

kualitas air layak minum pada air sumur terpasang kedalam perangkat mobile, maka langkah selanjutnya adalah mewujudkan rancangan sistem yang telah dibuat. Berikut ditunjukkan bagian terpenting yang akan diimplementasikan.

3.6.1 Implementasi Sistem

Dimulai sejak tahap penelitian hingga tahap implementasi dalam rancangan bangun aplikasi deteksi air sumur layak minum menggunakan perangkat komputer maupun mobile sebagai berikut :

A. *Hardware*

- *Central processing unit*
- *Random access memory* minimal 1GB
- Kamera minimal 13MP

B. *Software*

- *Matlab* 2014a
- *Microsoft excel*
- *Corel draw X7*
- *Photoshop*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

BAB IV

UJI COBA DAN PEMBAHASAN

Dalam bab ini akan dijelaskan tentang langkah-langkah dalam uji coba, hasil percobaan dan evaluasi dari penelitian yang telah dikerjakan. Hasil dari penelitain ini sendiri merupakan merupakan output dari sistem kualitas air layak konsumsi dan tidak. Dari penelitian diharapkan dapat memberikan sebuah pengetahuan serta memberikan kemudahan kepada seluruh masyarakat khususnya yang masih menggunakan air sumur sebagai air minum dan konsumsi keseharian mereka dalam menentukan air sumur layak dikonsumsi atau tidak secara parameter fisiknya.

4.1.Prosedur Pengujian

Prosedur pengujian dilakukan untuk mengintegrasikan metode yang telah dipilih dan digunakan dalam penelitian ini dengan langkah-langkah yang terstruktur sehingga akan diperoleh hasil uji coba yang bisa diukur tingkat akurasinya. Hal yang paling penting dalam strategi pengujian aplikasi adalah mendeskripsikan langkah-langkah yang akan dipakai dalam proses pengujian aplikasi. Langkah-langkah tersebut disusun secara baik dan terstruktur kemudian dijalankan sehingga dapat diperoleh nilai yang dibutuhkan untuk mengukur tingkat keberhasilan penelitian.

Berikut ini adalah beberapa tahapan dalam proses pengujian pada penelitian ini yang dapat diuraikan sebagai berikut.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

4.1.1 Persiapan Data

Tahap persiapan data ini merupakan tahapan dimana dan apa data yang akan digunakan. Data dalam penelitian ini menggunakan hasil pengambilan gambar obyek air sumur dan juga parameter fisik air sumur yang kemudian disimpan dalam direktori file pada perangkat komputer yang akan digunakan sebagai penguujian sistem dimana masing-masing data pengujian memiliki karakteristik masing-masing.

Total keseluruhan data yang akan digunakan dalam proses pengujian sistem ini adalah sebanyak 60 data yang terdiri dari dua kategori yaitu, 30 data air layak konsumsi dan 30 data air tidak layak konsumsi untuk proses *training*. Kemudian 10 data air layak konsumsi dan 10 data air tidak layak konsumsi untuk proses testing. Jadi jumlah data testing yaitu 20 data.

4.1.2 Preproses

Preproses merupakan tahapan paling awal untuk mempersiapkan pengolahan data pegujian (*testing)* yang kemudian akan dilakukan perhitungan dengan menggunakan metode *naive bayes classifier.* Secara garis besar dalam tahapan kali ini terdiri dari 3 proses. Yang pertama yaitu proses *cropping*, proses ini nantinya untuk menganbil bagian citra air yang tidak terpantulkan cahaya atau citra air yang sedikit pantulan cahayanya. Kemudian proses yang kedua yaitu *resize* citra, dalam proses ini nantinya citra air yang telah di potong untuk diambil bagian yang akan digunakan untuk uji data testing, dirubah ukuran *pixel* nya menjadi 200 x 200 sehingga akan memiliki ukuran yang sama. Yang ketiga adalah proses ekstraksi fitur, dalam hal ini fitur yang digunakan yaitu ekstraksi warna, dan model

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

ekstraksi warna yang digunakan yaitu RGB *(Red Green Blue)* sehingga citra air akan diambil nilai dari RGB *(Red Green Blue).*

Proses terakhir adalah estimasi parameter yaitu output dari tahapan ini menghasilkan nilai rata-rata (*mean)* dan *standart deviasi*dari RGB citra air, fitur ph dan fitur suhu, sedangkan untuk fitur bau dan rasa menggunakan probabilitas kemunculan dari setiap nilai. Nilai rata-rata, *standart deviasi,* dan probabilitas kemunculan tersebut nantinya akan digunakan sebagai landasan untuk menghitung peluang yang akan diperoleh masing-masing data kelas yang terdiri dari 2 kategori, yaitu kelas kualitas fisik air layak konsumsi dan kelas kualitas fisik air tidak layak konsumsi.

4.1.3 Proses Klasifikasi *Naive Bayes Classifier*

Hal utama dari proses klasifikasi adalah mengidentifikasi sebuah data citra air sumur dan pameter fisik air(suhu, ph, bau dan rasa) sumur sebagai anggota kelas air sumur yang layak dikonsumsi atau tidak layak untuk dikonsumsi melalui perhitungan probabilitas *naive bayes classifier*. Jika hasil probabilitas *naive bayes classifier* air sumur tersebut menunjukkan bahwa citra air sumur dan pameter fisik(suhu, ph, bau dan rasa) sumur tersebut termasuk dalam kelas air layak konsumsi maka akan teridentifikasi oleh sistem sebagai air sumur layak konsumsi, begitupun sebaliknya.

