

PENERAPAN DATA MINING UNTUK MENGLASIFIKASI TINGKAT BAHAYA POLUTAN PM10 DI KOTA BANJARBARU

Nur Arminarahmah¹⁾, Mirza Yogi Kurniawan²⁾, Al Fath Riza Kholdani³⁾

Fakultas Teknologi Informasi ,Universitas Islam Kalimantan Muhammad Arsyad Al Banjari Banjarmasin

Email¹⁾ : nur.armina@uniska-bjm.ac.id

Email ²⁾: mirza.yogy@gmail.com

Email ³⁾: kholdanialfath@gmail.com

Abstrak

Salah satu aplikasi Data Mining adalah klasifikasi, kami menggunakan algoritma Naive Bayes untuk mengklasifikasikan tingkat bahaya polutan PM 10 untuk menghasilkan model yang digunakan untuk mendapatkan informasi tentang kondisi polutan PM10. Aplikasi awal menggunakan data dari wilayah kota Banjarbaru karena ada kabut asap dalam beberapa tahun terakhir di Kalimantan Selatan dan studi kasus penelitian ini di kota Banjarbaru karena telah meningkat sehingga mengganggu visibilitas dan menghambat semua kegiatan baik kantor maupun sekolah. Proses klasifikasi dimulai dengan mengolah data PM10 menggunakan atribut suhu, kelembaban dan waktu kemudian data yang diperoleh diklasifikasikan menggunakan algoritma naive bayes untuk menghasilkan kelas tingkat bahaya dari 0 - 4 dengan tingkat akurasi keberhasilan klasifikasi 60%.

Kata Kunci : Datamining, Naive Bayes, PM10

1.PENDAHULUAN

Kejadian kabut asap pada beberapa tahun ini di Kalimantan selatan menimbulkan dampak yang sangat buruk bagi kondisi lingkungan serta bagi kesehatan sehingga diperlukan suatu peringatan dini serta pembelajaran tentang kondisi tingkat bahaya dari polusi asap yang ditimbulkan oleh pembakaran lahan.

Polutan memiliki dampak terhadap kesehatan dan lingkungan [1]. Dampak kesehatannya berupa infeksi saluran pernapasan selain berdampak pada pernapasan juga berbahaya bagi kesehatan jantung [1], [2], dan mengganggu pertumbuhan tulang pada anak-anak [3]. Terhadap lingkungan mengganggu jarak pandang [4] dan kenyamanan berada diluar ruangan.

Salah satu bentuk polutan adalah partikel padat yang mengambang diudara dan disebut sebagai *Particulate Matter* (PM).Salah satu PM yang rutin dilakukan pengukuran di banjarbaru adalah PM10 yaitu partikel mengambang dengan median ukuran diameter dibawah 10 µm.PM10 ini bisa berasal dari kegiatan

antropogenik seperti pembakaran hutan, debu dan asap akibat lalu lintas jalan, debu jalan, atau bisa juga berasal dari sumber alami seperti letusan gunung berapi, kebakaran hutan, hempasan ombak, ataupun berasal dari tanaman [4], [5] Ada banyak faktor yang mempengaruhi konsentrasi PM10 baik secara alami atau antropogenik, misalnya suhu, kelembaban, pembakaran hutan, kegiatan pabrik dan lalu lintas serta banyak lagi aktifitas lainnya.

Banjarbaru merupakan daerah yang pengaruh antropogeniknya sangat tinggi karena penduduknya padat serta lalu lintas cukup padat serta ada beberapa kawasan industri sehingga diperkirakan berpengaruh terhadap konsentrasi PM10 Banjarbaru. Dalam hal ini untuk mengetahui tingkat bahaya PM10 maka diterapkan metode klasifikasi menggunakan naive bayes untuk memperoleh hasil klasifikasi tingkat bahaya dari beberapa atribut dalam entitas di PM10.

Data mining merupakan proses menemukan informasi baru atau penggalian pengetahuan dengan menggunakan pola dengan aturan

tertentu dari sejumlah data yang besar untuk meningkatkan nilai tambah dari informasi tersebut.

