

**IMPLEMENTASI K-NEAREST NEIGHBORS UNTUK  
MEMPREDIKSI DELAY PENERBANGAN**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**MUHAMMAD ALDORA FAUSTA**

**19.11.2859**

Kepada

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS AMIKOM YOGYAKARTA**

**YOGYAKARTA**

**2023**

**IMPLEMENTASI K-NEAREST NEIGHBORS UNTUK  
MEMPREDIKSI DELAY PENERBANGAN**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**MUHAMMAD ALDORA FAUSTA**

**19.11.2859**

Kepada

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS AMIKOM YOGYAKARTA**

**YOGYAKARTA**

**2023**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**IMPLEMENTASI K-NEAREST NEIGHBORS UNTUK MEMPREDIKSI  
DELAY PENERBANGAN**

yang disusun dan diajukan oleh

**MUHAMMAD ALDORA FAUSTA**

**19.11.2859**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 16 Agustus 2023

**Dosen Pembimbing,**

**Kusrini, Prof., Dr., M.Kom  
NIK. 190301206**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**Implementasi K-Nearest Neighbors untuk Memprediksi Delay Penerbangan**

yang disusun dan diajukan oleh

**MUHAMMAD ALDORA FAUSTA**

**19.11.2589**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 18 oktober 2023

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Kusrini, Prof., Dr., M.Kom**  
**NIK. 190302106**

**Haryoko, S.Kom., M.Cs**  
**NIK. 190302286**

**Robert Marco, M.T.**  
**NIK. 190302228**

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 18 oktober 2023

**DEKAN FAKULTAS ILMU KOMPUTER**



**Hanif Al Fatta, S.Kom., M.Kom.**  
**NIK. 190302096**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : MUHAMMAD ALDORA FAUSTA  
NIM : 19.11.2859

Menyatakan bahwa Skripsi dengan judul berikut:

**Implementasi K-Nearest Neighbors untuk Memprediksi Delay Penerbangan**

Dosen Pembimbing : Kusrini, Prof., Dr., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 18 oktober 2023

Yang Menyatakan,



MUHAMMAD ALDORA FAUSTA

## HALAMAN PERSEMBAHAN

Dengan rasa syukur yang mendalam saya haturkan, penyelesaian skripsi ini saya persembahkan kepada :

1. Kepada orang tua dan keluarga terdekat saya yang senantiasa membantu saya selama proses pengerjaan skripsi ini.
2. Kepada dosen pembimbing saya yang sudah dengan tegas dan ketat selalu mengawasi skripsi saya ini, sehingga saya bisa menyelesaikan skripsi dengan tepat waktu.
3. Kepada teman-teman sekelas saya yang sudah menjadi penyemangat saya agar bisa menyelesaikan skripsi saya ini, baik secara langsung maupun tidak langsung

## KATA PENGANTAR

Dengan memanjatkan puja dan puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan judul **“Implementasi K-Nearest Neighbors untuk Memprediksi Delay Penerbangan”**, sebagai salah satu syarat untuk menyelesaikan studi Program Sarjana (S1) Jurusan Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Penulis menyadari bahwa skripsi ini tidak mungkin terselesaikan tanpa adanya dukungan dan bantuan dari berbagai pihak selama penyusunan skripsi ini. pada kesempatan kali ini izinkan penulis untuk menyampaikan rasa terima kasih yang setulus-tulusnya kepada :

1. Bapak Hanif Al Fatta, S.Kom., M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
2. Ibu Kusrini, Prof., Dr., M.Kom, selaku dosen pembimbing skripsi atas segala bimbingan serta saran yang diberikan kepada penulis sehingga skripsi ini dapat diselesaikan dengan baik
3. Bapak Asro Nasiri, Drs, M.Kom, selaku dosen wali penulis selama penulis berkuliah di Universitas Amikom Yogyakarta.
4. Seluruh staf pengajar Fakultas Ilmu Komputer Universitas Amikom Yogyakarta yang telah memberikan ilmu pengetahuan yang tak ternilai selama penulis menempuh pendidikan di Universitas Amikom Yogyakarta.
5. Kedua orang tua penulis, Akmal Siregar dan Yuliani Azwar, yang selalu memberikan kasih sayang, doa, nasehat, serta kesabarannya dalam setiap langkah hidup penulis, penulis berterima kasih dan bersyukur dengan apa yang telah penulis dapatkan selama ini.