4.2. Hasil Uji Coba

Strategi *training* dan *testing* dilakukan untuk mengintegrasikan metode yang digunakan dalam penelitian ini ke dalam langkah-langkah terencana yang

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

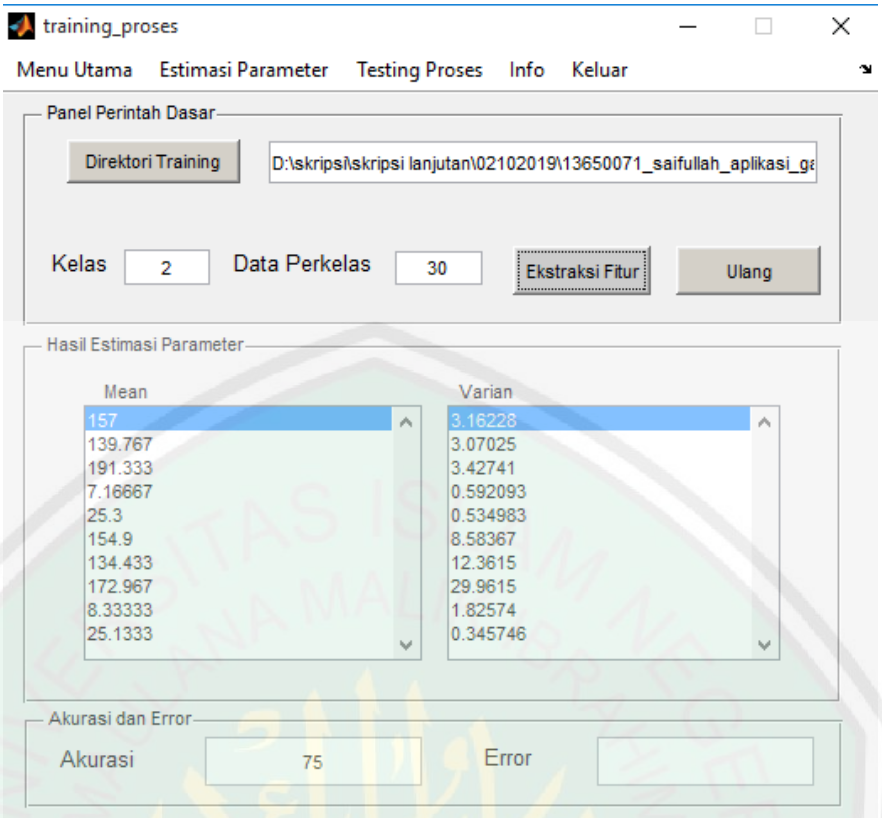
tersusun rapi sehingga diperoleh hasil yang dapat diukur tingkat akurasin ya. Yang terpenting dalam strategi pengujian aplikasi adalah mendeskripsikan langkah-langkah yang akan dipakai sebagai bagian dari proses pengujian. Langkah-langkah ini direncanakan dan kemudian dijalankan sehingga dapat diperoleh nilai yang dibutuhkan untuk mengukur tingkat keberhasilan penelitian

Setelah sistem selesai dirancang, selanjutnya harus diuji tingkat akurasi sistem dalam mengidentifikasi jenis kelayakan air sumur. Hal ini dilakukan untuk mengetahui sejauh mana sistem dapat bekerja dalam mendeteksi kelayaan fsik air sumur sebagai konsumsi. Sub bab ini akan membahas mengenai hasil dari sistem yang telah dirancang dan dibuat. Perincian akan diarahkan hasil dari proses *training* data air sumur dan proses *testing* data air sumur di aplikasi.

4.2.1. Hasil Proses *Training*

Dalam hasil proses *training* dilihatkan bagaimana pengaruh perubahan prosentase kesalahan dan akurasi yang dihasilkan oleh sistem ketika data dalam bentuk probabilitas kemunculan seperti bau dan rasa tidak diikutkan dalam perhitungan klasifikasi, dan ketika data dalam bentuk probabilitas kemunculan bau dan rasa dilibatkan dalam perhitungan klasifikasi. Dari 60 data yang telah di *training* untuk semua kelas, diperoleh hasil yang tidak sama dari hasil ujicoba tersebut. Untuk hasil *training* bisa dilihat pada Gambar 4.1.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG



Gambar 4.1 Hasil *Training* Data

Gambar 4.1 merupakan tampilan sistem *training* data yang memiliki tampilan sederhana. dalam sistem ini hanya perlu untuk memasukan folder gambar saja, sementara untuk beberapa parameter fisik lainya datanya langsung diambil dari media penyimpanan oleh sistem. Dalam tahap ini juga perlu memasukan jumlah kelas klasifikasi dan total seluruh data training. *Output* yang dikerluarkan dari proses *training* ini yaitu nilai *mean* dan *varian* dari setiap fitur yang ada. Dalam proses *training* data tersebut juga bisa dilihat akurasi data yang telah di *training.* Berikut adalah hasil ujicoba data *training* tanpa mengikutkan data probabilitas kemunculan fitur bau dan rasa ke perhitungan sistem bisa dilihat pada Table 4.1.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Tabel 4.1 Hasil uji coba data *training* tanpa fitur bau dan rasa

No. Data	Input	Output	Status
	Kualitas Air Sumur	**Kualitas Air Sumur**	
1	Layak Konsumsi	Layak Konsumsi	Benar
2	Layak Konsumsi	Layak Konsumsi	Benar
3	Layak Konsumsi	Layak Konsumsi	Benar
4	Layak Konsumsi	Layak Konsumsi	Benar
5	Layak Konsumsi	Layak Konsumsi	Benar
6	Layak Konsumsi	Tidak Layak Konsumsi	**Salah**
7	Layak Konsumsi	Layak Konsumsi	Benar
8	Layak Konsumsi	Layak Konsumsi	Benar
9	Layak Konsumsi	Layak Konsumsi	Benar
10	Layak Konsumsi	Layak Konsumsi	Benar
11	Layak Konsumsi	Layak Konsumsi	Benar
12	Layak Konsumsi	Layak Konsumsi	Benar
13	Layak Konsumsi	Tidak Layak Konsumsi	Salah
14	Layak Konsumsi	Layak Konsumsi	Benar
15	Layak Konsumsi	Layak Konsumsi	Benar
16	Layak Konsumsi	Layak Konsumsi	Benar
17	Layak Konsumsi	Layak Konsumsi	Benar
18	Layak Konsumsi	Layak Konsumsi	Benar
19	Layak Konsumsi	Layak Konsumsi	Benar
20	Layak Konsumsi	Layak Konsumsi	Benar
21	Layak Konsumsi	Layak Konsumsi	Benar
22	Layak Konsumsi	Layak Konsumsi	Benar
23	Layak Konsumsi	Layak Konsumsi	Benar
24	Layak Konsumsi	Layak Konsumsi	Benar
25	Layak Konsumsi	Layak Konsumsi	Benar
26	Layak Konsumsi	Layak Konsumsi	Benar
27	Layak Konsumsi	Layak Konsumsi	Benar
28	Layak Konsumsi	Layak Konsumsi	Benar
29	Layak Konsumsi	Layak Konsumsi	Benar
30	Layak Konsumsi	Layak Konsumsi	Benar
31	Tidak Layak Konsumsi	Layak Konsumsi	Salah
32	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
33	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
34	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
35	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
36	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
37	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
38	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
39	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

40	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
41	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
42	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
43	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
44	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
45	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
46	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
47	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
48	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
49	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
50	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
51	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
52	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
53	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
54	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
55	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
56	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
57	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
58	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
59	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
60	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**

Dari Tabel 4.1 bisa dilihat bahwa total data yang diuji berjumlah 60 data. dari 60 data tersebut 15 data mengalami kesalahan dalam proses akurasi sistem dan 45 data sukses diidentifikasi dengan benar. Untuk perhitungan akurasinya dilakukan menggunakan *confusion matrix.* Dapat dilihat pada Tabel 4.2.