Klasifikasi merupakan metode data mining yang dapat digunakan untuk menemukan model atau fungsi yang menjelaskan kelas data bertujuan untuk membedakan data kedalam kelas-kelas yang berbeda berdasarkan aturan tertentu dan kelas tersebut merupakan suatu objek yang labelnya tidak diketahui [6].

Proses klasifikasi pada penelitian ini bertujuan agar dapat melakukan pemetaan kelas – kelas bahaya polutan ke dalam kelas yang sudah didefinisikan. Metode Naive Bayes berpotensi baik dalam proses klasifikasi dokumen dibandingkan dengan metode pengklasifikasian yang yaitu dalam hal akurasi dan efisiensi komputasi [7]. Naive bayes memiliki keuntungan pembelajaran yang sangat cepat dan proses komputasi yang kecil [8].

Nilai ambang batas adalah batas konsentrasi polusi udara yang diperbolehkan berada dalam udara ambien. Peraturan Pemerintah RI No. 41 tahun 1999 tentang pengendalian pencemaran udara mengatur ambang baku mutu udara ambien nasional untuk PM10 adalah sebesar 150 µg/m³ dalam periode waktu 24 Jam.

Tabel 1. Kategori Tingkat Bahaya [9]

Level	Konsentrasi PM10	Tingkat Bahaya
0	0-50	Baik

1	51-150	Sedang
2	151-250	Tidak sehat
3	251-350	Sangat Tidak sehat
4	>350	Berbahaya

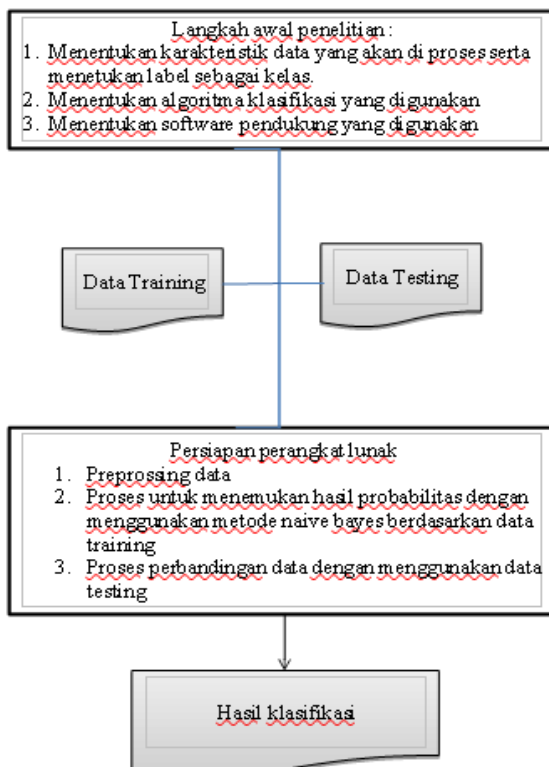
2.METODE PENELITIAN

Aliran perancangan penelitian yang akan dikerjakan, langkah awal dengan menentukan karakteristik data dan mengolah data training untuk diimplementasikan ke algoritma klasifikasi.

Gambar 2.1 Aliran Proses Klasifikasi

Tahapan - tahapan penelitian yaitu sebagai berikut :

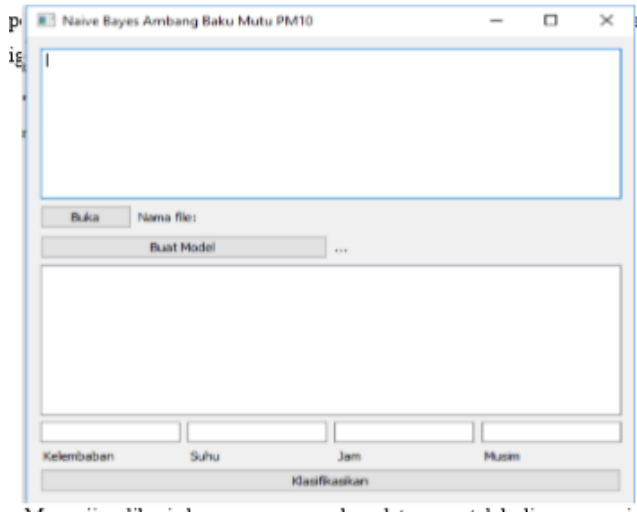
1. Proses preprocessing data PM10 dengan atribut Suhu, kelembaban dan



