

**IMPLEMENTASI METODE K-NEAREST NEIGHBOR DALAM
MEMPREDIKSI BAU, RASA DAN WARNA AIR SUMUR
DI KELURAHAN SELAT BARU BENGKALIS**

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Teknik Pada Fakultas Teknik
Universitas Islam Riau



OLEH:

CITRA YUNELA SARI
153510266

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS ISLAM RIAU
PEKANBARU

2019

LEMBAR PENGESAHAN PEMBIMBING SKRIPSI

Nama : Citra Yunela Sari
NPM : 153510266
Fakultas : Teknik
Program Studi : Teknik Informatika
Jenjang Pendidikan : Strata Satu (S1)
Judul Skripsi : Implementasi Metode K-Nearest Neighbor Dalam
Meprediksi Bau, Rasa dan Warna Air Sumur di Kelurahan
Selat Baru Bengkalis

Format sistematika dan pembahasan materi pada masing-masing bab dan sub bab dalam skripsi ini telah dipelajari dan dinilai relatif telah memenuhi ketentuan-ketentuan dan kriteria - kriteria dalam metode penulisan ilmiah. Oleh karena itu, skripsi ini dinilai layak dapat disetujui untuk disidangkan dalam ujian komprehensif.

Pekanbaru, 13 Desember 2019

Disetujui Oleh

Dosen Pembimbing

Ir. Hj. DES SURYANI, M.Sc

Disahkan Oleh :

Dekan Fakultas Teknik

Ketua Prodi Teknik Informatika
ari. selapudi

Ir. H. ABD. KUDUS ZAINI, MT.,MS., TR
NPK : 88 03 02 098

AUSE LABELLAPANSA, ST., M.Cs., M.Kom

**LEMBAR PENGESAHAN
TIM PENGUJI UJIAN SKRIPSI**

Nama : Citra Yunela Sari
NPM : 153510266
Fakultas : Teknik
Program Studi : Teknik Informatika
Jenjang Pendidikan : Strata Satu (S1)
Judul Skripsi : Implementasi Metode K-Nearest Neighbor Dalam
Memprediksi Bau, Rasa dan Warna Air Sumur di Kelurahan
Selat Baru Bengkalis

Skripsi ini secara keseluruhan dinilai telah memenuhi ketentuan-ketentuan dan kaidah-kaidah dalam penulisan penelitian ilmiah serta telah diuji dan dapat dipertahankan dihadapan tim penguji. Oleh karena itu, Tim Penguji Ujian Skripsi Fakultas Teknik Universitas Islam Riau menyatakan bahwa mahasiswa yang bersangkutan dinyatakan **Telah Lulus Mengikuti Ujian Komprehensif Pada Tanggal 13 Desember 2019** dan disetujui serta diterima untuk memenuhi salah satu syarat guna memperoleh gelar Sarjana Strata Satu Bidang Ilmu **Teknik Informatika.**

Pekanbaru, 13 Desember 2019

Tim Penguji

- | | | |
|--------------------------------|------------------------|---------|
| 1. Nesi Syafitri, S.Kom., M.Cs | Sebagai Tim Penguji I | (.....) |
| 2. Ana Yulianti, ST., M.Kom | Sebagai Tim Penguji II | (.....) |

**Disetujui Oleh
Dosen Pembimbing**


Ir. Hj. DES SURYANI, M.Sc

Disahkan Oleh :

Dekan Fakultas Teknik

Ketua Prodi Teknik Informatika
an. seherandi

Ir. H. ABD. KUDUS ZAINI, MT.,MS., TR
NPK : 88 03 02 098

AUSE LABELLAPANSA, ST., M.Cs., M.Kom

LEMBAR PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan dibawah ini:

Nama : Citra Yunela Sari
Tempat/Tgl Lahir : Pekanbaru, 6 November 1997
Alamat : Jln. Bandeng 1 No. 307 Perumnas Rumbai

Adalah mahasiswa Universitas Islam Riau yang terdaftar pada :

Fakultas : Teknik
Program Studi : Teknik Informatika
Jenjang Pendidikan : Strata-1 (S1)

Dengan ini menyatakan dengan sesungguhnya bahwa skripsi yang saya tulis adalah benar dan asli hasil dari penelitian yang telah saya lakukan dengan judul **"Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa dan Warna Air Sumur Di Kelurahan Selat Baru Bengkalis"**.

Apabila dikemudian hari ada yang merasa dirugikan dan atau menuntut karena penelitian ini menggunakan sebagian hasil tulisan atau karya orang lain tanpa mencantumkan nama penulis yang bersangkutan, atau terbukti karya ilmiah ini bukan karya saya sendiri atau **plagiat** hasil karya orang lain, maka saya bersedia menerima sanksi sesuai dengan peraturan perundangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya untuk dapat digunakan sebagaimana mestinya.

Pekanbaru, 18 Desember 2019

Yang membuat pernyataan,



Citra Yunela Sari

SURAT KEPUTUSAN DEKAN FAKULTAS TEKNIK UNIVERSITAS ISLAM RIAU
NOMOR : 0142 /KPTS/FT-UIR/2019
TENTANG PENGANGKATAN TIM PEMBIMBING PENELITIAN DAN PENYUSUNAN SKRIPSI

DEKAN FAKULTAS TEKNIK

- Membaca : Surat Ketua Program Studi Teknik Informatika Nomor : 009 /TA/TI/FT/2019 tentang persetujuan dan usulan pengangkatan Tim Pembimbing penelitian dan penyusunan Skripsi.
- Menimbang : 1. Bahwa untuk menyelesaikan perkuliahan bagi mahasiswa Fakultas Teknik perlu membuat Skripsi.
2. Untuk itu perlu ditunjuk Tim Pembimbing penelitian dan penyusunan Skripsi yang diangkat dengan Surat Keputusan Dekan.
- Mengingat : 1. Undang-Undang Nomor 20 Tahun 2003
2. UU Nomor 14 Tahun 2005 Tentang Guru Besar
3. UU Nomor 12 Tahun 2012 Tentang Pendidikan Tinggi
4. PP Nomor 4 Tahun 2014 Tentang Penyelenggaraan Pendidikan Tinggi dan Pengelolaan Perguruan Tinggi
5. Permenristek Dikti Nomor 44 Tahun 2015 Tentang Standar Nasional Pendidikan Tinggi
6. Permenristek Dikti Nomor 32 Tahun 2016 Tentang Akreditasi Prodi dan Perguruan Tinggi.
7. SK.Ban PT.Nomor : 2777/SK/BAN -- PT/Ared /S/X/2018
8. Statuta Universitas Islam Riau Tahun 2013
9. SK Rektor Universitas Islam Riau Nomor :112 /UIR/Kpts/2016

MEMUTUSKAN

- Menetapkan : 1. Mengangkat saudara-saudara yang namanya tersebut dibawah ini sebagai Tim Pembimbing Penelitian dan penyusunan Skripsi mahasiswa Fakultas Teknik Program Studi Teknik Informatika.

| No | Nama | Pangkat | Jabatan |
|----|---------------------|---------------|------------|
| 1 | Ir.Des Suryani.,MSc | Lektor Kepala | Pembimbing |

2. Mahasiswa yang akan dibimbing :

Nama : Citra Yuncia Sari
NPM : 153510266
Program Studi : Informatika
Jenjang Pendidikan : Strata Satu (S1)
Judul Skripsi : Implementasi Metode K- Nearest Neighbor Dalam Memprediksi Bau, Rasa dan Warna Air Sumur di Kelurahan Selat Baru Bengkalis .

3. Keputusan ini mulai berlaku pada tanggal ditetapkannya dengan ketentuan bila terdapat kekeliruan dikemudian hari segera ditinjau kembali.



Ditetapkan di : Pekanbaru
Pada Tanggal : 18 Jumadil Akhir 1440 H
23 Februari 2019 M

Ir. H. Abd. Kudus Zaini, MT., MS., TR
NPK : 88 03 02-098

Tembusan disampaikan :

1. Yth. Bapak Rektor UIR di Pekanbaru.
2. Yth. Sdr. Ketua Program Studi Teknik Informatika FT-UIR
3. Yang Bersangkutan .
4. Arsip



PROGRAM STUDI TEKNIK INFORMATIKA

Jln. Kaharuddin Nasution no.113, Perhentian Marpoyan, Pekanbaru-Riau 28284

Telp: 0761-674674, fax: 0761-674834

KARTU KONSULTASI BIMBINGAN SKRIPSI

Semester : Genap TA : 2019/2020

NAMA MAHASISWA : CITRA YUNELA SARI

NPM : 153510266

DOSEN PEMBIMBING : Ir. Des Suryani., M.Sc

TANGGAL PERSETUJUAN : 23 Februari 2019

JUDUL : Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa dan Warna Air Sumur di Kelurahan Selat Baru Bengkalis.

Tanda Tangan Mhs

Perpustakaan Universitas Islam Riau

Dokumen ini adalah Arsip Mhik :

| No | Tanggal | Keterangan | Tanda Tangan Pembimbing |
|-----|----------------|--|-------------------------|
| 1. | 05 - 03 - 2019 | Perbaikan Bab I & II | [Signature] |
| 2. | 20 - 03 - 2019 | Langitau perbaikan Bab I & II. Langitau Bab III | [Signature] |
| 3. | 02 - 04 - 2019 | Perbaikan DPD | [Signature] |
| 4. | 04 - 04 - 2019 | Perbaikan Disain I/O | [Signature] |
| 5. | 09 - 05 - 2019 | Langitau perbaikan disain I/O | [Signature] |
| 6. | 14 - 05 - 2019 | Langitau Propona skripsi | [Signature] |
| 7. | 15 - 05 - 2019 | Acc ^{ulang} ulang proposal skripsi I | [Signature] |
| 8. | 03 - 09 - 2019 | Perbaikan DPD - (tambahan) | [Signature] |
| 9. | 04 - 09 - 2019 | Perbaikan Logika program. | [Signature] |
| 10. | 08 - 10 - 2019 | Tambahan data training | [Signature] |
| 11. | 18 - 10 - 2019 | Perbaikan Bab IV | [Signature] |
| 12. | 01 - 10 - 2019 | Langitau ke kecermatan & Bab V | [Signature] |
| 13. | 06 - 11 - 2019 | Perbaikan Bab V | [Signature] |
| 14. | 07 - 11 - 2019 | Acc Laporan hasil. | [Signature] |
| | | | |
| | | | |

Pekanbaru, 1 Maret 2019

Dosen Pembimbing I,

[Signature]

(Ir. Des Suryani., M.Sc)

SURAT KEPUTUSAN DEKAN FAKULTAS TEKNIK UNIVERSITAS ISLAM RIAU
NOMOR : 1628/KPTS/FT-UIR/2019
TENTANG PENETAPAN DOSEN PENGUJI SKRIPSI MAHASISWA FAK. TEKNIK UNIV. ISLAM RIAU

DEKAN FAKULTAS TEKNIK

- Menimbang** : 1. Bahwa untuk menyelesaikan studi S.1 bagi mahasiswa Fakultas Teknik Univ. Islam Riau dilaksanakan Ujian Skripsi/Komprehensif sebagai tugas akhir. Untuk itu perlu ditetapkan mahasiswa yang telah memenuhi syarat untuk ujian dimaksud serta dosen penguji.
2. Bahwa penetapan mahasiswa yang memenuhi syarat dan dosen penguji yang bersangkutan perlu ditetapkan dengan Surat Keputusan Dekan.
- Mengingat** : 1. Undang-Undang Nomor : 20 tahun 2003 tentang Pendidikan Nasional
2. UU No. 14 Tahun 2005 Tentang Guru Besar
3. UU Nomor 12 Tahun 2012 Tentang Pendidikan Tinggi
4. PP Nomor 4 Tahun 2014 Tentang Penyelenggaraan Pendidikan Tinggi dan Pengelolaan Perguruan Tinggi
5. Permenristek Dikti Nomor 44 Tahun 2015 Tentang Standar Nasional Pendidikan Tinggi
6. Permenristek Dikti Nomor 32 Tahun 2016 Tentang Akreditasi Prodi dan Perguruan Tinggi
7. SK. BAN-PT Nomor : 2777/SK/BAN-PT/Ared/S/X/2018
8. Statuta Universitas Islam Riau Nomor : 112/UIR/kpts/2016

MEMUTUSKAN

- Menetapkan:**
- Mahasiswa Fakultas Teknik Universitas Islam Riau yang tersebut namanya dibawah ini :
Nama : Citra Yunela Sari
NPM : 153510266
Program Studi : Informatika
Jenjang Pendidikan : Strata Satu (S1)
Judul Skripsi : Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa dan Warna Air Sumur Dikelurahan Selat Baru Bengkalis
 - Penguji Skripsi Komprehensif mahasiswa tersebut terdiri dari :
 - Jr. Des Suryani, M.Sc Sebagai Ketua Merangkap Penguji
 - Nesi Syafitri, S.Kom., M.Cs Sebagai Anggota Merangkap Penguji
 - Ana Yulianti, ST., M.Kom Sebagai Anggota Merangkap Penguji
 - Laporan hasil ujian serta berita acara telah sampai kepada Pimpinan Fakultas selambat-lambatnya 1(satu) bulan setelah ujian dilaksanakan.
 - Keputusan ini mulai berlaku pada tanggal ditetapkannya dengan ketentuan bila terdapat kekeliruan dikemudian hari segera ditinjau kembali.
- KUTIPAN** : Disampaikan kepada yang bersangkutan untuk dapat dilaksanakan dengan sebaik-baiknya.



Ditetapkan di Pekanbaru
Pada Tanggal : 15 Rabi'ul Akhir 1441 H
12 Desember 2019 M

Dekan,

Ir. H. Abd. Kudus Zaini, MT., MS., TR., IPM
NPK. 88 03 02 098

Tembusan disampaikan :

- Yth. Rektor UIR di Pekanbaru.
- Yth. Ketua Program Studi Informatika FT-UIR.
- Yth. Pembimbing dan Penguji Skripsi.
- Mahasiswa yang bersangkutan.
- Arsip.



YAYASAN LEMBAGA PENDIDIKAN ISLAM (YLPI) RIAU
UNIVERSITAS ISLAM RIAU
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA

Jalan Kaharuddin Nasution No. 113 P. Marpoyan Pekanbaru Riau Indonesia – Kode Pos: 28284
 Telp. +62 761 674674 Website: www.eng.uir.ac.id Email: fakultas_teknik@uir.ac.id

BERITA ACARA UJIAN SKRIPSI

Berdasarkan Surat Keputusan Dekan Fakultas Teknik Universitas Islam Riau, Pekanbaru, tanggal 12 Desember 2019, Nomor: 1628/KPTS/FT-UIR/2019, maka pada hari Jumat, tanggal 13 Desember 2019, telah dilaksanakan Ujian Skripsi Program Studi Informatika Fakultas Teknik Universitas Islam Riau, Jenjang Studi S1, Tahun Akademik 2019/2020 berikut ini.

1. Nama : Citra Yunela Sari
2. NPM : 153510266
3. Judul Skripsi : Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa dan Warna Air Sumur Dikelurahan Selat Baru Bengkalis
4. Waktu Ujian : 11.00 WIB – Selesai
5. Tempat Pelaksanaan Ujian : Ruang Sidang Fakultas Teknik UIR

Dengan keputusan Hasil Ujian Skripsi:

~~Lulus*~~ / Lulus dengan Perbaikan* / Tidak Lulus*

* Coret yang tidak perlu.

Nilai Ujian:

Nilai Ujian Angka = ~~85,44~~ Nilai Huruf = ~~.....~~ **A-**

Tim Penguji Skripsi.

| No | Nama | Jabatan | Tanda Tangan |
|----|------------------------------|---------|--------------|
| 1 | Ir. Des Suryani, M.Sc | Ketua | 1. |
| 2 | Nesi Syaifitri, S.Kom., M.Cs | Anggota | 2. |
| 3 | Ana Yulianti, ST., M.Kom | Anggota | 3. |

Panitia Ujian
 Ketua,

Ir. Des Suryani, M.Sc
 NIDN. 102126801

Pekanbaru, 13 Desember 2019

Mengetahui,

Dekan Fakultas Teknik

Ir. H. Abd. Kudus Zamri, MT., MS., TR., IPM
 NIDN. 1011076202

LEMBAR IDENTITAS PENULIS

NPM : 153510266

Nama Lengkap : Citra Yunela Sari

Tempat, Tgl Lahir : Pekanbaru, 6 November 1997

Alamat : Jln. Bandeng 1 No 307 Perumnas Rumbai

Nama Ayah : Yurizal

Nama Ibu : Nelvisda

Nomor Handphone : 0812-7088-9058

Fakultas : Teknik

Program Studi : Teknik Informatika

Judul Skripsi : Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa dan Warna Air Sumur di Kelurahan Selat Baru Bengkalis

Tahun Masuk : 2015

Tahun Lulus : 2019

Pekanbaru, 18 Desember 2019

Citra Yunela Sari
153510266

DAFTAR RIWAYAT HIDUP

1. Data Personal

NPM : 153510266
Nama Lengkap : Citra Yunela Sari
Tempat, Tgl Lahir : Pekanbaru, 6 November 1997
Jenis Kelamin : Perempuan
Agama : Islam
Jenjang : Strata 1 (S1)
Program Studi : Teknik Informatika
Alamat : Jln. Bandeng 1 No 307 Perumnas Rumbai
Nomor Handphone : 0812-7088-9058
Email : citra.yunela@student.uir.ac.id

2. Pendidikan

| No | Jenjang | Nama Lembaga | Tahun |
|----|---------|------------------------|-------------|
| 1 | SD | SDN 021 Pekanbaru | 2003 - 2009 |
| 2 | SMP | SMPN 6 Pekanbaru | 2009 - 2012 |
| 3 | SMA | SMKN 1 Pekanbaru | 2012 - 2015 |
| 4 | PT | Universitas Islam Riau | 2015 - 2019 |

Demikian daftar riwayat hidup ini dibuat dengan sebenarnya.

Pekanbaru, 18 Desember 2019
Mahasiswa Ybs,

Citra Yunela Sari
153510266

HALAMAN PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Assalamu'alaikumWr.Wb.

Alhamdulillah, puji syukur penulis ucapkan kepada Allah SWT atas segala rahmat dan karunia-Nya yang telah diberikan kepada penulis sehingga dapat menyelesaikan skripsi yang berjudul *“Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa Dan Warna Air Sumur Di Kelurahan Selat Baru Bengkalis”* tepat pada waktunya.

Laporan skripsi ini disusun untuk memenuhi persyaratan untuk memperoleh gelar sarjana pada Fakultas Teknik Universitas Islam Riau. Penulis menyadari bahwa tanpa bantuan dari pihak-pihak lain, usaha yang penulis lakukan dalam menyelesaikan laporan skripsi ini tidak akan membuahkan hasil yang berarti. Dalam kesempatan ini penulis menyampaikan ucapan terimakasih kepada:

1. Yang teristimewa kedua orangtua, Papa (Yurizal) dan Mama (Nelvisda) yang tidak pernah lelah berkorban, memberi motivasi baik moril maupun materil, dan selalu mendo'akan anaknya agar menjadi orang yang berguna serta sukses dalam mewujudkan cita-cita.
2. Ibu Ir. Des Suryani, S.Kom., M.Sc., selaku Dosen Pembimbing Skripsi yang telah memberikan motivasi dan bimbingan sehingga penulis dapat menyelesaikan laporan skripsi ini dengan lancar.
3. Ibu Nesi Syafitri, S.Kom., M.Cs selaku Dosen Pembimbing Akademis dan Dosen Penguji 1 dan Ibu Ana Yulianti ST., M.Kom selaku Dosen Penguji 2 yang telah memberi masukan dan arahan dalam pembuatan skripsi ini.

4. Bapak Khairul Umam Syaliman, ST., M.Kom., selaku Dosen PCR yang telah banyak memberikan masukan dan ilmunya dalam penulisan skripsi ini.
5. Rekan-rekan *focused group discussion* Dinda Febrianti, Agus Yuliani, bang Amrizal, bang Novendra, kak Nani, Atika Martia, Bepriilia dan Mafia Hongkong yang selalu memotivasi, membantu dan mendukung penulis dalam menyelesaikan laporan skripsi ini.
6. Para asisten, khususnya asisten 2 yang siap siaga mengajar seorang diri disaat penulis sibuk menyelesaikan laporan skripsi ini dan para asisten yang tidak dapat disebutkan namanya, terimakasih atas bantuannya dalam pembuatan laporan skripsi ini.
7. Teman-teman terkhusus kepada kelas C angkatan 2015 TI UIR, terimakasih atas semangat motivasi dan kebersamaannya dan senior maupun junior yang tidak dapat disebutkan namanya satu persatu. Semoga kita bisa menjadi sarjana yang bermanfaat.

Akhir kata penulis mohon maaf atas kekeliruan dan kesalahan yang terdapat dalam skripsi ini dan berharap semoga skripsi ini dapat memberikan manfaat bagi pembaca.

Pekanbaru, 18 Desember 2019

Citra Yunela Sari
153510266

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Assalaamu'alaikum Wr.Wb.

Alhamdulillah, segala puji dan syukur penulis ucapkan kehadiran Allah SWT, karena atas rahmat dan hidayahnya penulis dapat menyelesaikan skripsi dengan judul “*Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa Dan Warna Air Sumur Di Kelurahan Selat Baru Bengkalis*” dengan tujuan untuk memenuhi salah satu syarat menyelesaikan pendidikan Strata Satu (S1) Teknik informatika di Universitas Islam Riau Pekanbaru.

Dalam penulisan skripsi ini, penulis banyak mendapat bantuan dan bimbingan dari berbagai pihak, baik secara langsung maupun tidak langsung. Pada kesempatan ini, penulis mengucapkan terima kasih kepada :

1. Ibu Ause Labellapansa, ST., M.Cs., M.Kom, selaku Ketua Program Studi Teknik Informatika Universitas Islam Riau.
2. Ibu Ir. Des Suryani, M.Sc., selaku Dosen Pembimbing yang selalu tabah dan sabar dalam membimbing penulis dalam penyusunan laporan skripsi.
3. Ibu Nesi Syafitri, S.Kom., M.Cs selaku Dosen Pembimbing Akademis.
4. Bapak Dewanra Bagus Eka Putra, B.Sc (HONS)., M.Sc., selaku pembimbing dari Program Studi Geologi yang telah membantu dan membimbing saya untuk menyelesaikan laporan skripsi ini.
5. Seluruh Dosen Teknik Informatika yang telah memberikan ilmunya selama penulis menduduki bangku perkuliahan.

Akhirnya penulis berharap semoga skripsi ini dapat bermanfaat bagi kita semua, khususnya bagi mahasiswa Teknik Informatika Universitas Islam Riau. Penulis menyadari masih banyak kekurangan dalam penyusunan skripsi ini. Oleh karena itu penulis mengharapkan adanya masukan dari semua pihak demi sempurnanya skripsi ini.

Pekanbaru, 18 Desember 2019

Citra Yunela Sari



Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa Dan Warna Air Sumur di Kelurahan Selat Baru Bengkalis

Citra Yunela Sari
Fakultas Teknik
Program Studi Teknik Informatika
Universitas Islam Riau
E-mail : citra.yunela@student.uir.ac.id

ABSTRAK

Kualitas air sumur pada umumnya lebih bersih dari air permukaan karena air yang merembes ke dalam tanah telah disaring oleh lapisan tanah yang dilewatinya. Namun berdasarkan survey yang dilakukan di Selat Baru Bengkalis, terdapat beberapa air sumur yang memiliki kualitas air yang kurang baik. Beberapa diantaranya mengeluarkan bau, rasa asin dan warna yang tidak jernih. Klasifikasi bau, rasa dan warna secara manual memiliki keterbatasan alokasi ruang dan waktu. Untuk menyelesaikan permasalahan tersebut, penelitian ini menawarkan implementasi metode *K-Nearest Neighbor* untuk memprediksi bau, rasa dan warna air sumur. Metode ini bekerja dengan cara mencari kedekatan jarak suatu data dengan data yang lain. Hasil dari pengujian terhadap nilai K terbaik diperoleh pada $K=11$ dan hasil pengujian akurasi sistem prediksi bau, rasa dan warna adalah 85% dan hasil implementasi yang didapat sistem ini sebesar 75,67% menunjukkan bahwa sistem prediksi bau, rasa dan warna air sumur ini dapat diterapkan.

Kata Kunci: *Data Mining, Metode K Nearest Neighbor, Air Sumur*

Implementation of K-Nearest Neighbor Method in Predicting Smell, Taste and Color of Well Water in Selat Baru Bengkalis

Citra Yunela Sari
Fakultas Teknik
Program Studi Teknik Informatika
Universitas Islam Riau
E-mail : citra.yunela@student.uir.ac.id

ABSTRACT

The quality of well water is generally cleaner than surface water because water that seeps into the soil has been filtered by a layer of soil through which it passes. However, based on a survey conducted in the Bengkalis Baru Strait, there are several wells that have poor water quality. Some of them emit odors, saltiness and color that is not clear. Classification of odor, taste and color manually has limited space and time allocation. To solve this problem, this study offers the implementation of the K-Nearest Neighbor method to predict the smell, taste and color of well water. This method works by finding the proximity of a data with other data. The results of testing the best K value obtained at $K = 11$ and the results of testing the accuracy of the odor, taste and color prediction system is 85% and the implementation results obtained by this system of 75.67% indicate that the smell, taste and color prediction system of this well water can be applied

Keywords: Data Mining, Method K Nearest Neighbor, Well Water

DAFTAR ISI

Hal

| | |
|---|----------|
| HALAMAN JUDUL | |
| LEMBAR PENGESAHAN PEMBIMBING SKRIPSI | |
| LEMBAR PENGESAHAN TIM PENGUJI UJIAN SKRIPSI | |
| LEMBAR PERNYATAAN BEBAS PLAGIARISME | |
| LEMBAR IDENTITAS PENULIS | |
| LEMBAR DAFTAR RIWAYAT HIDUP | |
| HALAMAN PERSEMBAHAN | i |
| KATA PENGANTAR..... | iii |
| ABSTRAK | v |
| ABSTRACT | vi |
| DAFTAR ISI..... | vii |
| DAFTAR TABEL | xi |
| DAFTAR GAMBAR..... | xiii |
| DAFTAR LAMPIRAN | xvi |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Identifikasi Masalah | 3 |
| 1.3 Batasan Masalah | 3 |
| 1.4 Rumusan Masalah | 4 |
| 1.5 Tujuan Penelitian | 4 |
| 1.6 Manfaat Penelitian | 5 |

| | |
|--|-----------|
| BAB II LANDASAN TEORI | 6 |
| 2.1 Studi Kepustakaan | 6 |
| 2.2 Dasar Teori | 9 |
| 2.2.1 Air Sumur | 9 |
| 2.2.2 Parameter Kualitas Air Sumur | 9 |
| 2.2.3 <i>Total Dissolved Solid (TDS)</i> | 10 |
| 2.2.4 Bau | 11 |
| 2.2.5 Rasa | 11 |
| 2.2.6 Warna | 12 |
| 2.2.7 <i>Data Mining</i> | 13 |
| 2.2.8 Klasifikasi | 14 |
| 2.2.9 Metode <i>K-Nearest Neighbor</i> | 18 |
| 2.2.10 <i>K-Fold Cross Validation</i> | 21 |
| 2.2.11 Perancangan Sistem | 22 |
| 2.2.11.1 <i>Use Case Diagram</i> | 22 |
| 2.2.11.2 <i>Data Flow Diagram</i> | 23 |
| 2.2.11.3 Program <i>Flowchart</i> | 24 |
| 2.2.12 <i>Hypertext Preprocessor (PHP)</i> | 25 |
| 2.2.13 MySQL | 26 |
| 2.2.14 Akurasi | 27 |
| BAB III METODOLOGI PENELITIAN | 29 |
| 3.1 Alat dan Bahan Penelitian yang Digunakan | 29 |

| | | |
|--|---|-----------|
| 3.1.1 | Spesifikasi Kebutuhan <i>Hardware</i> | 29 |
| 3.1.2 | Spesifikasi Kebutuhan <i>Software</i> | 29 |
| 3.2 | Metode Pengumpulan Data | 30 |
| 3.2.1 | Teknik Pengumpulan Data | 30 |
| 3.2.2 | Jenis Data | 30 |
| 3.3 | Analisa Sistem yang Sedang Berjalan | 31 |
| 3.4 | Perancangan Sistem | 31 |
| 3.4.1 | <i>Context Diagram</i> | 32 |
| 3.4.2 | <i>Hierarchy Chart</i> | 33 |
| 3.4.3 | <i>Data Flow Diagram (DFD) Level 0</i> | 33 |
| 3.4.4 | <i>Data Flow Diagram (DFD) Level 1 Proses 1</i> | 34 |
| 3.4.5 | <i>Data Flow Diagram (DFD) Level 1 Proses 2</i> | 35 |
| 3.5 | Rancangan Desain | 36 |
| 3.5.1 | Rancangan Desain <i>Output</i> | 36 |
| 3.5.2 | Rancangan Desain <i>Input</i> | 38 |
| 3.5.3 | Desain <i>Database</i> | 40 |
| 3.5.4 | Perhitungan Manual | 42 |
| 3.5.5 | Desain Antarmuka | 56 |
| 3.5.6 | Rancangan Logika Program | 57 |
| BAB IV HASIL DAN PEMBAHASAN | | 62 |
| 4.1 | Pengujian <i>Black Box</i> | 62 |
| 4.1.1 | Pengujian Login | 63 |
| 4.1.2 | Pengujian Pengolahan Data Pengguna | 65 |

| | | |
|-----------------------|--|-----------|
| 4.1.3 | Pengujian Pengolahan Data <i>Training</i> | 68 |
| 4.1.4 | Pengujian Pengolahan Data <i>Testing</i> | 73 |
| 4.1.5 | Pengujian Perhitungan <i>K-Fold Cross Validation</i> | 75 |
| 4.2 | Pengujian Akurasi Sistem | 77 |
| 4.3 | Pengujian Sistem Terhadap Pengguna | 83 |
| 4.3.1 | Hasil Persentase Jawaban Kuisisioner | 85 |
| BAB V | PENUTUP | 86 |
| 5.1 | Kesimpulan | 86 |
| 5.2 | Saran | 86 |
| DAFTAR PUSTAKA | | |
| LAMPIRAN | | |

DAFTAR TABEL

| | Hal |
|---|-----|
| Tabel 2.1 Pengelompokkan Bau Air | 11 |
| Tabel 2.2 Pengelompokkan Rasa Air | 12 |
| Tabel 2.3 Pengelompokkan Warna Air | 12 |
| Tabel 2.4 Simbol <i>Use Case Diagram</i> | 23 |
| Tabel 2.5 Simbol <i>Data Flow Diagram</i> | 24 |
| Tabel 2.6 Simbol dan Fungsi Program <i>Flowchart</i> | 24 |
| Tabel 2.7 <i>Confusion Matrix</i> 2 Kelas | 28 |
| Tabel 3.1 Tabel <i>User</i> | 40 |
| Tabel 3.2 Tabel <i>Training</i> | 40 |
| Tabel 3.3 Tabel <i>Testing</i> | 41 |
| Tabel 3.4 Tabel K Optimal | 42 |
| Tabel 3.5 Tabel Hasil Perhitungan $k = 3$ | 43 |
| Tabel 3.6 Tabel Hasil Perhitungan $k = 5$ | 43 |
| Tabel 3.7 Tabel Hasil Perhitungan $k = 7$ | 44 |
| Tabel 3.8 Tabel Hasil Perhitungan $k = 9$ | 44 |
| Tabel 3.9 Tabel Hasil Perhitungan $k = 11$ | 44 |
| Tabel 3.10 Tabel Hasil Perhitungan $k = 13$ | 45 |
| Tabel 3.11 Tabel Hasil Perhitungan $k = 15$ | 45 |
| Tabel 3.12 Tabel Hasil Perhitungan $k = 17$ | 46 |
| Tabel 3.13 Tabel Hasil Perhitungan $k = 19$ | 46 |
| Tabel 3.14 Data <i>Training</i> | 47 |

| | |
|---|----|
| Tabel 3.15 Data <i>Testing</i> | 48 |
| Tabel 3.16 Perhitungan Nilai Jarak | 51 |
| Tabel 3.17 Pengurutan Data dengan Nilai Jarak Terkecil ke Terbesar | 53 |
| Tabel 3.18 Hasil Klasifikasi Bau, Rasa dan Warna | 55 |
| Tabel 4.1 Rencana Pengujian | 62 |
| Tabel 4.2 Pengujian <i>Login</i> | 65 |
| Tabel 4.3 Pengujian Pengolahan Data Pengguna | 67 |
| Tabel 4.4 Pengujian Pengolahan Data <i>Training</i> | 72 |
| Tabel 4.5 Pengujian Pengolahan Data <i>Testing</i> | 75 |
| Tabel 4.6 Pengujian Nilai K Optimal | 76 |
| Tabel 4.7 Pengujian Akurasi Klasifikasi Bau Berdasarkan Nilai k Optimal | 77 |
| Tabel 4.8 <i>Confusion Matrix</i> Klasifikasi Bau | 78 |
| Tabel 4.9 Pengujian Akurasi Klasifikasi Rasa Berdasarkan Nilai k Optimal ... | 79 |
| Tabel 4.10 <i>Confusion Matrix</i> Klasifikasi Rasa | 80 |
| Tabel 4.11 Pengujian Akurasi Klasifikasi Warna Berdasarkan Nilai k Optimal | 81 |
| Tabel 4.12 <i>Confusion Matrix</i> Klasifikasi Warna | 82 |
| Tabel 4.13 Hasil Nilai Kuisisioner | 84 |
| Tabel 4.14 Hasil Nilai Persentase Pertanyaan Kuisisioner | 85 |

DAFTAR GAMBAR

| | Hal |
|---|-----|
| Gambar 2.1 Contoh Klasifikasi pada <i>K- Nearest Neighbor</i> | 18 |
| Gambar 2.2 Pembagian Data Pada Metode <i>10-Fold Cross Validation</i> | 22 |
| Gambar 3.1 Analisa Sistem yang Sedang Berjalan | 31 |
| Gambar 3.2 Analisa Sistem yang Diusulkan | 31 |
| Gambar 3.3 Konteks <i>Diagram</i> Sistem Prediksi Bau, Rasa dan Warna Air Sumur | 32 |
| Gambar 3.4 <i>Hierarchy Chart</i> Sistem Prediksi Bau, Rasa dan Warna Air Sumur | 33 |
| Gambar 3.5 <i>Data Flow Diagram (DFD)</i> Level 0 | 34 |
| Gambar 3.6 <i>Data Flow Diagram (DFD)</i> Level 1 Proses 1 | 35 |
| Gambar 3.7 <i>Data Flow Diagram (DFD)</i> Level 1 Proses 2 | 35 |
| Gambar 3.8 Desain <i>Output Data Training</i> | 36 |
| Gambar 3.9 Desain <i>Output Data Testing</i> | 37 |
| Gambar 3.10 Desain <i>Output</i> Perhitungan KNN | 37 |
| Gambar 3.11 Desain <i>Input Login</i> | 38 |
| Gambar 3.12 Desain <i>Input Data Training</i> | 38 |
| Gambar 3.13 Desain <i>Input Data Testing</i> | 39 |
| Gambar 3.14 Desain <i>Input K-Fold Cross Vaidation</i> | 39 |
| Gambar 3.15 Pembagian Data Pada Metode <i>10-Fold Cross Validation</i> | 42 |
| Gambar 3.16 Desain Menu Utama | 56 |
| Gambar 3.17 Flowchart Alur Program Login | 57 |
| Gambar 3.18 Flowchart Alur Program Menu Utama..... | 58 |

| | |
|--|----|
| Gambar 3.19 Flowchart Alur Program Data <i>Training</i> | 59 |
| Gambar 3.20 Flowchart Alur Program Data <i>Testing</i> | 60 |
| Gambar 3.21 Flowchart Alur Program Perhitungan <i>K-Fold Cross Validation</i> . | 61 |
| Gambar 4.1 Pengujian <i>Login</i> | 63 |
| Gambar 4.2 Pengujian <i>Login</i> “Gagal” | 64 |
| Gambar 4.3 Tampilan Menu Utama Setelah <i>Login</i> | 64 |
| Gambar 4.4 Pengujian Tambah Data <i>User</i> | 60 |
| Gambar 4.5 Pesan Tambah Data <i>User</i> “Berhasil” | 65 |
| Gambar 4.6 Pesan Tambah Data <i>User</i> “Gagal” | 66 |
| Gambar 4.7 Menu Profil..... | 66 |
| Gambar 4.8 Pengujian Edit Data <i>User</i> | 67 |
| Gambar 4.9 Menu Data <i>Training</i> | 68 |
| Gambar 4.10 Pengujian Tambah Data <i>Training</i> dengan Input Manual dan Upload | 69 |
| Gambar 4.11 Pesan Tambah Data <i>Training</i> “Berhasil” | 69 |
| Gambar 4.12 Pesan Kesalahan Data Kosong | 70 |
| Gambar 4.13 Pengujian Edit Data <i>Training</i> | 70 |
| Gambar 4.14 Pesan Edit Data <i>Training</i> “Berhasil” | 71 |
| Gambar 4.15 Pesan Pemberitahuan Untuk Hapus Data <i>Training</i> | 71 |
| Gambar 4.16 Pesan Hapus Data <i>Training</i> “Berhasil” | 71 |
| Gambar 4.17 Menu Data <i>Testing</i> | 73 |
| Gambar 4.18 Pengujian Simpan Data <i>Testing</i> dengan Input Manual dan Upload | 73 |

| | |
|---|----|
| Gambar 4.19 Pesan Simpan Data <i>Testing</i> “Berhasil” | 74 |
| Gambar 4.20 Hasil Perhitungan Data <i>Testing</i> | 74 |
| Gambar 4.21 Menu K Optimal | 75 |
| Gambar 4.22 Perhitungan K Optimal dari k=3 Sampai Dengan k=19..... | 76 |
| Gambar 4.23 Perhitungan Akurasi dari k=3 Sampai Dengan k=19 | 76 |
| Gambar 4.24 Grafik Hasil Kuisisioner..... | 84 |



DAFTAR LAMPIRAN

Lampiran 3.1 *Data Training*

Lampiran 3.2 *Data Testing*

Lampiran 3.3 Perhitungan Jarak

Lampiran 3.4 Pengurutan Data Nilai Jarak Terkecil ke Terbesar



BAB I

PENDAHULUAN

1.1 Latar Belakang

Air adalah suatu zat cair yang tidak mempunyai rasa, bau, warna dan terdiri dari hidrogen dan oksigen dengan rumus kimia H₂O. Karena air mempunyai sifat yang hampir bisa digunakan untuk apa saja, maka air merupakan zat yang paling penting bagi semua bentuk kehidupan (tumbuhan, hewan, dan manusia) sampai saat ini selain matahari yang merupakan sumber energi. Air merupakan suatu zat yang paling menentukan dalam menunjang kehidupan makhluk hidup. Sekitar 72% permukaan Bumi ditutupi oleh air dan 97% air tersebut merupakan air asin dan tidak dapat diminum sehingga perlu dilakukan langkah-langkah pengolahan air agar dapat dimanfaatkan dalam kehidupan sehari-hari (Triatmojo, 2008). Air yang digunakan untuk kebutuhan hidup sehari-hari khususnya untuk penyediaan air bersih harus memenuhi persyaratan yang diatur dalam Permenkes RI No. 416 tahun 1990 tentang syarat-syarat dan pengawasan kualitas air. Dalam peraturan tersebut, dijelaskan kualitas air bersih meliputi kualitas secara fisika, secara kimia, secara mikrobiologi dan kualitas secara radioaktivitas dan parameter-parameter yang harus terpenuhi meliputi :

1. Parameter fisika meliputi: Bau, Rasa, Warna, Daya Hantar Listrik dan Suhu.
2. Parameter kimia meliputi: kimia Anorganik seperti Air raksa, pH, timbal, dll.
3. Parameter Mikrobiologi meliputi: Total Caliform (MPN).

Kualitas bau, rasa dan warna pada air sumur dapat dilakukan secara manual menggunakan metode organoleptik, namun memiliki keterbatasan alokasi ruang

dan waktu. Di zaman yang modern ini tidak menutup kemungkinan bahwa teknologi atau sistem bisa ikut berperan dalam menentukan kualitas bau, rasa dan warna air pada sumur.

Sistem yang akan dibangun pada penelitian ini merupakan sistem yang menerapkan data mining. Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban, 2005). Adapun metode klasifikasi yang digunakan dalam sistem ini adalah *k-Nearest Neighbor* (k-NN), yaitu algoritma pengklasifikasian data sederhana dimana perhitungan jarak terpendek dijadikan ukuran untuk mengklasifikasikan suatu kasus baru berdasarkan ukuran kemiripan (Pandie, 2012). Dalam Algoritma k-NN terdapat salah satu nilai yang disebut k yaitu jumlah tetangga terdekat. Pemilihan nilai k menjadi hal yang penting karena akan mempengaruhi kinerja algoritma k-NN (Wu, 2009). Nilai k yang terlalu kecil, maka hasil klasifikasi akan lebih terpengaruh oleh *noise*. Di sisi lain, jika nilai k terlalu tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur.

Berdasarkan survey air sumur yang dilakukan di kelurahan Selat Baru Bengkalis, terdapat 6 parameter kualitas air yang datanya dapat diolah. 3 diantaranya berupa data angka yaitu suhu, pH dan EC dan 3 diantaranya adalah data huruf yaitu bau, rasa dan warna. Pada parameter bau, rasa dan warna terdapat 2 hingga 3 kategori pada setiap parameternya. Untuk parameter bau terdiri dari 2 kategori yaitu ada dan tidak bau. Parameter rasa terdiri dari 3 kategori yaitu tawar,

asin dan payau dan parameter warna terdiri dari berwarna dan tidak berwarna. Berbeda dari data angka, data huruf didapat melalui proses manual yaitu dengan menggunakan metode organoleptik. Lokasi yang jauh mengakibatkan proses membutuhkan waktu yang cukup lama dan biaya yang cukup besar.

Penelitian ini bertujuan untuk memudahkan mahasiswa, dosen, peneliti, dll dalam memprediksi bau, rasa dan warna pada air sumur dengan menggunakan metode k-NN berdasarkan perhitungan suhu, pH dan EC secara cepat dan tepat.

1.2 Identifikasi Masalah

Adapun identifikasi masalah yang dapat diambil dari latar belakang tersebut adalah sebagai berikut :

1. Belum adanya aplikasi *k-Nearest Neighbor* dalam memprediksi 3 (tiga) output parameter berupa bau, rasa dan warna air sumur menggunakan nilai k yang sama.
2. Klasifikasi parameter bau, rasa, dan warna air sumur masih dilakukan secara manual.

1.3 Batasan Masalah

Mengingat keterbatasan waktu, biaya, dan kemampuan peneliti, maka penelitian dibatasi dalam hal :

1. Data penelitian ini diperoleh dari hasil survey di Kelurahan Selat Baru, Bengkalis sebanyak 110 data.
2. Data yang dipakai dalam proses *training* sebanyak 90 data dan proses *testing* sebanyak 20 data.

3. Atribut yang digunakan dalam proses perhitungan adalah suhu, pH dan EC.
4. Penelitian ini menggunakan metode *k- Nearest Neighbor* dan *k - Fold Cross Validation*.
5. Penelitian ini menggunakan 10 Fold dan $k= 3, 5, 7, 9, 11, 13, 15, 17, 19$ pada metode *k - Fold Cross Validation*.
6. Penelitian ini menggunakan bahasa pemrograman PHP, HTML, CSS dan Database Mysql.
7. Keluaran yang dihasilkan berupa hasil perhitungan k-NN.

1.4 Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana membangun sistem yang mampu memprediksi bau, rasa dan warna air sumur secara tepat dan cepat.
2. Bagaimana mengimplementasikan metode *k-Nearest Neighbor* kedalam sistem sehingga menghasilkan perhitungan yang valid dan akurat.
3. Bagaimana merancang sebuah sistem yang mudah dimengerti oleh pengguna sistem (tenaga ahli, peneliti, pengamat air, dll).

1.5 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

1. Membangun sebuah sistem untuk memprediksi bau, rasa dan warna air sumur berdasarkan perhitungan suhu, pH dan EC yang baik.
2. Menghasilkan aplikasi yang menerapkan metode *k- Nearest Neighbor*.
3. Membuat aplikasi yang bersifat *user friendly*.

1.6 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah :

1. Menambah pengetahuan dalam memprediksi bau, rasa dan warna air sumur dengan menghitung suhu, pH dan EC.
2. Membantu pengguna (mahasiswa, dosen, tenaga ahli, peneliti, pengamat, dll) dalam memperkecil kesalahan data pada proses *input* data, *proses* dan *output* hasil.



BAB II

LANDASAN TEORI

2.1 Studi Kepustakaan

Untuk menyusun skripsi ini, penulis menggunakan bahan acuan kepustakaan yang bersumber pada penelitian-penelitian sebelumnya. Hal ini berguna sebagai bahan referensi.

Penelitian yang dilakukan oleh Muhamad Zanuardi (2018) menghasilkan sistem yang dapat membedakan citra batang tebu matang dengan citra batang tebu mentah berdasarkan fitur tekstur menggunakan metode k-NN. Metode k-NN digunakan dalam mencari nilai ketetanggan untuk menentukan kelompok batang tebu matang dengan citra batang tebu mentah berdasarkan tekstur. Akurasi yang didapatkan dengan metode k-NN dari tiap k adalah pada k=3 menghasilkan nilai akurasi sebesar 65%, dengan nilai k=5 menghasilkan nilai akurasi sebesar 62%, pada k=7 menghasilkan nilai akurasi sebesar 68% dan k=9 menghasilkan nilai akurasi sebesar 68%.

Persamaan penelitian ini dengan penelitian tersebut adalah menggunakan metode k-NN, yang menjadi perbedaan pada penelitian ini adalah penelitian tersebut membahas tentang batang tebu sedangkan penelitian yang dibahas adalah air sumur.

Penelitian yang dilakukan oleh M. Reza Noviansyah, Tedy Rismawan, Dwi Marisa Midyanti (2018) dalam menerapkan data mining dengan algoritma klasifikasi k-NN dalam mengklasifikasi tingkatan bahaya kebakaran hutan berdasarkan kelas Fire Weather Index (FWI). FWI merupakan sistem tingkatan

bahaya kebakaran hutan dan lahan yang terdiri dari rendah, sedang, tinggi dan ekstrim. Akurasi yang didapatkan dengan metode k-NN dengan menggunakan 252 data uji dan menghasilkan persentase keberhasilan 80,16% dengan nilai k=5.

Persamaan penelitian ini dengan penelitian tersebut adalah menggunakan metode k-NN dan yang menjadi perbedaan adalah pada penelitian tersebut membahas klasifikasi indeks cuaca kebakaran dan memiliki 1 parameter output yaitu FWI sedangkan penelitian ini membahas klasifikasi pada air sumur dan memiliki 3 parameter output yaitu bau, rasa dan warna.

Dalam penelitian Mutiara Ayu Banjarsari, H. Irwan Budiman, Andi Farmadi (2015), menggunakan metode k-NN yang merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data *training* yang jaraknya paling dekat dengan objek tersebut. Metode *k-Fold cross validation* dan Uji Akurasi digunakan untuk mengetahui nilai k-Optimal. Hasil yang didapat adalah nilai k=5 dengan tingkat akurasi sebesar 80.00% yang ditetapkan sebagai k-Optimal. Nilai k=5 diterapkan pada algoritma k-NN untuk prediksi kelulusan tepat waktu mahasiswa berdasarkan IP sampai dengan semester 4.

Persamaan penelitian ini dengan penelitian tersebut adalah menggunakan metode k-NN dan *k-Fold cross validation* dan yang menjadi perbedaan adalah pada penelitian tersebut membahas kelulusan tepat waktu mahasiswa berdasarkan IP sedangkan penelitian ini membahas klasifikasi pada air sumur berdasarkan suhu, pH, dan EC.

Selanjutnya penelitian Emerensye S. Y. Pandie (2012) yang menggunakan teknik data mining k-NN untuk menilai kemampuan nasabah berdasarkan data-

data masa lalu. Dengan menggunakan tahapan data mining, dihasilkan sejumlah 3.018 data nasabah yang dikategorikan dalam 3 kategori kredit, macet, tersendat dan lancar. Hasil filtrasi data tersebut diuji tingkat errornya menggunakan teknik cross validation dengan 20 fold dan 10 nilai k. Hasil pengujian menunjukkan persentase tingkat error data pada angka kurang dari 3.7% dan mencapai kestabilan data pada nilai k=3 sampai dengan nilai k=11.

Persamaan penelitian ini dengan penelitian tersebut adalah menggunakan metode k-NN dan Cross Validation dan yang menjadi perbedaan adalah pada penelitian tersebut mengangkat tema SPK untuk mengambil keputusan pengajuan kredit sedangkan penelitian ini membahas klasifikasi pada air sumur berdasarkan suhu, pH dan EC.

Kemudian dalam penelitian yang dilakukan oleh Diina Itsna Annisa, Rudy Ariyanto, Ariadi Tri Retno Hayati Ririd (2016) tentang klasifikasi kehamilan beresiko dengan menggunakan metode k-NN yang diharapkan mampu mendeteksi sejak dini dan mengurangi angka kematian ibu, janin dan bayi akibat kehamilan beresiko. Akurasi yang didapatkan dengan membandingkan data yang diperoleh dari dinas kesehatan (puskesmas) dengan sistem dan menghasilkan tingkat akurasi nilai sebesar 93 % dengan menggunakan nilai k = 5.

Persamaan penelitian ini dengan penelitian tersebut adalah menggunakan metode k-NN dan yang menjadi perbedaan adalah pada penelitian ini menambahkan metode *k-Fold cross validation* pada uji akurasi datanya.

2.2 Dasar Teori

2.2.1 Air Sumur

Menurut Ramdyasari (2014) air sumur adalah air yang berasal dari dalam tanah, air tersebut didapatkan dengan cara menggali tanah sehingga akan terbentuk sumur. Air sumur merupakan salah satu sumber air yang bermanfaat untuk kebutuhan sehari-hari bagi masyarakat dan biasanya mengandung bahan-bahan metal terlarut seperti Na, Mg, Ca, dan Fe. Air sumur pada umumnya lebih bersih dari air permukaan karena air yang merembes ke dalam tanah telah disaring oleh lapisan tanah yang dilewatinya (Dwijosaputro, 1981).

2.2.2 Parameter Kualitas Air Sumur

Parameter kualitas air merupakan parameter yang digunakan untuk mengukur tingkat kualitas air. Kualitas air adalah kondisi air yang diukur atau diuji berdasarkan parameter-parameter tertentu dan metode tertentu berdasarkan peraturan perundang-undangan yang berlaku (Peraturan Menteri Kesehatan Republik Indonesia Nomor 32 Tahun 2017). Kualitas air dapat dinyatakan dengan parameter kualitas air. Parameter ini meliputi parameter fisik, kimia, dan mikrobiologis (Masduqi, 2009).

- a. Parameter fisik, adalah parameter yang dapat diamati akibat perubahan fisika air seperti cahaya, suhu, kecerahan, kekeruhan, bau, rasa, warna, padatan tersuspensi, daya hantar listrik, hingga salinitas air.
- b. Parameter kimia terdiri dari salinitas, pH, alkalinitas, dissolved oxygen, biological Oxygen Demand, produktivitas primer, sedimen, oxidized layer.

- c. Parameter mikrobiologis merupakan parameter yang digunakan untuk mengetahui kepadatan biota di dalam air. Biota tersebut dapat berupa plankton, benthos, perifiton, bakteri maupun biota jenis lainnya.

2.2.3 *Total Dissolved Solid (TDS)*

TDS adalah salah satu parameter fisik air yang merupakan jumlah padatan yang berasal dari material-material terlarut yang dapat melewati filter yang lebih kecil daripada 2 μm (Djuhariningrum, 2005). Padatan ini terdiri dari senyawa-senyawa anorganik dan organik yang terlarut dalam air, mineral dan garam-garamnya. Penyebab tingginya nilai TDS adalah adanya kandungan bahan anorganik berupa ion-ion di perairan. Sebagai contoh air buangan sering mengandung molekul sabun, deterjen dan surfaktan yang larut air, misalnya pada air buangan rumah tangga dan industri pencucian. Metode *gravimetry* merupakan metode standar yang memiliki tingkat keakuratan yang tinggi dalam mengukur TDS, namun metode ini harus dilakukan di laboratorium dan pengukurannya membutuhkan waktu yang lama. Metode lain yang dapat digunakan untuk pengukuran nilai TDS melalui pengukuran konduktivitas listrik (Herlambang, 2006).

Berdasarkan Peraturan Menteri Kesehatan No 32 Tahun 2017 Tentang Standar Baku Mutu Kesehatan Lingkungan dan Persyaratan Kesehatan Air Minum untuk Keperluan Higiene Sanitasi, Kolam Renang, SPA dan Pemandian Umum, ditetapkan bahwa kadar maksimum TDS air untuk pemakaian *higiene sanitasi* adalah 1000 mg/l. Apabila kadar TDS melebihi kadar maksimum, efek negatif yang muncul adalah :

- Dari bau, air akan mengeluarkan gas-gas yang menimbulkan gejala mual.
- Dari rasa, air tersebut sudah tidak dapat diterima oleh tubuh.
- Terjadinya cardiac diseases serta toxaemia pada wanita-wanita hamil.

2.2.4 Bau

Bau adalah zat kimia yang tercampur di udara, umumnya dengan konsentrasi yang sangat rendah yang manusia terima dengan indra penciuman. Bau air dapat memberi petunjuk terhadap kualitas air, misalnya bau amis dapat disebabkan oleh adanya bakteri atau algae dalam air tersebut (Ramdyasari, 2014). Bau pada air dapat dipengaruhi oleh kadar TDS yang terdapat didalamnya. Air yang mengandung kadar TDS yang tinggi berarti memiliki kandungan senyawa-senyawa anorganik dan organik yang terlarut didalamnya. Salah satu contoh senyawa organik yang dapat mempengaruhi bau adalah molekul sabun deterjen dan surfaktan yang larut air. Kriteria bau air berdasarkan nilai TDS dapat dilihat ada tabel 2.1.

Tabel 2.1 Pengelompokan Bau Air

| TDS (mg/l) | Kriteria |
|------------|-----------|
| 0 – 1000 | Tidak Bau |
| >1000 | Bau |

2.2.5 Rasa

Rasa berarti tanggapan indra terhadap rangsangan saraf seperti manis, pahit, masam terhadap indra pengecap, atau panas, dingin, nyeri terhadap indra perasa. Air yang berasa menunjukkan kehadiran berbagai zat yang dapat membahayakan kesehatan. Efek yang dapat ditimbulkan terhadap kesehatan

manusia tergantung pada penyebab timbulnya rasa. Kriteria rasa air berdasarkan nilai TDS dapat dilihat ada tabel 2.2.

Tabel 2.2 Pengelompokan Rasa Air

| TDS (mg/l) | Kriteria |
|------------------|----------|
| 0 – 1.000 | Tawar |
| 1.001 – 10.000 | Payau |
| 10.001 – 100.000 | Asin |

2.2.6 Warna

Warna adalah kesan yang diperoleh mata dari cahaya yang dipantulkan oleh benda-benda yang dikenai cahaya tersebut. Warna pada air disebabkan oleh adanya partikel hasil pembusukan bahan organik, ion-ion metal alam (besi dan mangan), plankton, humus, buangan industri, dan tanaman air. Adanya oksida besi menyebabkan air berwarna kemerahan, sedangkan oksida mangan menyebabkan air berwarna kecoklatan atau kehitaman (Ramdyasari, 2014). Kelompok bakteri besi (misalnya *Crenothrix* dan *Sphaerotilus*) yang mampu mengoksidasi senyawa ferro menjadi ferri. Akibat kehadirannya, air sering berubah warna kalau disimpan lama yaitu warna kehitam-hitaman, kecoklat-coklatan, dan sebagainya. Kriteria warna air berdasarkan nilai TDS dapat dilihat ada tabel 2.3.

Tabel 2.3 Kriteria Warna Air

| TDS (mg/l) | Kriteria |
|------------|----------------|
| 0 - 1000 | Tidak Berwarna |
| >1000 | Berwarna |

2.2.7 Data Mining

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (*induction-based learning*) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari (Hermawati, 2013).

Data mining mempunyai fungsi yang penting untuk membantu mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi pengguna. Pada dasarnya, data mining mempunyai empat fungsi dasar yaitu:

1. Fungsi Prediksi (*prediction*).

Proses untuk menemukan pola dari data dengan menggunakan beberapa variabel untuk memprediksikan variabel lain yang tidak diketahui jenis atau nilainya.

2. Fungsi Deskripsi (*description*).

Proses untuk menemukan suatu karakteristik penting dari data dalam suatu basis data.

3. Fungsi Klasifikasi (*classification*).

Klasifikasi merupakan suatu proses untuk menemukan model atau fungsi untuk menggambarkan class atau konsep dari suatu data. Proses yang digunakan untuk mendeskripsikan data yang penting serta dapat meramalkan kecenderungan data pada masa depan.

4. Fungsi Asosiasi (*association*).

Proses ini digunakan untuk menemukan suatu hubungan yang terdapat pada nilai atribut dari sekumpulan data.

Menurut Pandie (2012) data mining merupakan penjelasan tentang masa lalu dan prediksi masa depan berdasarkan analisa pada sekelompok data. Dalam proses memprediksi masa depan ini, data mining menggunakan beberapa model yaitu pemodelan prediktif dan deskriptif.

- a) Pemodelan prediktif diawali dengan pembentukan model untuk memprediksi hasil. Jika hasil pemodelan dalam bentuk kategori (diskrit) maka disebut sebagai “klasifikasi” sedangkan jika hasil pemodelan bernilai kontinu maka disebut “regresi”.
- b) Pemodelan deskriptif atau lebih dikenal dengan istilah clustering, merupakan proses pengamatan terhadap kelompok data kemudian diikuti dengan pengelompokan cluster atau data terhadap data yang mempunyai kesamaan ciri.

2.2.8 Klasifikasi

Klasifikasi adalah proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui (Kamber, 2006). Klasifikasi termasuk dalam kelompok supervised learning, yaitu sebuah pendekatan dimana sudah terdapat data yang dilatih, dan terdapat variable yang ditargetkan sehingga tujuan dari pendekatan ini adalah

mengelompokan suatu data ke data yang sudah ada. Proses klasifikasi didasarkan pada empat komponen (Gorunescu, 2011):

- a. Kelas. Variabel yang berupa kategori yang merepresentasikan “label” yang terdapat pada objek.
- b. Predictor. Variabel yang direpresentasikan oleh karakteristik (atribut) data.
- c. Training dataset. Satu set data yang berisi nilai dari kedua komponen di atas yang digunakan untuk menentukan kelas yang cocok berdasarkan predictor.
- d. Testing dataset. Berisi data baru yang akan diklasifikasikan oleh model yang telah dibuat dan akurasi klasifikasi dievaluasi

Beberapa metode penilaian pada klasifikasi adalah seperti berikut:

- a. *Leave-one-out*, setiap data digunakan untuk pengujian sekali pada sebanyak model yang dikembangkan sehingga ada sejumlah titik data. Metode memakan waktu yang lama, namun untuk dataset yang kecil metode ini adalah pilihan yang tepat.
- b. *Bootstrapping*, sejumlah instans yang tetap dari data awal dijadikan sampel (dengan pergantian) untuk pelatihan dan dataset sisanya digunakan untuk pengujian. Proses ini diulang-ulang sebanyak yang diinginkan.
- c. *Jackknifing*, akurasi dihitung dengan mengeluarkan satu sampel pada setiap iterasi dari proses estimasi.
- d. *Area under the ROC curve*, adalah teknik penilaian secara grafis dimana true positive rate digambar pada sumbu Y dan false positive rate digambar pada sumbu X.

- e. *k-fold cross validation*, dataset yang utuh di pecah secara acak menjadi beberapa kelompok dengan *size* yang hampir sama dan semua data pada kelompok tersebut dilatih kecuali data pada kelompok pengujian.

Berikut ini adalah algoritma klasifikasi data mining yang paling populer menurut Gorunescu (2011):

a. *Decision Trees*.

Decision tree analysis (atau analisa pohon keputusan adalah suatu teknik yang termasuk keluarga machine-learning) bisa dibidang teknik klasifikasi yang paling populer pada area data mining.

b. *Bayesian classifiers/Naive Bayes classifiers*.

Naïve Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai teorema Bayes. Teorema tersebut dikombinasikan dengan "naive" dimana diasumsikan kondisi antar atribut saling bebas.

c. *Artificial Neural Networks*.

Jaringan syaraf tiruan merupakan simulasi dari otak biologis yang bertujuan untuk belajar mengenali pola data dan mensimulasikan proses belajar adaptif biologis, walau dalam skala yang sangat sederhana.

d. *Rough sets.*

Metode ini mempertimbangkan keanggotaan parsial dari label-label kelas ke kategori-kategori yang telah ditetapkan sebelumnya dalam membuat model-model (kumpulan aturan) untuk problem klasifikasi.

e. *Rule-Based methods.*

Adalah *advanced computer program* yang mencoba untuk meniru kemampuan manusia dalam membuat keputusan dan pemecahan masalah.

f. *Case Based Reasoning.*

Merupakan suatu metode penalaran komputer yang menggunakan pengetahuan lama untuk mengatasi masalah baru yang ditemui. Case Based Reasoning memberikan solusi terhadap kasus baru berdasarkan kedekatan antara kasus lama yang ada pada basis pengetahuan dengan kasus baru yang terjadi.

g. *Support Vector Machines.*

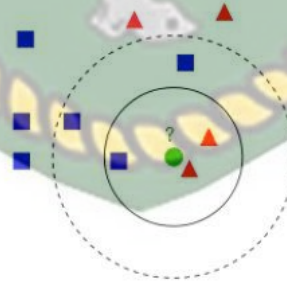
SVM adalah metode machine learning yang bekerja atas prinsip Structural Risk Minimization (SRM) dengan tujuan menemukan hyperplane terbaik yang memisahkan dua buah class pada input space. SVM merupakan sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur (feature space) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan learning bias yang berasal dari teori pembelajaran statistik.

h. *k-Nearest Neighbor*.

k-Nearest Neighbor (*k-NN*) adalah suatu metode yang menggunakan algoritma *supervised* untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Tujuan dari algoritma ini adalah mengklasifikasikan objek baru berdasarkan atribut dan *training sample*.

2.2.9 Metode *k-Nearest Neighbour*

Algoritma *k-Nearest Neighbor* merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. *k-NN* termasuk algoritma *supervised learning* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada *k-NN*. Kelas yang paling banyak muncul itu yang akan menjadi kelas hasil klasifikasi. Tujuan dari algoritma ini adalah mengklasifikasikan objek baru berdasarkan atribut dan *training sample*.



Gambar 2.1 Contoh Klasifikasi pada *k-Nearest neighbor*

Algoritma *k-Nearest Neighbor* menggunakan klasifikasi ketetanggaan (*neighbor*) sebagai nilai prediksi dari *query instance* yang baru. Algoritma ini sederhana, bekerja berdasarkan jarak terpendek dari *query instance* ke *training*

sample untuk menentukan ketetanggaannya (Rizal, 2013). Berikut beberapa formula perhitungan jarak yang digunakan dalam algoritma k-NN.

- **Euclidean Distance**

Jarak Euclidean adalah formula untuk mencari jarak antara 2 titik dalam ruang dua dimensi.

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

Keterangan:

x_i = Sampel data

y_i = Data uji atau data *testing*

d = Jarak

- **Manhattan Distance**

Manhattan Distance atau Taxicab Geometri adalah formula untuk mencari jarak d antar 2 vektor p, q pada ruang dimensi n .

$$d_e(x, y) = \sum_{i=1}^n |x_i - y_i| \dots\dots\dots(2.2)$$

- **Minkowski Distance**

Minkowski distance adalah formula pengukuran antar 2 titik pada ruang vektor normal yang merupakan hibridisasi yang mengeneralisasi euclidean distance dan mahattan distance.

$$d_e(x, y) = (\sum_{i=1}^n |x_i - y_i|^r)^{1/r} \dots\dots\dots(2.3)$$

Teknik pencarian tetangga terdekat disesuaikan dengan dimensi data, proyeksi, dan kemudahan implementasi oleh pengguna. Teknik pencarian tetangga terdekat pada umumnya dilakukan dengan menggunakan formula jarak euclidean.

Keuntungan algoritma *k-Nearest Neighbor* :

1. Sangat mudah diimplementasi dan dipahami.
2. Sangat non-linear.
3. Efektif untuk menghitung data dalam skala kecil.

Kelemahan algoritma *k-Nearest Neighbor* :

1. Memerlukan nilai *k* yang optimal.
2. Lambat saat proses prediksi.
3. Rentan terhadap perbedaan rentang variabel.

Algoritma klasifikasi dalam *k-Nearest Neighbor* :

1. Menentukan nilai *k* (jumlah tetangga paling dekat).
2. Menghitung jarak antara data yang akan dievaluasi dengan semua data pelatihan atau data sampel dengan menggunakan rumus *Euclidean Distance*.
3. Mengurutkan hasil jarak step '2' secara *ascending*.
4. Kumpulkan klasifikasi berdasarkan nilai *k*.
5. Memasangkan kategori atau kelas yang bersesuaian.
6. Mencari jumlah kelas terbanyak dari tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelas data yang akan dievaluasi.

Nilai *k* yang terbaik untuk algoritma ini tergantung pada data, Salah satu masalah yang dihadapi *k-NN* adalah dalam pemilihan nilai *k* yang tepat. Nilai *k* yang tinggi bisa mengurangi efek noise pada klasifikasi tetapi membuat batasan antara setiap kelas menjadi kabur sedangkan *k* yang terlalu kecil bisa menyebabkan algoritma terlalu *sensitive* terhadap *noise*. Nilai *k* terbaik dapat

dipilih dengan optimasi parameter, misalnya dengan menggunakan k-fold cross validation.

2.2.10 K-Fold Cross Validation

Menurut Pandie (2012) k-fold cross validation merupakan salah satu metode yang digunakan untuk mengetahui rata-rata keberhasilan dari suatu sistem dengan cara melakukan perulangan dengan mengacak atribut masukan sehingga sistem tersebut teruji untuk beberapa atribut input yang acak. K-fold cross validation diawali dengan membagi data sejumlah fold yang diinginkan. Dalam proses cross validation data akan dibagi dalam n buah partisi dengan ukuran yang sama D1,D2,D3..Dn selanjutnya proses testing dan training dilakukan sebanyak n kali. Dalam iterasi ke-i partisi Di akan menjadi data testing dan sisanya akan menjadi data training.

Menurut Hastie (2008), dengan k=5 atau k=10 dapat digunakan untuk memperkirakan tingkat kesalahan yang terjadi, sebab data training pada setiap fold cukup berbeda dengan data training yang asli. Secara keseluruhan, 5 atau 10-fold cross validation sama-sama direkomendasikan dan disepakati bersama.

Menghitung nilai akurasi dapat dilakukan dengan menggunakan persamaan :

$$Akurasi = \frac{Jumlah\ klasifikasi\ benar}{Jumlah\ data\ uji} \times 100\% \dots\dots\dots(2.4)$$

Langkah pertama untuk mendapatkan k-optimal dengan pengulangan 10 kali percobaan adalah membagi data training sebanyak 10 bagian atau kelompok data yang sama seperti gambar 2.2.



Gambar 2.2 Pembagian data pada metode *10-Fold Cross Validation*

Berdasarkan gambar diatas, pada P1 blok pertama terdiri dari data yang akan berperan sebagai data testing dan blok lainnya berperan sebagai data training. Setiap data testing dilakukan prediksi menggunakan algoritma k-NN. Hasil klasifikasi prediksi k-NN dibandingkan dengan klasifikasi data real dan dihitung jumlah prediksi yang cocok dengan data asli. Tingkat akurasi yang tinggi itulah yang terpilih menjadi nilai k terbaik (Banjarsari, 2015).

Perancangan Sistem

2.2.11.1 Use Case Diagram

Menurut Satzinger, dkk (2010), *use case diagram* adalah diagram yang menggambarkan berbagai peran dari pengguna untuk menggunakan sistem. Tujuan dari penggunaan dari penggunaan *use case diagram* adalah untuk mengidentifikasi bagaimana sistem tersebut akan digunakan. Adapun simbol *use case diagram* dapat dilihat pada tabel 2.4.

Tabel 2.4 Simbol *Use Case Diagram*

| Simbol | Nama | Fungsi |
|---|--------------------|--|
|  | Aktor | Menspesifikasikan himpunan peran yang pengguna mainkan ketika berinteraksi dengan <i>use case</i> . |
|  | <i>Use Case</i> | Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu aktor. |
|  | <i>Association</i> | Apa yang menghubungkan antara objek satu dengan objek lainnya. |
|  | <i>Include</i> | Memnspesifikasikan bahwa use case sumber secara eksplisit. |
|  | <i>Extend</i> | Menspesifikasikan bahwa use case target memperluas perilaku dari use case sumber pada suatu titik yang diberikan. |
|  | <i>System</i> | Menspesifikasikan paket yang menampilkan sistem secara terbatas. |

2.2.11.2 Data Flow Diagram (DFD)

Data flow diagram (DFD) menurut Jogiyanto (2005) adalah diagram yang menggunakan notasi simbol untuk menggambarkan arus data system. DFD sering digunakan untuk menggambarkan suatu sistem yang telah ada atau sistem yang baru yang akan dikembangkan secara logika dan menjelaskan arus data dari mulai pemasukan sampai dengan keluaran data tingkatan diagram arus data mulai dari diagram konteks yang menjelaskan secara umum suatu system atau batasan system dari level 0 dikembangkan menjadi level 1 sampai system tergambarkan secara rinci. Adapun simbol DFD dapat dilihat pada tabel 2.5.



Tabel 2.5 Simbol *Data Flow Diagram*

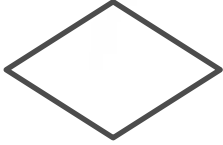






| Simbol | Nama | Fungsi |
|---|--------------------------------|---|
|  | Simbol entitas eksternal | Digunakan untuk menunjukkan tempat asal <i>data</i> . |
|  | Simbol proses | Digunakan untuk menunjukkan tugas atau proses yang dilakukan baik secara manual atau otomatis |
|  | Simbol penyimpanan <i>data</i> | Digunakan untuk menunjukkan gudang informasi atau <i>data</i> . |
|  | Simbol arus <i>data</i> | Digunakan untuk menunjukkan arus dari proses. |

2.2.11.3 Program *Flowchart*

Flowchart adalah representasi grafis dan langkah-langkah yang harus diikuti dalam menyelesaikan suatu permasalahan yang terdiri dari sekumpulan simbol, dimana masing-masing simbol merepresentasikan kegiatan tertentu. *Flowchart* membantu analis dan pemrogrammer untuk memecahkan masalah ke dalam segmen-segmen yang lebih kecil dan membantu dalam menganalisis alternatif-alternatif dalam pengoperasian (Yatina B, 2010). Adapun simbol program *flowchart* dapat dilihat pada tabel 2.6. diantaranya adalah:

Tabel 2.6 Simbol dan Fungsi Program *Flowchart*

| Nama | Simbol | Keterangan |
|---------------------|---|--|
| Terminator |  | Digunakan untuk menandai awal dan akhir dari suatu <i>flowchart</i> , simbol ini biasanya di beri label mulai dan selesai. |
| <i>Input-Output</i> |  | Digunakan untuk mempresentasikan fungsi I/O yang membuat sebuah dapat di proses (<i>input</i>) atau ditampilkan (<i>output</i>). |

| | | |
|-------------------|---|---|
| Kondisi |  | Digunakan untuk melakukan percabangan, yaitu pemeriksaan terhadap suatu kondisi. Dalam simbol ini, kita menuliskan keadaan yang harus dipenuhi. |
| Proses |  | Digunakan untuk kegiatan pemrosesan <i>input</i> , pada simbol ini kita dapat menuliskan operasi-operasi yang berkenaan pada <i>input</i> , atau operasi lainnya. |
| Persiapan |  | Digunakan untuk proses inisialisasi/ pemberian harga awal (misalnya dalam melakukan iterasi). |
| Sub Program |  | Digunakan untuk permulaan sub program/ proses menjalankan sub program. |
| Alir Data |  | Digunakan untuk menghubungkan setiap langkah dalam flowchart dan menunjukkan ke mana arah aliran diagram. Anak panah ini mempunyai arah kiri, kanan, atas, bawah. |
| Konektor On Page |  | Digunakan untuk menghubungkan satu langkah dengan langkah lain dalam keadaan masih satu halaman. |
| Konektor Off Page |  | Digunakan untuk menghubungkan satu langkah dengan langkah lain dalam halaman yang berbeda. |

Hypertext Preprocessor (PHP)

Menurut Wahyono (2009), PHP (dulu: *Personal Home Page*, sekarang PHP : *Hypertext Preprocessor*) merupakan script untuk membuat suatu aplikasi yang akan dapat terintegrasi ke dalam halaman HTML, sehingga suatu halaman web tidak lagi bersifat statis, namun menjadi bersifat dinamis. Menurut Anhar (2010), PHP singkatan dari *Hypertext Preprocessor* yaitu bahasa pemrograman web server-side yang bersifat *open source*.

Berdasarkan pengertian dari beberapa ahli di atas dapat disimpulkan bahwa PHP merupakan skrip yang bertempat dan di proses pada suatu server dengan keluaran yang dihasilkan dapat dilihat melalui browser, PHP juga merupakan salah satu bahasa pemrograman *open source* yang dapat digunakan pada berbagai sistem operasi seperti Linux, Unix, Macintosh, maupun Windows. Pada dasarnya PHP dirancang untuk pembuatan jenis web dinamis, yaitu web yang dalam pembuatannya dapat mengaplikasikan sesuai keinginan penggunanya. Salah satu kelebihan lain yang dimiliki PHP antara lain dapat terkoneksi pada beberapa database antara lain MySQL. Kelebihan dari PHP adalah :

PHP mudah dibuat dan dijalankan, maksudnya php dapat berjalan dalam web server apapun dan dapat dijalankan dengan sistem operasi yang berada seperti Windows dan Unix.

PHP bersifat efisien, karena hanya memerlukan *resource* sistem yang sangat sedikit dibandingkan dengan bahasa pemrograman lainnya.

MySQL

MySQL adalah sebuah sistem manajemen database relasi (*relational database management system*) yang bersifat *open source* (Arbie, 2004: 5).MySQL merupakan buah pikiran dari Michael “Monty” Widenius, David Axmark dan Allan Larson yang di mulai tahun 1995. Mereka bertiga kemudian mendirikan perusahaan bernama MySQL AB di Swedia. Pengertian MySQL menurut MySQL manual adalah sebuah open source software database SQL (*Search Query Language*) yang menangani sistem manajemen database dan sistem manajemen database relational.

MySQL adalah open source software yang dibuat oleh sebuah perusahaan Swedia yaitu MySQL AB. MySQL mempunyai fitur-fitur yang sangat mudah dipelajari bagi para penggunanya dan dikembangkan untuk menangani database yang besar dengan waktu yang lebih singkat. Kecepatan, konektivitas dan keamanannya yang lebih baik membuat MySQL sangat dibutuhkan untuk mengakses database di internet.

MySQL versi 1.0 di rilis pada Mei 1996 dan penggunaannya hanya terbatas di kalangan perusahaan saja. Barulah pada bulan Oktober 1996, MySQL versi 3.11.0 di rilis ke masyarakat luas. MySQL menggunakan bahasa standar SQL (Structure Query Language) sebagai bahasa interaktif dalam mengelola data. MySQL memiliki kinerja, kecepatan proses dan ketangguhan yang tidak kalah dibanding database-database besar lainnya yang komersil seperti ORACLE, Sybase, Unify dan sebagainya. MySQL dapat berjalan di atas banyak system operasi seperti Linux, Windows, Solaris, FreeBSD, Mac OS X, dan lain sebagainya.

Akurasi

Sebuah sistem yang melakukan pengujian diharapkan dapat melakukan pengujian semua set data dengan benar. Oleh karena itu, sebuah sistem pengujian harus diukur kinerjanya. Umum cara mengukur kinerja pengujian menggunakan *Confusion Matrix* (Prasetyo, 2014).

Confusion Matrix merupakan tabel yang mencatat hasil kinerja klasifikasi. Tabel 2.6 merupakan contoh Confusion Matrix yang melakukan pengujian masalah biner (dua kelas) untuk 2 kelas, misalnya kelas 0 dan 1. Setiap sel f_{ij} dalam matriks menyatakan jumlah record/data dari kelas i yang hasil prediksinya masuk ke kelas j . Adapun bentuk *confusion matrix* dapat dilihat pada tabel 2.7

Tabel 2.7 Confusion Matrix 2 kelas

| f_{ij} | | Kelas Hasil Prediksi | |
|--------------------|-----------|----------------------|----------|
| | | Kelas=1 | Kelas=0 |
| Kelas Asli (i) | Kelas = 1 | f_{11} | f_{10} |
| | Kelas = 0 | f_{01} | f_{00} |

Berdasarkan isi Confusion Matriks, maka dapat diketahui jumlah data dari masing-masing kelas yang di prediksi secara benar yaitu ($f_{11} + f_{00}$) dan data yang diprediksi secara salah yaitu yaitu ($f_{10} + f_{01}$). Kuantitas Confusion Matrix dapat dilakukan akurasi. Dengan mengetahui jumlah data yang diuji secara benar maka dapat diketahui akurasi hasil prediksi. Untuk menghitung akurasi digunakan formula sebagai berikut:

$$\text{Akurasi} = SV = \left(\frac{p}{p+n} \right) + sp \left(\frac{n}{n+p} \right) \dots\dots\dots (2.5)$$

BAB III

METODOLOGI PENELITIAN

3.1 Alat dan Bahan Penelitian yang Digunakan

Adapun kebutuhan spesifikasi perangkat keras dan perangkat lunak untuk perancangan pada penelitian ini adalah :

3.1.1 Spesifikasi Kebutuhan *Hardware*

Perangkat keras (*hardware*) yang digunakan dalam membangun sistem ini adalah:

Processor : *Intel Core i3-4030U*

Ram : 2 GB

Hardisk : 500 GB

3.1.2 Spesifikasi Kebutuhan *Software*

Perangkat lunak (*software*) yang digunakan dalam pembuatan sistem ini adalah :

Sistem Operasi : *Microsoft Windows 8.1 Pro*

Bahasa Pemrograman : HTML, PHP dan CSS

Database Management System

(DBMS) : *MySQL*

Web Browser : Google Chrome , Mozilla Firefox

Desain Logika Program : CorelDraw X4, DIA Graph Editor

Metode Pengumpulan Data

Teknik Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam rangkaian penelitian. Adapun metode pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut:

Studi Pustaka

Mengumpulkan data dengan cara mencari dan mempelajari dari berbagai sumber yang berkaitan dengan masalah yang diteliti dalam penyusunan tugas akhir ini, baik dari internet, buku, jurnal ilmiah dan dari bacaan lain yang dapat dipertanggungjawabkan.

Wawancara

Wawancara untuk pengumpulan informasi yang berkaitan tentang penelitian baik itu kepada mahasiswa bidang geologi, ataupun kepada Bapak Dewanra Bagus Eka Putra, B.Sc (Hons)., M.Sc yang mengetahui tentang air sumur dan tanah.

Jenis Data

Data primer adalah data yang diperoleh langsung dari Bapak Dewanra Bagus Eka Putra, B.Sc (Hons)., M.Sc melalui wawancara. Pada penelitian ini data yang didapatkan adalah data mengenai parameter yang digunakan untuk penentuan kualitas air sumur seperti suhu, pH, EC, bau, rasa dan warna.

Data sekunder adalah data yang diperoleh penulis dari teknik studi pustaka, dalam hal ini adalah data yang mengenai definisi, alat, cara kerja dan solusi penanganan air sumur yang berbau, berasa tak sedap dan berwarna.

3.3 Analisa Sistem yang Sedang Berjalan

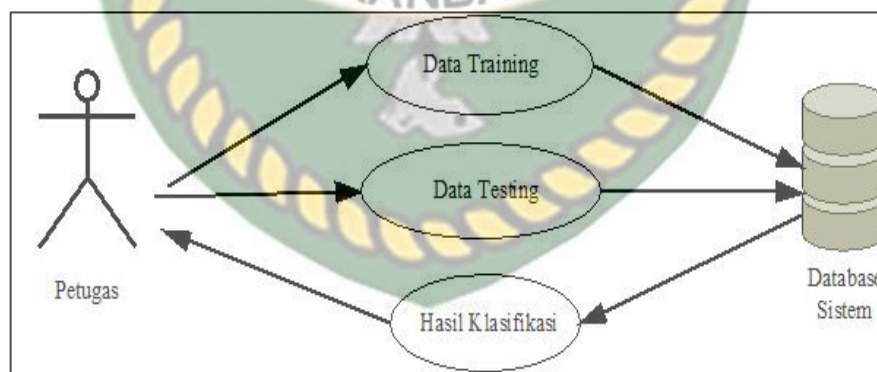
Sebelum sistem prediksi bau, rasa dan warna pada air sumur ini dirancang, sistem yang berjalan adalah narasumber menyimpan data sebagai dokumentasi laporan tahunan. Analisa sistem yang sedang berjalan bisa dilihat dalam bentuk *Use Case Diagram* pada gambar 3.1.



Gambar 3.1 Analisa Sitem yang Sedang Berjalan

3.4 Perancangan Sistem

Sistem yang akan dibangun digambarkan secara detail melalui perancangan sistem yang bisa dilihat pada gambar 3.2.



Gambar 3.2 Analisa Sistem yang Diusulkan

Berikut adalah gambaran umum sistem yang diusulkan :

User melakukan login ke sistem dan diarahkan ke menu utama.

Menu yang disediakan oleh sistem ini adalah menu penginputan data *training*, penginputan data *testing*, perhitungan *k-fold cross validation* menu profil dan menu panduan.

Pada menu data training, user melakukan penginputan data training berupa data suhu, pH dan EC. Data yang telah diinputkan akan diproses oleh sistem menggunakan algoritma *k-Nearest Neighbor* pada menu *testing*. Kemudian dari hasil perhitungan tersebut akan menghasilkan output berupa nilai klasifikasi untuk nilai bau, rasa dan warna air sumur.

3.4.1 Context Diagram

Context Diagram digunakan untuk menggambarkan hubungan input dan *output* antara sistem dengan entitas luar, suatu diagram konteks selalu memiliki satu proses yang mewakili seluruh sistem. Sistem ini memiliki satu buah eksternal *entity* yaitu petugas (admin). Perancangan *Context Diagram* dapat dilihat pada gambar 3.3.



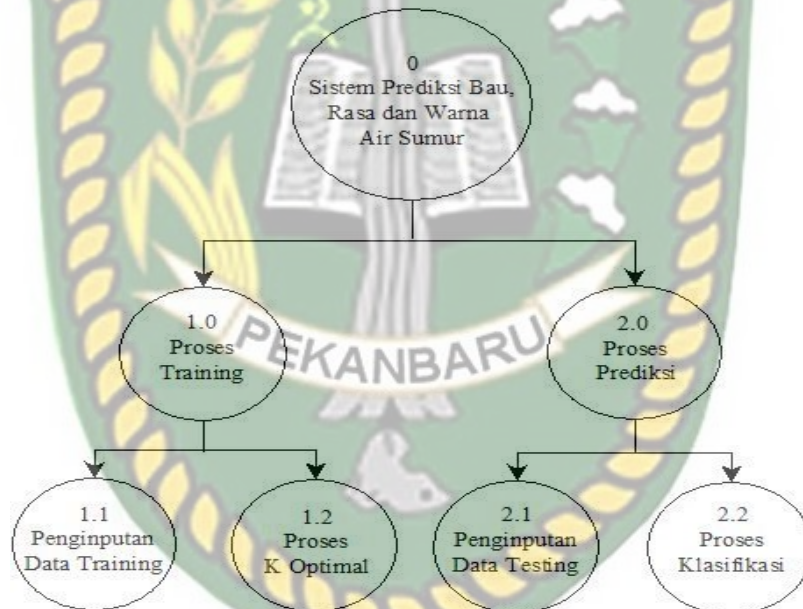
Gambar 3.3 Konteks Diagram Sistem Prediksi Bau, Rasa dan Warna Air Sumur

Berdasarkan gambar 3.3 seorang petugas (admin) akan menginputkan data training sebagai nilai yang akan diproses oleh sistem dan petugas juga dapat

melihat hasil perhitungan dari data tersebut. Data testing yang diinput akan diproses oleh sistem dan data tersebut akan menghasilkan sebuah keluaran berupa hasil klasifikasi dari bau, rasa dan warna.

3.4.2 *Hierarchy Chart*

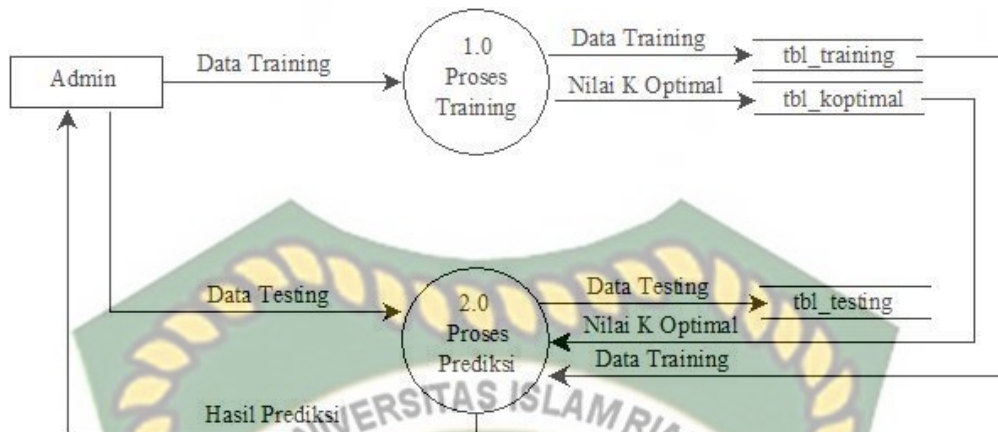
Hierarchy chart digunakan untuk memperlihatkan jenjang atau hirarki dari program yang akan dikembangkan. Dengan demikian dapat dijabarkan urutan kerja dari tiap program. *Hierarchy chart* sistem yang akan dibangun bisa dilihat pada gambar 3.4.



Gambar 3.4 *Hierarchy Chart* Sistem Prediksi Bau, Rasa dan Warna Air Sumur

3.4.3 *Data Flow Diagram (DFD) Level 0*

Data Flow Diagram (DFD) berfungsi untuk menggambarkan suatu sistem yang telah ada atau sistem baru yang akan dikembangkan secara logika tanpa memperhatikan lingkungan fisik dimana data tersebut mengalir. Rincian dari proses dapat akan diuraikan pada DFD Level 0 seperti pada gambar 3.5.

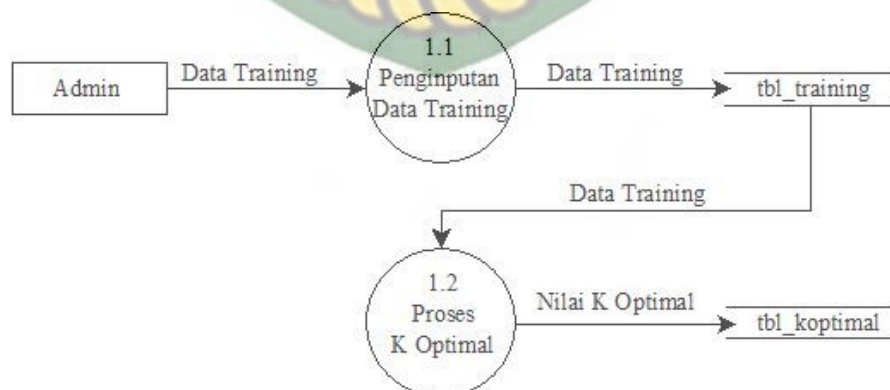


Gambar 3.5 Data Flow Diagram (DFD) Level 0

Berdasarkan gambar 3.5 DFD level 0 dapat dilihat aliran data pada sistem yang akan dibangun. Proses pertama yang dilakukan adalah proses training. Adapun data yang akan diproses yaitu data training dan melakukan proses k optimal untuk mendapatkan nilai k yang optimal.

3.4.4 Data Flow Diagram (DFD) Level 1 Proses 1

DFD level 1 digunakan untuk menggambarkan modul-modul yang ada dalam sistem yang akan dikembangkan. DFD level 1 merupakan *breakdown* DFD level 0 yang sebelumnya sudah dibuat. Rincian dari proses dapat akan diuraikan pada DFD Level 1 proses 1 seperti pada gambar 3.6.

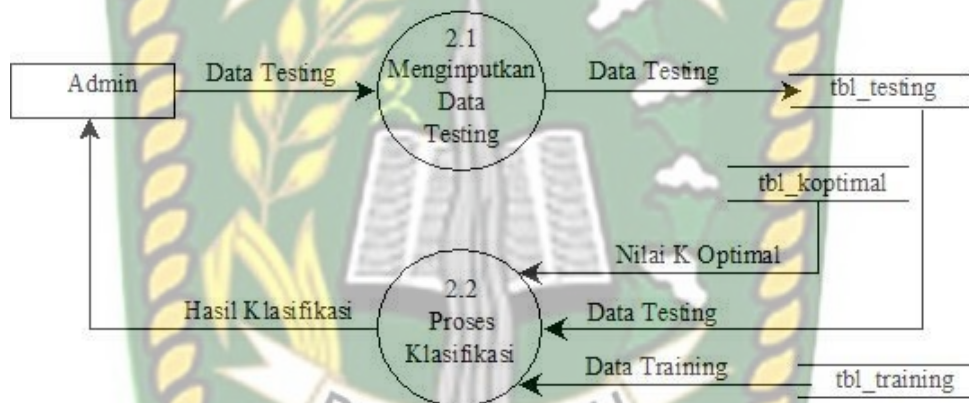


Gambar 3.6 Data Flow Diagram (DFD) Level 1 Proses 1

Berdasarkan gambar 3.6 Data flow diagram level 1 proses 1 menggambarkan proses pengelolaan data training. Pada proses 1.1 admin menginput data training ke dalam sistem. Selanjutnya akan dibentuk fold pada data training. Kemudian akan dilakukan proses knn pada setiap fold untuk mendapatkan nilai k optimal.

3.4.5 Data Flow Diagram (DFD) Level 1 Proses 2

Rincian dari proses dapat akan diuraikan pada DFD Level 1 proses 2 seperti pada gambar 3.7.



Gambar 3.7 Data Flow Diagram (DFD) Level 1 Proses 2

Berdasarkan gambar 3.7, pada proses 2.1 admin menginputkan data testing kedalam sistem. Kemudian pada proses 2.2 dilakukan proses knn pada data testing yang diinputkan.

3.5 Rancangan Desain

3.5.1 Rancangan Desain Output

Beberapa desain *output* yang dirancang dalam sistem yang dibangun :

Desain *Output* Data Training

Gambaran *output* untuk melihat data training yang sudah diinputkan seperti pada gambar 3.8.

| Hasil Perhitungan KNN | | | | | | | |
|-----------------------|-----------|-------|------|-----------|-----------|---------------|-------|
| No | Suhu | pH | EC | Bau | Rasa | Warna | Jarak |
| 999 | 999 | 999 | 999 | X (30) | X (30) | X (30) | 999 |
| Bau | | Rasa | | | Warna | | |
| Ada | Tidak Bau | Tawar | Asin | Payau | Bewarna | Tidak Bewarna | |
| 999 | 999 | 999 | 999 | 999 | 999 | 999 | |
| Hasil Kelas Tertinggi | | | | | | | |
| X(30) | | X(30) | | | X(30) | | |
| Kesimpulan: | | | | | | | |

Gambar 3.10 Desain *Output* Perhitungan k-NN

3.5.2 Rancangan Desain *Input*

Desain *input* adalah bentuk masukan pada sebuah sistem yang akan diproses untuk menghasilkan sebuah informasi.

Desain *Input Login*

Desain input login adalah rancangan yang dibuat untuk peneliti melakukan login terhadap sistem. Rancangan desain login dapat dilihat pada gambar 3.11.

**SISTEM PREDIKSI
BAU, RASA DAN WARNA AIR SUMUR**

Username:

Password:

Gambar 3.11 Desain *Input Login*

Desain *Input Data Training*

Desain input data training merupakan form yang dirancang untuk melakukan input data training. Desain input dapat dilihat pada gambar 3.12.

Nama Aplikasi / Sistem

Foto Username

Selamat datang Username, Hari ini tanggal : d M Y

Tambah Data Lihat Data Upload Data

Data Training Air Sumur

| | | | |
|--|-----------|-------|---|
| <input checked="" type="checkbox"/> Data Training | 1.No Data | 999 | |
| Tambah Data | 2. Suhu | 999 | |
| Lihat Data | 3. pH | 999 | |
| <input checked="" type="checkbox"/> Data Testing | 4. EC | 999 | |
| Tambah Data | 5. Bau | x(30) | ▼ |
| Lihat Data | 6. Rasa | x(30) | ▼ |
| <input checked="" type="checkbox"/> Nilai Akurasi | 7. Warna | x(30) | ▼ |
| <input checked="" type="checkbox"/> Panduan | | | |
| <input checked="" type="checkbox"/> Akun | | | |

Gambar 3.12 Desain *Input Data Training*

Desain *Input Data Testing*

Desain *input data testing* merupakan form yang dirancang untuk melakukan *input data training*, desain dapat dilihat pada gambar 3.13.

| Data Testing Air Sumur | |
|------------------------|-----|
| 1.No Data | 999 |
| 2. Suhu | 999 |
| 3. pH | 999 |
| 4. EC | 999 |

Gambar 3.13 Desain *Input* Data Testing

Desain *Input k-Fold Cross Validation*

Desain *input* k-fold merupakan form yang dirancang untuk melihat nilai k optimal dan akurasi, desain dapat dilihat pada gambar 3.14.

| Nilai Akurasi & K Optimal | |
|--------------------------------------|-----|
| Tampilkan Perhitungan Hingga Nilai K | 999 |

Gambar 3.14 Desain *Input* Perhitungan *K-Fold Cross Vaidation*

3.5.3 Desain *Database*

Dalam pembuatan sistem ini menggunakan sebuah database “data_sumur” yang terdiri dari beberapa file atau tabel sebagai berikut :

Tabel User

Tabel user merupakan tabel untuk menyimpan data login. Tabel ini mempunyai 4 field, dimana diantaranya username sebagai primary key, password, nama_lengkap dan foto. Rancangan tabel user dapat dilihat pada tabel 3.1.

Tabel 3.1 Tabel User

| No | Field | Data Type | Size | Description |
|----|--------------|-----------|------|---------------|
| 1 | Username | Varchar | 30 | Primary Key |
| 2 | Password | Varchar | 30 | Password User |
| 3 | nama_lengkap | Varchar | 30 | Nama User |
| 4 | Foto | Varchar | 200 | Foto User |

Tabel Data Training

Tabel training merupakan tabel untuk menyimpan data training. Tabel ini mempunyai 7 field, dimana diantaranya No_Data sebagai primary key, suhu, pH, EC, bau, rasa dan warna. Rancangan tabel training dapat dilihat pada tabel 3.2.

Tabel 3.2 Tabel Training

| No | Field | Data Type | Size | Description |
|----|---------|------------------------------|------|-------------------------|
| 1 | No_Data | Char | 10 | Primary Key |
| 2 | Suhu | Float | - | Suhu Data Training |
| 3 | pH | Float | - | Potensial Hidrogen |
| 4 | EC | Float | - | Electrical Conductivity |
| 5 | Bau | Enum (‘Ada’, ‘Tidak Bau’) | | Bau Data Training |
| 6 | Rasa | Enum | | Rasa Data Training |

| | | | |
|---|-------|--|----------------------------|
| | | (‘Tawar’, ‘Asin’, ‘Payau’) | |
| 7 | Warna | Enum (‘Berwarna’, ‘Tidak Berwarna’) | Warna Data <i>Training</i> |

z

Tabel Testing

Tabel testing merupakan tabel untuk menyimpan data testing sebelum disimpan ke tabel training. Tabel ini mempunyai 8 field, dimana diantaranya No_Data sebagai primary key, suhu, pH, EC, bau, rasa, warna dan ket. Rancangan tabel training dapat dilihat pada tabel 3.3.

Tabel 3.3 Tabel Testing

| No | Field | Data Type | Size | Description |
|----|---------|--|------|---------------------------|
| 1 | No_Data | Char | 10 | Primary Key |
| 2 | Suhu | Float | - | Suhu Data <i>Testing</i> |
| 3 | pH | Float | - | Potensial Hidrogen |
| 4 | EC | Float | - | Electrical Conductivity |
| 5 | Bau | Enum (‘Ada’, ‘Tidak Bau’) | | Bau Data <i>Testing</i> |
| 6 | Rasa | Enum (‘Tawar’, ‘Asin’, ‘Payau’) | | Rasa Data <i>Testing</i> |
| 7 | Warna | Enum (‘Berwarna’, ‘Tidak Berwarna’) | | Warna Data <i>Testing</i> |
| 8 | Ket | Varchar | 15 | Keterangan |

Tabel K Optimal

Tabel k optimal merupakan tabel untuk menyimpan nilai k optimal. Tabel ini mempunyai 4 field, yaitu id_k sebagai primary key, nilai_k, akurasi dan tgl. Rancangan tabel training dapat dilihat pada tabel 3.4.

Tabel 3.4 Tabel K Optimal

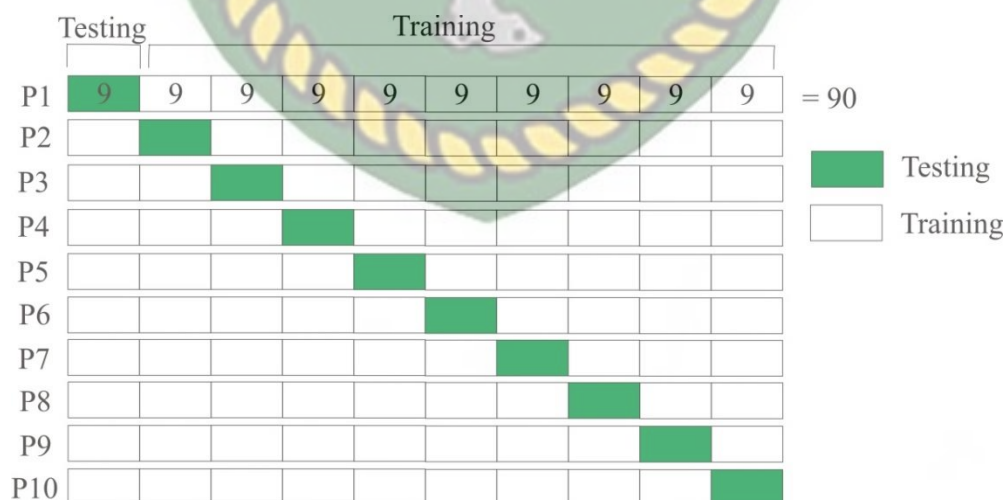
| No | Field | Data Type | Size | Description |
|----|---------|-----------|------|-------------------|
| 1 | id_k | Int | 5 | Primary Key |
| 2 | nilai_k | Int | 5 | Nilai K Optimal |
| 3 | akurasi | Float | - | Akurasi K Optimal |
| 4 | Tgl | Timestamp | - | Tanggal |

3.5.4 Perhitungan Manual

A. K Fold Cross Validation

Berikut adalah penjelasan perhitungan manual yang terjadi dalam system. Metode 10 fold cross validation digunakan dalam sistem ini untuk mendapatkan nilai k yang optimal. Pembagian data k-fold dapat dilihat pada gambar 3.14.

Data training : $90 / 10 \text{ fold} = 9$ data per blok.



Gambar 3.14 Pembagian data pada metode 10-Fold Cross Validation

Dilakukan prediksi menggunakan algoritma k-NN dengan setiap nilai k yang dimasukkan. Dalam penelitian ini, nilai k yang digunakan yaitu 3,5,7,9,11,13,15,17,19 dan hasil perhitungannya dapat dilihat pada tabel 3.5 sampai dengan tabel 3.13.

Tabel 3.5 Tabel Hasil Perhitungan k =3

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|------------------------------------|--------------|---------------|--------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 66.67 | 66.67 | 33.33 |
| 2 | 10-18 | 1-9, 19-90 | 66.67 | 88.89 | 44.44 |
| 3 | 19-27 | 1-20, 28-90 | 55.56 | 44.44 | 77.78 |
| 4 | 28-36 | 1-27, 37-90 | 55.56 | 88.89 | 55.56 |
| 5 | 37-45 | 1-36, 46-90 | 55.56 | 100 | 55.56 |
| 6 | 46-54 | 1-45, 55-90 | 88.89 | 66.67 | 66.67 |
| 7 | 55-63 | 1-54, 64-90 | 66.67 | 55.56 | 22.22 |
| 8 | 64-72 | 1-63, 73-90 | 66.67 | 88.89 | 66.67 |
| 9 | 73-81 | 1-72, 82-90 | 66.67 | 88.89 | 44.44 |
| 10 | 82-90 | 1-81 | 44.44 | 66.67 | 66.67 |
| Rata – Rata (%) | | | 63.33 | 75.56 | 53.33 |
| Rata – Rata Keseluruhan (%) | | | 64.07 | | |

Tabel 3.6 Tabel Hasil Perhitungan k =5

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|------------------------------------|--------------|---------------|--------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 66.67 | 88.89 | 66.67 |
| 2 | 10-18 | 1-9, 19-90 | 55.56 | 77.78 | 44.44 |
| 3 | 19-27 | 1-20, 28-90 | 88.89 | 44.44 | 77.78 |
| 4 | 28-36 | 1-27, 37-90 | 55.56 | 88.89 | 55.56 |
| 5 | 37-45 | 1-36, 46-90 | 33.33 | 100 | 33.33 |
| 6 | 46-54 | 1-45, 55-90 | 88.89 | 55.56 | 66.67 |
| 7 | 55-63 | 1-54, 64-90 | 77.78 | 66.67 | 22.22 |
| 8 | 64-72 | 1-63, 73-90 | 55.56 | 100 | 55.56 |
| 9 | 73-81 | 1-72, 82-90 | 55.56 | 88.89 | 55.56 |
| 10 | 82-90 | 1-81 | 66.67 | 88.89 | 55.56 |
| Rata – Rata (%) | | | 64.44 | 80 | 53.33 |
| Rata – Rata Keseluruhan (%) | | | 65.93 | | |

Tabel 3.7 Tabel Hasil Perhitungan k =7

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|------------------------------------|--------------|---------------|--------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 66.67 | 88.89 | 66.67 |
| 2 | 10-18 | 1-9, 19-90 | 55.56 | 77.78 | 44.44 |
| 3 | 19-27 | 1-20, 28-90 | 100 | 44.44 | 77.78 |
| 4 | 28-36 | 1-27, 37-90 | 77.78 | 88.89 | 55.56 |
| 5 | 37-45 | 1-36, 46-90 | 33.33 | 88.89 | 44.44 |
| 6 | 46-54 | 1-45, 55-90 | 77.78 | 44.44 | 77.78 |
| 7 | 55-63 | 1-54, 64-90 | 66.67 | 66.67 | 33.33 |
| 8 | 64-72 | 1-63, 73-90 | 55.56 | 88.89 | 55.56 |
| 9 | 73-81 | 1-72, 82-90 | 77.78 | 88.89 | 55.56 |
| 10 | 82-90 | 1-81 | 77.78 | 66.67 | 66.67 |
| Rata – Rata (%) | | | 68.89 | 74.44 | 57.78 |
| Rata – Rata Keseluruhan (%) | | | 67.04 | | |

Tabel 3.8 Tabel Hasil Perhitungan k =9

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|------------------------------------|--------------|---------------|--------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 66.67 | 88.89 | 55.56 |
| 2 | 10-18 | 1-9, 19-90 | 55.56 | 77.78 | 55.56 |
| 3 | 19-27 | 1-20, 28-90 | 77.78 | 44.44 | 77.78 |
| 4 | 28-36 | 1-27, 37-90 | 55.56 | 77.78 | 55.56 |
| 5 | 37-45 | 1-36, 46-90 | 33.33 | 88.89 | 44.44 |
| 6 | 46-54 | 1-45, 55-90 | 77.78 | 44.44 | 88.89 |
| 7 | 55-63 | 1-54, 64-90 | 66.67 | 66.67 | 44.44 |
| 8 | 64-72 | 1-63, 73-90 | 55.56 | 88.89 | 55.56 |
| 9 | 73-81 | 1-72, 82-90 | 66.67 | 88.89 | 66.67 |
| 10 | 82-90 | 1-81 | 77.78 | 66.67 | 66.67 |
| Rata – Rata (%) | | | 63.33 | 73.33 | 61.11 |
| Rata – Rata Keseluruhan (%) | | | 65.93 | | |

Tabel 3.9 Tabel Hasil Perhitungan k =11

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|-----------|--------------|---------------|-------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 88.89 | 88.89 | 55.56 |
| 2 | 10-18 | 1-9, 19-90 | 55.56 | 77.78 | 55.56 |

| | | | | | |
|------------------------------------|-------|-------------|--------------|--------------|--------------|
| 3 | 19-27 | 1-20, 28-90 | 100 | 44.44 | 77.78 |
| 4 | 28-36 | 1-27, 37-90 | 55.56 | 77.78 | 55.56 |
| 5 | 37-45 | 1-36, 46-90 | 44.44 | 88.89 | 66.67 |
| 6 | 46-54 | 1-45, 55-90 | 88.89 | 66.67 | 88.89 |
| 7 | 55-63 | 1-54, 64-90 | 66.67 | 66.67 | 55.56 |
| 8 | 64-72 | 1-63, 73-90 | 66.67 | 88.89 | 55.56 |
| 9 | 73-81 | 1-72, 82-90 | 55.56 | 88.89 | 77.78 |
| 10 | 82-90 | 1-81 | 77.78 | 66.67 | 66.67 |
| Rata – Rata (%) | | | 70 | 75.56 | 65.56 |
| Rata – Rata Keseluruhan (%) | | | 70.37 | | |

Tabel 3.10 Tabel Hasil Perhitungan k =13

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|------------------------------------|--------------|---------------|--------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 66.67 | 88.89 | 55.56 |
| 2 | 10-18 | 1-9, 19-90 | 44.44 | 77.78 | 44.44 |
| 3 | 19-27 | 1-20, 28-90 | 88.89 | 44.44 | 66.67 |
| 4 | 28-36 | 1-27, 37-90 | 55.56 | 77.78 | 66.67 |
| 5 | 37-45 | 1-36, 46-90 | 33.33 | 88.89 | 44.44 |
| 6 | 46-54 | 1-45, 55-90 | 77.78 | 77.78 | 77.78 |
| 7 | 55-63 | 1-54, 64-90 | 66.67 | 66.67 | 55.56 |
| 8 | 64-72 | 1-63, 73-90 | 66.67 | 88.89 | 44.44 |
| 9 | 73-81 | 1-72, 82-90 | 66.67 | 88.89 | 66.67 |
| 10 | 82-90 | 1-81 | 77.78 | 66.67 | 33.33 |
| Rata – Rata (%) | | | 64.44 | 76.67 | 55.56 |
| Rata – Rata Keseluruhan (%) | | | 65.56 | | |

Tabel 3.11 Tabel Hasil Perhitungan k =15

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|-----------|--------------|---------------|-------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 66.67 | 88.89 | 77.78 |
| 2 | 10-18 | 1-9, 19-90 | 55.56 | 77.78 | 55.56 |
| 3 | 19-27 | 1-20, 28-90 | 88.89 | 44.44 | 66.67 |
| 4 | 28-36 | 1-27, 37-90 | 55.56 | 77.78 | 66.67 |
| 5 | 37-45 | 1-36, 46-90 | 44.44 | 88.89 | 55.56 |
| 6 | 46-54 | 1-45, 55-90 | 77.78 | 77.78 | 77.78 |

| | | | | | |
|------------------------------------|-------|-------------|--------------|--------------|-----------|
| 7 | 55-63 | 1-54, 64-90 | 66.67 | 66.67 | 55.56 |
| 8 | 64-72 | 1-63, 73-90 | 66.67 | 88.89 | 44.44 |
| 9 | 73-81 | 1-72, 82-90 | 66.67 | 88.89 | 66.67 |
| 10 | 82-90 | 1-81 | 66.67 | 66.67 | 33.33 |
| Rata – Rata (%) | | | 65.56 | 76.67 | 60 |
| Rata – Rata Keseluruhan (%) | | | 67.41 | | |

Tabel 3.12 Tabel Hasil Perhitungan k =17

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|------------------------------------|--------------|---------------|--------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 66.67 | 88.89 | 66.67 |
| 2 | 10-18 | 1-9, 19-90 | 55.56 | 77.78 | 55.56 |
| 3 | 19-27 | 1-20, 28-90 | 66.67 | 44.44 | 66.67 |
| 4 | 28-36 | 1-27, 37-90 | 44.44 | 77.78 | 55.56 |
| 5 | 37-45 | 1-36, 46-90 | 44.44 | 88.89 | 55.56 |
| 6 | 46-54 | 1-45, 55-90 | 77.78 | 77.78 | 77.78 |
| 7 | 55-63 | 1-54, 64-90 | 66.67 | 66.67 | 55.56 |
| 8 | 64-72 | 1-63, 73-90 | 66.67 | 88.89 | 44.44 |
| 9 | 73-81 | 1-72, 82-90 | 66.67 | 88.89 | 55.56 |
| 10 | 82-90 | 1-81 | 55.56 | 66.67 | 33.33 |
| Rata – Rata (%) | | | 61.11 | 76.67 | 56.67 |
| Rata – Rata Keseluruhan (%) | | | 64.81 | | |

Tabel 3.13 Tabel Hasil Perhitungan k =19

| Pengujian | Data Testing | Data Training | Akurasi Bau | Akurasi Rasa | Akurasi Warna |
|-----------|--------------|---------------|-------------|--------------|---------------|
| 1 | 1-9 | 10-90 | 66.67 | 88.89 | 66.67 |
| 2 | 10-18 | 1-9, 19-90 | 66.67 | 77.78 | 44.44 |
| 3 | 19-27 | 1-20, 28-90 | 66.67 | 44.44 | 66.67 |
| 4 | 28-36 | 1-27, 37-90 | 55.56 | 88.89 | 66.67 |
| 5 | 37-45 | 1-36, 46-90 | 66.67 | 88.89 | 55.56 |
| 6 | 46-54 | 1-45, 55-90 | 77.78 | 77.78 | 88.89 |
| 7 | 55-63 | 1-54, 64-90 | 66.67 | 66.67 | 55.56 |
| 8 | 64-72 | 1-63, 73-90 | 55.56 | 88.89 | 44.44 |
| 9 | 73-81 | 1-72, 82-90 | 66.67 | 88.89 | 55.56 |
| 10 | 82-90 | 1-81 | 55.56 | 66.67 | 33.33 |

| | | | |
|------------------------------------|--------------|--------------|--------------|
| Rata – Rata (%) | 64.44 | 77.78 | 57.78 |
| Rata – Rata Keseluruhan (%) | 66.67 | | |

Berdasarkan rata-rata perhitungan pada tabel 3.4 sampai dengan tabel 3.13, maka k optimal yang didapatkan adalah 11.

B. Perhitungan k-Nearest Neighbor

Berikut proses perhitungan manual yang dilakukan dalam sistem ini. Terdapat beberapa data air sumur di daerah Selat Baru yang digunakan sebagai data training untuk diklasifikasikan dengan data testing menggunakan empat atribut yaitu Suhu, pH, EC sehingga dapat menentukan klasifikasi bau, rasa dan warna pada data testing tersebut. Data training yang digunakan dalam sistem ini dapat dilihat pada tabel 3.14 (data training selengkapnya dapat dilihat pada lampiran 3.1).

Tabel 3.14 Data Training

| No Data | Suhu (C) | pH | EC (mikroS/cm) | Klasifikasi Bau | Klasifikasi Rasa | Klasifikasi Warna |
|---------|----------|------|----------------|-----------------|------------------|-------------------|
| 1 | 28,2 | 6,55 | 125,8 | Tidak Bau | Tawar | Berwarna |
| 2 | 26,6 | 6,72 | 825 | Ada | Tawar | Berwarna |
| 3 | 26,6 | 6,71 | 597 | Ada | Tawar | Tidak Berwarna |
| .. | .. | .. | .. | .. | .. | .. |
| .. | .. | .. | .. | .. | .. | .. |
| 89 | 26,9 | 7,14 | 1201 | Ada | Payau | Tidak Berwarna |
| 90 | 29 | 6,35 | 11485 | Ada | Asin | Tidak Berwarna |

Setelah pembuatan data training maka perlu adanya data testing dengan atribut suhu, pH, EC. Data testing yang akan diuji dapat dilihat pada tabel 3.15.

Tabel 3.15 Data Testing

| No | Suhu (C) | pH | EC (MikroS/cm) |
|----|----------|------|----------------|
| 1 | 26,2 | 6,81 | 833 |
| 2 | 27,3 | 7,32 | 1236 |
| 3 | 27 | 7 | 226 |
| 4 | 27 | 6,77 | 718 |
| 5 | 27 | 7,3 | 5209 |
| 6 | 29,2 | 6,6 | 7220 |
| 7 | 27,2 | 7,34 | 1881 |
| 8 | 28 | 6,51 | 794 |
| 9 | 27 | 7,36 | 757 |
| 10 | 27,8 | 6,81 | 1067 |
| 11 | 27,8 | 6,81 | 822 |
| 12 | 27,6 | 7,25 | 1357 |
| 13 | 27,1 | 7,1 | 914 |
| 14 | 27 | 7,2 | 1032 |
| 15 | 28 | 7 | 478 |
| 16 | 27 | 6 | 1475 |
| 17 | 28 | 6 | 400 |
| 18 | 27 | 6 | 447 |
| 19 | 26 | 7 | 1645 |
| 20 | 26 | 6 | 861 |

Dokumen ini adalah Arsip Miik :

Perpustakaan Universitas Islam Riau

1. Penetapan Nilai k

Langkah awal metode k-Nearest Neighbor adalah menentukan nilai k (tetangga terdekat) berdasarkan hasil pencarian k-Optimal yang didapatkan yaitu k=11.

2. Perhitungan Nilai Jarak

Selanjutnya adalah mencari nilai jarak pada masing masing data training. Pertama, tentukan terlebih dahulu nilai jarak terdekat menggunakan rumus Euclidean(2.1).

Dalam kasus ini maka rumus akan berbentuk seperti ini :

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$$

Berdasarkan rumus diatas diketahui :

x1 = Suhu dari data training

y1 = Suhu dari data testing

x2 = pH dari data training

y2 = pH dari data testing

x3 = EC dari data training

y3 = EC dari data testing

Berikut contoh perhitungan jarak dengan menggunakan data testing no 1 pada tabel 3.15 :

Suhu = 26,2

pH = 6,81

EC = 833

$$D1 \sqrt{(28,2 - 26,2)^2 + (6,55 - 6,81)^2 + (125,8 - 833)^2} = 2068,82$$

$$D2 \sqrt{(26,6 - 26,2)^2 + (6,72 - 6,81)^2 + (825 - 833)^2} = 58,73$$

$$D3 \sqrt{(26,6 - 26,2)^2 + (6,71 - 6,81)^2 + (597 - 833)^2} = 585,24$$

.. ...

.. ..

$$D89 \sqrt{(26,9 - 26,2)^2 + (7,14 - 6,81)^2 + (1201 - 833)^2} = 439,06$$

$$D90 \sqrt{(29 - 26,2)^2 + (6,35 - 6,81)^2 + (11485 - 833)^2} = 666,77$$

Hasil perhitungan nilai jarak dapat dilihat pada tabel 3.16 (perhitungan nilai jarak selengkapnya dapat dilihat pada lampiran 3.3).



Tabel 3.16 Perhitungan Nilai Jarak

| No Data | Suhu (C) | pH | EC (mikroS/cm) | Klasifikasi Bau | Klasifikasi Rasa | Klasifikasi Warna | Jarak |
|---------|----------|------|----------------|-----------------|------------------|-------------------|--------|
| 1 | 28,2 | 6,55 | 125,8 | Tidak Bau | Tawar | Berwarna | 707,20 |
| 2 | 26,6 | 6,72 | 825 | Ada | Tawar | Berwarna | 8,01 |
| 3 | 26,6 | 6,71 | 597 | Ada | Tawar | Tidak Berwarna | 236 |
| .. | .. | .. | .. | .. | .. | .. | ... |
| .. | .. | .. | .. | .. | .. | .. | ... |
| 89 | 26,9 | 7,14 | 1201 | Ada | Payau | Tidak Berwarna | 368 |
| 90 | 29 | 6,35 | 11485 | Ada | Asin | Tidak Berwarna | 10652 |

3. Pengurutan Data

Setelah perhitungan jarak selesai, akan diurutkan data dari nilai jarak terkecil ke terbesar. Berdasarkan kasus diatas setelah terjadi proses perhitungan dan pengurutan hasilnya akan terlihat seperti tabel 3.17 (pengurutan data selengkapnya dapat dilihat pada lampiran 3.4).



Tabel 3.17 Pengurutan Data dengan Nilai Jarak Terkecil ke Terbesar

| Rank | Suhu (C) | pH | EC (mikroS/cm) | Klasifikasi Bau | Klasifikasi Rasa | Klasifikasi Warna | Jarak |
|------|----------|------|----------------|-----------------|------------------|-------------------|-------|
| 1 | 27,5 | 6,79 | 840 | Tidak Bau | Tawar | Tidak Berwarna | 7,12 |
| 2 | 26,6 | 6,72 | 825 | Ada | Tawar | Berwarna | 8,01 |
| 3 | 28,2 | 7,78 | 851 | Ada | Payau | Berwarna | 18,14 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 89 | 27 | 6,88 | 10987 | Ada | Asin | Berwarna | 10154 |
| 90 | 29 | 6,35 | 11485 | Ada | Asin | Tidak Berwarna | 10652 |

4. Penentuan Klasifikasi Bau, Rasa dan Warna

Setelah data diurut, data difilter sebanyak k optimal yaitu 11. Kemudian dari data yang sudah difilter, diklasifikasi bau, rasa dan warna dengan menggunakan frekuensi terbanyak dari masing-masing kelas. Hasil klasifikasi dapat dilihat pada tabel 3.18.



Tabel 3.18 Hasil Klasifikasi Bau, Rasa dan Warna

| Rank | Suhu (C) | pH | EC (mikroS/cm) | Klasifikasi Bau | Klasifikasi Rasa | Klasifikasi Warna | Jarak |
|------|----------|------|-------------------|-----------------|------------------|-------------------|-------|
| 1 | 27,5 | 6,79 | 840 | Tidak Bau | Tawar | Tidak Berwarna | 7,12 |
| 2 | 26,6 | 6,72 | 825 | Ada | Tawar | Berwarna | 8,01 |
| 3 | 28,2 | 7,78 | 851 | Ada | Payau | Berwarna | 18,14 |
| 4 | 27,7 | 6,63 | 869 | Ada | Tawar | Tidak Berwarna | 36,03 |
| 5 | 27 | 7 | 790 | Ada | Tawar | Berwarna | 43,01 |
| 6 | 26,6 | 6,57 | 877 | Tidak Bau | Tawar | Berwarna | 44 |
| 7 | 27 | 6 | 775 | Ada | Tawar | Berwarna | 58,01 |
| 8 | 27,2 | 7,61 | 893 | Ada | Payau | Berwarna | 60,01 |
| 9 | 26,9 | 7,02 | 772 | Ada | Tawar | Berwarna | 61 |
| 10 | 26,9 | 6,91 | 744 | Tidak Bau | Tawar | Berwarna | 89 |
| 11 | 27,2 | 7,1 | 724 | Tidak Bau | Tawar | Tidak Berwarna | 109 |

Kesimpulan:

Klasifikasi Bau “Ada” = 7 , “Tidak Bau” = 4.

Klasifikasi Rasa “Tawar” = 9 , “Payau” = 2 , “Asin” = 0.

Klasifikasi Warna “Berwarna” = 8 , “Tidak Berwarna” = 3.

Dari keseluruhan hasil perhitungan jarak dengan data training yang berjumlah 90 data, hasil perhitungan diurutkan mulai terkecil hingga terbesar. Setelah diurutkan, dilihat mayoritas klasifikasi yang muncul dari perhitungan jarak yang terkecil. Dapat dilihat pada tabel diatas, mayoritas klasifikasi yang muncul untuk parameter bau adalah “Ada”, parameter rasa adalah “Tawar” dan parameter Warna adalah “Berwarna”, sehingga data testing yang diklasifikasikan termasuk klasifikasi “Ada”, “Tawar” dan “Berwarna”.

3.5.5 Desain Antarmuka

Desain antarmuka merupakan bagian dari sistem yang akan digunakan sebagai media interaksi antara sistem dengan pengguna (user). Adapun desain antarmuka dari sistem terdapat pada gambar 3.16.

| Nama Aplikasi / Sistem | | Logout |
|---|----------|--|
| Foto | Username | Selamat datang Username, Hari ini tanggal : d M Y |
| <input checked="" type="checkbox"/> Data Training Tambah Data Lihat Data | | MENU UTAMA Ini adalah sistem prediksi kualitas air sumur yang dibangun untuk membantu peneliti dalam memprediksi bau, rasa dan warna air sumur Data yang diinput : Suhu, pH dan EC Silahkan klik tombol dibawah ini untuk melakukan proses prediksi <input type="button" value="Data Testing"/> |
| <input checked="" type="checkbox"/> Data Testing Tambah Data Lihat Data | | |
| <input checked="" type="checkbox"/> K-FOLD CV | | |
| <input checked="" type="checkbox"/> Panduan | | |
| <input checked="" type="checkbox"/> Akun | | |

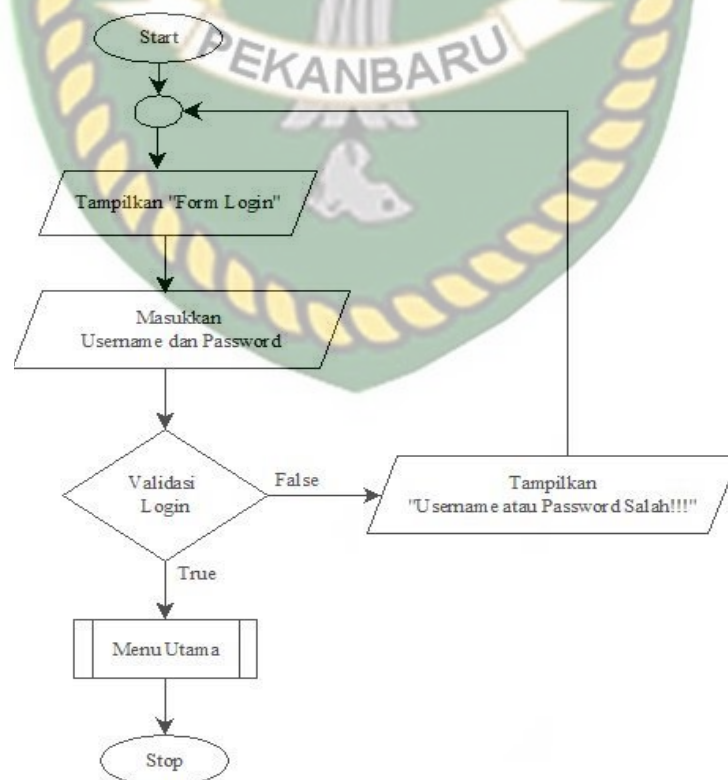
Gambar 3.16 Desain Menu Utama

3.5.6 Rancangan Logika Program

Desain logika program menggambarkan logika program yang akan dibuat menggunakan alat bantu *flowchart*. Desain logika program ini bertujuan untuk menunjukkan alur sistem maupun program, mulai dari data input sampai hasil output yang diproses oleh program. Logika program merupakan unsur penting untuk membangun sebuah sistem berbasis komputer, karena hal ini akan sangat membantu pengguna yang akan menggunakan sistem. Berikut desain logika program dari sistem ini.

1. Program Flowchart Login

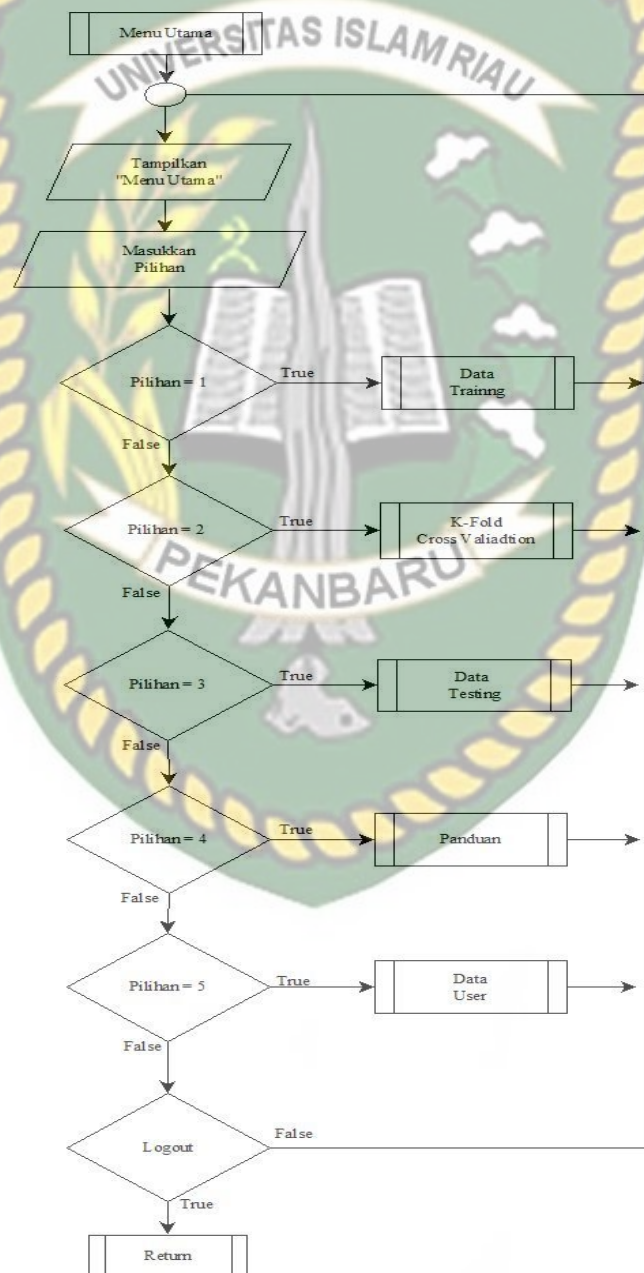
Program flowchart login adalah flowchart yang menjelaskan bagaimana proses admin untuk login ke menu utama sistem. Adapun desain program flowchart login dapat dilihat pada gambar 3.17.



Gambar 3.17 Flowchart Alur Program Login

2. Program Flowchart Menu Sistem

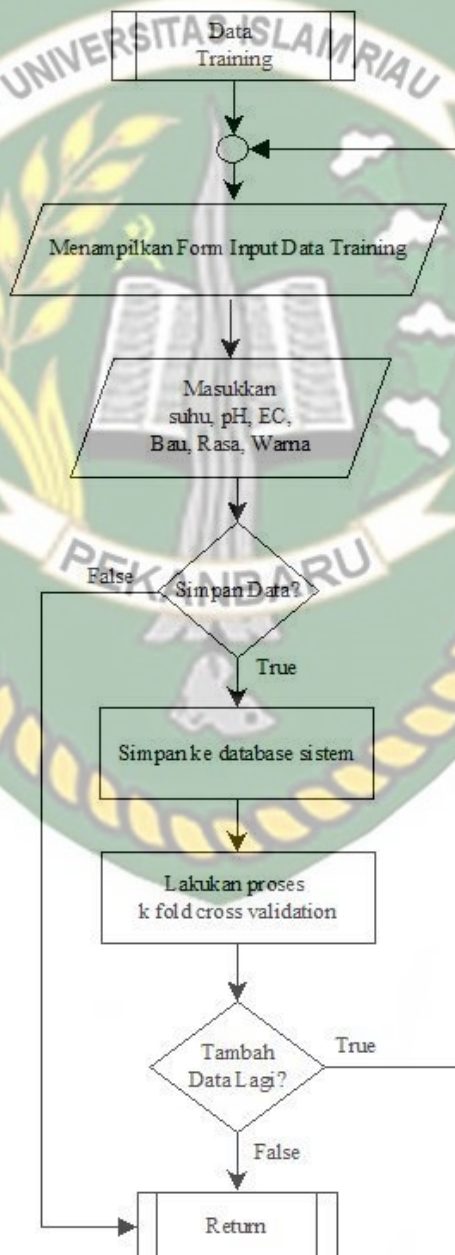
Flowchart ini menjelaskan menu-menu yang terdapat dalam sistem. Menu tersebut terdiri dari menu data training, menu data testing, menu k-fold cross validation, menu laporan dan menu user. Adapun desain program flowchart menu sistem dapat dilihat pada gambar 3.18.



Gambar 3.18 Flowchart Alur Program Menu Utama

3. Program *Flowchart* Input Data Training

Program *flowchart* input data training merupakan alur data program yang merancang input data training yang akan tersimpan kedalam sistem. Atribut yang dibutuhkan dalam proses ini adalah suhu, pH, EC, bau, rasa dan warna. Adapun desain program flowchart perhitungan data testing dapat dilihat pada gambar 3.19.



Gambar 3.19 *Flowchart* Alur Program Data Training

3. Program Flowchart Perhitungan Data Testing

Flowchart perhitungan data testing merupakan alur data program yang merancang input data testing yang akan di lakukan proses perhitungan k-NN dan menghasilkan output berupa hasil klasifikasi. Atribut yang dibutuhkan dalam proses ini adalah suhu, pH, EC dan nilai k optimal. Kemudian atribut tersebut diolah menggunakan metode k-NN. Adapun desain program flowchart perhitungan data testing dapat dilihat pada gambar 3.20.

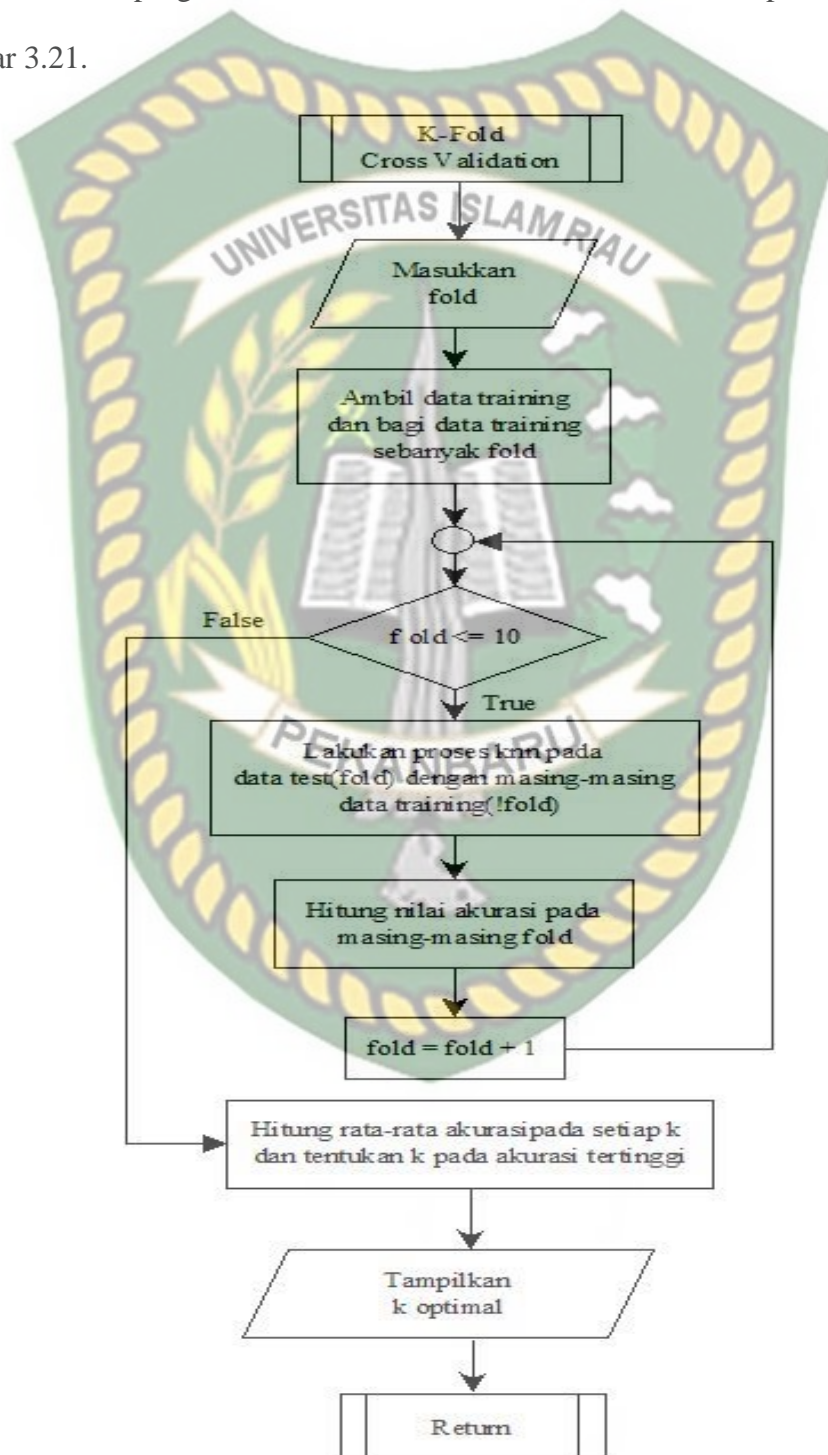


Gambar 3.20 *Flowchart* Alur Program Data Testing

4. Program Flowchart K Fold Cross Validation

Flowchart diatas menjelaskan alur perhitungan dalam mencari nilai k optimal.

Adapun desain program flowchart k-Fold Cross Validation dapat dilihat pada gambar 3.21.



Gambar 3.21 Flowchart Alur Program Perhitungan K-Fold Cross Validation

BAB IV HASIL DAN PEMBAHASAN

Pengujian *Black Box*

Setelah melakukan penelitian dan perancangan terhadap sistem yang berjalan maka dilakukan langkah selanjutnya yaitu pengujian *black box* yaitu pengujian yang berfokus kepada pengujian dengan melihat fungsi-fungsi yang ada dalam program tanpa harus mengetahui bagaimana fungsi tersebut dibuat programnya. Pengujian merujuk pada fungsi fungsi yang dimiliki sistem, kemudian membandingkan hasil program dengan hasil yang diharapkan. Bila hasil yang diharapkan sesuai dengan hasil pengujian, berarti perangkat lunak sesuai dengan desain yang telah ditentukan sebelumnya. Bila belum sesuai maka perlu dilakukan pengecekan lebih lanjut dan perbaikan. Rencana pengujian yang dilakukan pada sistem ini dapat dilihat pada tabel 4.1.

Tabel 4.1 Rencana Pengujian

| No | Kelas Uji | Butir Uji | Jenis Pengujian |
|----|-----------------------------|--|------------------|
| 1. | Login | Verifikasi nama dan kata sandi. | <i>Black Box</i> |
| 2. | Pengolahan Data Pengguna | Menambah dan mengedit data pengguna | <i>Black Box</i> |
| 3 | Pengolahan Data Training | Menambah, mengedit dan menghapus data <i>training</i> | <i>Black Box</i> |
| 4. | Pengolahan Data Testing | Menambah, menghitung dan mengklasifikasi data <i>testing</i> | <i>Black Box</i> |
| 5. | Nilai Akurasi dan K Optimal | Melihat nilai akurasi pada setiap k | <i>Black Box</i> |

Pengujian *Login*

Untuk dapat melakukan pengolahan data pada sebuah sistem, user harus *login* terlebih dulu kedalam sistem. User memasukkan *username* dan *password* yang telah didaftarkan. Apabila belum memiliki akun, user bisa membuat akun baru dengan mengklik tulisan “Klik Disini” yang tersedia pada halaman login. Berikut tampilan halaman *login* pada sistem, dapat dilihat pada gambar 4.1.



Gambar 4.1 Pengujian *Login*

Apabila data *login* tidak sesuai dengan data pengguna sistem, maka sistem informasi akan mengeluarkan pemberitahuan gagal atas *login* yang dilakukan, kemudian *user* akan diarahkan kembali ke form *login* seperti pada gambar 4.2.

SISTEM PREDIKSI BAU, RASA DAN WARNA AIR SUMUR



Gambar 4.2 Pengujian *Login* “Gagal”

Apabila *login* user berhasil dan data *username* dan *password* ditemukan, maka sistem akan menampilkan menu utama seperti gambar 4.3.

MENU UTAMA

Ini adalah sistem prediksi kualitas air sumur yang dibangun untuk membantu peneliti dalam memprediksi bau, rasa dan warna air sumur

Data yang diinput : Suhu, pH dan EC
Silahkan klik tombol dibawah ini untuk melakukan proses prediksi



Gambar 4.3 Tampilan Menu Utama Setelah *Login*

Penjelasan seluruh gambar selengkapnya bisa dilihat pada tabel 4.2.

Tabel 4.2 Pengujian *Login*

| Kasus dan Hasil Uji | | | | | |
|---------------------|---|---|--|---|------------|
| No | Skenario pengujian | Test kasus | Hasil yang diharapkan | Hasil Pengujian | Kesimpulan |
| 1. | Memasukkan data nama dan kata sandi yang benar, lalu mengklik tombol "Login". | Nama Pengguna : admin Kata Sandi : admin | Sistem menerima akses masuk dan kemudian menampilkan menu utama. | Tombol login dapat berfungsi sesuai yang diharapkan | Valid |
| 2 | Memasukkan nama dan kata sandi dengan kondisi salah satu data benar, lalu langsung mengklik tombol "Login". | Nama Pengguna : admin Kata Sandi : 1234 | Sistem akan menolak akses masuk dan menampilkan pesan "User name atau Password Salah!!!" | Sesuai diharapkan | Valid |

Pengujian Pengolahan Data Pengguna

Pada pengolahan data pengguna ini ada 2 (dua) pengujian yang dilakukan yaitu, penambahan dan edit data pengguna (user). Untuk penambahan data user dapat dilihat pada gambar 4.4.

[KEMBALI KE MENU LOGIN](#)

SISTEM PREDIKSI BAU, RASA DAN WARNA AIR SUMUR

Gambar 4.4 Pengujian Tambah Data User

Apabila data yang dimasukkan benar, maka data akan disimpan ke dalam database dan akan muncul pesan seperti gambar 4.5.



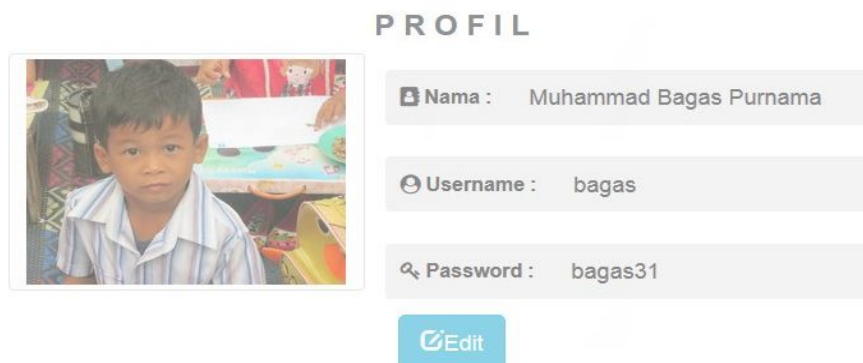
Gambar 4.5 Pesan Tambah Data User “Berhasil”

Sebaliknya, apabila salah satu kolom dikosongkan akan muncul pesan kesalahan seperti gambar 4.6.



Gambar 4.6 Pengujian Tambah Data User “Gagal”

Berikut adalah tampilan halaman profil yang terdapat dalam sistem. Pada halaman ini informasi yang ditampilkan yaitu nama, username, password dan foto. Tampilan halaman profil dapat dilihat pada gambar 4.7.



Gambar 4.7 Menu Profil

Apabila ada data user yang ingin diubah, user dapat melakukannya pada menu profil seperti gambar 4.8.

The image shows a web form titled "Edit Profil" with a background watermark of the Universitas Islam Riau logo. The form has the following fields and values:

- Username: bagas
- Nama: Muhammad Bagas Purnama
- Password: bagas31
- Foto: Browse... No file selected

At the bottom of the form, there are two buttons: "Reset" and "Simpan".

Gambar 4.8 Pengujian Edit Data User

Penjelasan seluruh gambar selengkapnya bisa dilihat pada tabel 4.3.

Tabel 4.3 Pengujian Pengolahan Data Pengguna

| Kasus dan Hasil Uji | | | | | |
|---------------------|----------------------|---|--|-------------------|------------|
| No | Fungsi yang diuji | Cara Menguji | Hasil yang diharapkan | Hasil Pengujian | Kesimpulan |
| 1. | Simpan data pengguna | Memasukkan seluruh data dengan benar, lalu mengklik tombol "Daftar". | Sistem akan langsung menyimpan data baru dan memasukkannya ke dalam <i>database</i> dan menampilkan pesan "Pendaftaran Anda Berhasil". | Sesuai diharapkan | Valid |
| 2. | Simpan data pengguna | Memasukkan data dengan salah satu dikosongkan, lalu mengklik tombol "Daftar". | Sistem tidak akan menyimpan data ke dalam <i>database</i> dan menampilkan pesan kesalahan "Please fill out this field" pada kolom yang kosong. | Sesuai diharapkan | Valid |
| 3. | Edit data pengguna | Mengklik tombol "Edit" pada menu | Sistem akan langsung menampilkan form | Sesuai diharapkan | Valid |

| | | | | | |
|--|--|---------|---|--|--|
| | | profil. | baru untuk melakukan edit pada data dan kemudian disimpan kembali ke dalam <i>database</i> dan menampilkan pesan “Data Berhasil Dirubah”. | | |
|--|--|---------|---|--|--|

Pengujian Pengolahan Data *Training*

Pada menu data *training* ada tiga pengujian yang dilakukan, diantaranya adalah penambahan, edit dan hapus data training. Tampilan menu data *training* dapat dilihat pada gambar 4.9 dan pengujian penambahan data *training* dapat dilihat pada gambar 4.10.

◀ Tambah Data Air Sumur
 📄 Tabel Data Air Sumur
 📄 Upload Data Air Sumur
 🆘 Bantuan

📄 Data Air Sumur

| | |
|------------|-------------------------|
| 1. No Data | 91 |
| 2. Suhu | Suhu Air Sumur |
| 3. pH | pH Air Sumur |
| 4. EC | EC Air Sumur |
| 5. Bau | --- Bau Air Sumur --- |
| 6. Rasa | --- Rasa Air Sumur --- |
| 7. Warna | --- Warna Air Sumur --- |

↻ Reset
 💾 Simpan

Gambar 4.9 Menu Data *Training*

Data Training Air Sumur

| | |
|------------|---------------|
| 1. No Data | 90 |
| 2. Suhu | 27.8 |
| 3. pH | 6.82 |
| 4. EC | 272.8 |
| 5. Bau | Tidak Bau |
| 6. Rasa | Tawar |
| 7. Warna | Tidak Bewarna |

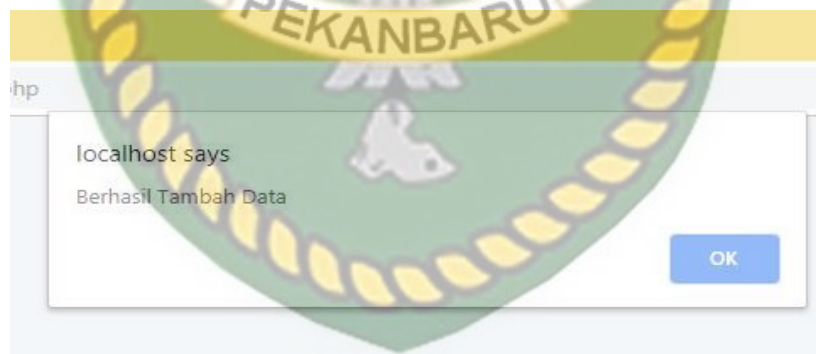
UPLOAD DATA TRAINING
 Pilih File:
 No file chosen

(A)

(B)

Gambar 4.10 Pengujian Tambah Data *Training* dengan Input Manual (A) dan Upload (B)

Apabila data yang dimasukkan benar, maka data akan disimpan ke dalam database dan akan muncul pesan seperti gambar 4.11.



Gambar 4.11 Pesan Tambah Data *Training* “Berhasil”

Sebaliknya, apabila salah satu kolom dikosongkan akan muncul pesan kesalahan seperti gambar 4.12.

Data Training Air Sumur

| | |
|------------|--|
| 1. No Data | 90 |
| 2. Suhu | 27.8 |
| 3. pH | 6.82 |
| 4. EC | EC Air Sumur |
| 5. Bau | Tidak Bau ! Please fill out this field. |
| 6. Rasa | Tawar |
| 7. Warna | Tidak Berwarna |

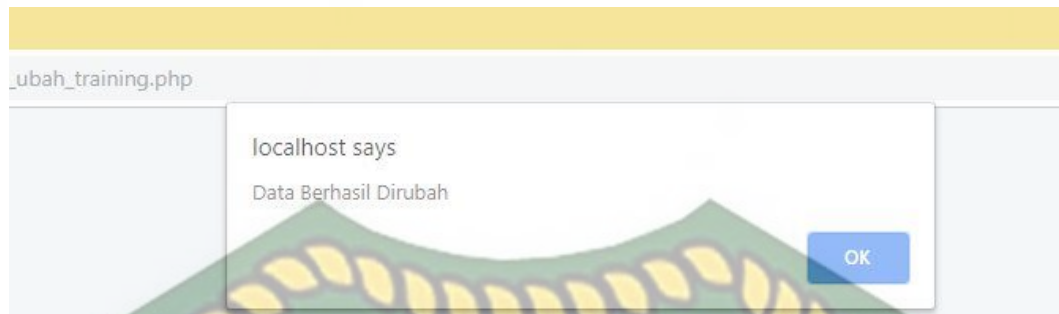
Gambar 4.12 Pesan Kesalahan Data Kosong

Apabila ada data *training* yang ingin diubah, user dapat melakukannya pada halaman edit data *training* seperti gambar 4.13 dan sistem akan memunculkan sebuah pesan seperti gambar 4.14.

Edit Data Training

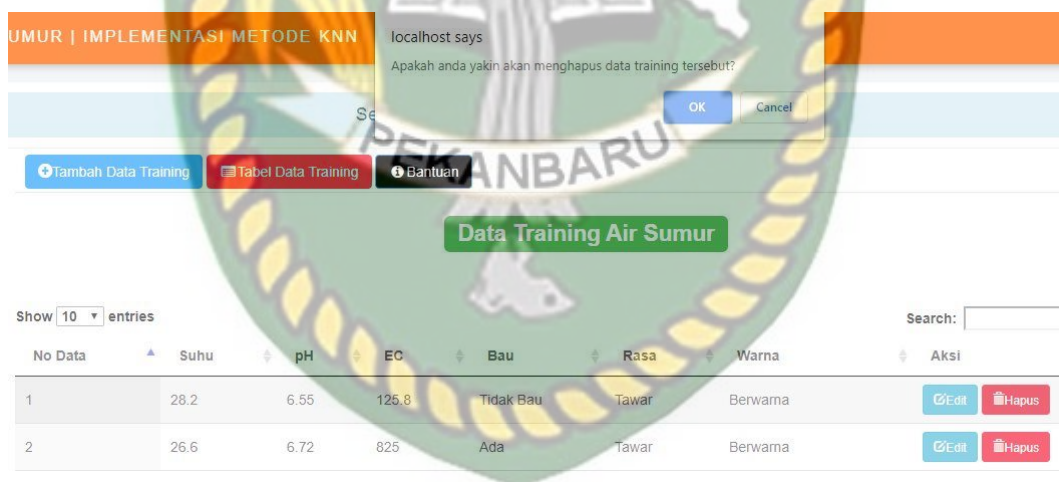
| | |
|------------|----------------|
| 1. No Data | 90 |
| 2. Suhu | 27.8 |
| 3. pH | 6.82 |
| 4. EC | 272.8 |
| 5. Bau | Tidak Bau |
| 6. Rasa | Tawar |
| 7. Warna | Tidak Berwarna |

Gambar 4.13 Pengujian Edit Data *Training*

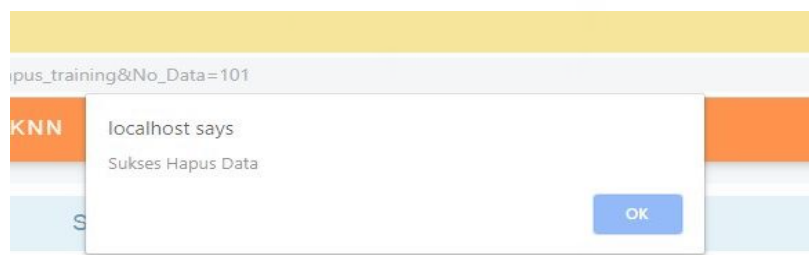


Gambar 4.14 Pesan Edit Data *Training* “Berhasil”

Sebelum menghapus data *training*, sistem akan menampilkan pesan pemberitahuan “Apakah anda yakin akan menghapus data training tersebut?” seperti gambar 4.15. Apabila diklik “OK”, sistem akan langsung menghapus data terpilih pada tabel maupun pada *database* dan menampilkan pesan “Sukses Hapus Data” seperti gambar 4.16.



Gambar 4.15 Pesan Pemberitahuan Untuk Hapus Data Training



Gambar 4.16 Pesan Hapus Data *Training* “Berhasil”

Penjelasan seluruh gambar selengkapnya bisa dilihat pada tabel 4.4.

Tabel 4.4 Pengujian Pengolahan Data *Training*

| Kasus dan Hasil Uji | | | | | |
|---------------------|-------------------------------|--|---|-------------------|------------|
| No | Fungsi yang diuji | Cara Menguji | Hasil yang diharapkan | Hasil Pengujian | Kesimpulan |
| 1. | Simpan data <i>training</i> . | Memasukkan seluruh data dengan benar, lalu mengklik tombol "Simpan". | Sistem akan langsung menyimpan data baru dan memasukkannya ke dalam <i>database</i> dan menampilkan pesan "Berhasil Tambah Data". | Sesuai diharapkan | Valid |
| 2. | Simpan data <i>training</i> . | Memasukkan data dengan salah satu <i>field</i> dikosongkan, lalu mengklik tombol "Simpan". | Sistem tidak akan menyimpan data ke dalam <i>database</i> dan menampilkan pesan kesalahan " <i>Please fill out this field</i> " pada kolom yang kosong. | Sesuai diharapkan | Valid |
| 3. | Edit data <i>training</i> . | Memilih salah satu data yang tampil pada tabel <i>training</i> , kemudian mengklik aksi "Edit". | Sistem akan langsung menampilkan form baru untuk melakukan edit pada data dan kemudian disimpan kembali ke dalam <i>database</i> dan menampilkan pesan "Data Berhasil Dirubah". | Sesuai diharapkan | Valid |
| 4. | Hapus data <i>training</i> . | Memilih salah satu data yang tampil pada tabel <i>training</i> , kemudian mengklik aksi "Hapus". | Sistem akan menampilkan pemberitahuan "Apakah anda yakin akan menghapus data <i>training</i> tersebut? ", apabila "OK" sistem akan langsung menghapus data terpilih pada tabel maupun pada <i>database</i> . dan menampilkan pesan "Sukses Hapus Data". | Sesuai diharapkan | Valid |

Pengujian Pengolahan Data *Testing*

Pada menu data *testing* ini ada 2 pengujian yang dilakukan, yaitu penambahan, dan proses perhitungan data *testing*. Tampilan menu data *testing* dapat dilihat pada gambar 4.17.

| Data Testing Air Sumur | |
|------------------------|----------------|
| 1. No Data | |
| 2. Suhu | Suhu Air Sumur |
| 3. pH | pH Air Sumur |
| 4. EC | EC Air Sumur |

Reset Simpan

Gambar 4.17 Menu Data *Testing*

Dalam pengujian data testing, ada 5 (lima) atribut yang wajib diisi ke dalam sistem, yaitu no data, suhu, pH, EC seperti gambar 4.18.

Testing Upload Data Testing Bantuan

| Data Testing Air Sumur | |
|------------------------|-----|
| 1. No Data | 20 |
| 2. Suhu | 26 |
| 3. pH | 6 |
| 4. EC | 861 |

Reset Simpan

UPLOAD DATA TESTING

Pilih File:

Choose File No file chosen

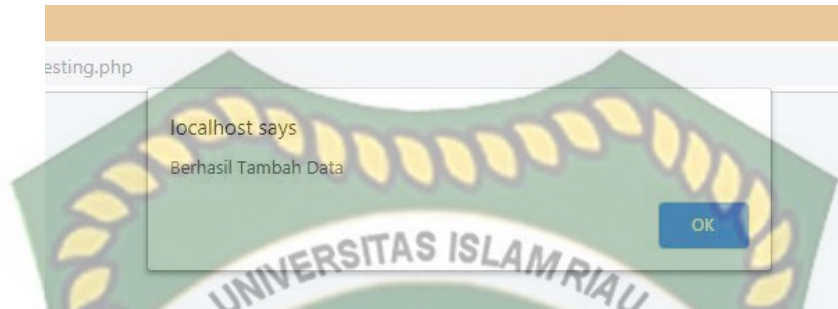
Upload

(B)

(A)

Gambar 4.18 Pengujian Simpan Data *Testing* dengan Input Manual (A) dan Upload (B)

Apabila data yang dimasukkan benar, maka data akan disimpan ke dalam database dan akan muncul pesan seperti gambar 4.19.



Gambar 4.19 Pesan Simpan Data *Testing* “Berhasil”

Apabila data yang diinput telah berhasil disimpan ke database, maka akan muncul hasil perhitungan dan klasifikasi dari data *testing* tersebut. Hasil klasifikasi dapat dilihat pada gambar 4.20.

Hasil Testing

| No | No Data | Suhu | pH | EC | Jarak | Termasuk Bau? | Termasuk Rasa? | Termasuk Warna? |
|-----------------------|-----------|-------|-------|-------|---------|---------------|----------------|-----------------|
| 1 | 58 | 27.5 | 6.79 | 840 | 7.12 | Tidak Bau | Tawar | Tidak Berwarna |
| 2 | 2 | 26.6 | 6.72 | 825 | 8.01 | Ada | Tawar | Berwarna |
| 3 | 80 | 28.2 | 7.78 | 851 | 18.14 | Ada | Payau | Berwarna |
| 4 | 63 | 27.7 | 6.63 | 869 | 36.03 | Ada | Tawar | Tidak Berwarna |
| 5 | 14 | 27 | 7 | 790 | 43.01 | Ada | Tawar | Berwarna |
| 6 | 8 | 26.6 | 6.57 | 877 | 44 | Tidak Bau | Tawar | Berwarna |
| 7 | 64 | 27 | 6 | 775 | 58.01 | Ada | Tawar | Berwarna |
| 8 | 6 | 27.2 | 7.61 | 893 | 60.01 | Ada | Payau | Berwarna |
| 9 | 33 | 26.9 | 7.02 | 772 | 61 | Ada | Tawar | Berwarna |
| 10 | 79 | 26.9 | 6.91 | 744 | 89 | Tidak Bau | Tawar | Berwarna |
| 11 | 45 | 27.2 | 7.1 | 724 | 109 | Tidak Bau | Tawar | Tidak Berwarna |
| BAU | | RASA | | | WARNA | | | |
| ADA | TIDAK BAU | TAWAR | ASIN | PAYAU | BEWARNA | TIDAK BEWARNA | | |
| 7 | 4 | 9 | 0 | 2 | 8 | 3 | | |
| HASIL NILAI TERTINGGI | | | | | | | | |
| Ada | | | Tawar | | | Berwarna | | |

KESIMPULAN :

Suhu = 26.2, pH = 6.81, EC = 833

Masuk Dalam Kategori Bau = Ada, Rasa = Tawar, Warna = Berwarna

VALIDASI DATA

Gambar 4.20 Hasil Perhitungan Data *Testing*

Penjelasan seluruh gambar selengkapnya bisa dilihat pada tabel 4.5.

Tabel 4.5 Pengujian Pengolahan Data Testing

| Kasus dan Hasil Uji | | | | | |
|---------------------|------------------------------|--|--|-------------------|------------|
| No | Fungsi yang diuji | Cara Menguji | Hasil yang diharapkan | Hasil Pengujian | Kesimpulan |
| 1. | Simpan data <i>testing</i> . | Memasukkan seluruh data dengan benar, lalu mengklik tombol “Simpan”. | Sistem akan langsung menyimpan data baru dan memasukkannya ke dalam <i>database</i> dan menampilkan pesan “Berhasil Tambah Data” dan menampilkan hasil klasifikasi data testing. | Sesuai diharapkan | Valid |
| 2. | Simpan data <i>testing</i> . | Memasukkan data dengan salah satu dikosongkan (kecuali keterangan), lalu mengklik tombol “Simpan”. | Sistem tidak akan menyimpan data ke dalam <i>database</i> dan menampilkan pesan kesalahan “ <i>Please fill out this field</i> ” pada kolom yang kosong. | Sesuai diharapkan | Valid |

Pengujian Nilai Akurasi

Pada menu nilai akurasi ini, dilakukan pengujian berupa menampilkan nilai akurasi sebanyak nilai k seperti gambar 4.21 sampai dengan gambar 4.23.

Nilai Akurasi & K Optimal

Tampilkan Perhitungan Hingga K :

20

Reset

Proses

Gambar 4.21 Menu Nilai Akurasi

K-FOLD CROSS VALIDATION

| Nilai k = 3 | | | |
|--------------|---------|----------|-----------|
| Percobaan | Bau (%) | Rasa (%) | Warna (%) |
| Percobaan-1 | 66.67 | 66.67 | 33.33 |
| Percobaan-2 | 66.67 | 88.89 | 44.44 |
| Percobaan-3 | 55.56 | 44.44 | 77.78 |
| Percobaan-4 | 55.56 | 88.89 | 55.56 |
| Percobaan-5 | 55.56 | 100 | 55.56 |
| Percobaan-6 | 88.89 | 66.67 | 66.67 |
| Percobaan-7 | 66.67 | 55.56 | 22.22 |
| Percobaan-8 | 66.67 | 88.89 | 66.67 |
| Percobaan-9 | 66.67 | 88.89 | 44.44 |
| Percobaan-10 | 44.44 | 66.67 | 66.67 |
| Rata - Rata | 63.33 | 75.56 | 53.33 |
| 64.07 | | | |
| Nilai k = 5 | | | |
| Percobaan | Bau (%) | Rasa (%) | Warna (%) |
| Percobaan-1 | 66.67 | 88.89 | 66.67 |
| Percobaan-2 | 55.56 | 77.78 | 44.44 |
| Percobaan-3 | 88.89 | 44.44 | 77.78 |
| Percobaan-4 | 55.56 | 88.89 | 55.56 |
| Percobaan-5 | 33.33 | 100 | 33.33 |
| Percobaan-6 | 88.89 | 55.56 | 66.67 |
| Percobaan-7 | 77.78 | 66.67 | 22.22 |
| Percobaan-8 | 55.56 | 100 | 55.56 |

Gambar 4.22 Perhitungan k Optimal dari k=3 Sampai Dengan k=19

AKURASI K-FOLD CROSS VALIDATION

| Nilai k | Bau Sama (%) | Rasa Sama (%) | Warna Sama (%) | Rata-Rata (%) |
|---------|--------------|---------------|----------------|---------------|
| k-3 | 65 | 85 | 65 | 71.67 |
| k-5 | 65 | 90 | 80 | 78.33 |
| k-7 | 65 | 90 | 80 | 78.33 |
| k-9 | 70 | 90 | 85 | 81.67 |
| k-11 | 80 | 95 | 80 | 85 |
| k-13 | 80 | 95 | 80 | 85 |
| k-15 | 75 | 95 | 70 | 80 |
| k-17 | 80 | 95 | 65 | 80 |
| k-19 | 65 | 95 | 65 | 75 |

Gambar 4.23 Perhitungan Akurasi dari k=3 Sampai Dengan k=19

Penjelasan seluruh gambar selengkapnya bisa dilihat pada tabel 4.6.

Tabel 4.6 Pengujian Nilai K Optimal

| Kasus dan Hasil Uji | | | | | |
|---------------------|-------------------|--|---|-------------------|------------|
| No | Fungsi yang diuji | Cara Menguji | Hasil yang diharapkan | Hasil Pengujian | Kesimpulan |
| 1. | Proses data. | Memasukkan nilai k, lalu mengklik tombol | Sistem akan langsung melakukan proses perhitungan k optimal dan | Sesuai diharapkan | Valid |

| | | | | | |
|----|--------------|--|--|-------------------|-------|
| | | “Proses”. | menampilkan akurasi. | | |
| 2. | Proses data. | Mengosongkan data lalu mengklik tombol “Simpan”. | Sistem tidak akan melakukan proses perhitungan dan menampilkan pesan kesalahan “ <i>Please fill out this field</i> ” pada kolom yang kosong. | Sesuai diharapkan | Valid |

Pengujian Akurasi Sistem

Pengujian akurasi klasifikasi yang dilakukan untuk mengetahui tingkat akurasi klasifikasi bau, rasa dan warna yang dilakukan secara manual dengan klasifikasi bau, rasa dan warna yang dilakukan oleh sistem dengan menggunakan metode k-NN. Pengujian dilakukan dengan menggunakan confusion matrix yaitu sebuah matrix dari prediksi yang akan dibandingkan dengan kelas yang asli dari data inputan. Pengujian dilakukan menggunakan 20 data testing dan menggunakan nilai k optimal yaitu 11. Data tersebut dibandingkan dengan hasil klasifikasi yang dilakukan oleh sistem. Hasil pengujian akurasi klasifikasi bau, rasa dan warna dapat dilihat pada tabel 4.7, tabel 4.9 dan tabel 4.11.