Tabel 4.2 Confusion matrix *hasil data* training *tanpa fitur bau dan rasa*

	Terdeteksi Fisik Air Layak	**Terdeteksi Fisik Air Tidak Layak**
Fisik Air Layak	28	2
Fisik Air Tidak Layak	13	17

Confusion matrix pada Tabel 4.2 menunjukkan besaran *training* yang diidentifikasi dengan benar oleh sistem. Jumlah data *training* yang dipakai sebanyak 60 data air sumur. data air sumur yang memiliki kualitas fisik layak

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

komsumsi dan mampu diidentifikasi dengan benar adalah sebanyak 28 data. Jadi hampir dari total keseluruan data air sumur yang layak konsumsi mampu diidentifikasi dengan akurat, sedangkan kualitas air sumur yang tidak layak konsumsi dapat diidentifikasi dengan benar sebanyak 17 data.

Uji coba yang ke dua yaitu ujicoba menggunakan seluruh fitur data yaitu RGB *(Red, Green, dan Blue)*, ph, suhu, bau dan rasa, jadi semua fitur yang ada dalam klasifikasi fisik air layak minum ini akan diikutsertakan dalam proses. Untuk hasil dari uji coba data training dengan gabungan data tersebut bisa dilihat pada Tabel 4.3.

Tabel 4.3 Hasil uji coba data t*raining* dengan fitur bau dan rasa

No. Data	Data Input	Output	Status
	Kualitas Air Sumur	Kualitas Air Sumur	
1	Layak Konsumsi	Layak Konsumsi	Benar
2	Layak Konsumsi	Layak Konsumsi	Benar
3	Layak Konsumsi	Layak Konsumsi	Benar
4	Layak Konsumsi	Layak Konsumsi	Benar
5	Layak Konsumsi	Tidak Layak Konsumsi	**Salah**
6	Layak Konsumsi	Layak Konsumsi	Benar
7	Layak Konsumsi	Layak Konsumsi	Benar
8	Layak Konsumsi	Layak Konsumsi	Benar
9	Layak Konsumsi	Layak Konsumsi	Benar
10	Layak Konsumsi	Layak Konsumsi	Benar
11	Layak Konsumsi	Layak Konsumsi	Benar
12	Layak Konsumsi	Layak Konsumsi	Benar
13	Layak Konsumsi	Layak Konsumsi	Benar
14	Layak Konsumsi	Layak Konsumsi	Benar
15	Layak Konsumsi	Layak Konsumsi	Benar
16	Layak Konsumsi	Layak Konsumsi	Benar
17	Layak Konsumsi	Layak Konsumsi	Benar
18	Layak Konsumsi	Layak Konsumsi	Benar
19	Layak Konsumsi	Layak Konsumsi	Benar
20	Layak Konsumsi	Layak Konsumsi	Benar
21	Layak Konsumsi	Layak Konsumsi	Benar
22	Layak Konsumsi	Layak Konsumsi	Benar
23	Layak Konsumsi	Layak Konsumsi	Benar

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

24	Layak Konsumsi	Layak Konsumsi	Benar
25	Layak Konsumsi	Layak Konsumsi	Benar
26	Layak Konsumsi	Layak Konsumsi	Benar
27	Layak Konsumsi	Layak Konsumsi	Benar
28	Layak Konsumsi	Layak Konsumsi	Benar
29	Layak Konsumsi	Layak Konsumsi	Benar
30	Layak Konsumsi	Layak Konsumsi	Benar
31	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
32	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
33	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
34	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
35	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
36	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
37	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
38	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
39	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
40	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
41	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
42	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
43	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
44	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
45	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
46	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
47	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
48	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
49	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
50	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
51	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
52	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
53	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
54	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
55	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
56	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
57	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
58	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
59	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
60	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar

Dari Tabel 4.3 bisa dilihat bahwa dari 60 data tersebut hanya 5 data yang tidak bisa diidentifikasi dengan benar, Sisanya yaitu 55 data *training* mampu

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

diidentifikasi dengan benar. Untuk perhitungan akurasinya dilakukan menggunakan *confusion matrix*. Dapat dilihat pada Tabel 4.4.

Tabel 4.4 Confusion matrix hasil data training tanpa fitur bau dan rasa

	Terdeteksi Fisik Air Layak	**Terdeteksi Fisik Air Tidak Layak**
Fisik Air Layak	29	1
Fisik Air Tidak Layak	4	26

Confusion matrix pada Tabel 4.4 menunjukkan besaran *training* yang diidentifikasi dengan benar oleh sistem. Jumlah data *training* yang dipakai sebanyak 60 data air sumur. data air sumur yang memiliki kualitas fisik layak komsumsi dan mampu diidentifikasi dengan benar adalah sebanyak 29 data. Jadi hampir dari total keseluruan data air sumur yang layak konsumsi mampu diidentifikasi dengan akurat, sedangkan kualitas air sumur yang tidak layak konsumsi dapat diidentifikasi dengan benar sebanyak 26 data. hasil ini jauh lebih bagus dari pada hasil *training* yang hanya menggunakan fitur yang memiliki data kontinu saja.

Hasil output yang dimunculkan oleh sistem kemudian akan di cocokkan dengan data real. Berikut ini merupakan kinerja hasil uji coba proses *training* yang dijelaskan dengan tabel akurasi. Berikut disajikan tabel 4.5. hasil akurasi uji coba data *training* pada 20 data.