20-Jan-15	02:20	10 0	23 0	16,51 7	10:20	8 8	25 8	10,37 6	18:20	93	24,8	2,78
20-Jan-15	02:30	10 0	23 1	16,43 4	10:30	8 8	26 0	10,63 8	18:30	94	24,6	3,426
20-Jan-15	02:40	10 0	23 1	15,38 15,38	10:40	8 7	26 1	10,47 7	18:40	94	24,5	3,962
20-Jan-15	02:50	10 0	23 0	14,97 6	10:50	8 8	26 0	10,33 7	18:50	95	24,4	6,911
20-Jan-15	03:00	99 1	23 8	14,21 8	11:00	9 9	26 1	11,43 6	19:00	95	24,2	9,056
20-Jan-15	03:10	99 0	23 3	14,11 3	11:10	9 0	25 1	13,20 1	19:10	96	24,2	10,41 1
20-Jan-15	03:20	99 0	23 7	13,21 7	11:20	9 9	26 0	14,47 3	19:20	96	24,1	11,13 5
20-Jan-15	03:30	98 0	23 9	13,32 9	11:30	9 3	26 4	15,31 4	19:30	96	24,0	13,60 8
20-Jan-15	03:40	98 0	23 3	13,08 3	11:40	9 0	26 4	15,70 4	19:40	97	24,0	14,57 3
20-Jan-15	03:50	98 0	23 6	13,48 6	11:50	9 7	25 6	15,30 6	19:50	97	23,9	15,17 15,89 3
20-Jan-15	04:00	98 1	23 5	13,80 5	12:00	8 9	26 1	14,95 2	20:00	98	23,8	17,05 3
20-Jan-15	04:10	98 0	23 1	14,28 1	12:10	7 1	26 1	14,42 1	20:10	98	23,7	18,51 8
20-Jan-15	04:20	98 0	23 4	14,44 4	12:20	8 0	26 2	14,25 2	20:20	98	23,7	19,67 19,89
20-Jan-15	04:30	98 1	23 2	14,61 2	12:30	8 6	26 1	13,17 1	20:30	99	23,8	19,80 4
20-Jan-15	04:40	97 2	23 6	14,90 6	12:40	8 8	26 1	11,72 1	20:40	99	23,7	20,09 19,78
20-Jan-15	04:50	98 1	23 5	14,63 5	12:50	9 9	25 1	11,39 1	20:50	99	23,6	21,18 18,98 3
20-Jan-15	05:00	98 1	23 8	14,08 8	13:00	9 0	25 2	11,18 2	21:00	99	23,6	22,69 18,98 3
20-Jan-15	05:10	98 1	23 6	13,80 6	13:10	8 8	25 4	10,59 4	21:10	99	23,6	23,6 18,98 3
20-Jan-15	05:20	97 1	23 9	14,23 9	13:20	8 7	25 7	9,906 7	21:20	99	23,5	24,28 18,98 3
20-Jan-15	05:30	98 2	23 3	14,30 3	13:30	8 8	25 5	9,854 5	21:30	99	23,6	25,80 19,55 5
20-Jan-15	05:40	97 2	23 9	13,31 9	13:40	8 5	25 9	9,906 9	21:40	99	23,6	26,14 19,55 5
20-Jan-15	05:50	98 1	23 3	12,75 3	13:50	8 8	25 5	8,888 5	21:50	99	23,6	27,69 20,86 7
20-Jan-15	06:00	99 1	23 2	11,88 2	14:00	8 8	25 3	8,48 3	22:00	99	23,6	28,69 22,69 9
20-Jan-15	06:10	99 1	23 8	12,22 8	14:10	9 4	25 4	8,64 4	22:10	99	23,6	29,2 24,28 2
20-Jan-15	06:20	10 0	23 2	12,17 2	14:20	9 2	25 3	8,342 3	22:20	99	23,5	29,80 25,80 3
20-Jan-15	06:30	10 0	23 9	12,34 9	14:30	9 1	25 4	7,334 4	22:30	99	23,5	30,3

