

Sistem Prediksi Tingkat Pencemaran Polusi Udara dengan Algoritma Naïve Bayes di Kota Makassar

Air Pollution Level Prediction System Using Naïve Bayes Algorithm in Makassar

Nurul Aini¹⁾, Rima Ruktiari²⁾, M Riyaldi Pratama³⁾, A. Fitrah Buana⁴⁾

^{1,2} Rekayasa Perangkat Lunak, ^{3,4} Teknik Informatika, STMIK Dipanegara Makassar
^{1,2,3,4}Jl. Perintis Kemerdekaan Km IX

nurulaini.m11@dipanegara.ac.id¹⁾, rimaruktiari@akane.waseda.jp²⁾, riyaldi@dipanegara.ac.id³⁾, fitrah.buana@gmail.com⁴⁾

Abstrak – Peningkatan kualitas hidup yang terjadi tidak diimbangi dengan peningkatan kualitas lingkungan. Berdasarkan data dari Badan Lingkungan Hidup Provinsi Sulawesi Selatan pada tahun 2018 menyatakan udara Kota Makassar termasuk daerah yang tercemar di banding daerah lain di Provinsi Sulawesi Selatan. Salah satu yang mempengaruhi hal ini adalah masifnya pertumbuhan kendaraan di kota Makassar. Untuk itu maka diperlukan satu sistem yang dapat memprediksi tingkat polusi udara di masa yang akan datang. Sehingga, masyarakat yang akan melakukan aktivitasnya di luar rumah pada keesokan harinya tidak perlu merasa khawatir karena sudah mendapatkan informasi tentang tingkat polusi udara. Klasifikasi dalam data mining bekerja pada data historis atau data sejarah. Data historis disebut data latihan atau training data. Historis data digunakan sebagai cara mendapatkan pengetahuan dan disebut data pengalaman. Naïve bayes merupakan kumpulan algoritma supervised learning yang berdasarkan Teorema Bayes dengan asumsi ‘naif’ saling bebas tidak tergantung antar fitur, metode klasifikasi dari Naïve Bayes memiliki tingkat akurasi lebih baik dari metode klasifikasi lainnya. Sistem ini mampu mengklasifikasi data polusi di kota makassar dan berdasarkan pengujian whitebox dan pengujian teknologi acceptance model dimana semua variabel memenuhi nilai syarat *cornbach* alpha diatas 0.60 dan ada beberapa faktor yang mempengaruhi tingkat polusi beberapa bagian di kota Makassar.

Abstract – Life quality improvement is not accompanied with environment quality improvement. Based on data of South Sulawesi Province Environment Agency in 2018, air quality in Makassar is categorized as polluted compared to other regions in South Sulawesi Province. One of the affecting factors is massive vehicle growth in Makassar. Therefore, we need a system to predict air pollution level in the future to make people feel save when they do their activities outside their houses. It can happen because they already have the information about air pollution level. Classification in data mining worked on historical data. Historical data is called training data. It is used as a way to gain knowledge and called as experience data. Naïve Bayes is a collection of supervised learning algorithms based on Bayes theorem with an assumption that 'naive' is independent among features. In addition, the Naïve Bayes classification method has a better accuracy than other methods. This system is able to classify pollution data in the city of Makassar and based on whitebox testing and acceptance model technology testing where all variables meet the Cornbach Aplha requirement values above 0.60. Further, based on the test results, there are several factors affecting pollution level in some regions of Makassar.

Keywords: Naïve Bayes, Prediksi, Polusi Udara

PENDAHULUAN

Untuk Kota Makassar polusi yang dihasilkan dari asap kendaraan seiring waktu makin bertambah dan berdasarkan data dari Samsat Kota Makassar, pada 2018 tercatat 1.425.151 unit atau bertambah 87.009 unit dibandingkan 2017. Adapun, pada 2016 jumlah kendaraan bermotor di Kota Makassar baru berkisar 1.252.755 unit. Artinya, dalam dua tahun tercatat pertambahan 172.395 unit.

Kondisi ini tentu sangat mengkhawatirkan mengingat Polutan yang dikeluarkan oleh kendaraan

bermotor seperti karbon monoksida (CO), nitrogen oksida (NOx), hidrokarbon (HC), Sulfur dioksida (SO2), timah hitam (Pb) dan karbon dioksida (CO2) sangat berbahaya bagi kesehatan manusia. Selain itu, informasi tentang kualitas udara ini bisa menjadi acuan bagi pemerintah untuk mengkaji kembali kebijakan yang berkaitan dengan pengaturan dan pembatasan jumlah kendaraan bermotor.

Pada penelitian ini, digunakan teknik data mining untuk menemukan pola polusi dari beberapa parameter ISPU di udara yang telah di hitung menggunakan

tingkat kadar ISPU oleh Dinas Pengelolaan Lingkungan Hidup (DPLH) Provinsi Sulawesi Selatan.