Tabel 4.5 Hasil akurasi data *training*

No	**Jenis Ujicoba**	**Akurasi**	***Error***
1	Fitur RGB*(Red, Green, Blue)*, ph, dan suhu	75.00%	25.00%
2	Fitur RGB*(Red, Green, Blue)*, ph, ,suhu, bau, dan rasa	91.66%	8.34%

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

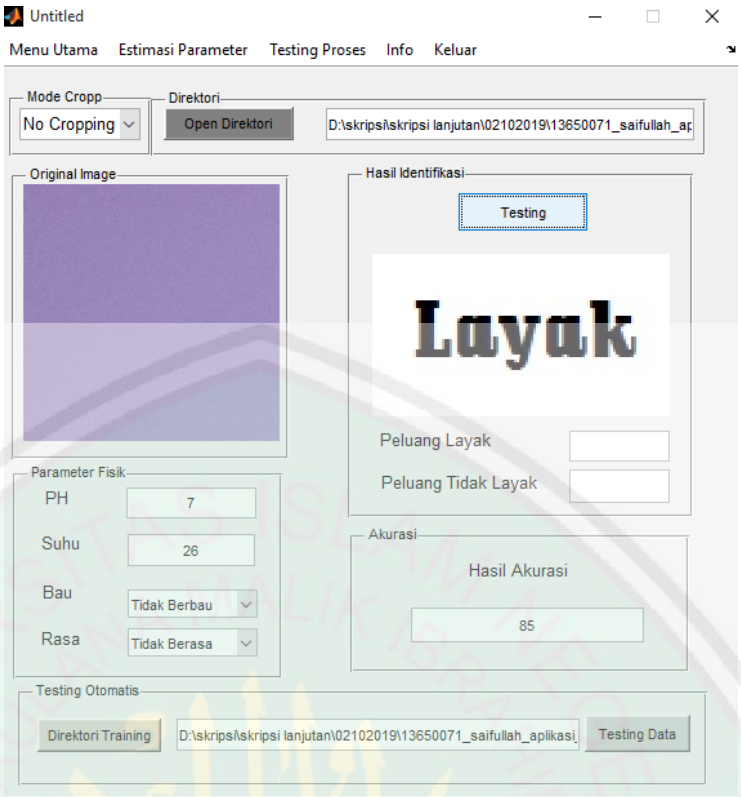
Dari Tabel 4.5 bisa dilihat hasil akurasi *training* dari sistem dengan tidak menggunakan fitur bau dan rasa dengan menggunakan semua fitur parameter fisik air memiliki perbedaan yang sangat tinggi sekali, akurasi dari penggunaan data fitur kontinu saja untuk klasifikasi kualitas fisik air hanya mampu sampai 75% dengan tingkat error 25% sedangkan gabungan antara semua data fitur memiliki tingkat akurasi 91.66% dan *error* 8.34%.

4.2.2. Hasil Proses *Testing*

Aplikasi sistem pendeteksi kualitasi fisik air untuk konsumsi sebagai alat pengidentifikasi kualitas fisik air sumur perlu dilakukan uji coba data *testing* untuk mengetahui seberapa besar kesalahan dan keakurasian perhitungannya. Jumlah data yang akan dipakai untuk proses *testing* ini berjumlah 20 data citra air beserta nilai parameter fisik air(suhu, ph, rasa, dan bau). Dari 20 data *testing* tersebut 10 diantaranya memiliki kelayakan fisik air untuk konsumsi dan 10 selebihnya tidak memiliki kelayakan fisik air untuk konsumsi. Setelah itu dilakukan proses *testing* dan dilihat seberapa tepat akurasi dan kesalahan pada sistem yang telah dibangun

Data yang diujikan untuk proses *testing* merupakan data yang belum dilatihkan pada sistem, jadi benar-benar data yang baru untuk pengujiaanya kali ini. Hasil identifikasi oleh sistem dipaparkan berdasarkan ujicoba menggunakan fitur dari data kontinu saja dan menggunakan fitur data gabungan yaitu kontinu dan diskrit. untuk hasil proses *testing* bisa dilihat pada Gambar 4.2.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG



Gambar 4.2 Hasil *testing* sistem

Dapat dilihat dari gambar 4.2 untuk proses *testing* memiliki dua cara yaitu cara yang pertama sistem manual dengan memasukkan data secara satu persatu kedalam sistem dan yang kedua yaitu cara otomatis dengan langsung mengambil data dari direktori penyimpanan sistem. Hasil uji coba data *testing* secara otomatis dengan uji coba data kontinu saja dapat dilihat pada Tabel 4.6.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Tabel 4.6 Hasil uji coba data *testing* tanpa fitur bau dan rasa

No. Data	Input	Output	Status
	Kualitas Air Sumur	**Kualitas Air Sumur**	
1	Layak Konsumsi	Layak Konsumsi	Benar
2	Layak Konsumsi	Layak Konsumsi	Benar
3	Layak Konsumsi	Layak Konsumsi	Benar
4	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
5	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
6	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
7	Layak Konsumsi	Layak Konsumsi	Benar
8	Layak Konsumsi	Layak Konsumsi	Benar
9	Layak Konsumsi	Layak Konsumsi	Benar
10	Layak Konsumsi	Layak Konsumsi	Benar
11	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
12	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
13	Layak Konsumsi	Layak Konsumsi	Benar
14	Layak Konsumsi	Layak Konsumsi	Benar
15	Layak Konsumsi	Tidak Layak Konsumsi	**Salah**
16	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
17	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
18	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
19	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
20	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**

Dapat dilihat dari tabel 4.6 dari 20 data yang dilakukan uji *testing* 12 data teruji dengan benar dan 8 data memiliki *output* yang salah. Untuk perhitungan akurasinya dilakukan menggunakan *confusion matrix*. Dapat dilihat pada tabel 4.7.

Tabel 4.7 *Confusion matrix* hasil data *testing* tanpa fitur bau dan rasa

	Terdeteksi Fisik Air Layak	**Terdeteksi Fisik Air Tidak Layak**
Fisik Air Layak	9	1
Fisik Air Tidak Layak	7	3

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Confusion matrix pada tabel 4.7 menunjukkan besaran *testing* yang diidentifikasi dengan benar oleh sistem. Jumlah data testing yang dipakai sebanyak 20 data air sumur. data air sumur yang memiliki kualitas fisik layak komsumsi dan mampu diidentifikasi dengan benar adalah sebanyak 9 data. Jadi dari total keseluruan data air sumur yang layak konsumsi hanya 1 data yang tidak mampu diidentifikasi dengan benar, sedangkan kualitas air sumur yang tidak layak konsumsi dapat diidentifikasi dengan benar sebanyak 3 data.

Uji coba *testing* yang ke dua yaitu uji coba yang melibatkan semua fitur fisik air termasuk bau dan rasa Hasil uji coba data *testing* secara otomatis dengan uji coba data semua fitur parameter fisik air dapat dilihat pada tabel 4.8.

