

**PENERAPAN *DATA MINING* UNTUK MENGETAHU
POLA KLASIFIKASI TINGKAT PENCEMARAN
UDARA DI KOTA JAKARTA BULAN JANUARI 2020
MENGUNAKAN ALGORITMA *NAÏVE BAYES***

SKRIPSI

Oleh:

JOVA AGUNG NUGRAHA

201710225188



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2021**

LEMBAR PERSETUJUAN PEMBIMBING

LEMBAR PERSETUJUAN PEMBIMBING

Judul Proposal Skripsi : Penerapan *Data Mining* Untuk Mengetahui Pola Klasifikasi Tingkat Pencemaran Udara Bulan Januari 2020 Di Kota Jakarta Menggunakan Algoritma *Naïve Bayes*.

Nama Mahasiswa : Jova Agung Nugraha

Nomor Pokok Mahasiswa : 201710225188

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 16 Juli 2021

Bekasi, 22 Juli 2021

MENYETUJUI,

Pembimbing I



Wowon Priatna, ST., MTi
NIDN. 0429118007

Pembimbing II



Tri Dharma Putra, ST., M.SC
NIDN. 0302117101

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul Skripsi : Penerapan *Data Mining* Untuk Mengetahui Pola Klasifikasi Tingkat Pencemaran Udara Bulan Januari 2020 Di Kota Jakarta Menggunakan Algoritma *Naïve Bayes*.

Nama Mahasiswa : Jova Agung Nugraha

Nomor Pokok Mahasiswa : 201710225188

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 16 Juli 2021

Bekasi, 22 Juli 2021

Mengesahkan,

Ketua Tim Penguji : Mukhlis, S.Kom., M.T
NIDN. 0312116802

Penguji (I) : Mugiarso, S.Kom., M.Kom
NIDN. 0420117403

Penguji (II) : Wowon Priatna, ST., M.Ti
NIDN. 0429118007

Mengetahui,

Ketua Program Studi

Informatika

Rakhmat Purnomo, S.Pd., S.Kom., M.Kom.

NIDN. 0322108201

Dekan

Fakultas Ilmu Komputer

Herlawati, S.Si., M.M., M.Kom.

NIDN. 0311097302

LEMBAR PERNYATAAN BUKAN PLAGIASI

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Jova Agung Nugraha
NPM : 201710225188
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Penerapan *Data Mining* Untuk Mengetahui Pola
Klasifikasi Tingkat Pencemaran Udara Bulan Januari
2020 Di Kota Jakarta Menggunakan Algoritma *Naïve
Bayes*

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 24 Juni 2021

Penulis



Jova Agung Nugraha

ABSTRAK

Jova Agung Nugraha, 201710225188. Penerapan Data Mining Untuk Mengetahui Pola Klasifikasi Pencemaran Udara Bulan Januari 2020 Di Kota Jakarta Menggunakan Algoritma Naïve Bayes.

Kualitas udara di beberapa kota Indonesia saat ini juga tidak terlau baik atau tergolong tidak sehat. Pada tahun 2020 Dinas Lingkungan Hidup juga melakukan pemantauan melalui beberapa titik stasiun yang ada di Jakarta dengan hasil pola perubahan tingkat udara setiap harinya. Berdasarkan hasil data pemantauan Dinas Lingkungan Hidup selama bulan Januari tahun 2020 melalui beberapa titik stasiun yang ada di kota Jakarta. Penelitian ini bertujuan untuk mengetahui tingkat pencemaran udara di kota Jakarta bulan Januari tahun 2020, dan mengetahui tingkat akurasi yang didapatkan dalam penerapan algoritma Naïve Bayes. Pada penelitian ini *Data Mining* digunakan untuk mengolah data dan Naïve Bayes sebagai algoritma yang bertujuan untuk mengklasifikasikan data yang telah diolah dalam proses *Data Mining*. Dari data Indeks Standar Pencemar Udara (ISPU) di kota Jakarta pada bulan Januari yang di uji dengan konsep klasifikasi menghasilkan akurasi sebesar **67%**. Dapat dikatakan bahwa algoritma *Naïve Bayes* bisa memberikan hasil yang baik dalam pemecahan untuk mengetahui kualitas udara di kota Jakarta selama bulan Januari tahun 2020.

Kata kunci: Pencemaran Udara, Klasifikasi, *Naïve Bayes*, *Data Mining*, ISPU.

ABSTRACT

Jova Agung Nugraha, 201710225188. *Application of Data Mining to determine the January 2020 Air Pollution Classification Pattern in the City of Jakarta Using the Naïve Bayes Algorithm*

Air quality in several Indonesian cities is currently not too good or classified as unhealthy. In 2020 the Environment Agency will also monitor through several station points in Jakarta with the results of a pattern of changes in air levels every day. Based on the results of monitoring data from the Environmental Service during January 2020 through several station points in the city of Jakarta. This study aims to determine the level of air pollution in the city of Jakarta in January 2020, and determine the level of accuracy obtained in the application of the Naïve Bayes algorithm. In this study Data Mining is used to process data and Naïve Bayes as an algorithm that aims to classify data that has been processed in the Data Mining process. From the Air Pollution Standard Index (ISPU) data in the city of Jakarta in January which was tested with the classification concept, it produced an accuracy of 67%. It can be said that the Naïve Bayes algorithm can give good results in solving the air quality in Jakarta during January 2020.