Data mining proses menemukan hubungan dalam data yang tidak diketahui oleh pengguna dan menyajikannya dengan cara yang dapat dipahami sehingga hubungan tersebut dapat menjadi dasar pengambilan keputusan. Teknik *data mining* yang akan digunakan dalam penelitian ini adalah algoritma NBC yang merupakan sebuah pengklasifikasian probabilitas sederhana yang mengaplikasikan teorema bayes. Algoritma NBC dapat mengolah data kuantitatif dan data diskrit yang hanya memerlukan sejumlah kecil data pelatihan untuk perhitungan estimasi peluang yang dibutuhkan untuk klasifikasi.

METODOLOGI PENELITIAN

Ukuran Ketepatan Metode Peramalan

Dalam pemodelan deret berkala, sebagian data yang diketahui dapat digunakan untuk meramalkan data berikutnya. Selisih besaran (ukuran kesalahan peramalan) data peramalan terhadap data aktual yang terjadi merupakan suatu data penting untuk menilai ketepatan suatu metode peramalan. Dengan membandingkan ukuran kesalahan beberapa metode peramalan, akan diperoleh metode mana yang mempunyai ukuran kesalahan terkecil, sehingga nilai peramalan dapat dipakai sebagai acuan dalam menentukan kebutuhan-kebutuhan dimasa yang akan datang.

Indeks Kualitas Udara

Indeks Kualitas Udara (*Air Quality Indeks*) adalah standar yang digunakan dalam pengukuran pencemaran udara. Setiap suatu Negara memiliki standar yang berbeda dalam pengukuran pencemaran udara di Indonesia ISPU (Indeks standar Pencemaran Udara), di Beijing *Air Quality Indeks (AQI)*, di Singapore *Pollutant Standards Index (PSI)*, di Korea Selatan *Comprehensive Air Quality Indeks (CSI)*, di United Kingdom *Air Pollutan Banding (APB)*, di Eropa *Common Air Quality Index (CAQI)* dan lainnya. Adapun material dari pencemaran atau polusi udara dibedakan atas partikel dan gas. Material partikel diantaranya (PM_{10} , $PM_{2.5}$) pada gas (CO , NO_2 , SO_2 , O_3). Dari beberapa material pencemar udara tersebut sangat berdampak buruk terhadap pada kesehatan jadi jika orang terpapar oleh material tersebut dapat menyebabkan gangguan pernapasan, ispa, keracunan gas, dan lainnya (Hermawan, Hananto, & Lasut, 2016).

Di Indonesia ISPU diatur berdasarkan Keputusan Badan Pengendalian Dampak Lingkungan (Bapedal) Nomor KEP-107/Kabapedal/11/1997. Berikut tabel Batas Indeks Standart Pencemar Udara (Dalam Satuan SI) dalam proses pengukurannya masih dilakukan secara konvensional bukan secara digital, pengukuran dilakukan dengan berdasarkan periode waktu tertentu. Adapun standar dari dan batas kadar konsentrasi gas pencemaran udara dapat dilihat pada tabel sesuai dengan kategori (Akhmad, Vitianingsih, & Wijaya, 2017).

Naïve Bayes

Pengklasifikasian menggunakan Teorema Bayes ini membutuhkan biaya komputasi yang mahal (waktu processor dan ukuran *memory* yang besar) karena kebutuhan untuk menghitung nilai probabilitas untuk tiap nilai dari perkalian kartesius untuk tiap nilai atribut dan tiap nilai kelas. Data latih untuk Teorema Bayes membutuhkan paling tidak perkalian kartesius dari seluruh kelompok atribut yang mungkin, jika misalkan ada 16 atribut yang masing-masingnya berjenis boolean tanpa missing value, maka data latih minimal yang dibutuhkan oleh Teorema bayes untuk digunakan dalam klasifikasi adalah $2^{16} = 65.536$ data. Untuk mengatasi kekurangan tersebut maka digunakan *Naive Bayes*.

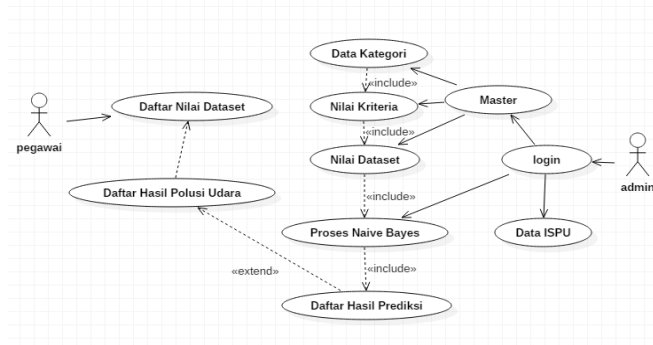
Naive Bayes Classifier merupakan sebuah metoda klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari Naive Bayes Classifier ini adalah asumsi yang sangat kuat (naïf) akan independensi dari masing-masing kondisi atau kejadian (Syukri-Mustafa & Wayan-Simpem, 2019).

Menurut Olson dan Delen , menjelaskan Naïve Bayes untuk setiap kelas keputusan, menghitung probabilitas dengan syarat bahwa kelas keputusan adalah benar, mengingat vektor informasi obyek. Algoritma ini mengasumsikan bahwa atribut obyek adalah independen. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dari "master" tabel keputusan (Olson & Delen, 2008).

Naive Bayes Classifier bekerja sangat baik dibanding dengan model *classifier* lainnya. Hal ini dibuktikan oleh Xhemali, Hinde dan Stone dalam

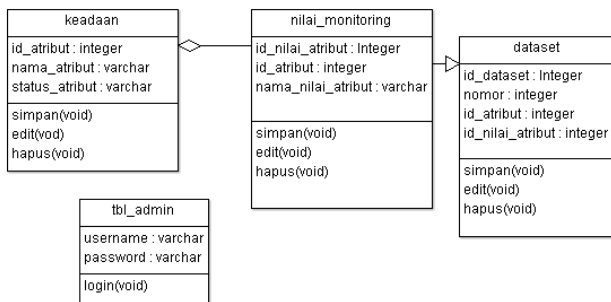
jurnalnya “*Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages*” mengatakan bahwa “*Naive Bayes Classifier* memiliki tingkat akurasi yang lebih baik dibanding model *classifier* lainnya”.

Rancangan sistem merupakan suatu sistem kegiatan yang dilakukan untuk mendesain suatu sistem yang mempunyai tahapan-tahapan kerja yang tersusun secara logis, dimulai dari pengumpulan data yang diperlukan guna pelaksanaan perancangan tersebut. *Use Case Diagram* menggambarkan fungsionalitas yang diharapkan dari sebuah sistem serta menggambarkan kebutuhan sistem dari sudut pandang *user*.



Gambar 1 Uses Case yang di usulkan

Class diagram merupakan gambaran objek yang mewakili sistem yang digunakan sebagai dasar membangun sistem dalam sisi pemrograman dan basis data yang terdiri dari beberapa *class*



Gambar 2 Class Diagram

Data Training

Data training digunakan menyusun pola klasifikasi yang membentuk informasi baru dalam sejumlah data yang banyak. Dalam proses mining terdapat sejumlah 646 data ISPU terhitung dari tanggal 1 januari 2018 hingga 1 oktober 2019 yang diperoleh dari Badan Pengelolaan Lingkungan Hidup , titik pengambilan data di Jln. R.A kartini Baru ,Kota Makassar. Terdapat data polusi kondisi Baik = 596 , kondisi Sedang = 49 , dan kondisi Tidak Sehat = 1.

Tabel 1 Data unsur udara dari 5 parameter.

KODE STASIUN	PM10	SO2	CO	O3	NO2
KLHK-MAKASSAR	0	0	0	0	0
KLHK-MAKASSAR	11	0	0	0	0
KLHK-MAKASSAR	0	0	0	0	0
KLHK-MAKASSAR	12	0	0	0	0
KLHK-MAKASSAR	0	0	0	0	0
KLHK-MAKASSAR	11	0	0	0	0
KLHK-MAKASSAR	0	0	0	0	0
KLHK-MAKASSAR	11	0	0	0	0
KLHK-MAKASSAR	0	0	0	0	0
KLHK-MAKASSAR	11	0	0	0	0
KLHK-MAKASSAR	0	0	0	0	0

Tabel 2 Data ISPU dari setiap parameter.

KODE STASIUN	ISPU PM10	ISPU SO2	ISPU CO	ISPU O3	ISPU NO2
KLHK-MAKASSAR	10	15	11	25	0
KLHK-MAKASSAR	11	0	0	0	0
KLHK-MAKASSAR	11	15	11	25	0
KLHK-MAKASSAR	12	0	0	0	0
KLHK-MAKASSAR	10	15	11	25	0
KLHK-MAKASSAR	11	0	0	0	0
KLHK-MAKASSAR	10	15	11	25	0
KLHK-MAKASSAR	11	0	0	0	0
KLHK-MAKASSAR	10	15	11	25	0
KLHK-MAKASSAR	11	0	0	0	0
KLHK-MAKASSAR	11	15	11	25	0

Data Testing

Data *testing* atau data uji yang digunakan sejauh mana algoritma naïve bayes dapat menghasilkan klasifikasi data yang benar dan akurat. Data ini akan digunakan pengujian berupa data ISPU dari 5 parameter untuk bulan Oktober . Setelah proses *mining*, data ini akan memiliki kelas berdasarkan probabilitas yang diperoleh dari data *training* .

Tabel 3 Data Testing yang akan di uji kedalam sistem

No	Tgl	ISPU PM10	ISPU SO2	ISPU CO	ISPU O3	ISPU NO2
1	14 Okt	10	25	30	85	10
2	16 Okt	10	14	0	0	0

Data Testing diatas akan di implementasikan ke dalam algortima Naïve Bayes dengan beberapa tahapan.

HASIL DAN PEMBAHASAN

Langkah awal dari Algoritma Naïve Bayes adalah menemukan nilai probabilitas dari setiap atribut, jika data dari atribut/parameter berupa data variatif/diskrit maka data tersebut diproses dengan analisis statistik menentukan nilai Mean dan Standar Deviasi .

Menentukan Prior Probability (P).

Tabel 4 Hasil Klasifikasi Data Training

Kondisi	Jumlah
Baik	596
Sedang	49
Tidak Sehat	1
Jumlah	646

$$P(\text{Baik}) = \frac{596}{646} = 0,922 \quad P(\text{Sedang}) = \frac{49}{646} = 0,075$$

$$P(\text{Tidak Sehat}) = \frac{1}{646} = 0,001$$

Menentukan *Mean* dan *Standar Deviasi* untuk parameter PM10 (X1), SO2(X2), CO(X3), O3(X4) NO2(X5) . Dikarenakan setiap nilai dari parameter bersifat variatif , maka akan dimasukkan kedalam perhitungan *Distibusi Gauss* untuk memudahkan dalam menentukan nilai probability dari tiap parameter.

Tabel 5 Nilai Mean dan Standar Deviasi PM10 (X1)

	PM10(X1)		
	Baik	Sedang	Tidak Sehat
10	13	18	
11	28	0	
....		
24	14	0	
Means	12.24	12.03	16
Stdv	7.65	9.34	2

Tabel 6 Nilai Mean dan Standar Deviasi SO2 (X2)

	SO2(X2)		
	Baik	Sedang	Tidak Sehat
15	53	19	
0	14	13	
....	
2	0	0	
Means	1.97	9.67	16.00
Stdv	6.42	21.35	3.00

Tabel 7 Nilai Mean dan Standar Deviasi CO (X3)

	CO(X3)		
	Baik	Sedang	Tidak Sehat
11	5	16	
0	1	1	
....	
0	0	0	
Means	0.42	6.46	8.50
Stdv	2.41	19.38	7.50

Tabel 8 Nilai Mean dan Standar Deviasi CO3 (X4)

	PM10(X1)		
	Baik	Sedang	Tidak Sehat
25	55	129	
0	82	109	
....	
2	0	0	
Means	6.90	62.21	119.00
Stdv	16.26	22.99	10.00

Tabel 9 Nilai Mean dan Standar Deviasi NO2 (X5)

	PM10(X1)		
	Baik	Sedang	Tidak Sehat
0	0	0	
0	0	0	
....	
1	0	0	
Means	0.00	0.00	0.00
Stdv	0.04	0.00	0.00

Menghitung nilai probabilitas dari setiap atribut/parameter dari Data Testing menggunakan Distribusi Gauss..

Untuk Data Testting yang ke-1.

Dimana nilai X1 = 10 , X2 = 25 , X3 = 30 , X4=85 ,X5 = 10 .

1. Probability PM10
 $P(X1| 10 = \text{“ Baik “}) = 0,61517$
 $P(X1| 10 = \text{“ Sedang “}) = 0,58594$
 $P(X1| 10 = \text{“ Tidak Sehat “}) = 0,99865$
2. Probability SO2
 $P(X2| 25 = \text{“ Baik “}) = 0,00017$
 $P(X2| 25 = \text{“ Sedang “}) = 0,23637$
 $P(X2| 25 = \text{“ Tidak Sehat “}) = 0,00135$
3. Probability CO
 $P(X3| 30 = \text{“ Baik “}) = 0$
 $P(X3| 30 = \text{“ Sedang “}) = 0,11225$
 $P(X3| 30 = \text{“ Tidak Sehat “}) = 0,00207$
4. Probability O3
 $P(X4| 85 = \text{“ Baik “}) = 0$
 $P(X4| 85 = \text{“ Sedang “}) = 0,16077$
 $P(X4| 85 = \text{“ Tidak Sehat “}) = 0,99966$
5. Probability NO2
 $P(X5| 10 = \text{“ Baik “}) = 0$
 $P(X5| 10 = \text{“ Sedang “}) = 0$

$$P(X5|10 = \text{“ Tidak Sehat “}) = 0$$

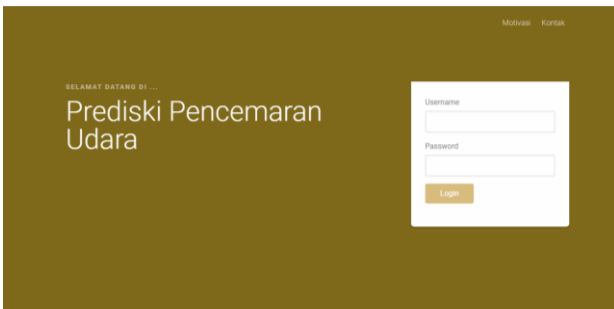
Sehingga nilai probabilitas setiap atribut dikalikan untuk menentukan nilai Likelihood, sebagai berikut :

Tabel 10 Perkalian nilai Likelihood atribut.

	PM10	SO2	CO	O3	Hasil Kali
Baik	0.61517	0.00017	0.00000	0.00000	0
Sedang	0.58594	0.23637	0.11225	0.16077	0.002499406
Tidak Sehat	0.99865	0.00135	0.00207	0.99966	0.00000279

Kesimpulan prediksi polusi adalah **SEDANG** karena nilai Likelihood Sedang lebih tinggi dibandingkan nilai likelihood Baik dan Tidak Sehat.

Hasil perancangan sistem dituangkan dalam bentuk system WEB untuk memudahkan *user* dalam mengolah data prediksi polusi udara di Makassar. Sistem ini dibangun sesuai dari perancangan diagram UML



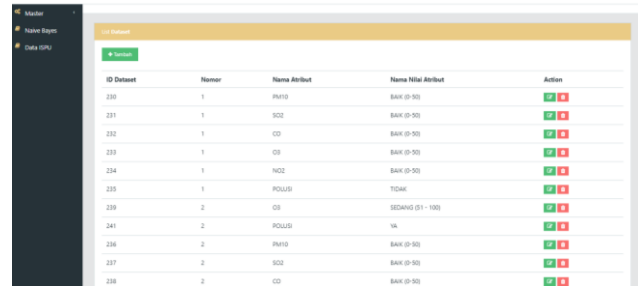
Gambar 3 Tampilan sistem untuk halaman Login Admin

Pada halaman Admin ini disediakan beberapa akses untuk menginput data kriteria , data ISPU harian , dan olah data Naïve Bayes.



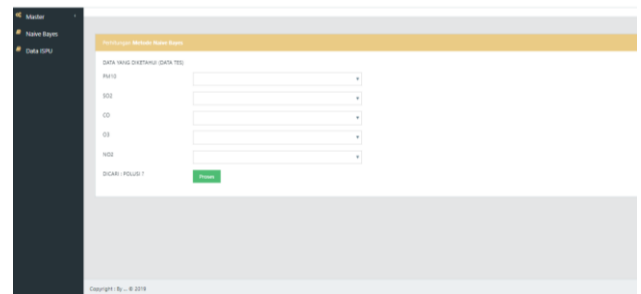
Gambar 4 Tampilan Sistem untuk atribut penilaian

Pada tampilan olah data kriteria / atribut , user dapat menambah data baru, edit data kriteria dan menghapus data kriteria. Data yang telah diinputkan sejumlah 646 data yang diperoleh dari Badan Pengelolaan Lingkungan Hidup di Makassar mulai dari tanggal 1 januari 2019 hingga 30 oktober 2019. Pada halaman ini data input melalui proses manual dengan cara menginputkan satu per satu yang akan dikerjakan oleh admin.



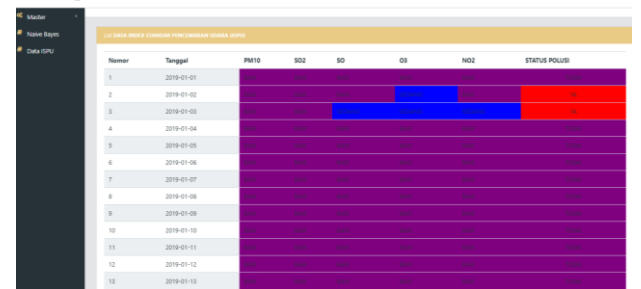
Gambar 5 Tampilan sistem untuk Data Training.

Pada halaman ini user dapat menginput data yang lama atau data yang telah memiliki klasifikasi .User dapat menambah data , mengubah data dan menghapus data.



Gambar 6 Tampilan sistem untuk menginput nilai atribut polusi prose Naïve Bayes.

Pada tampilan ini user dapat menginput data yang akan diprediksi oleh algoritma Naïve Bayes. Setelah user menekan tombol proses hasil perhitungan akan ditampilkan untuk melihat perincian perhitungan.



Gambar 7 Tampilan dataset ISPU.

Halaman ini user dapat melihat data training dengan nilai ISPU setiap atribut. Sesuai data yang diperoleh di Badan Pengelolaan Lingkungan Hidup rata-rata data polusi di Kota Makassar dalam status Baik dan sangat jarang ditemui dalam status Tidak Sehat.

Pengujian Teknologi Acceptance Model (TAM)

Reaksi dan persepsi akan mempengaruhi sikap seseorang dalam menerima Teknologi Informasi. Salah satu tolak ukur yang mempengaruhi adalah persepsi pengguna terhadap kemanfaatan dan kemudahan menggunakan teknologi sebagai suatu tindakan yang berupa alasan dalam konteks lingkungan pengguna IT.

Pada penelitian ini pengujian sistem menggunakan metode Teknologi Acceptance Model dimana model ini dapat mengukur kebermanfaatan sistem setelah di implemmentasikan. TAM mengasumsikan bahwa penerimaan seseorang atas teknologi informasi dipengaruhi oleh 2 variabel utama yaitu *Perceived Usefulness* (Persepsi Kebermanfaatan) dan *Perceived Ease Of Use* (Persepsi Kemudahan Penggunaan) (Sensuse & Widiatmika, 2008).

Kuesioner dibagikan secara acak , dari hasil perhitungan kuesioner diperoleh data bahwa responden dominan adalah kategori pegawai. Hasil pengelompokan data responden dapat dilihat sebagai berikut :

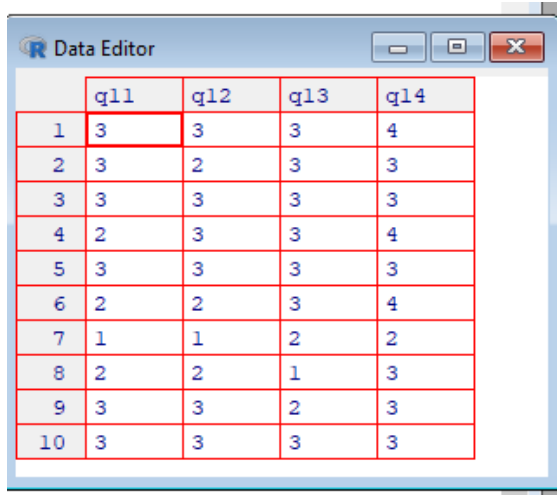
Tabel 11. Pengelompokan jumlah responden

No	Kategori	Jumlah
1	Masyarakat Umum	2
2	Pegawai	5
3	Pelajar	3
Jumlah		10

Pengolahan data terhadap 10 sample responden menggunakan 3 variabel TAM yang berisi 10 pertanyaan. Setiap pertanyaan menggunakan skala likert, yaitu :

- Angka 1 = Sangat Tidak Setuju
- Angka 2 = Tidak Setuju
- Angka 3 = Setuju
- Angka 4 = Sangat Setuju

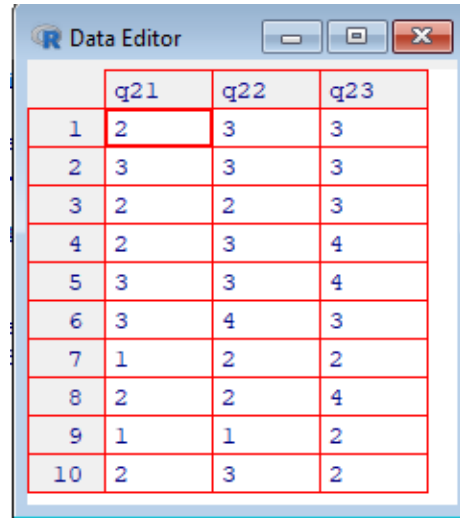
Untuk mempermudah pengolahan data , pada penelitian ini menggunakan bahasa R untuk mengolah instrument. Data responden di input manual kedalam program R GUI sebagai data frame , maka akan tampil R Editor sebagai berikut.



Gambar 8 Data Variabel TAM X1(Persepsi Kebermanfaatan)

Untuk variable TAM X1 pertanyaan yang tercantum dalam tabel adalah sebagai berikut :

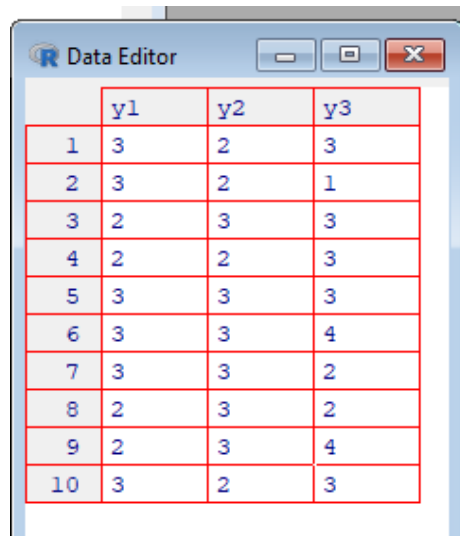
- X1.1 Penggunaan Sistem memberi informasi pada pekerja.
- X1.2 Penggunaan Sistem meningkatkan kinerja.
- X1.3 Penggunaan Sistem mempermudah pekerjaan.
- X1.4 Penggunaan Sistem bermanfaat bagi masyarakat.



Gambar 9 Data Variabel TAM X2 (Persepsi Kemudahan)

Untuk variabel TAM X2 pertanyaan yang tercantum dalam tabel adalah sebagai berikut :

- X2.1 Sistem yang digunakan User friendly
- X2.2 Sistem yang digunakan mudah menambah data
- X2.3 Sistem yang digunakan bersifat dinamis



Gambar 10 Data Variabel TAM Y (Penerimaan)

Untuk variable TAM Y pertanyaan yang tercantum dalam tabel adalah sebagai berikut :

- Y1 Saya Menggunakan Sistem Prediksi Polusi
- Y2 Saya merasa dengan adanya sistem ini bermanfaat
- Y3 Saya akan merekomendasikan Sistem ini ke teman lainnya

Berdasarkan hasil pengumpulan data dari 10 responden, semua instrument akan di uji realibilitasnya untuk memastikan bawah instrument yang digunakan sudah akurat dan handal. Pengujian realibilitas pada penelitian ini dilakukan dengan *Cronbach Alpha*, bertujuan mengukur konsistensi jawaban responden, pengujian ini dapat dilakukan bersama terhadap seluruh seluruh pertanyaan yang diajukan . Instrument dapat dikatakan memenuhi realibilitas jika nilai *Cronbach Alpha* lebih besar dari 0,60. Berikut hasil pengujian realibilitas menggunakan bahasa R.

```

R Console
Error in psych::alpha(ujirealiber) : object 'ujirealiber' not found
> psych::alpha(ujirealiber)
Reliability analysis
Call: psych::alpha(x = ujirealiber)

raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
 0.79 0.79 0.8 0.48 3.7 0.11 2.7 0.54 0.45

lower alpha upper 95% confidence boundaries
0.57 0.79 1.01

Reliability if an item is dropped:
raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
q11 0.72 0.72 0.64 0.47 2.6 0.15 0.00071 0.45
q12 0.65 0.65 0.57 0.38 1.9 0.19 0.01366 0.45
q13 0.76 0.76 0.77 0.51 3.1 0.13 0.07013 0.50
q14 0.79 0.79 0.76 0.56 3.8 0.12 0.03593 0.45

Item statistics
n raw.r std.r r.cor r.drop mean sd
q11 10 0.80 0.79 0.76 0.62 2.5 0.71
q12 10 0.88 0.87 0.87 0.75 2.5 0.71
q13 10 0.75 0.75 0.60 0.55 2.6 0.70
q14 10 0.69 0.70 0.57 0.47 3.2 0.63

Non missing response frequency for each item
1 2 3 4 miss
q11 0.1 0.4 0.6 0.0 0
q12 0.1 0.3 0.6 0.0 0
q13 0.1 0.2 0.7 0.0 0
    
```

Gambar 11 Realibilitas Instrument X1

Berdasarkan gambar diatas hasil pengujian bahasa R menunjukkan bahwa variabel TAM X1 yaitu persepsi kebermanfaatan dikatakan telah reliable. Hal ini dibuktikan dengan *cornbach alpha* sebagai index relibilitas variable persepsi kebermanfaatan sebesar 0,79.

```

R Console
> ujirealiberX2 <- data.frame(q21=numeric(0),q22=numeric(0),q23=numeric(0))
> ujirealiberX2 <- edit(ujirealiberX2)
> psych::alpha(ujirealiberX2)
Reliability analysis
Call: psych::alpha(x = ujirealiberX2)

raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
 0.78 0.79 0.79 0.55 3.7 0.13 2.6 0.67 0.55

lower alpha upper 95% confidence boundaries
0.53 0.78 1.03

Reliability if an item is dropped:
raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
q21 0.49 0.49 0.32 0.32 0.95 0.324 NA 0.32
q22 0.71 0.71 0.55 0.55 2.49 0.152 NA 0.55
q23 0.88 0.88 0.79 0.79 7.33 0.077 NA 0.79

Item statistics
n raw.r std.r r.cor r.drop mean sd
q21 10 0.92 0.93 0.92 0.83 2.1 0.74
q22 10 0.84 0.84 0.78 0.62 2.6 0.84
q23 10 0.75 0.75 0.54 0.46 3.6 0.82

Non missing response frequency for each item
1 2 3 4 miss
q21 0.2 0.5 0.3 0.0 0
q22 0.1 0.3 0.5 0.1 0
q23 0.0 0.3 0.4 0.3 0
    
```

Gambar 12 Realibilitas Instrument X2

Berdasarkan gambar diatas hasil pengujian bahasa R menunjukkan bahwa variabel TAM X2 yaitu persepsi kemudahan dikatakan telah reliable. Hal ini dibuktikan dengan *cornbach alpha* sebagai index relibilitas variable persepsi kemudahan sebesar 0,78.

Jika masing – masing variabel memiliki nilai dibawah raw_alpha dari pengolahan bahasa R , dapat dikatakan bahwa variable tersebut perlu dilakukan perbaikan

```

R Console
> psych::alpha(ujirealiberY)
Some items ( y1 ) were negatively correlated with the total scale and
probably should be reversed.
To do this, run the function again with the 'check.keys=TRUE' option
Reliability analysis
Call: psych::alpha(x = ujirealiberY)

raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
-0.81 -0.96 -0.38 -0.2 -0.49 0.92 2.7 0.33 -0.25

lower alpha upper 95% confidence boundaries
-2.6 -0.81 0.99

Reliability if an item is dropped:
raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
y1 0.16 0.18 0.098 0.098 0.22 0.47 NA 0.098
y2 -0.66 -0.68 -0.254 -0.254 -0.41 1.02 NA -0.254
y3 -1.40 -1.51 -0.431 -0.431 -0.60 1.44 NA -0.431

Item statistics
n raw.r std.r r.cor r.drop mean sd
y1 10 0.26 0.23 NaN -0.42 2.4 0.70
y2 10 0.30 0.49 NaN -0.22 2.6 0.52
y3 10 0.75 0.62 NaN -0.19 3.1 0.88

Non missing response frequency for each item
1 2 3 4 miss
y1 0.1 0.4 0.5 0.0 0
y2 0.0 0.4 0.6 0.0 0
y3 0.0 0.3 0.3 0.4 0
Warning messages:
    
```

Gambar 13 Realibilitas Instrument Y

Berdasarkan gambar diatas hasil pengujian bahasa R menunjukkan bahwa variabel TAM Y yaitu penerimaan (Acceptance) dikatakan telah reliable. Hal ini dibuktikan dengan *cornbach alpha* sebagai index relibilitas variable penerimaan sebesar 0,81.

Hasil *cronbach alpha* ini lebih besar dari 0.60 sebagai syarat bahwa instrument dapat dikatakan reliable. Dengan demikian dalam penelitian ini valid dan reliable karena telah memenuhi persyaratan.

KESIMPULAN

Berdasarkan hasil yang diperoleh pada pembangunan sistem dan pengujian perangkat lunak maka dapat disimpulkan bahwa Aplikasi data mining ini dapat mengklasifikasikan kondisi pencemaran polusi udara di Kota Makassar dengan menggunakan Algoritma Naive Bayes Classifier. Semakin banyak data training yang digunakan, maka kecerdasan sistem akan semakin baik, serta akan meningkatkan akurasi penelitian. Berdasarkan hasil pengujian akurasi, bahwa ada beberapa daerah yang memiliki tingkat pencemaran yang tidak sama dengan lokasi pengambilan data sehingga untuk mengambil keputusan status polusi untuk Kota Makassar tidak dapat merata.

Dari hasil pengujian Teknologi Acceptance Model diperoleh instrument variabel TAM untuk X1 nilai *cornbach alpha* sebesar 0.79 , X2 sebesar 0.78 dan Y sebesar 0.81 , dimana semua variable memenuhi syarat yaitu lebih besar dari 0.60. Sehingga dapat dikatakan instrument penilaian yang digunakan reliable.

Saran untuk pengembangan penelitian selanjutnya berupa data training atau data set dapat di peroleh di beberapa titik Kota Makassar tentunya dukungan ini sangat diperlukan dari Badan Pengelolaan Lingkungan Hidup . Pengembangan sistem ini masih dibutuhkan dari segi user interface agar lebih memudahkan user dalam berinteraksi dengan sistem , pengolahan database dimana data yang digunakan untuk algoritma Naïve Bayes sangat banyak , dan penggunaan metode data mining lainnya untuk dapat dibandingkan hasilnya. Dan kedepannya di harapkan dapat memberikan informasi yang layak kepada masyarakat luas mengenai tingkat polusi udara di kota Makassar khususnya para turis yang akan datang berkunjung .

Biasa Menggunakan Weighted Product Method determination Application Quality Schools Using Weighted Product Method. In *SISITI: Seminar Ilmiah Sistem Informasi dan Teknologi Informasi* (Vol. 5, No. 1)

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada STMIK Dipanegara atas dukungannya pada penelitian ini dan Badan Pengelolaan Lingkungan Hidup Makassar atas dukungannya dalam memperlancar penelitian ilmiah . Penulis juga berterima kasih kepada KOMINFO Makassar atas kesempatan yang diberikan untuk mempublish karya ilmiah ini di Seminar Nasional Komunikasi dan Informatika yang ke 3.

DAFTAR PUSTAKA

- Akhmad, M. W., Vitianingsih, A. V., & Wijaya, T. A. (2017). Pemetaan Tingkat Polusi Udara di Kota Surabaya Berbasis Android. *Inform: Jurnal Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 1(1).
- Hermawan, A., Hananto, M., & Lasut, D. (2016). Peningkatan Indeks Standar Pencemaran Udara (ISPU) dan kejadian gangguan saluran pernapasan di kota Pekanbaru. *Jurnal Ekologi Kesehatan*, 15(2), 76–86.
- Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques*. Springer Science & Business Media.
- Sensuse, D. I., & Widiatmika, I. M. A. A. (2008). Pengembangan model penerimaan teknologi internet oleh pelajar dengan menggunakan konsep Technology Acceptance Model (TAM). *Jurnal Sistem Informasi*, 4(2), 81–92.
- Syukri-Mustafa, M., & Wayan-Simpen, I. (2019). Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier (Studi Kasus STMIK Dipanegara). *SISITI: Seminar Ilmiah Sistem Informasi Dan Teknologi Informasi*, 7.
- Tamsir, N., Arie, J., Rezky, Y., & Purnama, M. I. (2017, August). Aplikasi Penentuan Kualitas Sekolah Luar