Yogyakarta, 24 Agustus 2023

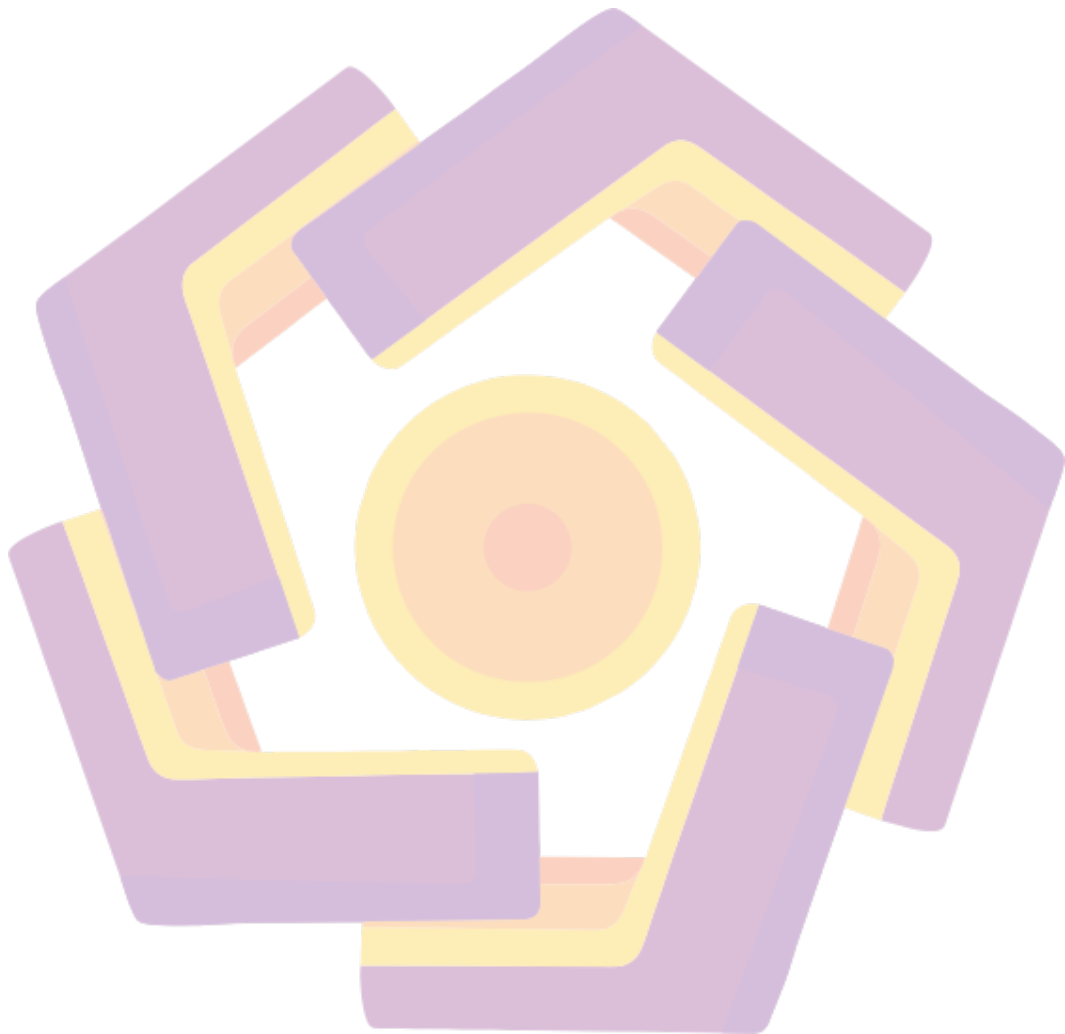
Penulis, MUHAMMAD ALDORA FAUSTA

## DAFTAR ISI

HALAMAN PERSETUJUAN	I
HALAMAN PENGESAHAN	II
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	III
HALAMAN PERSEMBAHAN	IV
KATA PENGANTAR	V
DAFTAR ISI	VI
DAFTAR GAMBAR	VIII
DAFTAR TABEL	IX
DAFTAR RUMUS	XI
DAFTAR LAMPIRAN	XII
DAFTAR LAMBANG DAN SINGKATAN	XIII
DAFTAR ISTILAH	XIV
INTISARI	XV
ABSTRACT	XVI
BAB I	
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Metode Penelitian	3
1.7 Sistematika Penulisan	4
BAB II	
TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori	9
BAB III	
METODE PENELITIAN	29
3.1 Objek Penelitian	29
3.2 Alur Penelitian	29
3.3 Alat dan Bahan	31
BAB IV	
HASIL DAN PEMBAHASAN	42
4.1 Perancangan Model	42
4.2 Hasil Eksperimen	60
4.3 Pengujian Aplikasi	61