Tabel 4.8 Hasil uji coba data *testing* dengan fitur bau dan rasa

No. Data	Data Input	Output	Status
	Kualitas Air Sumur	Kualitas Air Sumur	
1	Layak Konsumsi	Layak Konsumsi	Benar
2	Layak Konsumsi	Layak Konsumsi	Benar
3	Layak Konsumsi	Layak Konsumsi	Benar
4	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
5	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
6	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
7	Layak Konsumsi	Layak Konsumsi	Benar
8	Layak Konsumsi	Layak Konsumsi	Benar
9	Layak Konsumsi	Layak Konsumsi	Benar
10	Layak Konsumsi	Layak Konsumsi	Benar
11	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
12	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
13	Layak Konsumsi	Layak Konsumsi	Benar
14	Layak Konsumsi	Tidak Layak Konsumsi	**Salah**
15	Layak Konsumsi	Layak Konsumsi	Benar
16	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
17	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**
18	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
19	Tidak Layak Konsumsi	Tidak Layak Konsumsi	Benar
20	Tidak Layak Konsumsi	Layak Konsumsi	**Salah**

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Dapat dilihat dari Tabel 4.8 dari 20 data yang dilakukan uji *testing* 17 data teruji dengan benar dan 3 data memiliki *output* yang salah. Hal ini tentu saja memiliki perbedaan yang sangat jauh sekali dengan dengan testing yang tidak melibatkan fitur bau dan rasa. Untuk perhitungan akurasinya dilakukan menggunakan *confusion matrix.* Dapat dilihat pada tabel 4.9.

Tabel 4.9 *Confusion matrix* hasil data *testing* dengan fitur bau dan rasa

	Terdeteksi Fisik Air Layak	**Terdeteksi Fisik Air Tidak Layak**
Fisik Air Layak	**9**	1
Fisik Air Tidak Layak	**2**	8

Confusion matrix pada Tabel 4.9 menunjukkan besaran testing yang diidentifikasi dengan benar oleh sistem. Jumlah data testing yang dipakai sebanyak 20 data air sumur. data air sumur yang memiliki kualitas fisik layak komsumsi dan mampu diidentifikasi dengan benar adalah sebanyak 9 data dari 10 data. Jadi dari total keseluruan data air sumur yang layak konsumsi hanya 1 data yang tidak mampu diidentifikasi dengan benar, sedangkan kualitas air sumur yang tidak layak konsumsi dapat diidentifikasi dengan benar sebanyak 8 data dari 10 data. ini merupakan hasil terbaik selama uji coba testing ini dilakukan.

Hasil output yang dimunculkan oleh sistem kemudian akan di cocokkan dengan data real. Berikut ini merupakan kinerja hasil uji coba proses *testing* yang dijelaskan dengan tabel akurasi. Berikut disajikan tabel 4.10. hasil akurasi uji coba data *testing* pada 20 data.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Tabel 4.10 Hasil akurasi data *testing*

No	Jenis Ujicoba	Akurasi	*Error*
1	Fitur RGB*(Red, Green, Blue)*, ph, dan suhu	60.00%	40.00%
2	Fitur RGB*(Red, Green, Blue)*, ph, suhu, bau, dan rasa	85.00%	15.00%

Dari Tabel 4.10 bisa dilihat hasil akurasi testing dari sistem dengan menggunakan 20 data, ketika tidak menggunakan fitur bau dan rasa hanya memperoleh 60% akurasi sedangkan dengan menggunakan fitur bau dan rasa akurasi sangat tinggi mencapai sampai 85%.

4.3. Pembahasan

Berdasarkan hasil uji coba data *testing* menunjukkan bahwa secara umum data *image* air dan parameter fisik air dengan proses pengolahan data tidak menggunakan nilai probabilitas kemunculan bau dan rasa dengan menggunakan nilai probabilitas kemunculan bau dan rasa memiliki tingkat akurasi yang sangat berbeda. Terbukti dari Tabel 4.5 dan 4.6 sangat terlihat jelas perbedaan akurasinnya sangat tinggi, mencapai 10% dari masing-masing data *training* dan *testing*. Beberapa model data proses *testing* menghasilkan tingkat akurasi yang kurang optimal, hal ini disebabkan oleh berbagai faktor utama dalam proses akuisisi data. pertama adalah tidak fokusnya kamera dalam pengambilan citra air sehingga citra yang dihasilkan menjadi kabur/blur.

Kedua adalah masalah pencahayaan, dimana ini akan mempengaruhi intensitas cahaya yang dihasilkan dalam proses peneropongan/candling yaitu citra yang dihasilkan tidak sesuai dengan hasil yang diharapkan. Citra yang redup akibat dari penca//hayaan yang kurang akan mempengaruhi hasil dari proses tensting.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Pencahayaanya sendiri ada dari keempat sisi wadah, sehingga ketika salah satu pencahayaan mati akan menimbulkan sisi fokus gelap sebagian, karena proses pencahayaannya tidak langsung dari atas wadah. Tapi menggunakan pantulan cahaya dari ke empat sisi tersebut. Sehingga ke empat cahaya tersebut harus menyala dan memiliki kualitas terang yang sama.

Terakhir adalah masalah kebersihan wadah citra, ketika selesai menggunakan wadah untuk pengambilan gambar, wadah air harus dibersihkan dengan sangat bersih, karena kalau masih meninggalkan air yang terdahulu maka akan bercampur dengan sampel air yang baru dimasukan. Sehingga akan mempengarui kualitas warna citra yang baru dimasukan.

Kurangnya pengetahuan masyarakat dalam menentukan kualitas air sumur yang bersih dan bisa untuk dikomsumsi sehari-hari baik untuk memasak dan juga untuk minum memungkinkan tubuh dimasuki oleh senyawa yang berbahaya sehingga dapat menimbulkan penyakit dalam tubuh kita.