Tabel 4.7 Pengujian Akurasi Klasifikasi Bau Berdasarkan Nilai k Optimal

| Data | Suhu | pH | EC | Manual | Sistem | Akurasi Manual dan Sistem |
|------|------|------|------|-----------|-----------|---------------------------|
| 1 | 26,2 | 6,81 | 833 | Tidak Bau | Ada | 0 |
| 2 | 27,3 | 7,32 | 1236 | Tidak Bau | Tidak Bau | 1 |
| 3 | 27 | 7 | 226 | Tidak Bau | Tidak Bau | 1 |
| 4 | 27 | 6,77 | 718 | Ada | Ada | 1 |
| 5 | 27 | 7,3 | 5209 | Ada | Ada | 1 |
| 6 | 29,2 | 6,6 | 7220 | Ada | Ada | 1 |
| 7 | 27,2 | 7,34 | 1881 | Ada | Ada | 1 |

| | | | | | | |
|----|------|------|------|-----------|-----------|---|
| 8 | 28 | 6,51 | 794 | Ada | Ada | 1 |
| 9 | 27 | 7,36 | 757 | Ada | Ada | 1 |
| 10 | 27,8 | 6,81 | 1067 | Tidak Bau | Tidak Bau | 1 |
| 11 | 27,8 | 6,81 | 822 | Ada | Ada | 1 |
| 12 | 27,6 | 7,25 | 1357 | Ada | Ada | 1 |
| 13 | 27,1 | 7,1 | 914 | Ada | Ada | 1 |
| 14 | 27 | 7,2 | 1032 | Tidak Bau | Tidak Bau | 1 |
| 15 | 28 | 7 | 478 | Tidak Bau | Tidak Bau | 1 |
| 16 | 27 | 6 | 1475 | Tidak Bau | Ada | 0 |
| 17 | 28 | 6 | 400 | Tidak Bau | Tidak Bau | 1 |
| 18 | 27 | 6 | 447 | Tidak Bau | Tidak Bau | 1 |
| 19 | 26 | 7 | 1645 | Tidak Bau | Ada | 0 |
| 20 | 26 | 6 | 861 | Tidak Bau | Ada | 0 |

Keterangan :

Kelas Ada = 9

Kelas Tidak Bau = 11

Tabel 4.8 menggambarkan hasil klasifikasi bau berdasarkan *confusion matrix* :

Tabel 4.8 *Confusion Matrix* Klasifikasi Bau

| | | Hasil Prediksi Sistem | |
|------------|-----------|-----------------------|-----------|
| | | Ada | Tidak Bau |
| Hasil Real | Ada | 9 | 0 |
| | Tidak Bau | 4 | 7 |

Setelah sistem melakukan klasifikasi, lalu hitung nilai akurasinya. Rumus akurasi yaitu:

$$\text{Akurasi} = \frac{\text{jumlah data yang diprediksi secara benar}}{\text{jumlah prediksi yang dilakukan}} \times 100\%$$

$$= \frac{9+7}{20} \times 100\%$$

$$= \frac{16}{20} \times 100\%$$

$$= 80 \%$$

$$\text{Error Rate} = \frac{\text{jumlah data yang diprediksi salah}}{\text{jumlah prediksi yang dilakukan}} \times 100\%$$

$$= \frac{4+0}{20} \times 100 \%$$

$$= \frac{4}{20} \times 100\%$$

$$= 20 \%$$

Tabel 4.9 Pengujian Akurasi Klasifikasi Rasa Berdasarkan Nilai k Optimal

| Data | Suhu | pH | EC | Manual | Sistem | Akurasi Manual dan Sistem |
|------|------|------|------|--------|--------|---------------------------|
| 1 | 26,2 | 6,81 | 833 | Tawar | Tawar | 1 |
| 2 | 27,3 | 7,32 | 1236 | Tawar | Tawar | 1 |
| 3 | 27 | 7 | 226 | Tawar | Tawar | 1 |
| 4 | 27 | 6,77 | 718 | Tawar | Tawar | 1 |
| 5 | 27 | 7,3 | 5209 | Payau | Payau | 1 |
| 6 | 29,2 | 6,6 | 7220 | Payau | Payau | 1 |
| 7 | 27,2 | 7,34 | 1881 | Payau | Payau | 1 |
| 8 | 28 | 6,51 | 794 | Tawar | Tawar | 1 |
| 9 | 27 | 7,36 | 757 | Tawar | Tawar | 1 |
| 10 | 27,8 | 6,81 | 1067 | Tawar | Tawar | 1 |
| 11 | 27,8 | 6,81 | 822 | Tawar | Tawar | 1 |
| 12 | 27,6 | 7,25 | 1357 | Tawar | Tawar | 1 |
| 13 | 27,1 | 7,1 | 914 | Tawar | Tawar | 1 |
| 14 | 27 | 7,2 | 1032 | Tawar | Tawar | 1 |
| 15 | 28 | 7 | 478 | Tawar | Tawar | 1 |
| 16 | 27 | 6 | 1475 | Tawar | Tawar | 1 |
| 17 | 28 | 6 | 400 | Tawar | Tawar | 1 |
| 18 | 27 | 6 | 447 | Tawar | Tawar | 1 |

| | | | | | | |
|----|----|---|------|-------|-------|---|
| 19 | 26 | 7 | 1645 | Tawar | Payau | 0 |
| 20 | 26 | 6 | 861 | Tawar | Tawar | 1 |

Keterangan :

Kelas Tawar = 17

Kelas Asin = 0

Kelas Payau = 3

Tabel 4.10 menggambarkan hasil klasifikasi rasa berdasarkan *confusion matrix* :

Tabel 4.10 *Confusion Matrix* Klasifikasi Rasa

| | | Hasil Prediksi Sistem | | |
|------------|-------|-----------------------|-------|-------|
| | | Asin | Tawar | Payau |
| Hasil Real | Asin | 0 | 0 | 0 |
| | Tawar | 0 | 16 | 1 |
| | Payau | 0 | 0 | 3 |

Setelah sistem melakukan klasifikasi, lalu hitung nilai akurasinya. Rumus akurasi

yaitu:

$$\text{Akurasi} = \frac{\text{jumlah data yang diprediksi secara benar}}{\text{jumlah prediksi yang dilakukan}} \times 100\%$$

$$= \frac{16+3+0}{20} \times 100\%$$

$$= \frac{19}{20} \times 100\%$$

$$= 95\%$$

$$\begin{aligned}
 \text{Error Rate} &= \frac{\text{jumlah data yang diprediksi salah}}{\text{jumlah prediksi yang dilakukan}} \times 100\% \\
 &= \frac{0+0+1}{20} \times 100\% \\
 &= \frac{1}{20} \times 100\% \\
 &= 5\%
 \end{aligned}$$

Tabel 4.11 Pengujian Akurasi Klasifikasi Warna Berdasarkan Nilai k Optimal

| Data | Suhu | pH | EC | Manual | Sistem | Akurasi Manual dan Sistem |
|------|------|------|------|----------------|----------------|---------------------------|
| 1 | 26,2 | 6,81 | 833 | Tidak Berwarna | Berwarna | 0 |
| 2 | 27,3 | 7,32 | 1236 | Tidak Berwarna | Tidak Berwarna | 1 |
| 3 | 27 | 7 | 226 | Tidak Berwarna | Tidak Berwarna | 1 |
| 4 | 27 | 6,77 | 718 | Berwarna | Berwarna | 1 |
| 5 | 27 | 7,3 | 5209 | Tidak Berwarna | Tidak Berwarna | 1 |
| 6 | 29,2 | 6,6 | 7220 | Tidak Berwarna | Tidak Berwarna | 1 |
| 7 | 27,2 | 7,34 | 1881 | Berwarna | Bewarna | 1 |
| 8 | 28 | 6,51 | 794 | Tidak Berwarna | Bewarna | 0 |
| 9 | 27 | 7,36 | 757 | Berwarna | Bewarna | 1 |
| 10 | 27,8 | 6,81 | 1067 | Tidak Berwarna | Tidak Berwarna | 1 |
| 11 | 27,8 | 6,81 | 822 | Berwarna | Berwarna | 1 |
| 12 | 27,6 | 7,25 | 1357 | Berwarna | Bewarna | 1 |
| 13 | 27,1 | 7,1 | 914 | Tidak Berwarna | Bewarna | 0 |
| 14 | 27 | 7,2 | 1032 | Berwarna | Tidak Berwarna | 0 |
| 15 | 28 | 7 | 478 | Berwarna | Bewarna | 1 |
| 16 | 27 | 6 | 1475 | Berwarna | Bewarna | 1 |
| 17 | 28 | 6 | 400 | Berwarna | Bewarna | 1 |

| | | | | | | |
|----|----|---|------|----------|---------|---|
| 18 | 27 | 6 | 447 | Berwarna | Bewarna | 1 |
| 19 | 26 | 7 | 1645 | Berwarna | Bewarna | 1 |
| 20 | 26 | 6 | 861 | Berwarna | Bewarna | 1 |

Keterangan :

Kelas Berwarna = 12

Kelas Tidak Berwarna = 8

Tabel 4.12 menggambarkan hasil klasifikasi warna berdasarkan *confusion matrix* :

Tabel 4.12 *Confusion Matrix* Klasifikasi Warna

| | | Hasil Prediksi Sistem | |
|------------|----------------|-----------------------|----------------|
| | | Berwarna | Tidak Berwarna |
| Hasil Real | Berwarna | 11 | 1 |
| | Tidak Berwarna | 3 | 5 |

Setelah sistem melakukan klasifikasi, lalu hitung nilai akurasinya. Rumus akurasi yaitu:

$$\text{Akurasi} = \frac{\text{jumlah data yang diprediksi secara benar}}{\text{jumlah prediksi yang dilakukan}} \times 100\%$$

$$= \frac{11+5}{20} \times 100\%$$

$$= \frac{16}{20} \times 100\%$$

$$= 80\%$$

$$\text{Error Rate} = \frac{\text{jumlah data yang diprediksi salah}}{\text{jumlah prediksi yang dilakukan}} \times 100\%$$

$$= \frac{3+1}{20} \times 100\%$$

$$= \frac{4}{20} \times 100\%$$

$$= 20 \%$$

Karena ada 3 klasifikasi yaitu bau, rasa dan warna, maka nilai akurasi yang didapat adalah:

$$\begin{aligned} \text{Akurasi} &= \frac{\text{akurasi bau} + \text{akurasi rasa} + \text{akurasi warna}}{3} \\ &= \frac{80 + 95 + 80}{3} \\ &= \frac{255}{3} \\ &= 85 \% \end{aligned}$$

Berdasarkan hasil dari pengujian yang telah dilakukan dengan menggunakan *confusion matrix*, maka diperoleh kesimpulan bahwa sistem yang dibangun sudah sesuai dengan yang diharapkan karena tingkat akurasi yang diperoleh sebesar 85%.

Pengujian Sistem Terhadap Pengguna

Salah satu pengujian sistem terhadap pengguna yang dilakukan yaitu dengan membagikan 1 lembar kuisiner kepada mahasiswa atau dosen yang bersangkutan dengan penelitian ini sebanyak 20 orang responden, yang masing-masing lembar kuisiner tersebut berisikan lima pertanyaan. Adapun lima pertanyaan tersebut adalah sebagai berikut:

Bagaimana tanggapan anda tentang sistem ini?

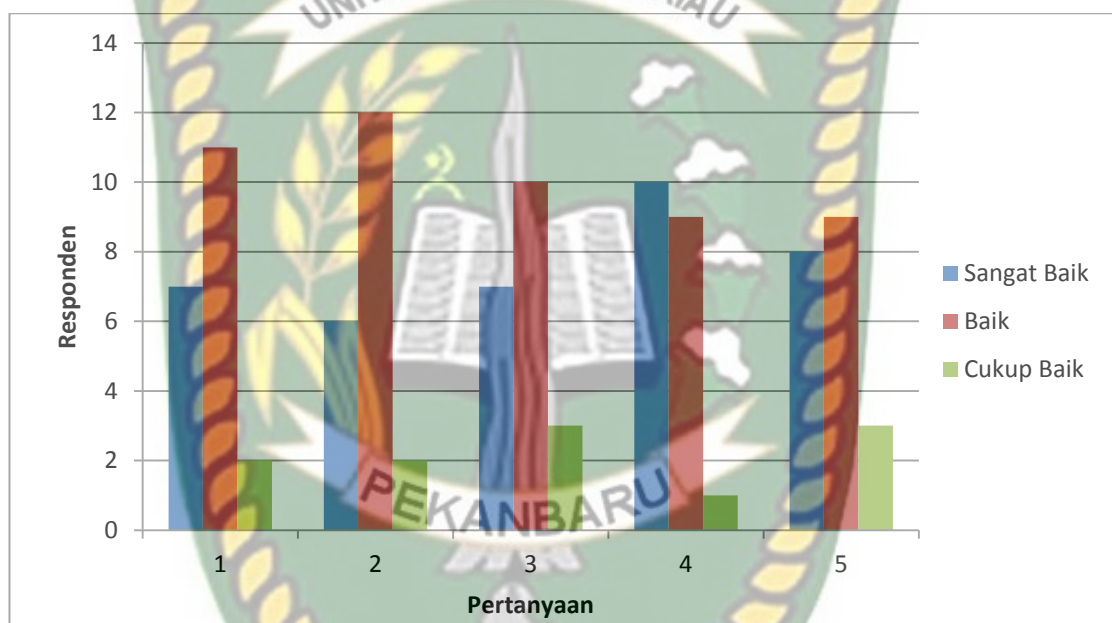
Apakah sistem ini mudah dimengerti untuk digunakan?

Apakah tata letak tampilan atau interface mudah dikenali?

Apakah sistem ini dapat mempermudah peneliti atau pengamat dalam memprediksi bau, rasa dan warna air sumur?

Apakah kedepannya sistem ini layak untuk digunakan?

Tanggapan dari responden terhadap kinerja atau performance dari sistem berdasarkan pertanyaan yang diajukan adalah sebagai berikut :



Gambar 4.24 Grafik Hasil Kuisisioner

Pada tabel 4.13 akan dijelaskan grafik hasil kuisisioner yang menunjukkan nilai untuk setiap pertanyaan-pertanyaan diatas adalah sebagai berikut :

Tabel 4.13 Hasil Nilai Kuisisioner

| Pertanyaan | Sangat Baik | Baik | Cukup Baik |
|------------|-------------|------|------------|
| 1 | 7 | 11 | 2 |
| 2 | 6 | 12 | 2 |
| 3 | 7 | 10 | 3 |
| 4 | 10 | 9 | 1 |

| | | | |
|-------|----|----|----|
| 5 | 8 | 9 | 3 |
| Total | 38 | 51 | 11 |

4.3.1 Hasil Persentase Jawaban Kuisisioner

Berdasarkan hasil kuisisioner yang dilakukan terhadap 20 responden, maka dapat disimpulkan dengan menggunakan skala likert yang telah dimodifikasi, yaitu responden memilih 3 jawaban yang tersedia dengan bobot masing-masing, yakni SB (Sangat Baik) = 3, B (Baik) = 2, CB (Cukup Baik) = 1.

Tabel 4.14 Hasil Nilai Persentase Pertanyaan Kuisisioner

| Pertanyaan | SB | B | CB | Skor SB | Skor B | Skor CB | TOTAL |
|--|----|----|----|---------|--------|---------|---------------|
| 1. Bagaimana tanggapan anda tentang sistem ini? | 7 | 11 | 2 | 21 | 22 | 2 | 45 |
| 2. Apakah sistem ini mudah dimengerti untuk digunakan? | 6 | 12 | 2 | 18 | 24 | 2 | 44 |
| 3. Apakah tata letak tampilan atau interface mudah dikenali? | 7 | 10 | 3 | 21 | 20 | 3 | 44 |
| 4. Apakah sistem ini dapat mempermudah peneliti atau pengamat dalam memprediksi bau, rasa dan warna air sumur? | 10 | 9 | 1 | 30 | 18 | 1 | 49 |
| 5. Apakah kedepannya sistem ini layak untuk digunakan? | 8 | 9 | 3 | 24 | 18 | 3 | 45 |
| Total Skor | | | | | | | 227 |
| Total Skor Tertinggi (Bobot Tertinggi * Jumlah Kuisisioner * Total Pertanyaan) | | | | | | | 300 |
| Rata -Rata (Total Skor / Total Skor Tertinggi) | | | | | | | 75,67% |

Dari hasil persentase tabel 4.14 Implementasi metode *k-Nearest Neighbor* dalam memprediksi bau, rasa dan warna air sumur ini memiliki nilai persentase rata-rata sebesar 75,67% dan dinilai sangat baik sehingga sistem ini layak untuk diimplementasikan.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil perancangan sistem implementasi metode *K-Nearest Neighbor* dalam memprediksi bau, rasa dan warna air sumur yang telah berhasil dilakukan, serta melakukan serangkaian pengujian pada sistem yang dibangun untuk menguji kemampuan sistem, dapat disimpulkan sebagai berikut :

Sistem ini dapat memprediksi potensi bau yaitu bau dan tidak bau, rasa yaitu asin, tawar dan payau dan warna yaitu berwarna dan tidak berwarna pada air sumur.

Dari hasil pengujian sistem, penerapan metode *k-Nearest Neighbor* dalam memprediksi bau, rasa dan warna air sumur memiliki tingkat akurasi sebesar 85% dengan nilai $k=11$.

Persentase nilai jawaban kuisioner dalam sistem ini sebesar 75,67% atau diinterpretasikan sangat baik untuk di implementasikan.

Sistem ini dibangun sebagai alat bantu untuk membantu pengguna baik itu peneliti, mahasiswa atau pelajar, dosen, dll dalam menentukan potensi bau, rasa dan warna pada air sumur.

5.2 Saran

Saran dari penulis untuk sistem implementasi metode *K-Nearest Neighbor* dalam memprediksi bau, rasa dan warna air sumur adalah sebagai berikut :

Penelitian berikutnya disarankan dapat menggunakan metode lain untuk memprediksi bau, rasa dan warna air sumur, seperti metode regresi, *Hold*

Out , *Support Vector Machine* (SVM) ataupun metode jaringan saraf tiruan.

Mengembangkan sistem supaya dapat memprediksi parameter selain bau, rasa dan warna air sumur.

Mengembangkan sistem sekarang menjadi sistem aplikasi berbasis *mobile*.



Dokumen ini adalah Arsip Miik :

Perpustakaan Universitas Islam Riau

DAFTAR PUSTAKA

- Anhar., 2010, *PHP & MySql Secara Otodidak*, Jakarta, PT TransMedia.
- Annisa, Itsna, Diina, Rudy Ariyanto, Ariadi Tri Retno Hayati Ririd., 2016, *Klasifikasi Kehamilan Beresiko Dengan Menggunakan Metode K-Nearest Neighbor (Studi Kasus Dinas Kesehatan Kabupaten Malang)*, Jurnal, Politeknik Negeri Malang, Jawa Timur.
- Arbie., 2004, *Manajemen Database dengan MySQL*, Andi, Yogyakarta.
- Banjarsari, Mutiara A., 2015, *Pencarian k-Optimal pada Algoritma kNN untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Berdasarkan IP Sampai Dengan Semester 4*, FMIPA Unlam, Banjarbaru.
- Djuhariningrum, T., 2005, *Pusat Pengembangan Geologi Nuklir-Batan*, Jakarta.
- Dwijosaputro., 1981, *Dasar – Dasar Mikrobiologi*, Jakarta, Djambatan.
- Freeze, R.A & Cherry, J.A., 1979, *Groundwater*, Prentice-Hall, Inc, USA.
- Gorunescu, Florin., 2011, *Data Mining: Concepts, Models, and Technique*, Verlag, Berlin Heidelberg, Springer.
- Han, J & M Kamber., 2006, *Data Mining Concepts and Techniques Second Edition*, San Francisco, Morgan Kaufmann.
- Hastie, T R Tibshirani, and J. Friedman., 2008, *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*, New York, Springer.
- Herlambang, A., 2006, *Pencemaran Air dan Strategi Penanggulangannya*, JAI, Volume 2, Nomor 1, Peneliti Pusat Teknologi Lingkungan, BPPT, halaman 16-28.
- Hermawati Fajar Astuti., 2013, *Data Mining*, Penerbit Andi, Yogyakarta.
- Jogiyanto, Hartono., 2005, *Analisis & Desain Sistem Informasi Pendekatan, Terstruktur Teori dan Praktek Aplikasi Bisnis*, Andi, Yogyakarta.
- Kusrini., 2018, *Sebaran Air Tanah Dangkal Di Permukaan Sekitar Pantai Kecamatan Ternate Utara Kota Ternate*, Universitas Khairun, Ternate Utara.

Masduqi, Ali., 2009, *Parameter Kualitas Air*, Surabaya.

Mestika, Indah, Ali, Manalu., 2014, *Perancangan Alat Ukur Konduktivitas Air (Conductivity Meter) Digital Dengan Sensor Resistif*, Fakultas Matematika Dan Ilmu Pengetahuan Alam (Fmipa), Universitas Sumatera Utara, Sumatera Utara.

Noviansyah, Reza, Muhammad, Tedy Rismawan, Dwi Marisa Midyanti., 2018, *Penerapan Data Mining Menggunakan Metode K-Nearest Neighbor Untuk Klasifikasi Indeks Cuaca Kebakaran Berdasarkan Data Aws (Automatic Weather Station) (Studi Kasus: Kabupaten Kubu Raya)*, Jurnal, Universitas Tanjungpura, Kalimantan Barat.

Pandie, Emerensye S Y., 2012, *Implemetasi Algoritma Data Mining K-Nearest Neighbor (KNN) Dalam Pengambilan Keputusan Pengajuan Kredit. Jurusan Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana, Kupang.*

Peraturan Menteri Kesehatan Republik Indonesia Nomor 416 Tahun 1990, *Syarat-Syarat dan Pengawasan Kualitas Air*, Jakarta.

Peraturan Menteri Kesehatan Republik Indonesia Nomor 32 Tahun 2017, *Standar Baku Mutu Kesehatan Lingkungan Dan Persyaratan Kesehatan air Untuk Keperluan Higiene Sanitasi, Kolam Renang, Solus Per Aqua, Dan Pemandian Umum*, Jakarta.

Prasetyo, E., 2014, *Data mining mengolah data menjadi informasi menggunakan matlab*, Yogyakarta, Andi Offset.

Ramdyasari, Intan., 2014, *Pengolahan Air Sumur Menjadi Air Siap Minum Melalui Proses Reverse Osmosis*, Other thesis, Politeknik Negeri Sriwijaya.

Rizal, Azwar., 2013, *Perbandingan Performa antara Imputasi Metode Konvensional dan Imputasi dengan Algoritma Mutual Nearest Neighbor*, Jurusan Teknik Informatika, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember. Surabaya.

Satzinger, Gary., Rosenblatt, Harry J., 2012, *System Analysis and Design in a Changing World*, Boston, MA, Course Technology.

Triatmojo, Bambang., 2008, *Hidrologi Terapan*, Beta Offset, Yogyakarta.

Turban, E., 2005, *Decision Support Systems and Intelligent Systems Edisi Bahasa Indonesia Jilid 1*, Andi, Yogyakarta.

Wahyono, Teguh., 2009, *Practice Guide PHP on Windows*, Jakarta, PT Elex Media.

Wu X., Kumar V., 2009, *The Top Ten Algorithms in Data Mining*, New York, CRC Press.

Yatina, B., Indra., 2010, *Flowchart, Algoritma, dan Pemrograman Menggunakan Bahasa C++ Builder*, Graha Ilmu.

Zanuardi, Muhammad., 2018, *Klasifikasi Kematangan Tebu Berdasarkan Tekstur Batang Menggunakan Metode K-Nearest Neighbor (Knn)*, Skripsi, Universitas Muhammadiyah Gresik, Jawa Timur.



Dokumen ini adalah Arsip Miik :

Perpustakaan Universitas Islam Riau