Keywords: Air Pollution, Classification, Nave Bayes, Data Mining, ISPU.

LEMBAR PERNYATAAN PUBLIKASI

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini:

Nama : Jova Agung Nugraha
NPM : 201710225188
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul:

Penerapan *Data Mining* Untuk Mengetahui Pola Klasifikasi Tingkat Pencemaran Udara Bulan Januari 2020 Di Kota Jakarta Menggunakan Algoritma *Naïve Bayes*.

Beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 24 Juni 2021
Yang Menyatakan


Jova Agung Nugraha

KATA PENGANTAR

Penulis menyadari bahwa dalam penulisan skripsi ini masih jauh dari kata sempurna. Karena itu, saran dan kritik yang bersifat membangun sangat penulis harapkan. Penulis berharap semoga penulisan skripsi ini dapat bermanfaat dan jadi penambah ilmu khususnya bagi penulis dan umumnya bagi pembaca.

Bekasi, 21 Juli 2021


Penulis



DAFTAR ISI

Halaman

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	3
1.3 Rumusan Masalah.....	3
1.4 Batasan Masalah.....	3
1.5 Tujuan Penelitian	3
1.6 Manfaat Peneliti	3
1.7 Waktu Penelitian	4
1.8 Sistematika Penulisan.....	4
BAB II LANDASAN TEORI	5
2.1 Data Mining	5

2.1.1	Fungsi Utama <i>Data Mining</i>	6
2.3	Tahapan <i>Data Mining</i>	7
2.3.1	<i>Data Selection</i>	8
2.3.2	Preprocessing atau Cleaning.....	8
2.3.3	Transformation	8
2.3.4	<i>Data Mining</i>	8
2.3.5	Interpretation atau Evaluation	8
2.4	Teknik-Teknik <i>Data Mining</i>	8
2.5	Klasifikasi	9
2.6	Algoritma <i>Naïve Bayes</i>	9
2.7	Pencemaran Udara	10
2.8	Google Colab	11
2.9	Alat Pengukur Kadar Dalam Polutan Udara	11
BAB III METODOLOGI PENELITIAN		12
3.1	Obyek Penelitian	12
3.2	Kerangka Penelitian	12
3.3	Metodologi Penelitian	14
BAB IV PERANCANGAN SISTEM DAN IMPLEMENTASI		15
4.1	Pemahaman Data.....	15
4.2	Modelling	15
4.3	Menghitung Class Probabilitas	16
4.4	Menghitung <i>Conditional</i> Probabilitas	16
4.4.1	Probabilitas Atribut Critical	18
4.5	Perhitungan Manual Data Testing.....	21

4.6	Pengujian Algoritma <i>Naive Bayes</i>	24
BAB V PENUTUP		29
5.1	Kesimpulan	29
5.2	Saran.....	29
DAFTAR PUSTAKA		30
LAMPIRAN.....		31



DAFTAR TABEL

	Halaman
Table 4. 1 Probabilitas Atribut Crit	17
Table 4. 2 Mean dan Std. Deviasi PM10	20
Table 4. 3 Mean dan Std. Deviasi SO2	20
Table 4. 4 Mean dan Std. Deviasi CO	20
Table 4. 5 Mean dan Std. Deviasi O3	21
Table 4. 6 Mean dan Std. Deviasi NO2	21
Table 4. 7 Sample Perhitungan	22



DAFTAR GAMBAR

	Halaman
Gambar 2. 1 Tahapan Data Mining	7
Gambar 2. 2 Papan Indeks Standar Pencemar Udara (ISPU)	11
Gambar3. 1 Kerangka Penelitian	12
Gambar 4. 1Library Yang Dibutuhkan	24
Gambar 4. 2Memanggil Dataset	24
Gambar 4. 3Mengubah String Menjadi Float	24
Gambar 4. 4Membagi Dataset	25
Gambar 4. 5Membagi Dataset Tergantung Class	25
Gambar 4. 6Function Nilai Mean dan Std. Deviasi	26
Gambar 4. 7Meringkas Dataset	26
Gambar 4. 8Meringkas Atribut Class	26
Gambar 4. 9Menghitung Nilai Probabilitas Class	27
Gambar 4. 10Memprediksi Menggunakan Nilai Input Vector	27
Gambar 4. 11Hasil dan Akurasi	28

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Plagiarisme	1
Lampiran 2 Biodata Mahasiswa	2
Lampiran 3 Kartu Bimbingan Skripsi 1	3
Lampiran 4 Kartu Bimbingan Skripsi 2	4
Lampiran 5 Dataset.....	4