Sistem pendeteksi kualitas fisik air menggunakan beberapa fitur untuk pengklasifikasiannya salah satunya yaitu berdasarkan fitur warna. Dalam firman Allah surat *Ar-rum* ayat 22 yang berbunyi :

وَمِنْ ءَايَٰتِهِۦ خَلْقُ ٱلسَّمَٰوَٰتِ وَٱلْأَرْضِ وَٱخْتِلَٰفُ أَلْسِنَتِكُمْ وَأَلْوَٰنِكُمْ ۚ إِنَّ فِي ذَٰلِكَ لَءَايَٰتٍ لِّلْعَٰلِمِينَ ٢٢

Artinya : *"Dan di antara tanda-tanda kekuasaan-Nya ialah menciptakan langit dan bumi dan berlain-lainan bahasamu dan warna kulitmu. Sesungguhnya pada yang demikan itu benar-benar terdapat tanda-tanda bagi orang-orang yang mengetahui(QS.Ar-Rum/[30]:22)"(Al-Kalam, 2009).*

Al-Zuhaili (2009) dan Ali Al-Shoubuni(1891) ketika menafsirkan ayat di atas menyebutkan,bahwa salah satu tanda kekuasaan Allah. Dapat dilihat dari

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

berbagai macam ciptaanya (makhluk-Nya) yang semuanya berbeda-beda, mulai dari bentuk bahasa dan warna kulit, baik hitam, putih,coklat dan lain sebagainya. Dari keterangan ini dapat disimpulkan, bahwa pebedaan-perbedaan tersebut, termasuk warna berimplikasi terhadap perbedaan karakter, sifat dan lain sebagainya.

Maka dari ayat tersebut dapat dijadikan pijakan skripsi yang peneliti angkat bahwa perubahan warna pada air sumur menjadi salah satu indikator berkenaan kontaminasi zat-zat yang dapat membahayakan bagi tubuh manusia.

Selain itu, Allah SWT. berfirman dalam Al-Qur'an surat Al-An'am ayat 96 yang berbunyi :

فَالِقُ ٱلۡإِصۡبَاحِ وَجَعَلَ ٱلَّيۡلَ سَكَنٗا وَٱلشَّمۡسَ وَٱلۡقَمَرَ حُسۡبَانٗاۚ ذَٰلِكَ تَقۡدِيرُ ٱلۡعَزِيزِ ٱلۡعَلِيمِ ٩٦

Artinya : "Dia menyingsingkan pagi dan menjadikan malam untuk beristirahat, dan (Menjadikan) matahari dan bulan untuk perhitungan. Itulah ketetapan Allah yang maha perkasa, Maha mengetahui (Al-An'am[6]:96)" (Al-Kalam, 2009).

Maksut dari arti ayat *"dan (Menjadikan) matahari dan bulan untuk perhitungan"* dalam tafsir Ibnu Katsir adalah keduanya berjalan menurut perhitungan yang sempurna, terukur, tidak berubah, dan beraturan. Masing-masing dari keduanya memiliki orbit yang dilaluinya pada musim panas dan musim dingin,sehingga perjalanan itu menghasilkan pergantian malam dan siang berikut panjang dan pendeknya.

Kata *husbaanan* berarti perhitungan. Berdasarkan perspektif sains, untuk melakukan sebuah perhitungan selalu dikaitkan dengan metode yang digunakan. Fungsi dari metode tersebut adalah untuk mendapatkan hasil perhitungan yang

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

sempurna, akurat, dan beraturan. Metode *naive bayes classifier* merupakan metode yang digunakan untuk menghitung ketidak pastian data menjadi data yang jelas dengan cara membandingkan dengan data lainnya.

Adapun kekurangan sistem,antara lain yaitu metode *naive bayes classifier* merupakan sistem yang dapat berfungsi sebagai alat bantu.di dalam pengambilan keputusan masih terpengaruh terhadap faktor pendukung atau kebijakan-kebijakan lain. Selain itu hasil *testing* kualitas fisik air belum sama persis dengan data *real.*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

BAB V

PENUTUP

5.1 Kesimpulan

Dari hasil penelitian dan pembahasan tentang sistem yang menggunakan metode *naive bayes classifier* dalam mendeteksi kualitas fisik air dapat ditarik kesimpulan sebagai berikut.

Berdasarkan hasil uji coba pada data *training* fisik air, data yang digunakan adalah sebanyak 60 data dimana kesemua data tersebut dibagi menjadi 2 yakni 30 data air dengan parameter layak konsumsi dan 30 data air dengan parameter fisik air tidak layak konsumsi didapat tingkat akurasi terbaik adalah sebesar 91.66% dengan proses pengolah fitur warna RGB (*Red*, *Green* dan *Blue*), ph, suhu, bau dan rasa. Sedangkan uji coba dengan penggunaan fitur dengan data kontinyu saja memiliki akurasi sebesar 75%. Hasil uji coba data *testing* parameter fisik air dengan jumlah data 20 data air dengan parameter fisik layak konsumsi dan 30 data air dengan parameter fisik tidak layak konsumsi didapat tingkat akurasi terbaik adalah sebesar 85% dengan proses pengolah fitur warna RGB (*Red*, *Green* dan *Blue*), ph, suhu, bau dan rasa. Sedangkan uji coba dengan penggunaan fitur dengan data kontinyu saja memiliki akurasi sebesar 60%. Dari kedua percobaan tersebut diketahu bahwa gabungan antara fitur data kontinyu dan diskrit dalam perhitungan *naive bayes* untuk klasifikasi kualitas fisik air layak konsumsi sangat berpengaruh karena masing-masing dari parameter fisik air sangat penting.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

5.2 Saran

Untuk pengembangan sistem pendeteksi kelayakan fisik air untuk konsumsi, diperlukan beberapa perbaikan untuk mencapai hasil yang lebih optimal, diantaranya:

1. Memperbaiki data training pada data citra sehingga didapat akurasi yang semakin tinggi.
2. Menambahkan data training air sumur sehingga menambah keakuratan data.
2. Memperbaiki *prototype* akuisisi data yang lebih inovatif.
3. Implementasi sistem pendeteksi menggunakan sistem berbasis *mobile* yang lebih praktis.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

DAFTAR PUSTAKA

Achmad., R. 2004. *Kimia Lingkungan.* Andi, Yogyakarta.

Chandra, Budiman, Dr. 2005. *Pengantan Kesehatan Lingkungan*. Jakarta : Penerbit Buku Kedokteran EGC.

Dameshwari Sahu, Dewangan Chitesh. 2016. *Identification and Classification of Mango Fruits Using Image Processing*, International Journal of Scientific Research in Computer Science, Engineering and Information Technology Vol. 2. No. 2. ISSN. 2456-3307

Depkes, Permenkes RI, No. 492/MenKes/Per/IV/2010, *Persyaratan Kualitas Air Minum*.

Dong-Chul Park. 2016. *Image Classification Using Naive Bayes Classifier*, International Journal of Computer Science and Electronics Engineering Vol. 4. No. 3. ISSN. 2320-4028

Visa Sofia, Dkk. 2011. *Confusion Matrix-Based Feature Selection.* Conference Paper, ResearchGate.