BAB V	
PENUTUP	65
5.1 Kesimpulan	65
5.2 Saran	65
REFERENSI	66
LAMPIRAN	68



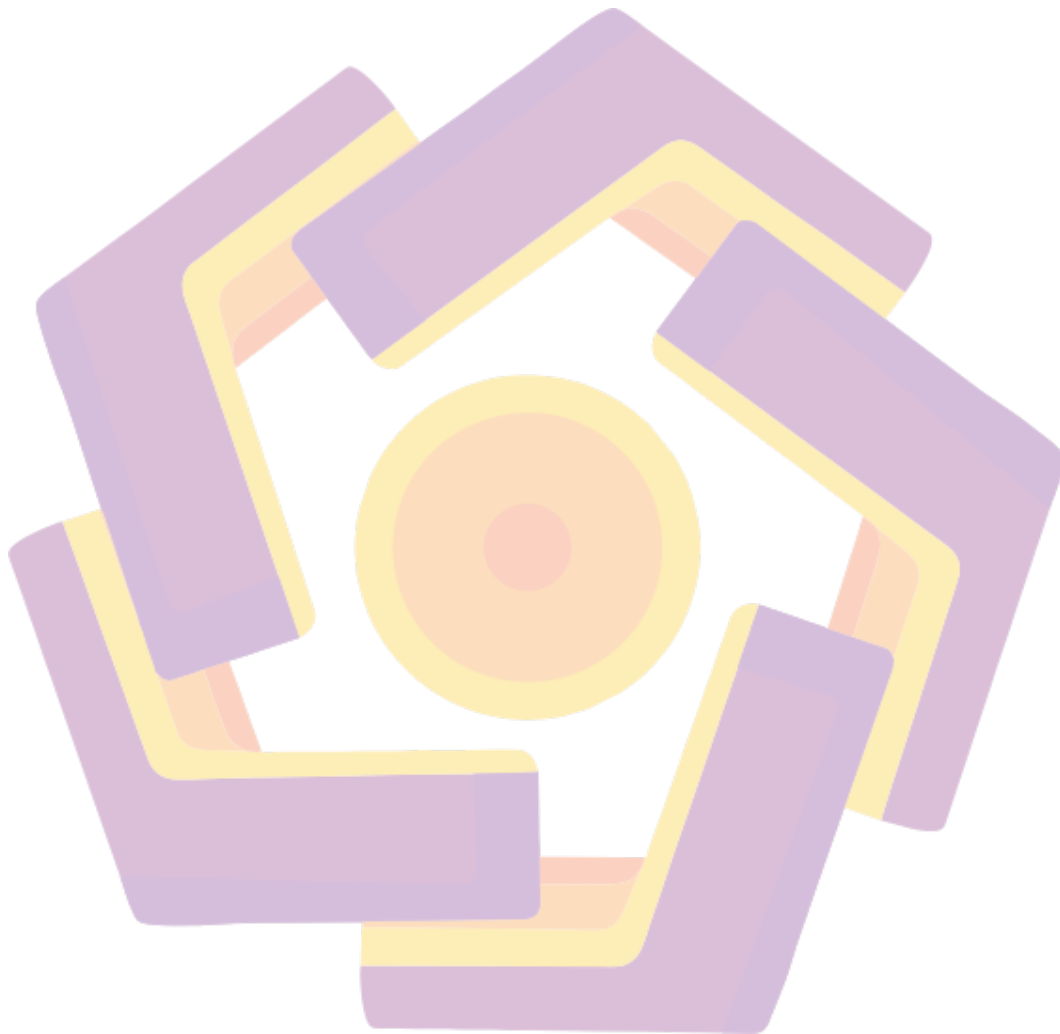
## DAFTAR GAMBAR

Gambar 2.1. Proses Pengujian K-Fold Cross Validation	21
Gambar 3.1. Flowchart Alur Penelitian	28
Gambar 4.1. Akuisisi Data	42
Gambar 4.2. Perubahan Type Data	42
Gambar 4.3. Perubahan Atribut Data Menjadi Skala Likert	43
Gambar 4.4. Hasil Perubahan Atribut Data	44
Gambar 4.5. Pemisahan Data Label dan Data Latih	44
Gambar 4.6. Hasil Pemisahan Data Label dan Data Latih	45
Gambar 4.7. Pemisahan Data Training dan Data Testing, Serta Implementasi Algoritma K-Neighbors dan Perhitungan Performa Model	45
Gambar 4.8. Pemilihan Fitur Menggunakan Metode Chi Square	47
Gambar 4.9. Implementasi K-Means Pada Kolom Airline	48
Gambar 4.10. Perhitungan Metode Elbow Pada Kolom Airline	49
Gambar 4.11. Hasil Implementasi Algoritma K-Means Pada Kolom Airline	50
Gambar 4.12. Proses Normalisasi Data Latih	56
Gambar 4.13. Proses Pemisahan Data Training dan Data Testing	57
Gambar 4.14. Implementasi Model Pengembangan	59
Gambar 4.15. Antarmuka Aplikasi Web	60
Gambar 4.16. Data Latih Pada Pengujian Aplikasi	61
Gambar 4.17. Label Data Pada Pengujian Aplikasi	61

## DAFTAR TABEL

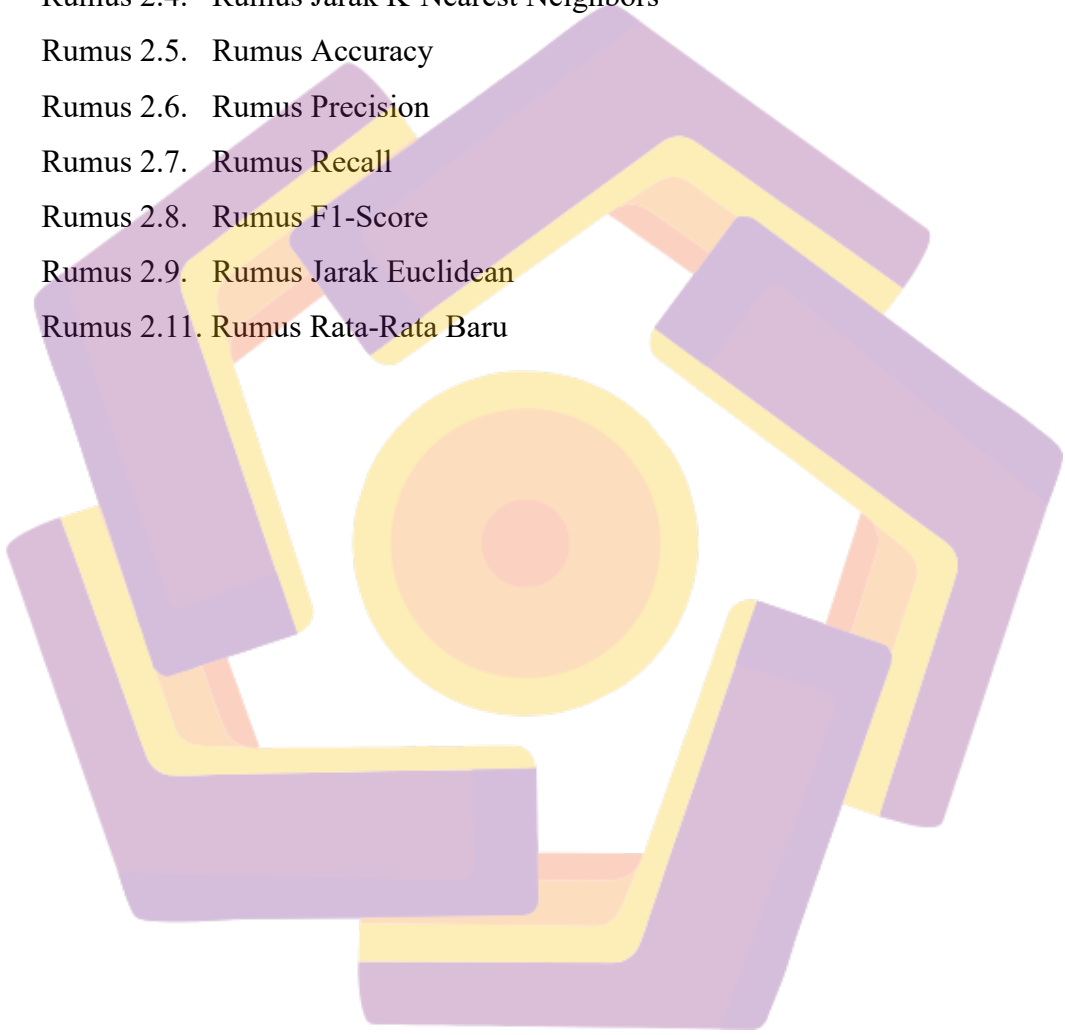
Tabel 2.1.	Profil obyek Penelitian	9
Tabel 2.2.	Dokumentasi Penelitian	12
Tabel 2.3.	Keterlambatan Siswa Pada Kurun Waktu 6 Bulan	13
Tabel 2.4.	Tabel Pembantu Perhitungan Chi Square Hitung	13
Tabel 2.5.	Nilai Mahasiswa Pada Satu Kelas	15
Tabel 2.6.	Tabel Pembantu Perhitungan Data Scaling	15
Tabel 2.7.	Berat Anak Laki-Laki Berdasarkan Indeks Bmi	17
Tabel 2.8.	Tabel Bantu Hitung Jarak K-Nearest Neighbors	18
Tabel 2.9.	Tabel Penentuan Kelas Baru	18
Tabel 2.10.	Tabel Confusion Matrix	19
Tabel 2.11.	Hasil Klasifikasi Model Machine Learning	21
Tabel 2.12.	Data Berat Anak Perempuan SMA	24
Tabel 2.13.	Tabel Pembantu Perhitungan Cluster	25
Tabel 3.1.	Tabel Histogram Kolom Airline	30
Tabel 3.2.	Tabel Histogram Kolom Flight	32
Tabel 3.3.	Tabel Histogram Kolom AirportFrom	33
Tabel 3.4.	Tabel Histogram Kolom AirportTo	35
Tabel 3.5.	Tabel Histogram Kolom DayOfWeek	36
Tabel 3.6.	Tabel Histogram Kolom Time	36
Tabel 3.7.	Tabel Histogram Kolom Length	37
Tabel 4.1.	Metode Elbow Pada Kolom Flight	51
Tabel 4.2.	Tabel Histogram Hasil Klasifikasi K-Means Pada Kolom Flight	51
Tabel 4.3.	Metode Elbow Pada Kolom AirportTo	52
Tabel 4.4.	Hasil Implementasi Algoritma K-Means Pada Kolom AirportTo	53
Tabel 4.5.	Metode Elbow Pada Kolom Time	54
Tabel 4.6.	Hasil Implementasi K-Means Pada Kolom Time	54
Tabel 4.7.	Metode Elbow Pada Kolom Length	55
Tabel 4.8.	Hasil Implementasi K-Means Pada Kolom Length	56
Tabel 4.9.	Visualisasi Perkembangan Nilai K	58

Tabel 4.10. Perbandingan Performa Model Awal dan Penembangan	59
Tabel 4.11. Hasil Prediksi Aplikasi	61
Tabel 4.12. Confusion Matrix Hasil Pengujian	62



## DAFTAR RUMUS

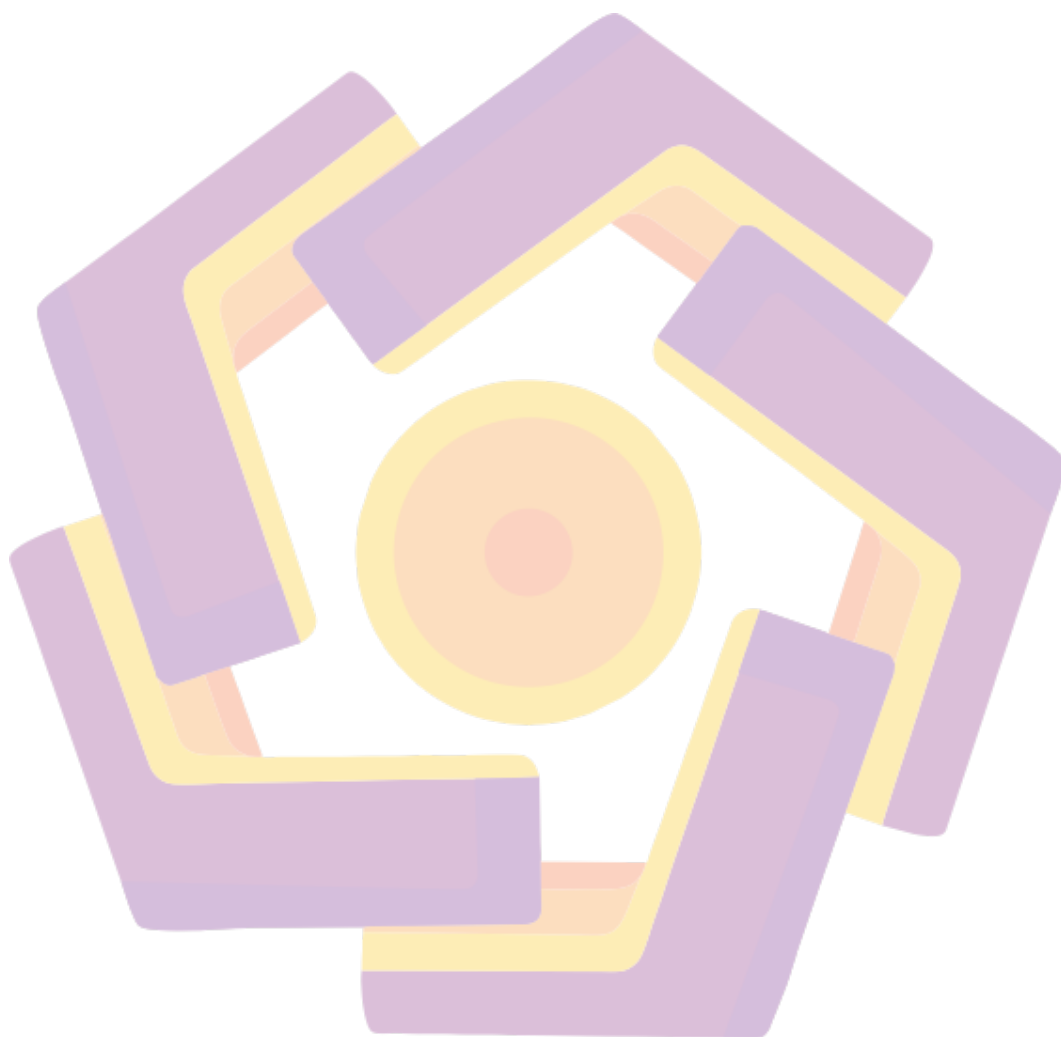
Rumus 2.1. Rumus Chi Square	11
Rumus 2.2. Rumus Derajat Kebebasan	11
Rumus 2.3. Rumus Min-Max Scaling	14
Rumus 2.4. Rumus Jarak K-Nearest Neighbors	16
Rumus 2.5. Rumus Accuracy	19
Rumus 2.6. Rumus Precision	20
Rumus 2.7. Rumus Recall	20
Rumus 2.8. Rumus F1-Score	20
Rumus 2.9. Rumus Jarak Euclidean	23
Rumus 2.11. Rumus Rata-Rata Baru	23



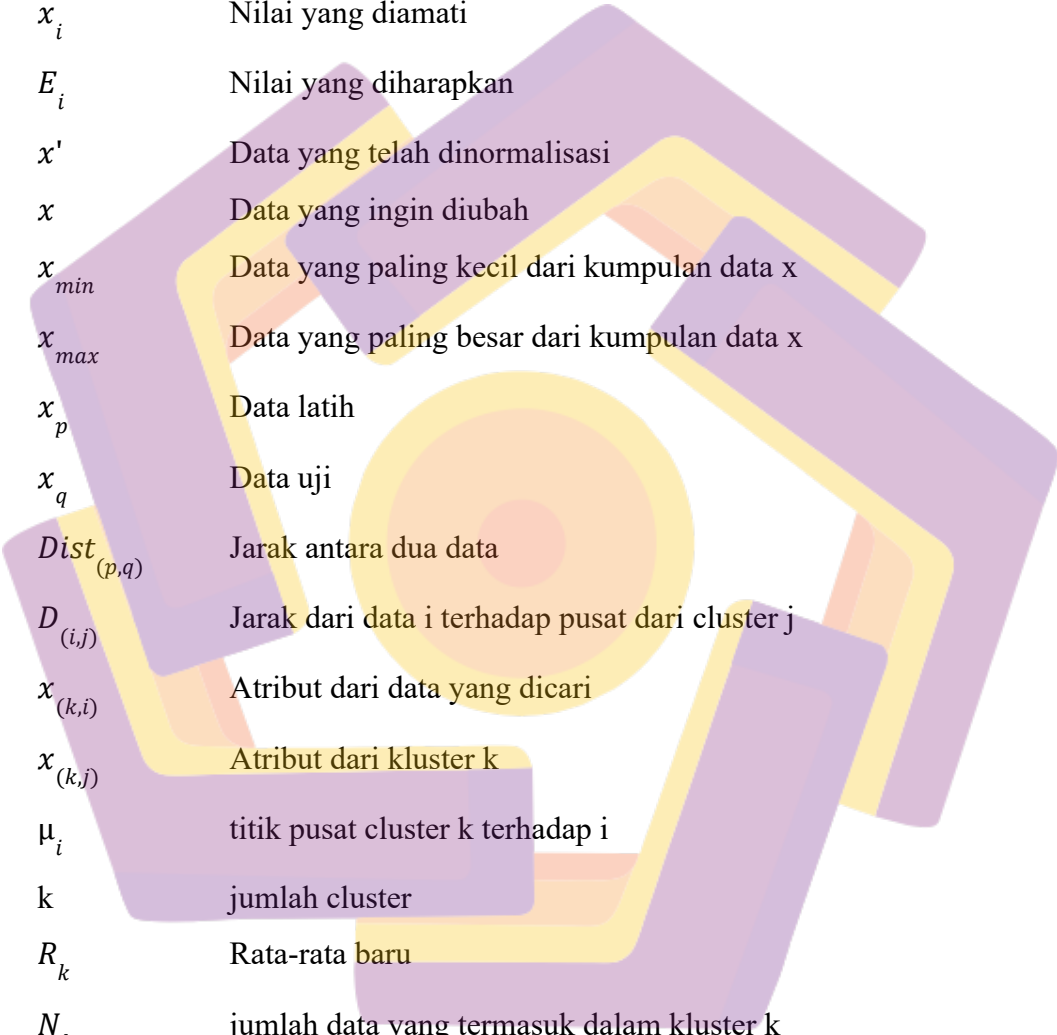
## DAFTAR LAMPIRAN

Lampiran 1. Laman Website Sumber Data Penelitian

68

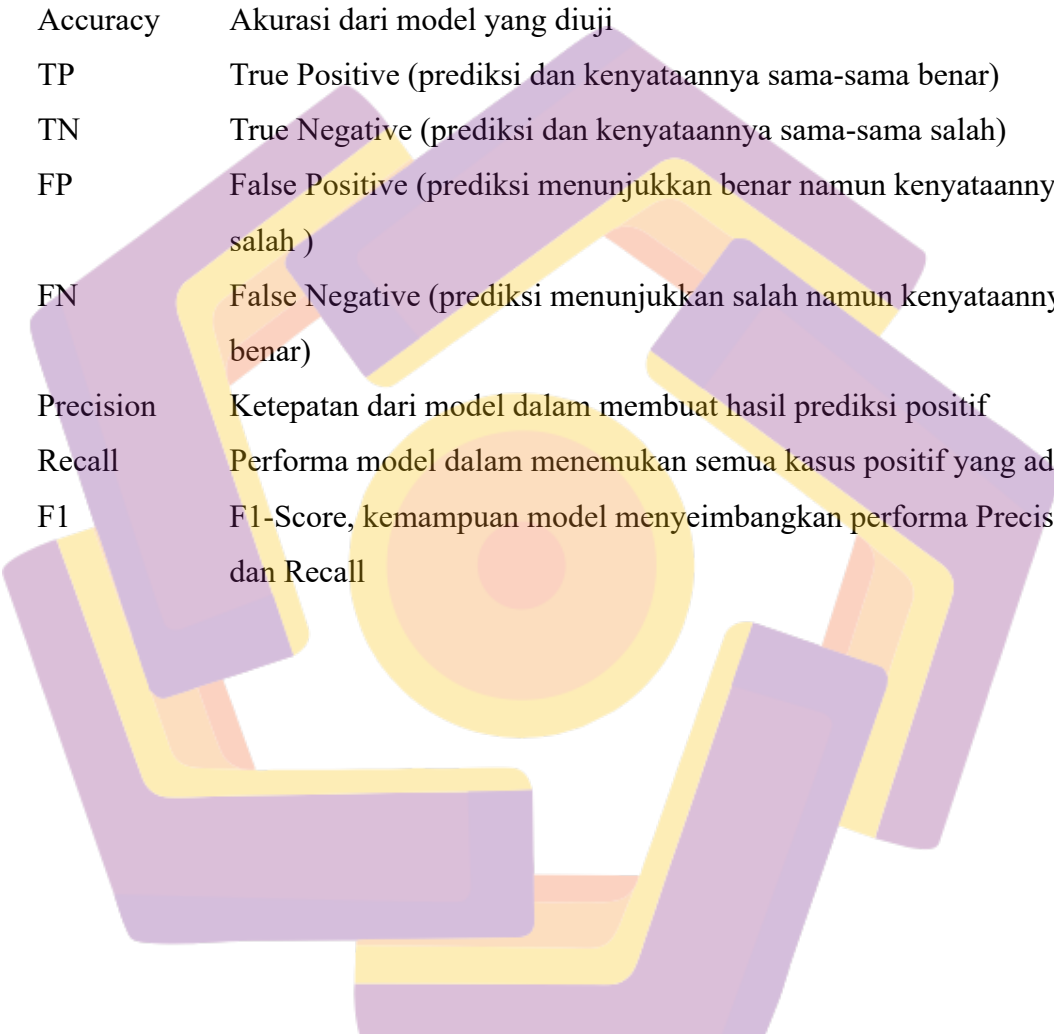


## DAFTAR LAMBANG DAN SINGKATAN



$\chi^2$	Chi Square
$\Sigma$	Jumlah Dari
$n$	Jumlah Sampel
$x_i$	Nilai yang diamati
$E_i$	Nilai yang diharapkan
$x'$	Data yang telah dinormalisasi
$x$	Data yang ingin diubah
$x_{min}$	Data yang paling kecil dari kumpulan data x
$x_{max}$	Data yang paling besar dari kumpulan data x
$x_p$	Data latih
$x_q$	Data uji
$Dist_{(p,q)}$	Jarak antara dua data
$D_{(i,j)}$	Jarak dari data i terhadap pusat dari cluster j
$x_{(k,i)}$	Atribut dari data yang dicari
$x_{(k,j)}$	Atribut dari kluster k
$\mu_i$	titik pusat cluster k terhadap i
$k$	jumlah cluster
$R_k$	Rata-rata baru
$N_k$	jumlah data yang termasuk dalam kluster k
$x_{nk}$	pola ke n dalam kluster k

## DAFTAR ISTILAH



<i>Df</i>	Derajat kebebasan (degree of freedom)
WCSS	Within-Circle sum of square (jumlah jarak data dalam suatu kluster terhadap titik pusat cluster)
Accuracy	Akurasi dari model yang diuji
TP	True Positive (prediksi dan kenyataannya sama-sama benar)
TN	True Negative (prediksi dan kenyataannya sama-sama salah)
FP	False Positive (prediksi menunjukkan benar namun kenyataannya salah )
FN	False Negative (prediksi menunjukkan salah namun kenyataannya benar)
Precision	Ketepatan dari model dalam membuat hasil prediksi positif
Recall	Performa model dalam menemukan semua kasus positif yang ada
F1	F1-Score, kemampuan model menyeimbangkan performa Precision dan Recall



## INTISARI

Delay sudah menjadi masalah yang lumrah terjadi dalam jadwal penerbangan pesawat terbang. Delaynya sebuah pesawat terbang dapat merugikan berbagai pihak, baik dari pihak pengguna maskapai penerbangan yang dirugikan waktunya, serta pihak maskapai yang harus mengganti rugi kepada pengguna maskapai penerbangan. hal ini yang menjadi latar belakang dari penelitian skripsi ini, dengan tujuan untuk memprediksi apakah akan terjadi delay pesawat pada satu jadwal penerbangan pesawat terbang.

Data yang akan digunakan dalam prediksi kejadian delay ini adalah data penerbangan yang tercatat dalam berbagai airport di amerika serikat. Data ini digunakan sebagai purwarupa dari apa yang sekiranya akan diperlukan dalam implementasi prediksi delay penerbangan pesawat terbang di indonesia. Jenis data yang digunakan pada penelitian skripsi ini seluruhnya berjenis kuantitatif. algoritma yang akan digunakan dalam prediksi delay pesawat terbang ini adalah algoritma k-nearest neighbors.

hasil penelitian ini menunjukkan bahwa algoritma k-nearest neighbors belum cukup baik dalam memprediksi delay pada penerbangan pesawat terbang, hal ini dibuktikan dengan performa prediksi algoritma k-nearest neighbors yang tetap berkutat pada 65%. Dalam hal ini peneliti menyarankan kepada peneliti selanjutnya untuk menggunakan algoritma yang berbeda, menambahkan data yang akan digunakan seperti data cuaca, serta melakukan klasifikasi ulang pada data yang digunakan pada penelitian skripsi ini.

**Kata kunci:** Delay, K-Nearest Neighbors, Kuantitatif.

## ABSTRACT

The delay has become a common problem in airplane flight schedules. The delay of an airplane can be detrimental to various parties, both from the airline users who are losing time, as well as the airline that must compensate the airline users. This has become the background of this thesis research, with the aim of predicting whether there will be an airplane delay on an airplane flight schedule.

The data that will be used in the prediction of delay events is flight data that is recorded in various airports in the United States. This data is used as a prototype of what will be needed for the implementation of airplane flight delay prediction in Indonesia. The type of data used in this thesis research is entirely quantitative. The algorithm that will be used in the prediction of airplane delay is the k-nearest neighbors algorithm.

The results of this study indicate that the k-nearest neighbors algorithm is not quite good enough in predicting delays in airplane flights, this is evidenced by the prediction performance of the k-nearest neighbors algorithm which remains at 65%. In this case, the researcher suggests to future researchers to use different algorithms, add data to be used such as weather data, and reclassify the data used in this thesis research.

**Keyword:** Delay, K-Nearest Neighbors, Quantitative