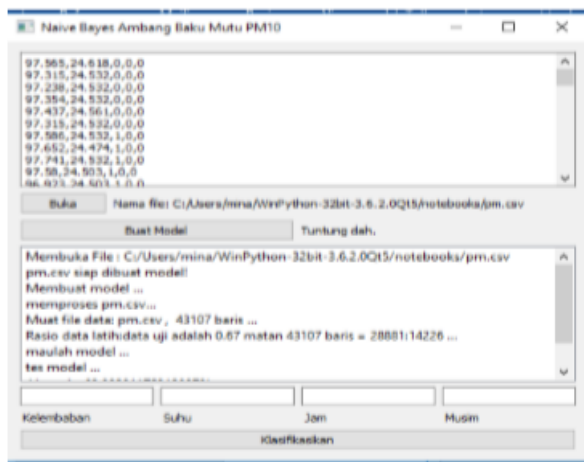
waktu dengan menggunakan aplikasi pengolah angka.dipreprocessing

Gambar 2.2. Sample Data PM10
 2. Mendesain antar muka (GUI) serta membuat program menggunakan software opensource Python dengan menerapkan prinsip kerja dari klasifikasi yang digunakan yaitu algoritma Naive Bayes.



Gambar 2.2. GUI dengan PyQt

3. Menguji aplikasi dengan menggunakan data yang telah dipreprocessing serta membuat model klasifikasi dengan menggunakan data PM10.



Gambar 2.3 Hasil Ujicoba

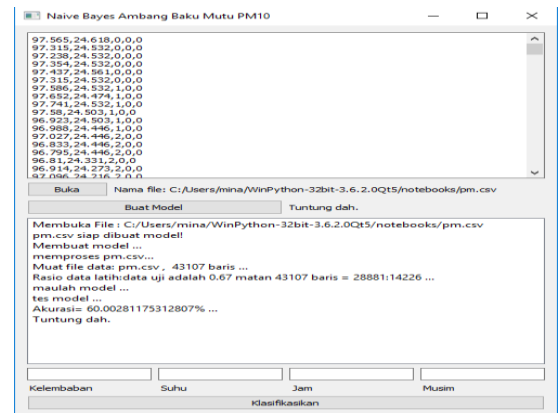
Tahapan penelitian yang telah dikerjakan dengan menerapkan algoritma Naive Bayes sebagai algoritma klasifikasi yang

menghasilkan model untuk klasifikasi tingkat bahaya sehingga dapat bermanfaat sebagai media informasi yang mudah dipahami.

3.HASIL DAN PEMBAHASAN

Pengelompokan yang dihasilkan dari penerapan metode klasifikasi naive bayes adalah model yang digunakan jika kondisi tidak adanya data PM 10 dikarenakan kendala teknis seperti pemadaman listrik atau karena data rusak maka digunakanlah model pengelompokan ini untuk menentukan tingkat bahaya berdasarkan klasifikasi menggunakan algoritma naive bayes dan menghasilkan model.

Tahapan hasil yang telah dikerjakan dapat dijelaskan dengan gambar dibawah ini



Gambar 2. Analisis menggunakan Python

Python PyQt

Dari gambar diatas dapat dilihat hasil analisis yang ditampilkan dengan menggunakan interface yang berbasis Python. Pertama dengan membuka file pada analisis ini menggunakan data PM10 dengan mengarah ke drive penyimpanan data PM10 yang memiliki atribut suhu, kelembaban, jam, musim serta tingkat bahaya. Setelah proses buka file maka lanjutan proses berikutnya adalah buat model dan menghasilkan akurasi klasifikasi tingkat bahaya yang berkisar dinilai 60 dari klasifikasi menggunakan algoritma klasifikasi naive bayes.

Analisis program yang telah dikerjakan dapat dilihat pada listing program dibawah ini

```

import bayeskan as bas
from PyQt5.QtGui import *
from PyQt5.QtCore import *
from PyQt5.QtWidgets import *
class MainForm(QWidget):
    def __init__(self):
        super().__init__()
        self.setupUi()
    def setupUi(self):
        self.resize(500, 450)
        self.move(300, 300)
        self.setWindowTitle('Naive Bayes Ambang
Baku Mutu PM10')
        self.textEdit = QTextEdit()
        self.prosesEdit = QListWidget()
        self.openButton = QPushButton('&Buka')
        self.fileLabel = QLabel('Nama file: ')
        hbox = QHBoxLayout()
        hbox.addWidget(self.openButton)

```

4.KESIMPULAN

Kesimpulan dari penerapan algoritma naive bayes untuk klasifikasi tingkat bahaya polutan di kota banjarbaru adalah :

1.Implementasi metode Naive Bayes untuk mengklasifikasi tingkat bahaya polutan pada suatu aplikasi berbasis desktop dengan menggunakan GUI (*Graphical User Interface*) Python mampu menjalankan data 43.107 dengan kecepatan memadai.

2.Menghasilkan model klasifikasi yang dapat digunakan ketika atribut PM 10 kosong yaitu dengan mengganti atribut yang selalu tersedia seperti kelembabab, suhu ,jam kejadian menggunakan Klasifikasi Naïve bayes.

3.Menggunakan Klasifikasi Naïve Bayes mampu menghasilkan akurasi 60% untuk mewakili tingkat bahaya saat tidak tersedianya data PM10.

4.Model Naïve Bayes bisa digunakan untuk mendapatkan peringatan tingkat bahaya.

Saran untuk pengembangan penelitian berikutnya adalah membandingkan dengan metode klasifikasi yang lain sehingga bisa dibandingkan kemampuan klasifikasi yang dihasilkan.

5.REFERENSI

[1]World Health Organization, "Public Health, Environmental and Social Determinants of Health (PHE)," *World Health Organization*, p. http://www.who.int/phe/air_quality_q&a, 2016.

[2]H. S. Huboyo and E. Sutrisno, "Analisis Konsentrasi Particulate Matter 10 (PM10) pada Udara Diluar Ruang (Studi Kasus : Stasiun Tawang - Semarang)," *TEKNIK*, vol. 30, no. 1, pp. 44-48, 2009.

[3]C. Liu, E. Fuertes, C. Flexeder, L. C. Hofbauer, D. Berdel, B. Hoffmann and e. al., "Associations between Ambient Air Pollution and Bone Turnover Markers in 10-year old children: Result from the GINplus and LISplus Studies," *International Journal of Hygiene and Environmental Health*, vol. 218, pp. 58-65, 2015.

[4]N. A. Abdullah, S. H. Shuhaimi, T. Y. Ying and A. H. Shapee, "The study of seasonal variation of PM10 concentration in Peninsular, Sabah and Sarawak," *Malaysian Meteorol. Dep.*, vol. 9, pp. 1-28, 2011.

[5]Asean Specialised Meteorological Centre, "Asean Specialised Meteorological Centre," *Asean Specialised Meteorological Centre*, pp. <http://asmc.asean.org/asmc-haze-air-quality/>, 2015.

[6]Bustami, "Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi," *TECHSI:Jurnal Penelitian Tekni Informatika*, pp. 128-146, 2013.

[7]T. SL, W. Ip and A. H. Tsang, "Is Naive Bayes a Good Classifier for Document Classification," *International Journal of Software Engineering and Its Applications*, vol. 5, no. 3, pp. 37-46, 2011.

[8]M. K. Stern, J. E. Beck and B. P. Woolf, "Naïve Bayes Classifiers for User Modeling," *Proceedings of the Conference on User Modeling.*, 1999.

[9]Peraturan Pemerintah RI No. 41, *Pengendalian Pencemaran Udara*, 1999.