Nugroho A dan Subahar. 2013. *Klasifikasi Naïve Bayes Untuk Prediksi Kelahiran Pada Data Ibu Hamil*. Berkala Mipa, Vol 23, Ed 3.

Fais, Sukma Nur, Muhammad Aditya D, Satria Mulya I, Donny Ramadien, and Askia Sani. 2013. *Klasifikasi Calon Pendonor Darah Dengan Metode Naive Bayes Clasifier*. Jurusan Teknologi Informasi dan Ilmu Komputer, Malang: Universitas Brawijaya.

Misigo Ronald. 2016. *Classification Of Selected Apple Fruit Varieties Using Naive Bayes*. Indian Journal of Computer Science and Engineering. Vol. 7. No. 1. ISSN. 0976-5166. pp. 13-19.

M.Zainuddin, Erna. *Penerapan Algoritma Naive Bayes Untuk Mengklasifikasikan Kelayakan Calon Nasabah Asuransi Di Pt.Bni Semarang.*

Mulia, R. M. 2005. *Kesehatan Lingkungan.* Yogyakarta : Graha Ilmu

Nugroho A dan Subahar. 2013. *Klasifikasi Naïve Bayes Untuk Prediksi Kelahiran Pada Data Ibu Hamil*. Berkala Mipa, Vol 23, Ed 3.

Rismawati Laila, Husaini, dan Khairiyati Laily. 2016. *Efektifitas Pengolahan Air Minum Ditinjau Dari Kualitas Air Minum Berdasarkan Parameter Fisik, Kimia, Dan Biologi di Ipa II Pinus Pdam Intan Banjar*, Jurnal Publikasi Kesehatan Masyarakat Indonesia, Vol. 3 No. 2.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Santra, A K, and C Josephine Christy. 2012. *Genetic Algorithm and Confusion Matrix for Document Clustering*. International Journal of Computer Science Vol. 9. No. 2. ISSN. 1694-0814

Santoso, Irwan Budi.2015. *Deteksi Boraks pada Bakso Berbasis Image dengan Menggunakan Gaussian Classifier,* Jurusan Teknik Informatika, Malang: Universitas Islam Negeri Maulana Malik Ibrahim.

Sasongko Endar Budi, Widyastuti Endang, dan Priyono Rawuh Edy. 2014. *Kajian Kualitas Air Dan Penggunaansumur Gali Oleh Masyarakatdisekitar Sungai Kaliyasa Kabupaten Cilacap,* Universitas Diponegoro. ISSN. 1829-8907.

Sutrisno, T., 2004. *Teknologi Penyediaan Air Bersih*, PT Rineka Cipta, Jakarta.

Webb, Andrew R, dan Copsey Keith D. (2011), *Statistical Pattern Recognition, Third Edition,* Mathematics and Data Analysis Consultancy Malvern, United Kingdom: John Wiley & Sons, Ltd

Wibowo, Jati Sasongko. 2011. *Deteksi dan Klasifikasi Citra Berdasarkan Warna Kulit Menggunakan HSV,* Jurnal Teknologi Informasi DINAMIK Vol. 16. No. 2. ISSN. 0854-9524. Pp. 118-123.

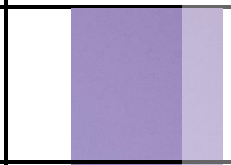
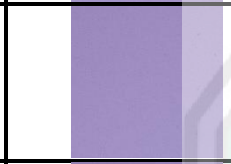
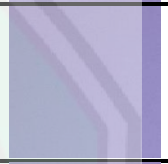
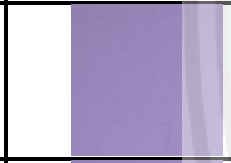
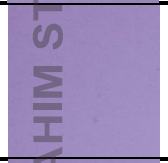
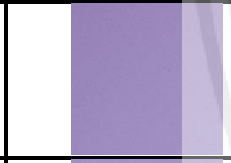
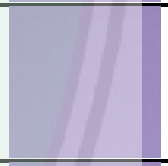
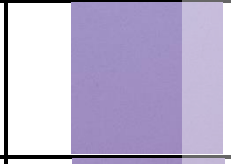
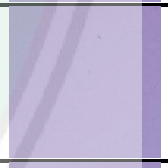
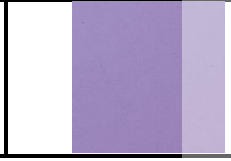
Yustika Fuzy, Saputra Kana. *Klasifikasi Blimbing Menggunakan Naive Bayes Berdasarkan Fitur Warna RGB*, STMIK Kaputama. ISSN. 1978-1520

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

LAMPIRAN - LAMPIRAN

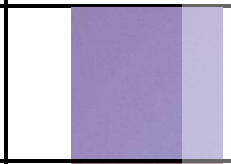
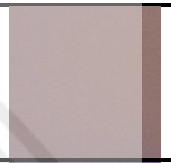
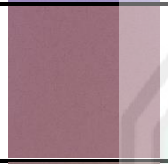
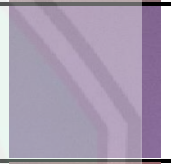
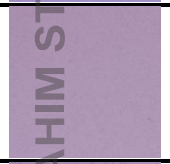
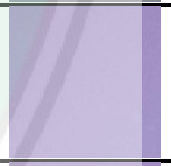
LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF M

1. **Data training Image air layak secara fisik**

No. Data	Citra	No. Data	Citra	No. Data	Citra	No. Data	Citra	No. Data	Citra
1		7		13		19		25	
2		8		14		20		26	
3		9		15		21		27	
4		10		16		22		28	
5		11		17		23		29	
6		12		18		24		30	

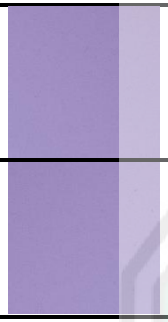
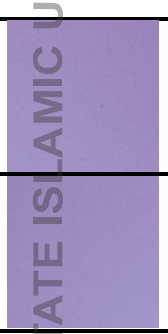
LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF

2. Data training Image air tidak layak secara fisik

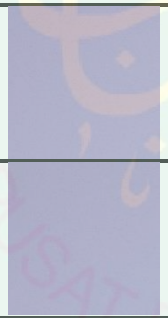
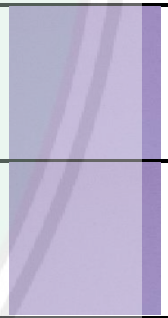
No. Data	Citra	No. Data	Citra	No. Data	Citra	No. Data	Citra	No. Data	Citra
1		7		13		19		25	
2		8		14		20		26	
3		9		15		21		27	
4		10		16		22		28	
5		11		17		23		29	
6		12		18		24		30	

LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF M

3. **Data testing Image air layak secara fisik**

No. Data	Citra	No. Data	Citra	No. Data	Citra	No. Data	Citra	No. Data	Citra
1		3		5		7		9	
2		4		6		8		10	

4. **Data testing Image air tidak layak secara fisik**

No. Data	Citra	No. Data	Citra	No. Data	Citra	No. Data	Citra	No. Data	Citra
1		3		5		7		9	
2		4		6		8		10	

LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF

5. Data training parameter fisik air layak konsumsi

No	Data	Bau	Rasa	PH	Suhu
1	Data 1	Tidak Berbau	Tidak Berasa	7	25
2	Data 2	Tidak Berbau	Tidak Berasa	7	25
3	Data 3	Tidak Berbau	Tidak Berasa	7	25
4	Data 4	Tidak Berbau	Tidak Berasa	7	25
5	Data 5	Berbau	Berasa	7	25
6	Data 6	Tidak Berbau	Tidak Berasa	8	25
7	Data 7	Tidak Berbau	Tidak Berasa	7	25
8	Data 8	Tidak Berbau	Tidak Berasa	7	25
9	Data 9	Tidak Berbau	Tidak Berasa	8	25
10	Data 10	Tidak Berbau	Tidak Berasa	8	26
11	Data 11	Tidak Berbau	Tidak Berasa	7	25
12	Data 12	Tidak Berbau	Tidak Berasa	7	25
13	Data 13	Tidak Berbau	Tidak Berasa	7	25
14	Data 14	Tidak Berbau	Tidak Berasa	7	25
15	Data 15	Tidak Berbau	Tidak Berasa	7	25
16	Data 16	Tidak Berbau	Tidak Berasa	7	25
17	Data 17	Tidak Berbau	Tidak Berasa	7	25
18	Data 18	Tidak Berbau	Tidak Berasa	8	25
19	Data 19	Tidak Berbau	Tidak Berasa	7	25
20	Data 20	Tidak Berbau	Tidak Berasa	7	25
21	Data 21	Tidak Berbau	Tidak Berasa	8	25
22	Data 22	Tidak Berbau	Tidak Berasa	7	25
23	Data 23	Tidak Berbau	Tidak Berasa	7	25
24	Data 24	Tidak Berbau	Tidak Berasa	7	25
25	Data 25	Tidak Berbau	Tidak Berasa	7	25
26	Data 26	Tidak Berbau	Tidak Berasa	7	26
27	Data 27	Tidak Berbau	Tidak Berasa	7	25
28	Data 28	Tidak Berbau	Tidak Berasa	7	25
29	Data 29	Tidak Berbau	Tidak Berasa	7	25
30	Data 30	Tidak Berbau	Tidak Berasa	7	24

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

6. Data training parameter fisik air tidak layak konsumsi

No	Data	Bau	Rasa	Ph	Suhu
31	Data 31	Berbau	Tidak Berasa	8	25
32	Data 32	Berbau	Tidak Berasa	9	25
33	Data 33	Tidak Berbau	Berasa	9	25
34	Data 34	Berbau	Berasa	11	25
35	Data 35	Berbau	Berasa	12	26
36	Data 36	Berbau	Berasa	10	26
37	Data 37	Berbau	Berasa	13	25
38	Data 38	Tidak Berbau	Tidak Berasa	7	25
39	Data 39	Tidak Berbau	Tidak Berasa	7	25
40	Data 40	Tidak Berbau	Tidak Berasa	7	25
41	Data 41	Tidak Berbau	Tidak Berasa	7	25
42	Data 42	Tidak Berbau	Tidak Berasa	7	25
43	Data 43	Tidak Berbau	Tidak Berasa	7	25
44	Data 44	Berbau	Berasa	10	25
45	Data 45	Berbau	Berasa	8	25
46	Data 46	Tidak Berbau	Berasa	8	25
47	Data 47	Tidak Berbau	Berasa	8	26
48	Data 48	Tidak Berbau	Berasa	7	26
49	Data 49	Tidak Berbau	Berasa	8	25
50	Data 50	Tidak Berbau	Berasa	7	25
51	Data 51	Berbau	Berasa	8	25
52	Data 52	Berbau	Berasa	8	25
53	Data 53	Berbau	Berasa	7	25
54	Data 54	Berbau	Berasa	7	25
55	Data 55	Berbau	Berasa	7	25
56	Data 56	Berbau	Berasa	7	25
57	Data 57	Tidak Berbau	Tidak Berasa	7	25
58	Data 58	Tidak Berbau	Tidak Berasa	7	25
59	Data 59	Berbau	Berasa	13	25
60	Data 60	Berbau	Berasa	9	25

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

7. Data testing parameter fisik air layak konsumsi

No	Data	Bau	Rasa	PH	Suhu
1	Data 1	Tidak Berbau	Tidak Berasa	7	25
2	Data 2	Tidak Berbau	Tidak Berasa	7	25
3	Data 3	Tidak Berbau	Tidak Berasa	7	25
4	Data 4	Tidak Berbau	Tidak Berasa	7	25
5	Data 5	Berbau	Berasa	7	25
6	Data 6	Tidak Berbau	Tidak Berasa	8	25
7	Data 7	Tidak Berbau	Tidak Berasa	7	25
8	Data 8	Tidak Berbau	Tidak Berasa	7	25
9	Data 9	Tidak Berbau	Tidak Berasa	8	25
10	Data 10	Tidak Berbau	Tidak Berasa	8	26

8. Data testing parameter fisik air tidak layak konsumsi

No	Data	Bau	Rasa	PH	Suhu
1	Data 1	Tidak Berbau	Tidak Berasa	7	25
2	Data 2	Tidak Berbau	Tidak Berasa	7	25
3	Data 3	Tidak Berbau	Tidak Berasa	8	25
4	Data 4	Berbau	Berasa	7	25
5	Data 5	Tidak Berbau	Tidak Berasa	7	25
6	Data 6	Berbau	Berasa	8	25
7	Data 7	Berbau	Berasa	8	26
8	Data 8	Berbau	Berasa	7	26
9	Data 9	Tidak Berbau	Berasa	8	25
10	Data 10	Tidak Berbau	Berasa	7	25

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG