



ILKOM JURNAL ILMIAH

TERAKREDITASI PERINGKAT 4 SK. NO. 10/E/KPT/2019



DITERBITKAN OLEH:

TEKNIK INFORMATIKA - FAKULTAS ILMU KOMPUTER - UNIVERSITAS MUSLIM INDONESIA

DOI: <https://doi.org/10.33096/ilkom> | email: jurnal.ilkom@umi.ac.id



p-ISSN 2087-1716
e-ISSN 2548-7779

No. 200/M/KPT/2020

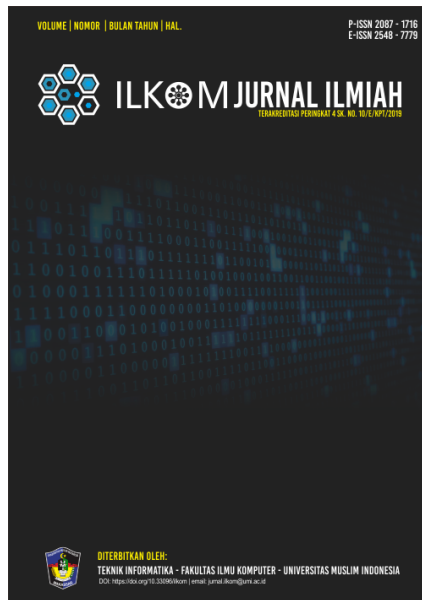
Alamat Redaksi: Jl. Urip Sumoharjo KM. 05 Makassar

HOME ABOUT USER HOME SEARCH CURRENT ARCHIVES ANNOUNCEMENTS PUBLICATION FEES

Home > Vol 15, No 2 (2023)

ILKOM Jurnal Ilmiah

Journal title	ILKOM Jurnal Ilmiah
Initials	ILKOM
Abbreviation	ilk. J. Ilm
Frequency	3 issues per year
DOI	prefix 10.33096
Online ISSN	2548-7779
Editor-in-chief	Yulita Salim
Managing Editor	Ramdan Satra
Publisher	Teknik Informatika Fakultas Ilmu Komputer Universitas Muslim Indonesia
Citation Analysis	Google Scholar / Dimenssion
Frequency published	ILKOM Jurnal Ilmiah is issued three times a year in April, August, and December.



Focus and Scope, ILKOM Jurnal Ilmiah encompasses all aspects of the latest outstanding research and developments in the field of Computer science including:

- Artificial intelligence,
- Data science,
- Databases,
- Computer performance analysis,
- Computer security and cryptography,
- Computer networks,
- Parallel and distributed systems,
- Microcontroller,
- Internet of Things,
- Software engineering.

Before Submission,

You have to make sure that your paper is prepared using the **paper TEMPLATE**, **has been carefully proofread and polished**, and **conformed to the author guidelines**. The Journal invites original articles and not simultaneously submitted to another journal or conference. Make sure you have read and understand the ethics of publication for authors.

Online Submissions

- Already have a Username/Password for ILKOM Jurnal Ilmiah? **GO TO LOGIN**
- Need a Username/Password? **GO TO REGISTRATION**

Registration and login are required to submit items online and to check the status of current submissions.

Announcements

The 5th 2023 East Indonesia Conference on Computer and Information Technology

The 2023 5th EAST INDONESIA CONFERENCE ON COMPUTER AND INFORMATION TECHNOLOGY (EIConCIT) is an international scientific conference for research in the field of advanced computational science and technology, to be held during September, 2023, online and onsite from Bali, Indonesia. The conference is aimed at keeping abreast of the current development and innovation in the advanced of research area on Engineering, Computing and Applied Technology as well as providing an engaging forum for participants to share knowledge and expertise in related issues.

Citedness in Scopus



USER

You are logged in as...

dadang

- » My Journals
- » My Profile
- » Log Out

QUICK MENU

Editorial Team

Peer Reviewer

Peer Review Process

Focus and Scope

Publication Ethics

Acceptance Rate

Author Guidelines

Open Access Policy

Screening for Plagiarism

Online Submission

Author Declaration Form

Journal Workflow

Author Fee

Contact

JOIN WHATSAPP GROUP



TOOLS

We invite you to The 2023 5th EAST INDONESIA CONFERENCE ON COMPUTER AND INFORMATION TECHNOLOGY (EIconCIT). The theme is "Green Technology Innovation in ICT Sustainability for Developing Tourism and Cultural Heritage"

Posted: 2023-07-30

More...

Sinta 2 accreditation

ILKOM has been awarded Sinta 2 accreditation, read more

Posted: 2021-01-28

More Announcements...

Vol 15, No 2 (2023)

The author's countries are coming from Indonesia, Malaysia, India, UAE, Polandia

Table of Contents

Artikel

Diabetes Mellitus Early Detection Simulation using The K-Nearest Neighbors Algorithm with Cloud-Based Runtime (COLAB)

PDF
215-221

Mohamad Jamil^(1*); Budi Warsito⁽²⁾; Adi Wibowo⁽³⁾; Kiswanto Kiswanto⁽⁴⁾

- (1) Universitas Khairun
- (2) Universitas Diponegoro
- (3) Universitas Diponegoro
- (4) Universitas Diponegoro
- (*) Corresponding Author

10.33096/ilkom.v15i2.1510.215-221 Abstract view: 77 times

Building The Prediction of Sales Evaluation on Exponential Smoothing using The OutSystems Platform

PDF
222-228

Sasa Ani Arnomo^(1*); Yulia Yulia⁽²⁾; Ukas Ukas⁽³⁾

- (1) Universitas Putera Batam
- (2) STIE Nagoya Indonesia
- (3) Universitas Putera Batam
- (*) Corresponding Author

10.33096/ilkom.v15i2.1529.222-228 Abstract view: 57 times

Application of General Regression Neural Network Algorithm in Data Mining for Predicting Glass Sales and Inventory Quantity

PDF
229-239

Suryani Suryani^(1*); Indo Intan⁽²⁾; Farhan Mochtar Yunus⁽³⁾; Adammas Haris⁽⁴⁾; Faizal Faizal⁽⁵⁾; Nurdiansah Nurdiansah⁽⁶⁾

- (1) Univeristas Dipa Makassar
- (2) Univeristas Dipa Makassar
- (3) Univeristas Dipa Makassar
- (4) Univeristas Dipa Makassar
- (5) Univeristas Dipa Makassar
- (6) Univeristas Dipa Makassar
- (*) Corresponding Author

10.33096/ilkom.v15i2.1562.229-239 Abstract view: 44 times

The Effect of The Prediction of The K-Nearest Neighbor Algorithm on Surviving COVID-19 Patients in Indonesia

PDF
240-249

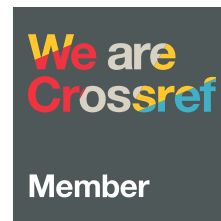
Aris Martono⁽¹⁾; Henderi Henderi^(2*); Giandari Maulani⁽³⁾

- (1) Universitas Raharja
- (2) Universitas Raharja
- (3) Universitas Raharja
- (*) Corresponding Author



TUTORIAL MENDELEY

CURRENT MEMBER



SPONSORSHIP



ISSN NUMBER



TEMPLATE



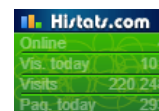
History Template

- » Template 2016
- » Template 2017-2019
- » Template 2020
- » Template 2021 - now

DOWNLOAD AWARD



VISITOR



View Visitor

Visitors

	266,361		420
	8,393		414
	3,246		375
	1,071		271
	568		256



10.33096/ilkom.v15i2.1234.240-249 Abstract view: 121 times

Evaluating The Application of Library Information System Technology using The PIECES Method in Remote Areas

PDF
250-261

KEYWORDS

Android Augmented Reality
Classification Clustering
Data Mining Decision Support
System Deep Learning GLCM
Information System K-Means K-
NN K-Nearest Neighbor k-
Nearest Neighbors KNN Machine
Learning Naive Bayes Naive
Bayes SAW SVM Sentiment
Analysis Support Vector Machine

Anton Yudhana^(1*); Herman Herman⁽²⁾; Suwanti Suwanti⁽³⁾; Muhammad Kunta Biddinika⁽⁴⁾

- (1) Universitas Ahmad Dahlan
- (2) Universitas Ahmad Dahlan
- (3) Universitas Ahmad Dahlan
- (4) Universitas Ahmad Dahlan
- (*) Corresponding Author

10.33096/ilkom.v15i2.1539.250-261 Abstract view: 44 times

Naïve Bayes and K-Nearest Neighbor Algorithm Approach in Data Mining Classification of Drugs Addictive Diseases

PDF
262-270

FONT SIZE

Dadang Priyanto^(1*); Ahmad Robbiul Iman⁽²⁾; Deny Jollyta⁽³⁾

- (1) Universitas Bumigora
- (2) Universitas Bumigora
- (3) Institut Bisnis Dan Teknologi Pelita Indonesia
- (*) Corresponding Author

10.33096/ilkom.v15i2.1544.262-270 Abstract view: 38 times

Rectified Linear Units and Adaptive Moment Estimation Optimizer on ANN with Saved Model Prediction to Improve The Stock Price Prediction Framework Performance

PDF
271-282

CURRENT ISSUE

ATOM	1.0
RSS	2.0
RSS	1.0

Sekhudin Sekhudin⁽¹⁾; Yuli Purwati^(2*); Fandy Setyo Utomo⁽³⁾; Mohd Sanusi Azmi⁽⁴⁾; Pungkas Subarkah⁽⁵⁾

- (1) Universitas AMIKOM Purwokerto
- (2) Universitas AMIKOM Purwokerto
- (3) Universitas AMIKOM Purwokerto
- (4) Universiti Teknikal Malaysia Melaka
- (5) Universitas AMIKOM Purwokerto
- (*) Corresponding Author

10.33096/ilkom.v15i2.1586.271-282 Abstract view: 55 times

Fuzzy Logic Algorithm of Sugeno Method for Controlling Line Follower Mobile Robot

PDF
283-289

Bayu Aji⁽¹⁾; Sutikno Sutikno^(2*)

- (1) Universitas Diponegoro
- (2) Universitas Diponegoro
- (*) Corresponding Author

10.33096/ilkom.v15i2.1558.283-289 Abstract view: 13 times

Sentiment Analysis for Online Learning using the Lexicon-Based Method and the Support Vector Machine Algorithm

PDF
290-302

M. Khairul Anam^(1*); Triyani Arita Fitri⁽²⁾; Agustin Agustin⁽³⁾; Lusiana Lusiana⁽⁴⁾; Muhammad Bambang Firdaus⁽⁵⁾; Agus Tri Nurhuda⁽⁶⁾

- (1) STMIK Amik Riau
- (2) STMIK Amik Riau
- (3) STMIK Amik Riau
- (4) STMIK Amik Riau
- (5) Universitas Mulawarman
- (6) STMIK Amik Riau
- (*) Corresponding Author

10.33096/ilkom.v15i2.1590.290-302 Abstract view: 70 times

Design and Build of IoT Based Flood Prone Monitoring System at Semani's Pump House Drainage System

PDF
303-316

'Aisyah 'Aisyah⁽¹⁾; Aji Ery Burhandenny^(2*); Happy Nugroho⁽³⁾; Didit Suprihanto⁽⁴⁾


(1) Mulawarman University

(1) Mulawarman University
 (2) Mulawarman University
 (3) Mulawarman University
 (4) Mulawarman University
 (*) Corresponding Author

 10.33096/ilkom.v15i2.1581.303-316  Abstract view: 78 times

Combination of YOLOv3 Algorithm and Blob Detection Technique in Calculating Nile Tilapia Seeds

PDF
317-325

 Diana Tri Susetianingtias⁽¹⁾; Eka Patriya^(2*); Rini Arianty⁽³⁾

(1) Universitas Gunadarma
 (2) Universitas Gunadarma
 (3) Universitas Gunadarma
 (*) Corresponding Author

 10.33096/ilkom.v15i2.1634.317-325  Abstract view: 25 times

Combination of The MADM Model Yager and k-NN to Group Single Tuition Payments

PDF
326-334

 Alders Paliling⁽¹⁾; Muh Nurtanzis Sutoyo^(2*)

(1) Universitas Sembilabelas November
 (2) Universitas Sembilabelas November
 (*) Corresponding Author

 10.33096/ilkom.v15i2.1349.326-334  Abstract view: 43 times

Development ETL (Extract, Transform and Load) Module in Indonesian Agricultural Commodities OLAP System

PDF
335-343

 Aditia Yudhistira^(1*); Imas Sukaesih Sitanggang⁽²⁾; Hari Agung Adrianto⁽³⁾

(1) Institut Pertanian Bogor
 (2) Institut Pertanian Bogor
 (3) Institut Pertanian Bogor
 (*) Corresponding Author

 10.33096/ilkom.v15i2.1758.335-343  Abstract view: 51 times

A Soft Voting Ensemble Classifier to Improve Survival Rate Predictions of Cardiovascular Heart Failure Patients

PDF
344-352

 Arif Munandar⁽¹⁾; Wiga Maulana Baihaqi^(2*); Ade Nurhopipah⁽³⁾

(1) Universitas Amikom Purwokerto
 (2) Universitas Amikom Purwokerto
 (3) Universitas Amikom Purwokerto
 (*) Corresponding Author

 10.33096/ilkom.v15i2.1632.344-352  Abstract view: 69 times

Feature Space Augmentation for Negation Handling on Sentiment Analysis

PDF
353-357


 Lutfi Budi Ilmawan^(1*); Muladi Muladi⁽²⁾; Didik Dwi Prasetya⁽³⁾

(1) Universitas Negeri Malang, Universitas Muslim Indonesia
 (2) Universitas Negeri Malang
 (3) Universitas Negeri Malang
 (*) Corresponding Author

 10.33096/ilkom.v15i2.1695.353-357  Abstract view: 26 times

Classifying BISINDO Alphabet using TensorFlow Object Detection API

PDF
358-364

 Lilis Nur Hayati^(1*); Anik Nur Handayani⁽²⁾; Wahyu Sakti Gunawan Irianto⁽³⁾; Rosa Andrie Asmara⁽⁴⁾; Dolly Indra⁽⁵⁾; Muhammad Fahmi⁽⁶⁾


(1) Universitas Negeri Malang, Universitas Muslim Indonesia
 (2) Universitas Negeri Malang
 (3) Universitas Negeri Malang
 (4) Politeknik Negeri Malang
 (5) Universitas Muslim Indonesia

(6) Universitas Muslim Indonesia
(*) Corresponding Author

 10.33096/ilkom.v15i2.1692.358-364  Abstract view: 37 times

K-Nearest Neighbors Analysis for Public Sentiment towards Implementation of Booster Vaccines in Indonesia

PDF
365-372

 Ihwana As'ad^(1*); Muhammad Arfah Asis⁽²⁾; Hariani Ma'tang Pakka⁽³⁾; Randi Mursalim⁽⁴⁾; Yusnita binti Muhamad Noor⁽⁵⁾

- (1) Universitas Muslim Indonesia
- (2) Universitas Muslim Indonesia
- (3) Universitas Muslim Indonesia
- (4) Universitas Muslim Indonesia
- (5) Universiti Malaysia Pahang Al Sultan Abdullah
- (*) Corresponding Author

 10.33096/ilkom.v15i2.1561.365-372  Abstract view: 35 times

Decision Tree C4.5 Performance Improvement using Synthetic Minority Oversampling Technique (SMOTE) and K-Nearest Neighbor for Debtor Eligibility Evaluation

PDF
373-381

 Edi Priyanto⁽¹⁾; Enny Itje Sela⁽²⁾; Luther Alexander Latumakulita^(3*); Noourul Islam⁽⁴⁾

- (1) Universitas Teknologi Yogyakarta
- (2) Universitas Teknologi Yogyakarta
- (3) Universitas Sam Ratulangi
- (4) Kanpur Institute of Technology
- (*) Corresponding Author

 10.33096/ilkom.v15i2.1676.373-381  Abstract view: 24 times

Comparative Study of Herbal Leaves Classification using Hybrid of GLCM-SVM and GLCM-CNN

PDF
382-389

 Purnawansyah Purnawansyah^(1*); Aji Prasetya Wibawa⁽²⁾; Triyanna Widyaningtyas⁽³⁾; Haviluddin Haviluddin⁽⁴⁾; Cholisah Erman Hasihi⁽⁵⁾; Ming Foey Teng⁽⁶⁾; Herdianti Darwis⁽⁷⁾

- (1) Universitas Negeri Malang, Universitas Muslim Indonesia
- (2) Universitas Negeri Malang
- (3) Universitas Negeri Malang
- (4) Universitas Mulawarman
- (5) Universitas Muslim Indonesia
- (6) American University of Sharjah
- (7) Universitas Muslim Indonesia
- (*) Corresponding Author

 10.33096/ilkom.v15i2.1759.382-389  Abstract view: 24 times

The Support Vector Regression Method Performance Analysis in Predicting National Staple Commodity Prices

PDF
390-397

 Huzain Azis^(1*); Purnawansyah Purnawansyah⁽²⁾; Nirwana Nirwana⁽³⁾; Felix Andika Dwiyanto⁽⁴⁾

- (1) Universitas Muslim Indonesia
- (2) Universitas Muslim Indonesia
- (3) Universitas Muslim Indonesia
- (4) AGH University of Science and Technology
- (*) Corresponding Author

 10.33096/ilkom.v15i2.1686.390-397  Abstract view: 18 times

ILKOM Jurnal Ilmiah indexed by



ILKOM Jurnal Ilmiah

ISSN 2548-7779

Published by Prodi Teknik Informatika FIK Universitas Muslim Indonesia

W : <https://fikom.umi.ac.id/>

E : jurnal.ilkom@umi.ac.id



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0



p-ISSN 2087-1716
e-ISSN 2548-7779

No. 200/M/KPT/2020

Alamat Redaksi: Jl. Urip Sumoharjo KM. 05 Makassar

HOME ABOUT USER HOME SEARCH CURRENT ARCHIVES ANNOUNCEMENTS PUBLICATION FEES

Home > About the Journal > Editorial Team

Editorial Team

Editor in Chief

Yulita Salim, SCOPUS ID: 57202237115 - Computer Science - Universitas Muslim Indonesia, Indonesia

Managing Editor

Ramdan Satra, SCOPUS ID: 57211712874 - Computer Science - Universitas Muslim Indonesia, Indonesia

Editor

- Andri Pranolo, SCOPUS ID : 56572821900 - Universitas Ahmad Dahlan / Association for Scientific Computing Electrical and Engineering (ASCEE), Indonesia
- Aji Prasetya Wibawa, SCOPUS ID: 56012410400 - Universitas Negeri Malang / Association for Scientific Computing Electrical and Engineering (ASCEE), Indonesia
- Muhammad Iqbal, SCOPUS ID: 57211874405 - Universitas Negeri Semarang, Indonesia
- Imam Ahmad, SCOPUS ID: 56572821900 - Universitas Teknokrat Indonesia
- Andri Hidayat, SCOPUS ID: 57210682634 - Politeknik Negeri Sambas, Indonesia
- Mr. Mohamad Iqbal Suriansyah, SCOPUS ID: 57193504922 - University Pakuan, Indonesia
- Tasrif Hasanuddin, SCOPUS ID: 57211711050 - Universitas Muslim Indonesia, Indonesia, Indonesia
- Huzain Azis, SCOPUS ID: 57211407624 - Universitas Muslim Indonesia, Indonesia
- Herdianti Darwis, SCOPUS ID: 56167533100 - Universitas Muslim Indonesia, Indonesia, Indonesia
- Nia Kurniati, SCOPUS ID: 57211335623 - Universitas Muslim Indonesia, Indonesia

ILKOM Jurnal Ilmiah indexed by



ILKOM Jurnal Ilmiah

ISSN 2548-7779
Published by Prodi Teknik Informatika FIK Universitas Muslim Indonesia
W : <https://fikom.umi.ac.id/>
E : jurnal.ilkom@umi.ac.id

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0

Citedness in Scopus



USER

You are logged in as...

- dadang**
- » My Journals
- » My Profile
- » Log Out

QUICK MENU

- Editorial Team
- Peer Reviewer
- Peer Review Process
- Focus and Scope
- Publication Ethics
- Acceptance Rate
- Author Guidelines
- Open Access Policy
- Screening for Plagiarism
- Online Submission
- Author Declaration Form
- Journal Workflow
- Author Fee

Contact

JOIN WHATSAPP GROUP

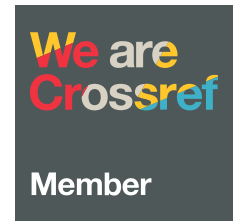


TOOLS



TUTORIAL MENDELEY

CURRENT MEMBER



SPONSORSHIP



ISSN NUMBER



TEMPLATE



History Template

- » Template 2016
- » Template 2017-2019
- » Template 2020
- » **NEW!** Template 2021 - now

DOWNLOAD AWARD



VISITOR

HiStat.com	
Vis. today	113
Visits	220 260
Pag. today	323
Pages	555 251



View Visitor

Visitors

	266,367		420
	8,393		414
	3,284		375
	1,071		271
	568		256



KEYWORDS

Android Augmented Reality
Classification Clustering
Data Mining Decision Support
System Deep Learning GLCM
Information System K-Means K-
NN K-Nearest Neighbor κ -
Nearest Neighbors KNN Machine
Learning Naive Bayes Naïve
Bayes SAW SVM Sentiment
Analysis Support Vector Machine

FONT SIZE

CURRENT ISSUE

ATOM	1.0
RSS	2.0
RSS	1.0



p-ISSN 2087-1716
e-ISSN 2548-7779

No. 200/M/KPT/2020

Alamat Redaksi: Jl. Urip Sumoharjo KM. 05 Makassar

[HOME](#) [ABOUT](#) [USER HOME](#) [SEARCH](#) [CURRENT](#) [ARCHIVES](#) [ANNOUNCEMENTS](#) [PUBLICATION FEES](#)

[Home](#) > [User](#) > [Author](#) > [Submissions](#) > #1544 > **Summary**

#1544 Summary

[SUMMARY](#) [REVIEW](#) [EDITING](#)

Submission

Authors	Dadang Priyanto, Ahmad Robbiul Iman, Deny Jollyta
Title	Naïve Bayes and K-Nearest Neighbor Algorithm Approach in Data Mining Classification of Drugs Addictive Diseases
Original file	1544-4104-1-SM.DOCX 2022-12-05
Supp. files	None
Submitter	Mr Dadang Priyanto
Date submitted	December 6, 2022 - 10:30 AM
Section	Artikel
Editor	Ramdan Satra
Author comments	Mohon di Proses
Abstract Views	38

Author Fees

Article Submission	0.00 IDR	PAY NOW
Article Publication	Paid August 3, 2023 - 11:28 AM	

Status

Status	Published Vol 15, No 2 (2023)
Initiated	2023-08-16
Last modified	2023-09-13

Submission Metadata

Authors

Name	Dadang Priyanto
ORCID iD	http://orcid.org/0000-0001-9240-056X
Affiliation	Universitas Bumigora
Country	Indonesia
Bio Statement	—
Principal contact for editorial correspondence.	
Name	Ahmad Robbiul Iman
Affiliation	Universitas Bumigora
Country	Indonesia
Bio Statement	—
Name	Deny Jollyta
Affiliation	Institut Bisnis Dan Teknologi Pelita Indonesia
Country	Indonesia
Bio Statement	—

Title and Abstract

Title	Naïve Bayes and K-Nearest Neighbor Algorithm Approach in Data Mining Classification of Drugs Addictive Diseases
Abstract	Indonesia, with its very large population, is a potential market for drugs trafficking. Hence, seriousness

Citedness in Scopus



USER

You are logged in as...

dadang

- » [My Journals](#)
- » [My Profile](#)
- » [Log Out](#)

QUICK MENU

[Editorial Team](#)

[Peer Reviewer](#)

[Peer Review Process](#)

[Focus and Scope](#)

[Publication Ethics](#)

[Acceptance Rate](#)

[Author Guidelines](#)

[Open Access Policy](#)

[Screening for Plagiarism](#)

[Online Submission](#)

[Author Declaration Form](#)

[Journal Workflow](#)

[Author Fee](#)

[Contact](#)

JOIN WHATSAPP GROUP



TOOLS

is needed in cracking down or preventing drug trafficking. Narcotics are substances or drugs that can cause dependence or addicted and other negative impacts on users. The problem is that drug users do not realize and even ignore diseases caused by drug addiction. The diseases can be life-threatening for users, such as inflammation of the liver, heart disease, hypertension, stroke, and others. The prevalence rate of drug abuse in West Nusa Tenggara (NTB) is included in the high category, reaching 292 cases or around 37.24% cases. This study aimed to create an application that can classify various diseases of drug users using the naive bayes and KNN methods. The results of this study indicated that there was a very close relationship between drug users and various deadly diseases. The prediction results showed that the naive bayes method provided a prediction accuracy of 94.5% while the KNN showed a prediction accuracy of 92.5%. This shows that the naive bayes method provides better predictive performance than the KNN in the data set of drug addicts in NTB.

Indexing

Keywords Drug Addiction; Drug Addiction Disease; Naive Bayes; Narkoba; K-Nearest Neighbor
Language en

Supporting Agencies

Agencies —

OpenAIRE Specific Metadata

ProjectID —

References

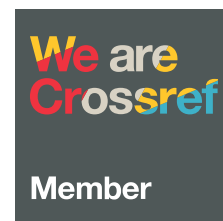
- References
- [1] B. Priambodo, Y. Jumaryadi, and U. Salamah, "Penggunaan Aplikasi Deteksi Pecandu Narkoba Di Meruya Utara," J. Pasopati Pengabd. Masy. dan Inov. Pengemb. Teknol., vol. 4, no. 3, 2022, doi: 10.14710/pasopati.2022.14468.
- [2] M. Mustaqim, H. Dafitri, and D. Dharmawati, "Edukasi Digital Pengenalan Bahaya Narkoba Bagi Anak Usia Dini Berbasis 3D Dan Augmented Reality," Djtechno J. Teknol. Inf., vol. 2, no. 2, pp. 170–176, 2021, doi: 10.46576/djtechno.v2i2.1623.
- [3] B. C. Laksono and N. W. K. Projo, "Pemodelan Analisis Rantai Markov untuk Mengestimasi Potensi Kasus Narkoba di Indonesia," Semin. Nas. Off. Stat., vol. 2021, no. 1, pp. 715–722, 2021, doi: 10.34123/semnasoffstat.v2021i1.1016.
- [4] A. Sinjar and T. Sahuri, "Bahaya Narkoba Terhadap Masa Depan Generasi Muda," vol. 2, no. 2, pp. 154–160, 2021.
- [5] D. Priyanto, M. Zarlis, H. Mawengkang, and S. Efendi, "Analysis of earthquake hazards prediction with multivariate adaptive regression splines," Int. J. Electr. Comput. Eng., vol. 12, no. 3, pp. 2885–2893, 2022, doi: 10.11591/ijece.v12i3.pp2885-2893.
- [6] H. Hairani, G. S. Nugraha, M. N. Abdillah, and M. Innuddin, "Komparasi Akurasi Metode Correlated Naive Bayes Classifier dan Naive Bayes Classifier untuk Diagnosis Penyakit Diabetes," InfoTekjar (Jurnal Nas. Inform. dan Teknol. Jaringan), vol. 3, no. 1, pp. 6–11, 2018, doi: 10.30743/infotekjar.v3i1.558.
- [7] Y. F. Safri, R. Arifudin, and M. A. Muslim, "K-Nearest Neighbor and Naive Bayes Classifier Algorithm in Determining The Classification of Healthy Card Indonesia Giving to The Poor," Sci. J. Informatics, vol. 5, no. 1, p. 18, 2018, doi: 10.15294/sji.v5i1.12057.
- [8] R. M. Abarca, "Sistem Pakar Diagnosis Penyakit Mata," Nuevos Sist. Comun. e Inf., no. sistem pakar penyakit mata, pp. 2013–2015, 2021, [Online]. Available: https://repository.dinamika.ac.id/id/eprint/982/6/BAB_II.pdf.
- [9] A. H. Bisri Merluarini, Diah Safitri, "Perbandingan Analisis Klasifikasi menggunakan Metode K-Nearest Neighbor (K-Nn) dan Multivariate Adaptive Regression Spline (Mars) pada Data Akreditasi Sekolah Dasar Negeri di Kota Semarang," vol. 3, pp. 313–322, 2014.
- [10] P. I. Lestari, D. E. Ratnawati, and L. Muflikhah, "Implementasi Algoritme K-Means Clustering Dan Naive Bayes Classifier Untuk Klasifikasi Diagnosa Penyakit Pada Kucing," J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya, vol. 3, no. 1, pp. 968–973, 2019.
- [11] A. Rajkumar, "Comparison of Fuzzy Diagnosis with K-Nearest Neighbor and Naive Bayes-2010.pdf," Glob. J. Comput. Sci. Technol., vol. 10, no. 10, pp. 38–43, 2010.
- [12] A. Desiani, "Perbandingan Implementasi Algoritma Naive Bayes dan K-Nearest Neighbor Pada Klasifikasi Penyakit Hati," Simkom, vol. 7, no. 2, pp. 104–110, 2022, doi: 10.51717/simkom.v7i2.96.
- [13] A. B. Wibisono and A. Fahrurrozi, "Perbandingan Algoritma Klasifikasi Dalam Pengklasifikasian Data Penyakit Jantung Koroner," J. Ilm. Teknol. dan Rekayasa, vol. 24, no. 3, pp. 161–170, 2019, doi: 10.35760/tr.2019.v24i3.2393.
- [14] S. Sulastri, K. Hadiono, and M. T. Anwar, "Analisis Perbandingan Klasifikasi Prediksi Penyakit Hepatitis Dengan Menggunakan Algoritma K-Nearest Neighbor, Naive Bayes Dan Neural Network," Dinamik, vol. 24, no. 2, pp. 82–91, 2020, doi: 10.35315/dinamik.v24i2.7867.

1151 A. Suragala, P. Venkateswarlu, and M. China Raiu, "A Comparative Study of Performance Metrics of



TUTORIAL MENDELEY

CURRENT MEMBER



SPONSORSHIP



ISSN NUMBER



TEMPLATE



History Template

- » Template 2016
- » Template 2017-2019
- » Template 2020
- » ~~NEW~~ Template 2021 - now

DOWNLOAD AWARD



VISITOR

Histat.com	
Online	3
Vis. today	98
Visits	220 245
Pag. today	274



View Visitor

Visitors	
	266,359
	8,393
	3,227
	1,071
	568
	420
	414
	374
	271
	256

[15] A. Subregate, P. Venkateshwar, and M. Chinn Raja, "A Comparative Study of Performance Metrics of Data Mining Algorithms on Medical Data," Lect. Notes Electr. Eng., vol. 698, no. February 2021, pp. 1549-1556, 2021, doi: 10.1007/978-981-15-7961-5_139.

[16] V. Biksham, V. Srujana, I. Meghana, B. Harshath, and G. Tarun, "Heart Disease Prediction Using Machine Learning," Ymer, vol. 21, no. 4, pp. 489-494, 2022, doi: 10.37896/YMER21.04/48.

[17] E. Wijaya, "Implementation Analysis of GLCM and Naive Bayes Methods in Conducting Extractions on Dental Image," IOP Conf. Ser. Mater. Sci. Eng., vol. 407, no. 1, 2018, doi: 10.1088/1757-899X/407/1/012146.

[18] J. N. Peksi, B. Yuwono, and Y. M. Florestiyanto, "Classification of Anemia with Digital Images of Nails and Palms using the Naive Bayes Method," J. Inform. dan Teknol. Inf., vol. 18, no. 1, pp. 118-130, 2021, doi: 10.31515/telematika.v18i1.4587.

[19] P. Chalekar, S. Shroff, S. Pise, and S. Panicker, "Use of K-Nearest Neighbor in Thyroid Disease Classification," Int. J. Curr. Eng. Sci. Res., vol. 1, no. 2, pp. 2394-0697, 2014.

[20] J. M. Martínez-Otzeta and B. Sierra, "Analysis of the iterated probabilistic weighted k nearest neighbor method, a new distance-based algorithm," ICEIS 2004 - Proc. Sixth Int. Conf. Enterp. Inf. Syst., no. January, pp. 233-240, 2004, doi: 10.5220/0002605402330240.

[21] A. K. Saputro, K. A. Wibisono, and F. P. Pratiwi, "Identification of Disease Types on Tea - Plant Varieties Based Image Processing with K-Nearest Neighbor Method," J. Phys. Conf. Ser., vol. 1569, no. 3, 2020, doi: 10.1088/1742-6596/1569/3/032078.

ILKOM Jurnal Ilmiah indexed by



ILKOM Jurnal Ilmiah

ISSN 2548-7779

Published by Prodi Teknik Informatika FIK Universitas Muslim Indonesia

W : <https://fikom.umi.ac.id/>

E : jurnal.ilkom@umi.ac.id



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0



KEYWORDS

Android Augmented Reality
Classification Clustering
 Data Mining Decision Support
 System Deep Learning GLCM
 Information System K-Means K-
 NN K-Nearest Neighbor k-
 Nearest Neighbors KNN Machine
 Learning Naive Bayes Naive
 Bayes SAW SVM Sentiment
 Analysis Support Vector Machine

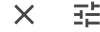
FONT SIZE

CURRENT ISSUE

ATOM	1.0
RSS	2.0
RSS	1.0



ilkomjurnal@gmail.com



99+

Compose

Mail

Inbox

182



ILKOM Jurnal Ilmiah <ilkomjurnal@gmail.com>

Tue, Dec 6, 2022, 9:30 AM

to me

Mr Dadang Priyanto:

Thank you for submitting the manuscript, "Naive Bayes And K-Nearest Neighbor Algorithm Approach In Data Mining Classification Of Drugs Addictive Diseases" to ILKOM Jurnal Ilmiah. With the online journal management system that we are using, you will be able to track its progress through the editorial process by logging in to the journal web site:

Manuscript URL:

<https://jurnal.fikom.umi.ac.id/index.php/ILKOM/author/submission/1544>

Username: dadang

If you have any questions, please contact me. Thank you for considering this journal as a venue for your work.

ILKOM Jurnal Ilmiah

Labels

3 of 4



p-ISSN 2087-1716
e-ISSN 2548-7779

No. 200/M/KPT/2020

Alamat Redaksi: Jl. Urip Sumoharjo KM. 05 Makassar

HOME ABOUT USER HOME SEARCH CURRENT ARCHIVES ANNOUNCEMENTS PUBLICATION FEES

Home > User > Author > Submissions > #1544 > Review

#1544 Review

SUMMARY REVIEW EDITING

Submission

Authors	Dadang Priyanto, Ahmad Robbiul Iman, Deny Jollyta
Title	Naïve Bayes and K-Nearest Neighbor Algorithm Approach in Data Mining Classification of Drugs Addictive Diseases
Section	Artikel
Editor	Ramdan Satra

Peer Review

Round 1

Review Version	1544-4108-1-RV.DOCX 2022-12-06
Initiated	2023-01-24
Last modified	2023-05-24
Uploaded file	None

Editor Decision

Decision	Accept Submission 2023-06-09
Notify Editor	Editor/Author Email Record 2023-05-15
Editor Version	None
Author Version	1544-4503-1-ED.DOCX 2023-04-02 DELETE
Upload Author Version	<input type="button" value="Choose File"/> No file chosen <input type="button" value="Upload"/>

ILKOM Jurnal Ilmiah indexed by



ILKOM Jurnal Ilmiah

ISSN 2548-7779
Published by Prodi Teknik Informatika FIK Universitas Muslim Indonesia
W : <https://fikom.umi.ac.id/>
E : jurnal.ilkom@umi.ac.id

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0

Citedness in Scopus



USER

You are logged in as...

dadang

- » My Journals
- » My Profile
- » Log Out

QUICK MENU

Editorial Team

Peer Reviewer

Peer Review Process

Focus and Scope

Publication Ethics

Acceptance Rate

Author Guidelines

Open Access Policy

Screening for Plagiarism

Online Submission

Author Declaration Form

Journal Workflow

Author Fee

Contact

JOIN WHATSAPP GROUP

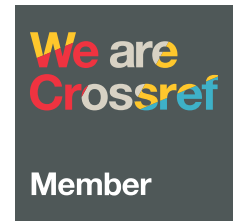


TOOLS



TUTORIAL MENDELEY

CURRENT MEMBER



SPONSORSHIP



ISSN NUMBER



TEMPLATE



History Template

- » Template 2016
- » Template 2017-2019
- » Template 2020
- » **NEW!** Template 2021 - now

DOWNLOAD AWARD



VISITOR

HiStat.com	
Online	2
Vis. today	98
Visits	220 245
Pag. today	284



View Visitor

Visitors

	266,359		420
	8,393		414
	3,240		374
	1,071		271
	568		256



KEYWORDS

Android Augmented Reality
Classification Clustering
Data Mining Decision Support
System Deep Learning GLCM
Information System K-Means K-
NN K-Nearest Neighbor κ -
Nearest Neighbors KNN Machine
Learning Naive Bayes Naïve
Bayes SAW SVM Sentiment
Analysis Support Vector Machine

FONT SIZE

CURRENT ISSUE

ATOM	1.0
RSS	2.0
RSS	1.0



Dadang Priyanto <dadang.priyanto@universitasbumigora.ac.id>

[ILKOM] Editor Decision

1 message

Ramdan Satra <ilkomjurnal@gmail.com>

Fri, Mar 24, 2023 at 4:43 PM

To: Mr Dadang Priyanto <dadang.priyanto@universitasbumigora.ac.id>

Mr Dadang Priyanto:

We have reached a decision regarding your submission to ILKOM Jurnal Ilmiah, "Naive Bayes And K-Nearest Neighbor Algorithm Approach In Data Mining Classification Of Drugs Addictive Diseases".

Our decision is: Revisions Required

Ramdan Satra
SCOPUS ID: 57211712874 - Computer Science - Universitas Muslim Indonesia
ramdan@umi.ac.id

Reviewer A:

Quality:

Does the paper reflect excellence of research work? Is it technically sound?

:

Questionable content

Originality:

Does the paper provide a novel approach? Does it discuss new/unpublished work? Does it contain significant additional material to that already published? (For research articles only, not applicable for review articles, tutorials and case studies):

Not really innovative

Clarity:

Does the paper clearly communicate its message? Are tables and graphics clear to read and labeled appropriately?:

Readable

Relevance:

How well does the content fit the ILKOM Jurnal Ilmiah objectives (interdisciplinarity, bridge between theory and application, international dialogue) and the issue's theme? Is the paper of interest to the readership of the ILKOM Jurnal Ilmiah?:

Close enough

References:

Does the paper contain the appropriate referencing to provide adequate context for the present work?:

Moderately adequate

Summary of evaluation:

What is your global evaluation for this paper?:

Fair

Overall recommendation:

If "Accept with minor revisions", write the comments to the author on the section below and/or upload a file.

If "Accept after major revisions", write the comments to the author on the section below and/or upload a file.

If "Reject", justify below.

:

Accept after major revisions, provided that the author makes the major changes suggested

Reviewer familiarity:

How familiar are you with the topic? (will be used to weight the evaluation):

High

Explanation on rejection:

If your overall recommendation is to "Reject" this paper, explain why.

Otherwise, write "N/A".:

N/A

Comments to the author:

Insert here the comments to the author on how to improve the paper. You can also upload a file (see above "Upload review attachment"). If you have no comments write "N/A":

results are the texts presenting the meaning of your research data. One mistake that some authors often make is to use text to direct the reader to find a specific table or figure without further explanation. This can confuse the readers when they interpret the meaning of the data completely different from what the authors had in mind.

Figures and tables present information about your research data visually.

The use of these visual illustrations is necessary so the readers can summarize, compare, and interpret large data at a glance. You can use graphs or figures to compare groups or patterns.

Naive Bayes And K-Nearest Neighbor Algorithm Approach In Data Mining Classification Of Drugs Addictive Diseases

By Dadang Priyanto



Naive Bayes And K-Nearest Neighbor Algorithm Approach In Data Mining Classification Of Drugs Addictive Diseases

Dadang Priyanto ^{a,1,*}; Ahmad Robbiul Iman ^{a,2}; Deny Jollyta ^{b,3}

^a Department of Computer Science, Universitas Bumigora, Ismail Marzuki Street, Mataram 83127, Indonesia.

^b Department of Informatics Engineering, Institut Bisnis Dan Teknologi Pelita Indonesia, Pekanbaru Riau 28127, Indonesia

¹ dang.priyanto@universitasbumigora.ac.id ; ² robbishimozuki@gmail.com ; ³ deny.jollyta@lecturer.pelitaindonesia.ac.id

* Corresponding author

Article history: Received Month xx, 2021; Revised Month xx, 2021; Accepted Month xx, 2021; Available online Month xx, 2021 (Times 8pt)

Abstract

Indonesia, with a very large population, is a potential market for drug trafficking, so seriousness is needed in cracking down on drug trafficking or preventing it. Narcotics are substances or drugs that can cause dependence and other negative impacts on users. The problem is that drug users do not realize and even ignore the diseases caused by drug addiction, these diseases can be life-threatening for users, such as diseases such as inflammation of the liver, heart disease, hypertension, stroke and others. The prevalence rate of drug abuse in West Nusa Tenggara (NTB) is included in the high category, reaching 292 cases or around 37.24% of cases. This study aims to create an application that can classify various diseases of drug users using the Naive Bayes and K-Nearest Neighbor methods. The results of this study indicate that there is a very strong relationship between drug users and various deadly diseases that accompany them. The prediction results show that the Naive Bayes method provides a prediction accuracy of 94.5% while the K-Nearest Neighbor shows a prediction accuracy of 92.5%. This shows that the Naive Bayes method provides better predictive performance than the K-Nearest Neighbor in the data set of drug addicts in NTB.

Keywords: Narkoba; drug addiction; Naive Bayes; Drug Addiction Disease; K-Nearest Neighbor.

Introduction

Indonesia as an archipelagic country with a population of more than 276 million people, this is a very large population and has great potential for the marketing and distribution of narcotics. Narkoba is an abbreviation of Narcotics, Psychotropics and other addictive substances. The term narcotics itself is a substance or drug that is natural, synthetic, or semi-synthetic which when consumed by the human body can affect the central nervous system (SPP), namely the brain and the acuity of a person's analysis and thinking [1], [2]. Drug or narcotic users will feel addicted and addicted to these illegal drugs. Drug users can feel excessive anxiety, increased sexual appetite, paranoia, experience delusions, behavior changes towards aggressiveness. Trafficking of narcotics is prohibited because it can cause dependence and various diseases caused [3], [4]. The problem is that drug users do not realize and even ignore the diseases caused by drug addiction, these diseases can be life-threatening for users, such as diseases such as inflammation of the liver, heart disease, hypertension, stroke and others. The prevalence rate of drug abuse in West Nusa Tenggara is in the high category, reaching 292 cases or around 37.24% of drug abuse cases, out of a population of approximately 5.1 million people.

This research aims to create an application that can classify various diseases of drug users based on the Naive Bayes and K-Nearest Neighbor methods. Classification is a systemic arrangement in groups or groups according to established rules or standards. Literally classification can be said as the division of something according to classes. According to science, classification is the process of grouping objects or data based on their characteristics, similarities, and differences. Processing data with the classification of data mining processing categories, and many methods that can be used in data mining. Data mining is grouped into descriptive data mining and predictive data

mining [5]. The Naïve Bayes and K-Nearest Neighbor methods can be used in classification. The Naïve Bayes Classifier is based on the static Bayes probability theorem put forward by the English scientist Thomas Bayes, namely improving future opportunities based on past experience with the main characteristic of a very strong (naive) assumption of the independence of each condition or event. [6]–[8]. The K-Nearest Neighbor (KNN) method is an algorithm used to classify an object, based on the k training data that is closest to the object. The condition for the value of k is that it cannot be greater than the number of training data, and the value of k must be odd and more than one. Near or far the distance of the training data that is closest to the object to be classified can be calculated using the Euclidean Distance [9], [10].

Previous studies have used a variety of data mining methods such as disease diagnosis by using the Fuzzy, Naïve Bayes and K-Nearest Neighbor comparative methods for smallpox, dengue fever and flu [11]. Another study using the Naïve Bayes and K-Nearest Neighbor methods for liver disease classification, which gave the best performance was the KNN algorithm [12], another comparison of classification algorithms in coronary heart disease showed that the Random Forest algorithm provided better performance in terms of accuracy [13], then a comparative analysis of the classification method for Hepatitis shows that the Naïve Bayes method has better performance [14], Other comparative studies of data mining algorithms on medical data performance Support Vector Machine (SVM) provide the best accuracy [15], and the use of genetic algorithms and machines learning in the classification and prediction of heart disease [16].

Previous studies have shown that various classification methods provide different best performance results based on the data set used, therefore this study provides a novelty in classification using the Naïve Bayes and K-Nearest Neighbor methods on drug addict disease data sets in the Nusa Tenggara region. West. From the results of the comparison will show the performance of the best method which differs in accuracy from previous studies.

Method

The data mining method used in this study is a classification category and is a grouping of an object into a particular class. Various cases related to grouping objects can be solved by applying classification techniques. In general, classification performance is carried out using the confusion matrix, and this study uses the Naive Bayes and K-Nearest Neighbor methods.

A. Naive Bayes method

Naive Bayes is a simple classifying probabilistic method based on Bayes' Theorem where classification is carried out through training sets of a number of data efficiently [17]. Naive Bayes assumes that the value of an input attribute in a given class does not depend on the values of other attributes [18]. Bayes' theorem itself was put forward by British scientist Thomas Bayes, namely predicting future opportunities based on previous experience so that it is known as Bayes' Theorem. Where the Bayes Theory equation is:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (1)$$

Where :

- X : Data with unknown class
- C : Hypothesis data X is a specific class
- P (C|X) : Probability of hypothesis C based on condition X (posterior probability)
- P(C) : Probability of hypothesis C (prior probability)
- P(X|C) : Probability of X based on conditions in hypothesis C
- P(X) : Probability of X

B. K-Nearest Neighbor method

K-Nearest Neighbor is a method using a supervised learning algorithm in which a number of new attributes whose class is unknown can be searched for by looking for similarities based on the majority of K nearest neighbors as a reference in determining the class of an attribute [19]–[21]. The K-Nearest Neighbor method can be calculated using the euclidean distance in determining the distance between datasets. The Euclidean distance formula can be seen in formula 2 below:

$$d = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + \dots + (N_1 - N_2)^2} \quad (2)$$

Where :

- d = distance between the two attributes
- X1 = actual data
- X2 = Data testing
- Y1 = nth actual data
- Y2 = nth testing data
- N1 = first nth attribute
- N2 = second nth attribute

44

The flowchart of the calculation process using the K-Nearest Neighbor method can be seen in Figure 1.

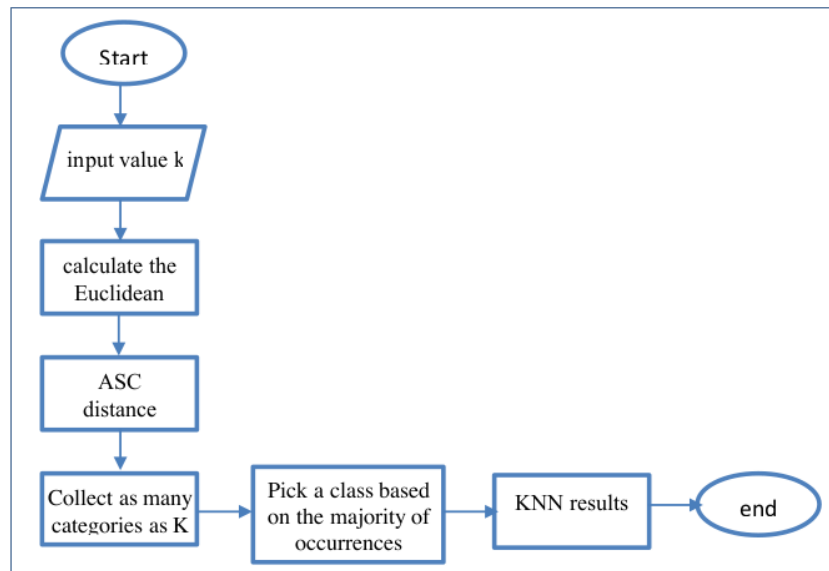


Figure 1. Flowchart of K-Nearest Neighbor

C. Data Set

This study uses data taken from the West Nusa Tenggara Provincial National Narcotics Agency (BNNP), with data collection using interview techniques and in the form of data files. From the results of the interviews, information was obtained about the symptoms experienced by drug addicts, the types of drugs consumed, the level of addiction, and history of drug addict disease. The data obtained from the results of a survey regarding the types of drug addict disease obtained 336 data records, and can be partially presented in tabular form as in table 1 below:

Table 1. Drug Addiction Disease Dataset Table

Drug Type	Addiction Level	G001	...	G024	Disease History	Disease
Kokain	Rekreasional	Yes	...	No	Stomach Inflammation	GERD
Ganja	Try Use	Yes	...	Yes	Heart disease	Hepatitis
Sabu-sabu	Active dependency	No	...	No	Asthma	Acute Asthma
Morfin	Try Use	Yes	...	No	Hypertension	Kista
Gorilla tobacco	recreational	Yes	...	No	Malaria	Thypoid
...
Gorilla tobacco	recreational	Yes		Yes	Diabetes	Hepatitis

From table 1 it can be explained the description of the specified symptom variables as in table 2 below:

Table 2. Description of symptom variables

Symptom Code	Symptom
G001	Nausea and vomiting
G002	Sore throat
G003	Expulsion of stomach contents involuntarily
G004	Bad breath
G005	coughs
G006	hard to breathe
G007	Hoarseness
G008	Speech disorder
G009	Hard to sleep
G010	Heart rate increases
G011	Dizzy
G012	Easily Tired
G013	Chest feels tight
G014	Redness of the skin around the cyst area
G015	Blood or pus that smells bad from the lump
G016	Fever
G017	Mild or severe pain in the lower abdomen
G018	Stomach ache
G019	Loss of appetite
G020	Joint pain
G021	Diarrhea
G022	Dry Cough
G023	Jaundice
G024	Dark urine

Results and Discussion

A. Preprocessing

Data before being processed with the Algorithm needs pre-processing so that the data is easier to read, reduces the burden of representation in the data, reduces the duration of data processing, and simplifies the process of data analysis by the system. The dataset that is formed must be a dataset that is free from missing values and Unconsistent Attributes so as not to interfere with the process at the modeling stage. The initial stage of preprocessing [4] by initializing attributes using the variables "X_train" and "Y_train" as class labels. In the [4] annual calculation of Naïve Bayes and K-Nearest Neighbor the attribute must have a numeric value. This is because [4] Naïve Bayes and K-Nearest Neighbor utilize Probability and statistical calculations. So the data preprocessing stage for the Naive Bayes and K-Nearest Neighbor [4] algorithms is done by changing the value of the numeric "X_train" attribute. The results of the preprocessing for the Naive Bayes and K-Nearest Neighbor algorithms can be partially seen in table 3 below:

Table 3 Preprocessing Results

Drugs	Addicted	G001	...	G024	History	Disease
0.84	1.29	1	...	-1	0.25	GERD
0.18	0.06	1	...	1	1.75	Hepatitis
1.54	1.18	-1	...	-1	0.72	Acute Asthma
0.48	0.06	1	...	-1	0.96	Cyst
1.15	1.29	1	...	-1	0.35	Thypoid
...
1.54	1.29	1	...	-1	0.96	Cyst
0.48	0.06	-1	...	-1	0.81	Acute Asthma
1.15	1.29	1	...	1	0.81	Hepatitis

Based on table 3, it can be seen that the results of data processing describe the condition of the disease based on the type of drug used, the level of addiction, the symptoms experienced in the range of symptoms 1 to 24, as well as a history of previous illnesses. For example, record data set 1 has the type of addiction 0.84 as the type of drug 'Cocaine', with an addiction level of 1.29 as 'Recreational' addiction, and symptoms that are felt according to code G001 are 'Nausea and Vomiting' and so on until symptom 24, and combined with a history of previous illness it can be concluded as 'GERD' disease. And so on the same thing in reading the next record in table 3 above. Information about the attributes used in reading table 3 can be seen v[11] the attributes of the type of drug, the level of addiction, symptoms of the disease and history of the disease which can be explained in the following table 4-7. :

Table 4. Description of Drug Type Attributes

Initial Attribute Value	Numerical Attribute Values
Crystal meth	1.54
Cocaine	0.84
Morphine	0.48
Gorilla Tobacco	1.15
Marijuana	0.18

Table 4 Provides information regarding the attribute values for the types of drugs used and numerical values according to the types of drugs used. For example, narcotics of the methamphetamine type have a numerical attribute value of 1.54 and so on.

Table 5. Addiction Level Attribute Description

Initial Attribute Value	Numerical Attribute Values
Try Use	0.06
recreational	1.29
Active Dependence	1.18

Table 5 Provides a description of the level of addiction experienced by users such as a value of 0.06 indicating addiction at the 'Try Using' stage and so on.

Table 6. Symptom Description

Initial Attribute Value	Numerical Attribute Values
Yes	1
Not	2

Table 6 Describes information on the symptoms experienced by addicts with a determination value of 1 'Yes' experience' and 2 'No' experience', while the symptoms experienced refer to table 2 above.

Table 7. Description of Disease History

Initial Attribute Value	Numerical Attribute Value
Asthma	0.72
Stomach Inflammation	0.25
Heart disease	1.75
Malaria	0.35
Pneumonia	1.66
tuberculosis	0.44
Diabetes	0.81
Hypertension	0.96
There isn't any	0

Table 7 describes the previous history of the disease experienced by the addict as an additional variable in determining the conclusion of the illness.

B. Calculation results of the K-Nearest Neighbor method

The results of calculations with the K-Nearest Neighbor from the amount of existing data, 10 testing data are needed to be calculated with the results as in table 8 below:

Table 8. Data Testing

Drug Type	Addiction Level	G001	...	G024	Disease History	Disease
0.48	1.18	1	...	-1	0	Thypoid
1.54	1.29	1	...	-1	1.75	GERD
1.15	0.06	-1	...	-1	0.96	Hepatitis
0.48	1.29	-1	...	-1	0	GERD
0.84	1.29	1	...	-1	0.81	Cyst
0.18	0.06	-1	...	1	1.75	Cyst

0.18	1.29	1	...	-1	0.81	Acute Asthma
1.15	0.06	1	...	1	1.75	Hepatitis
1.15	1.29	-1	...	-1	1.75	Cyst
1.54	0.06	1	...	-1	0.35	GERD

As seen in table 8, it can be concluded that the results of testing from 10 data, with varying levels of disease experienced by addicts such as typhoid, gerd, and others according to the drugs consumed and the level of addiction they experience. From table 8 it can be calculated to determine 49 closest majority K value similarity or based on the distance between data sets with the Euclidean distance formula and the results are as in table 9 below:

Table 9. Euclidean distance result

Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
0.677495387	1.18156676	1.64693655	1.834557167	1.4066272	1.731964203	0.43104524	1.614218077	1.0694391	1.33574
1.094257739	1.34331679	1.52715422	1.698293261	1.5270887	2.021113554	0.67749539	1.771242502	0.8258935	1.22793
1.923174459	0.88977525	1.09535382	0.849352695	0.8713208	2.141144554	1.11919614	1.345696846	0.5374942	0.47223
1.718924082	0.44922155	1.83338485	1.697910481	0.4114608	2.255393536	0.94175368	1.67002994	0.9393615	0.54452
1.479932431	1.70120545	0.66272166	1.04517941	1.8645643	1.638566447	1.21268298	1.315028517	1.1157509	1.59706
1.706399719	0.27037012	1.53590364	1.3713497	0.2014944	2.021014597	0.91869473	1.336674979	0.868159	0.50359
1.263566381	1.59176003	1.56626945	1.762498227	1.7652762	2.167694628	0.93648278	1.934760967	0.9404786	1.39893
0.992219734	1.22711043	1.16357209	1.380543371	1.4259383	1.683627037	0.55506756	1.401784577	0.7598684	1.19537
1.158015544	1.3095419	1.66679333	1.824883558	1.4974645	0.671490879	1.23798223	0.820548597	1.8714967	1.91348
2.067147793	1.10498869	0.93193347	0.624579859	1.0901835	1.801887899	1.43405021	0.872811549	1.1983322	1.13723
0.517010638	1.05763888	1.42369238	1.637101096	1.3042622	1.288953064	0.37802116	1.206067991	1.1540364	1.38351
0.933380951	1.31244047	1.20307938	1.44934468	1.5182226	1.65562677	0.60712437	1.434747365	0.8673523	1.30407
1.741895519	0.53037722	1.77476759	1.634441801	0.4987986	2.298543017	0.93455872	1.693664666	0.8190849	0.41833
1.293290377	1.40730238	0.74040529	1.04890419	1.5836666	0.881873007	1.1319894	0.516623654	1.3944533	1.6533
...
1.443502684	1.64972725	0.52316345	0.962756459	1.8177184	1.251958466	1.257417989	0.940638081	1.3234047	1.70678

From table 9 it can be determined or the selection of the smallest distance value that is included in the ranking of a number of K values. In table 9 the value of K is 10, so you have to find the smallest distance value of 10 values. Furthermore, the results of the highest number of classes from a predetermined K value can be obtained as predicted in table 10 below :

Table 10. The highest number of classes from the value of K

Nth Test Data	Original Class	Acute Asthma	Cyst	Gred	Hepatitis	Typoid
1	Thypoid	False	False	False	False	True
2	Hepatitis	True	False	False	False	False
3	Asma Akut	False	False	False	True	False
4	Kista	False	False	False	True	False
5	Asma Akut	True	False	False	False	False
6	Hepatitis	False	False	False	False	False
7	Kista	True	False	False	False	False
8	Asma Akut	False	False	False	False	False
9	GERD	False	False	True	False	False
10	GERD	False	False	True	False	False

Table 10 Explains that the test results indicate a 'True' value, so addicts are diagnosed with a disease according to the predicted results as presented in table 9 test results.

C. Calculation results with Naive Bayes

The calculation process with the Naive Bayes method begins by retrieving training data from the data that has been obtained. The variables that will be used in the classification of drug addicts are disease, type of drug, level of drug addiction, history of illness, and the results of the calculation show that the average probability of the disease class is 20%. while for the probability of drug type attributes as shown in table 11 below:

Table 11. Attribute Probability of Drug Types

Drug Type	Acute Asthma	Gred	Typhoid	Cyst	Hepatitis
Crystal meth	10%	30%	30%	30%	10%
Cocaine	50%	20%	20%	20%	20%
Morphine	0%	10%	0%	10%	10%
Gorilla Tobacco	30%	30%	40%	10%	0%
Marijuana	10%	10%	10%	30%	60%
Total number	100%	100%	100%	100%	100%

Table 11. Attribute Probability of Drug Types

Addiction Disease	Acute Asthma	Gred	Typoid	Cyst	Hepatitis
Recreational	40%	30%	20%	30%	40%
Active dependency	0%	30%	50%	30%	40%
Try it on	60%	40%	30%	40%	20%
Total	100%	100%	100%	100%	100%

The resulting model will be evaluated with the aim of assessing the performance of the Naive Bayes algorithm and the K-Nearest Neighbor algorithm. Evaluation is done by calculating the accuracy, precision, and recall of each algorithm used using the Confusion Matrix Technique. If the evaluation results show good performance, then the two algorithms are feasible in making predictions on new data whose labels are unknown.

D. Testing the Naïve Bayes

At this stage the test is carried out on the performance of the Naive Bayes classification model that has been made when making predictions on data testing, using the confusion matrix and the confusion matrix plot. The test results can be seen in Figure 2 of the Confusion chart below:

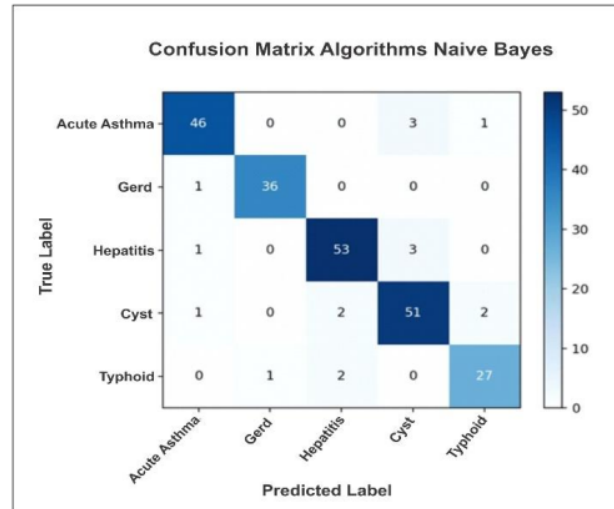


Figure 2. Naive Bayes Confusion Matrix Graph

In the Confusion Matrix f above, it can be seen that the prediction for the Acute Asthma class has a true value of 46, the prediction for the Gerd class has a value of 36, the prediction for the Hepatitis class has a true value of 53, the prediction for the true class of Cyst class is 51, and the prediction for the Typhoid class with a true value of 27 , wrong prediction in the Acute Asthma class by 4, Gerd class by 1, Hepatitis class by 4, Cyst class by 2, and Typhoid class by 3. The results of the calculation of the Naive Bayes confusion matrix can be seen in table 11 below:

Table 12. Confusion Matrix Naïve Bayes

		Acute Asthma	Gerd	Typhoid	Cyst	Hepatitis
Predictions	Acute Asthma	46	0	0	3	1
	GERD	1	36	0	0	0
	Hepatitis	1	0	53	3	0
	Cyst	1	0	2	51	0
	Typhoid	0	1	2	0	27

Based on Table 12, prediction accuracy can be calculated using the Naïve Bayes algorithm which produces a prediction accuracy of 93% as in the following formula 3 calculation:

$$\text{Prediction Accuracy} = \frac{46+36+53+51+27}{230} = 0,93\%$$

E. Comparison Results of the Naïve Bayes and K-Nearest Neighbor Methods

This test was carried out by dividing training data and data testing into 4 variations of data, and testing was carried out on Naïve Bayes and K-Nearest Neighbor modeling on variation of the percentage of training data and testing data. The results of testing with the Confusion Matrix can be seen in the comparison of accuracy, precision and recall of the two algorithms as shown in Table 13 below:

Table 13. Results of Comparison of Accuracy, Precision and Recall

Data Training	Data Testing	Accuracy (%)		Precision (%)		Recall (%)	
		Naïve Bayes	K-NN	Naïve Bayes	K-NN	Naïve Bayes	K-NN
335	201	93	93	93	93	93	93
335	180	96	90	95	90	95	90
335	230	93	91	93	91	93	91
335	181	96	96	96	96	96	96
Average		94.5	92.5	94.25	92.5	94.25	92.5

Based on table 13, the calculations after being validated show that the Naïve Bayes method has a higher accuracy than the K-NN, with a value of 94.5%, while K-NN after being validated has an accuracy value of 92.5%.

Conclusion

Narcotics are in the category of drugs that are prohibited from being widely distributed, because they can lead to addiction for their users and can cause various deadly diseases. This research was conducted to classify various diseases of Narcotic addicts in the form of application programs using two methods, namely Naïve Bayes and K-Nearest Neighbor. This research has shown that there is a very close relationship between drug users and various deadly diseases that accompany them. The prediction results show that the Naïve Bayes method provides a prediction accuracy of 94.5% while the K-Nearest Neighbor shows a prediction accuracy of 92.5%. This shows that the Naïve Bayes method provides better predictive performance than the K-Nearest Neighbor with the data set used in this study.

References

- [1] B. Priambodo, Y. Jumaryadi, and U. Salamah, "Penggunaan Aplikasi Deteksi Pecandu Narkoba Di Meruya Utara," *J. Pasopati Pengabd. Masy. dan Inov. Pengemb. Teknol.*, vol. 4, no. 3, 2022, doi: 10.14710/pasopati.2022.14468.
- [2] M. Mustaqim, H. Dafitri, and D. Dharmawati, "Analisis Dan Pengenalan Bahaya Narkoba Bagi Anak Usia Dini Berbasis 3D Dan Augmented Reality," *Djtechno J. Teknol. Inf.*, vol. 2, no. 2, pp. 170–176, 2021, doi: 10.46576/djtechno.v2i2.1623.
- [3] B. C. Laksono and N. W. K. Projo, "Pemodelan Analisis Rantai Markov untuk Mengestimasi Potensi Kasus Narkoba di Indonesia," *Semin. Nas. Off. Stat.*, vol. 2021, no. 1, pp. 715–722, 2021, doi: 10.34123/semnasoffstat.v2021i1.1016.
- [4] A. Sinjar and T. Sahuri, "BAHAYA NARKOBA TERHADAP MASA DEPAN GENERASI MUDA," vol.

- 2, no. 2, pp. 154–160, 2021.
- [5] D. Priyanto, M. Zarlis, H. Mawengkang, S. Efendi, “Analysis of earthquake hazards prediction with multivariate adaptive regression splines,” *Int. J. Electr. Comput. Eng.*, vol. 12, no. 3, pp. 2885–2893, 2022, doi: 10.11591/ijece.v12i3.pp2885-2893.
- [6] H. Hairani, G. S. Nugraha, M. N. Abdillah, and M. Innuddin, “Komparasi Akurasi Metode Correlated Naive Bayes Classifier dan Naive Bayes Classifier untuk Diagnosis Penyakit Diabetes,” *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 3, no. 1, pp. 6–11, 2018, doi: 10.30743/infotekjar.v3i1.558.
- [7] Y. F. Safri, R. Arifudin, and M. A. Muslim, “K-Nearest Neighbor and Naive Bayes Classifier Algorithm in Determining The Classification of Healthy Card Indonesia Giving to The Poor,” *Sci. J. Informatics*, vol. 5, no. 1, p. 18, 2018, doi: 10.15294/sji.v5i1.12057.
- [8] R. M. Abarca, “Sistem Pakar Diagnosis Penyakit Mata,” *Nuevos Sist. Comun. e Inf.*, no. sistem pakar penyakit mata, pp. 2013–2015, 2021, [Online]. Available: https://repository.dinamika.ac.id/id/eprint/982/6/BAB_II.pdf.
- [9] A. H. Bisri Merluarini, Diah Safitri, “PERBANDINGAN ANALISIS KLASIFIKASI MENGGUNAKAN METODE K-NEAREST NEIGHBOR (K-NN) DAN MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS) PADA DATA AKREDITASI SEKOLAH DASAR NEGERI DI KOTA SEMARANG,” vol. 3, pp. 313–322, 2014.
- [10] P. I. Lestari, D. E. Ratnawati, and L. Muflikhah, “Implementasi Algoritma K-Means Clustering Dan Naive Bayes Classifier Untuk Klasifikasi Diagnosa Penyakit Pada Kucing,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 3, no. 1, pp. 968–973, 2019.
- [11] A. Rajkumar, “Comparison of Fuzzy Diagnosis with K-Nearest Neighbor and Naive Bayes-2010.pdf,” *Glob. J. Comput. Sci. Technol.*, vol. 10, no. 10, pp. 38–43, 2010.
- [12] A. Desiani, “Perbandingan Implementasi Algoritma Naive Bayes dan K-Nearest Neighbor Pada Klasifikasi Penyakit Hati,” *Simkom*, vol. 7, no. 2, pp. 104–110, 2022, doi: 10.51717/simkom.v7i2.96.
- [13] A. B. Wibisono and A. Fahriz, “Perbandingan Algoritma Klasifikasi Dalam Pengklasifikasian Data Penyakit Jantung Koroner,” *J. Ilm. Teknol. dan Rekayasa*, vol. 24, no. 3, pp. 161–170, 2019, doi: 10.35760/tr.2019.v24i3.2393.
- [14] Sulastris, K. Hadiono, and M. T. Anwar, “Analisis Perbandingan Klasifikasi Prediksi Penyakit Hepatitis Dengan Menggunakan Algoritma K-Nearest Neighbor, Naive Bayes Dan Neural Network,” *Dinamik*, vol. 24, no. 2, pp. 82–91, 2020, doi: 10.35315/dinamik.v24i2.7867.
- [15] A. Suragala, P. Venkateswarlu, and M. China Raju, “A Comparative Study of Performance Metrics of Data Mining Algorithms on Medical Data,” *Lect. Notes Electr. Eng.*, vol. 698, no. February 2021, pp. 1549–1556, 2021, doi: 10.1007/978-981-15-7961-5_139.
- [16] V. Biksham, V. S. Na, I. Meghana, B. Harshath, and G. Tarun, “Heart Disease Prediction Using Machine Learning,” *Ymer*, vol. 21, no. 4, pp. 489–494, 2022, doi: 10.37896/YMER21.04/48.
- [17] E. Wijaya, “Implementation Analysis of GLCM and Naive Bayes Methods in Conducting Extractions on Dental Image,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 407, no. 1, 2018, doi: 10.1088/1757-899X/407/1/012146.
- [18] J. N. Peksi, B. Yuwono, and Y. M. Florestiyanto, “Classification of Anemia with Digital Images of Nails and Palms using the Naive Bayes Method,” *J. Inform. dan Teknol. Inf.*, vol. 18, no. 1, pp. 118–130, 2021, doi: 10.31515/telematika.v18i1.4587.
- [19] P. Chalekar, S. Shroff, S. Pise, and S. Panicker, “Use of K-Nearest Neighbor in Thyroid Disease Classification,” *Int. J. Curr. Eng. Sci. Res.*, vol. 1, no. 2, pp. 2394–0697, 2014.
- [20] J. M. Martínez-Otzeta and B. Sierra, “Analysis of the iterated probabilistic weighted k nearest neighbor method, a new distance-based algorithm,” *ICEIS 2004 - Proc. Sixth Int. Conf. Enterp. Inf. Syst.*, no. January, pp. 233–240, 2004, doi: 10.5220/0002605402330240.
- [21] A. K. Saputro, K. A. Wibisono, and F. P. Pratiwi, “Identification of Disease Types on Tea - Plant Varieties Based Image Processing with K-Nearest Neighbor Method,” *J. Phys. Conf. Ser.*, vol. 1569, no. 3, 2020, doi: 10.1088/1742-6596/1569/3/032078.

Naive Bayes And K-Nearest Neighbor Algorithm Approach In Data Mining Classification Of Drugs Addictive Diseases

ORIGINALITY REPORT

17%

SIMILARITY INDEX

PRIMARY SOURCES

- 1 dokumen.dinus.ac.id
Internet 33 words — 1%
- 2 Ida Setiani, Meilany Nonsi Tentua, Sunggito Oyama. "Prediction of Banking Stock Prices Using Naïve Bayes Method", Journal of Physics: Conference Series, 2021
Crossref 31 words — 1%
- 3 F. Rouberty, J. Fournier. "Optimization of HPLC Separation of Carbamate Insecticides (Carbofuran, Hydroxycarbofuran and Aldicarb) by Experimental Design Methodology", Journal of Liquid Chromatography & Related Technologies, 2006
Crossref 30 words — 1%
- 4 Yoshiyuki Yamashita. "Supervised learning for the analysis of process operational data", Computers & Chemical Engineering, 2000
Crossref 30 words — 1%
- 5 iocscience.org
Internet 29 words — 1%
- 6 proceeding.unpkediri.ac.id
Internet 26 words — 1%
- 7 jurnal.fikom.umi.ac.id
Internet

		24 words — < 1%
8	www.grafiati.com Internet	24 words — < 1%
9	e-jurnal.stmikbinsa.ac.id Internet	23 words — < 1%
10	Denni Kurniawan, Muhammad Yasir, Farah Chikita Venna. "Optimization of Sentiment Analysis using Naive Bayes with Features Selection Chi-Square and Information Gain for Accuracy Improvement", 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2022 Crossref	22 words — < 1%
11	journal.ikopin.ac.id Internet	22 words — < 1%
12	nlistsp.inflibnet.ac.in Internet	22 words — < 1%
13	ojs.unm.ac.id Internet	22 words — < 1%
14	www.scitepress.org Internet	22 words — < 1%
15	downloads.hindawi.com Internet	21 words — < 1%
16	www.mdpi.com Internet	20 words — < 1%
17	Hafiz Abbad Ur Rehman, Chyi-Yeu Lin, Zohaib Mushtaq, Shun-Feng Su. "Performance Analysis	19 words — < 1%

of Machine Learning Algorithms for Thyroid Disease", Arabian Journal for Science and Engineering, 2021

Crossref

-
- 18 journal.universitasbumigora.ac.id 19 words — < 1%
Internet
-
- 19 Fadhillah Azmi, M Khalil Gibran, Achmad Ridwan. "Enhancing Water Potability Assessment Using Hybrid Fuzzy-Naïve Bayes", Indonesian Journal of Computer Science, 2023 17 words — < 1%
Crossref
-
- 20 Mira Chandra Kirana, Maidel Fani, Tri Shella Kartikasari, Muhammad Nashrullah. "Downtime Data Classification Using Naïve Bayes Algorithm on 2008 ESEC Engine", 2020 3rd International Conference on Applied Engineering (ICAE), 2020 16 words — < 1%
Crossref
-
- 21 eprints.umm.ac.id 16 words — < 1%
Internet
-
- 22 insightsociety.org 16 words — < 1%
Internet
-
- 23 jurnal.murnisadar.ac.id 16 words — < 1%
Internet
-
- 24 ojs.stmikpringsewu.ac.id 16 words — < 1%
Internet
-
- 25 telkomnika.uad.ac.id 16 words — < 1%
Internet
-
- 26 Irfan Pratama, Putri Taqwa Prasetyaningrum. "Pemetaan Profil Mahasiswa Untuk Peningkatan 15 words — < 1%

Strategi Promosi Perguruan Tinggi Menggunakan Predictive Apriori", Jurnal Eksplora Informatika, 2021

Crossref

27 jurnal.dharmawangsa.ac.id 15 words — < 1%
Internet

28 repository.ub.ac.id 15 words — < 1%
Internet

29 wiki.control.fel.cvut.cz 14 words — < 1%
Internet

30 Alaa Menshawi, Mohammad Mehedi Hassan, Nasser Allheeb, Giancarlo Fortino. "A Hybrid Generic Framework for Heart Problem Diagnosis Based on a Machine Learning Paradigm", Sensors, 2023
Crossref

31 openaccess.altinbas.edu.tr 13 words — < 1%
Internet

32 Allan D Alexander, Ratna Salkiawati, Hendarman Lubis, Fathur Rahman, Herlawati Herlawati, Rahmadya Trias Handayanto. "Local Binary Pattern Histogram for Face Recognition in Student Attendance System", 2020 3rd International Conference on Computer and Informatics Engineering (IC2IE), 2020
Crossref

33 kc.umn.ac.id 12 words — < 1%
Internet

34 Yohana Christela Oktaviani, Yulia Wahyuningsih. "FACE EXPRESSION RECOGNIZER DENGAN CONVOLUTIONAL NEURAL NETWORK UNTUK MEMBANTU PENDERITA AUTISME MENGENALI EKSPRESI WAJAH

-
- 35 conference.upnvj.ac.id 11 words — < 1%
Internet
-
- 36 ejournal.raharja.ac.id 11 words — < 1%
Internet
-
- 37 www.onesearch.id 11 words — < 1%
Internet
-
- 38 etheses.uinsgd.ac.id 10 words — < 1%
Internet
-
- 39 Abdulaziz Y.I. Abushawish, Ali Bou Nassif. "Prediction Of Early-Stage Diabetes using machine learning", 2023 Advances in Science and Engineering Technology International Conferences (ASET), 2023 9 words — < 1%
Crossref
-
- 40 Rika Ismayanti, Faza Alameka, Dedy Mirwansyah, Nariza Wanti Wulan Sari, Abdul Rahim, Riyayatsyah. "Identification of Pests on Black Orchid Plants Using Naïve Bayes Method Based on Leaf Image Texture", 2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA), 2022 9 words — < 1%
Crossref
-
- 41 Sabrina Aulia Rahmah, Jovi Antares. "Application of K-Means Clustering Algorithm in Determining Prospective Students Receiving Foundation Scholarship", 2021 International Conference on Computer Science and Engineering (IC2SE), 2021 9 words — < 1%
Crossref

42	garuda.kemdikbud.go.id Internet	9 words — < 1%
43	Devie Rosa Anamisa, Muhammad Yusuf, Wahyudi Agustiono, Mohammad Syarief, Muhammad Ali Syakur, Husna. "Identification of Acute Respiratory Infections in Toddlers Based on the Chi-Square And Naive Bayes Methods", 2021 IEEE 7th Information Technology International Seminar (ITIS), 2021 Crossref	8 words — < 1%
44	addi.ehu.es Internet	8 words — < 1%
45	adoc.pub Internet	8 words — < 1%
46	biblio.vub.ac.be Internet	8 words — < 1%
47	e-journal.hamzanwadi.ac.id Internet	8 words — < 1%
48	ijair.id Internet	8 words — < 1%
49	repository.up.ac.za Internet	8 words — < 1%
50	thuvientailieu.org Internet	8 words — < 1%
51	www.researchgate.net Internet	8 words — < 1%

52 M. Soranamageswari. "Statistical Feature Extraction for Classification of Image Spam Using Artificial Neural Networks", 2010 Second International Conference on Machine Learning and Computing, 02/2010 7 words — < 1%

Crossref

53 Asrul Sani, Samuel, Djaka Suryadi, Firman Noor Hasan, Ade Davy Wiranata, Siti Aisyah. "Predicting the Success of Garment Sales on Transaction Data using the Classification Method with the Naïve Bayes Algorithm", 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), 2023 6 words — < 1%

Crossref

54 Hairani Hairani, Ahmad Zuli Amrullah. "Pelatihan Pengenalan Data Science untuk Meningkatkan Kemampuan dalam Pengolahan Data", Jurnal Abdidas, 2020 6 words — < 1%

Crossref

55 Song, Xiaomu, and Nan-kuei Chen. "A SVM-based quantitative fMRI method for resting-state functional network detection", Magnetic Resonance Imaging, 2014. 6 words — < 1%

Crossref

EXCLUDE QUOTES OFF

EXCLUDE SOURCES OFF

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE MATCHES OFF



Naïve Bayes and K-Nearest Neighbor Approaches in Data Mining Classification of Drugs Addictive Diseases

Dadang Priyanto ^{a,1,*}; Ahmad Robbiul Iman ^{a,2}; Deny Jollyta ^{b,3}

^a Department of Computer Science, Universitas Bumigora, Ismail Marzuki Street, Mataram, 83127, Indonesia.

^b Department of Informatics Engineering, Institut Bisnis Dan Teknologi Pelita Indonesia, Pekanbaru Riau, 28127, Indonesia

¹ dadang.priyanto@universitasbumigora.ac.id; ² robbishimozuki@gmail.com; ³ deny.jollyta@lecturer.pelitaindonesia.ac.id

* Corresponding author

Article history: Received December 06, 2022; Revised January 24, 2023; Accepted June 09, 2023; Available online August 16, 2023

Abstract

Indonesia, with its very large population, is a potential market for drugs trafficking. Hence, seriousness is needed in cracking down or preventing drug trafficking. Narcotics are substances or drugs that can cause dependence or addicted and other negative impacts on users. The problem is that drug users do not realize and even ignore diseases caused by drug addiction. The diseases can be life-threatening for users, such as inflammation of the liver, heart disease, hypertension, stroke, and others. The prevalence rate of drug abuse in West Nusa Tenggara (NTB) is included in the high category, reaching 292 cases or around 37.24% cases. This study aimed to create an application that classify various diseases of drug users using the naïve bayes and KNN methods. The results of this study indicated that there was a very close relationship between drug users and various deadly diseases. The prediction results showed that the naïve bayes method provided a prediction accuracy of 94.5% while the KNN showed a prediction accuracy of 92.5%. This shows that the naïve bayes method provides superior predictive performance than the KNN in the data set of drug addicts in NTB.

Keywords: Drug Addiction; Drug Addiction Disease; Naïve Bayes; Narcotics; K-Nearest Neighbor.

Introduction

Indonesia as an archipelagic country with more than 276 million people. This is a very large population and has great potential for the marketing and distribution of narcotics. *Narkoba* in Indonesia is an abbreviation of narcotics, psychotropics and other addictive substances. The term narcotics itself is a substance or drug that is natural, synthetic, or semi-synthetic if consumed by the human, it can affect the central nervous system, namely the brain, and influence the analysis and thinking ability [1], [2]. Drug or narcotic users will be addicted to these illegal drugs. Drug users can feel excessive anxiety, increased sexual appetite, paranoia, experience delusions, behavior changes towards aggressiveness. Trafficking of narcotics is prohibited because it can cause dependence and various diseases [3], [4]. The problem is that drug users do not realize and even ignore the diseases caused by drug addiction, these diseases can be life-threatening for users, such as inflammation of the liver, heart disease, hypertension, stroke and others. The prevalence rate of drug abuse in West Nusa Tenggara is in the high category, reaching 292 cases or around 37.24% of drug abuse cases, out of a population of approximately 5.1 million people.

This research aimed to create an application that can classify various diseases of drug users based on the naïve bayes and K-Nearest Neighbor (KNN) methods. Classification is a systemic arrangement in groups according to established rules or standards. Literally classification can be said as the division of something according to classes. According to science, classification is the process of grouping objects or data based on their characteristics, similarities, and differences. Processing data with the classification of data mining processing categories, and many methods that can be used in data mining. Data mining is grouped into descriptive data mining and predictive data mining [5]. The naïve bayes and KNN methods can be used in classification. The naïve bayes classifier is based on the static bayes probability theorem which was put forward by the English scientist Thomas bayes. It improves future opportunities based on past experience with the main characteristic of a very strong (naïve) assumption of the independence of each condition or event. [6]–[8]. The KNN method is an algorithm used to classify an object, based

on the k training data that is closest to the object. The condition for the value of k is that it cannot be greater than the number of training data, and the value of k must be odd and more than one. The distance of an object determine its classification, a close distance of an object classified in one group by calculating using the Euclidean Distance [9], [10].

Previous studies have used a variety of data mining methods such as disease diagnosis by using the fuzzy, naïve bayes and KNN comparative methods for smallpox, dengue fever and flu [11]. Another study using the naïve bayes and KNN methods for liver disease classification, which gave the best performance was the KNN algorithm [12], another comparison of classification algorithms in coronary heart disease showed that the random forest algorithm provided better performance in terms of accuracy [13], then a comparative analysis of the classification method for Hepatitis shows that the naïve bayes method has better performance [14], Other comparative studies of data mining algorithms on medical data performance Support Vector Machine (SVM) provide the best accuracy [15], and the use of genetic algorithms and machines learning in the classification and prediction of heart disease [16].

Previous studies have shown that various classification methods provide different best performance results based on the data set used, therefore this study provides a novelty in classification using the naïve bayes and KNN methods on drug addict disease data sets in West Nusa Tenggara region. From the results of the comparison showed the performance of the best method which differed in accuracy from previous studies.

Method

The data mining methods used in this study were a classification category and a grouping of an object into a particular class. Various issues related to grouping objects can be solved by applying classification techniques. In general, classification performance was carried out using the confusion matrix, and this study used the naïve bayes and KNN methods.

A. Naïve Bayes method

Naïve bayes is a simple classifying probabilistic method based on Bayes' theorem where classification is carried out through training sets of a number of data efficiently [17]. Naïve Bayes assumes that the value of an input attribute in a given class does not depend on the values of other attributes [18]. Bayes' theorem itself was put forward by British scientist Thomas bayes, namely predicting future opportunities based on previous experience so that it is known as Bayes' theorem. Where the Bayes theory Equation 1 is.

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (1)$$

Where :

- X : Data with unknown class
- C : Hypothesis data X is a specific class
- $P(C|X)$: Probability of hypothesis C based on condition X (posterior probability)
- $P(C)$: Probability of hypothesis C (prior probability)
- $P(X|C)$: Probability of X based on conditions in hypothesis C
- $P(X)$: Probability of X

B. K-Nearest Neighbor method

KNN is a method using a supervised learning algorithm in which a number of new attributes whose class is unknown can be searched for by looking for similarities based on the majority of KNN as a reference in determining the class of an attribute [19]–[21]. The KNN method can be calculated using the Euclidean distance in determining the distance between datasets. The Euclidean distance formula can be seen in Equation 2.

$$d = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + \dots + (N_1 - N_2)^2} \quad (2)$$

Where:

- d : distance between the two attributes
- X_1 : actual data
- X_2 : data testing
- Y_1 : n^{th} actual data
- Y_2 : n^{th} testing data
- N_1 : first n^{th} attribute
- N_2 : second n^{th} attribute

The flowchart of the calculation process using the K-NN method can be seen in Figure 1.

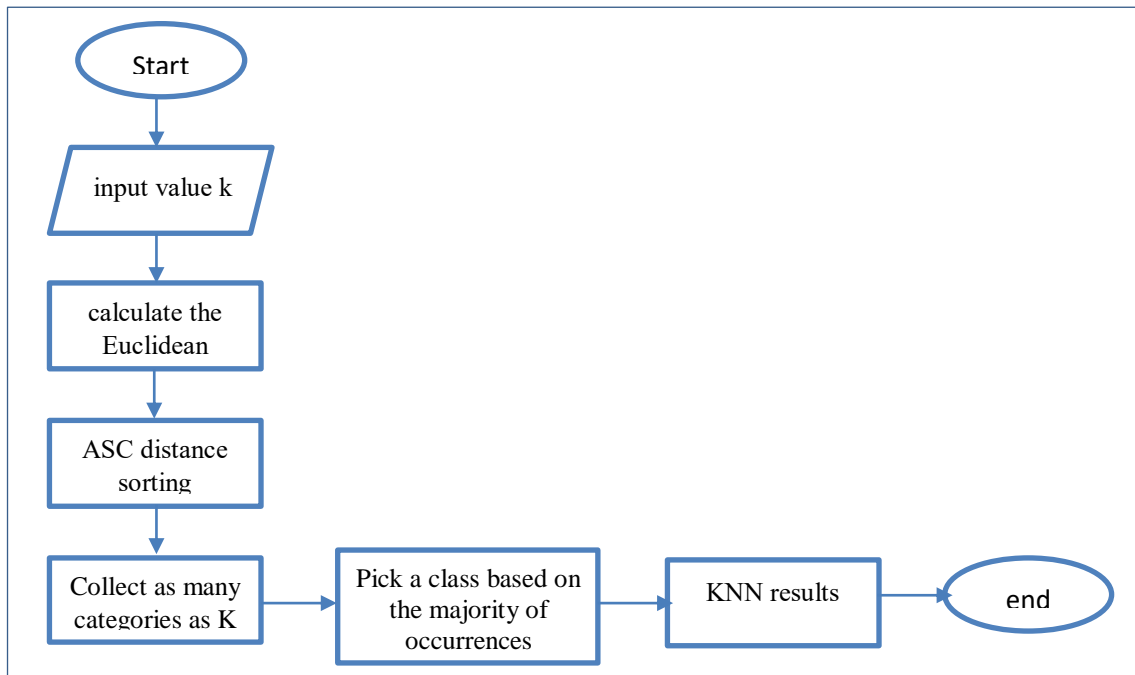


Figure 1. Flowchart of K-Nearest Neighbor

C. Data Set

This study used data taken from the National Narcotics Agency of West Nusa Tenggara Province(BNNP), with data collection using interview techniques and in the form of data files. From the results of the interviews, information was obtained about the symptoms experienced by drug addicts, the types of drugs consumed, the level of addiction, and history of drug addict disease. The data obtained from the results of a survey regarding the types of drug addict disease obtained 336 data, and can be partially presented in tabular form as in Table 1.

Table 1. Drug Addiction Disease Dataset Table

Drug Type	Addiction Level	G001	...	G024	Disease History	Disease
Cocaine	Recreational	Yes	...	No	Stomach Inflammation	GERD
Marijuana	Trying	Yes	...	Yes	Heart disease	Hepatitis
Crystal Meth	Active dependency	No	...	No	Asthma	Acute Asthma
Morfin	Trying	Yes	...	No	Hypertension	Cyst
Gorilla tobacco	Recreational	Yes	...	No	Malaria	Thypoid
...
Gorilla tobacco	Recreational	Yes		Yes	Diabetes	Hepatitis

From Table 1 it can be explained the description of the specified symptom variables as in Table 2 below.

Table 2. Description of symptom variables

Symptoms Code	Symptoms
G001	Nausea and vomiting
G002	Sore throat
G003	Expulsion of stomach contents involuntarily
G004	Bad breath
G005	Coughs
G006	Hard to breathe
G007	Hoarseness
G008	Speech disorder
G009	Hard to sleep
G010	Heart rate increases
G011	Dizzy
G012	Easily Tired
G013	Chest feels tight
G014	Redness of the skin around the cyst area
G015	Blood or pus that smells bad from the lump
G016	Fever

Symptoms Code	Symptoms
G017	Mild or severe pain in the lower abdomen
G018	Stomach ache
G019	Loss of appetite
G020	Joint pain
G021	Diarrhea
G022	Dry Cough
G023	Jaundice
G024	Dark urine

Results and Discussion

A. Preprocessing

Before processing data with the Algorithm, pre-processing was a must. This was to make the data is easier to read, reduce the burden of representation in the data, reduce the duration of data processing, and simplify the process of data analysis by the system. The dataset that was formed must be a dataset that was free from missing values and inconsistent attributes so there would not be interference with the process at the modeling stage. The initial stage of preprocessing was by initializing attributes using the variables "X_train" and "Y_train" as class labels. In the manual calculation of naïve bayes and KNN the attribute must have a numeric value. This is because naïve bayes and KNN utilize Probability and statistical calculations. So the data preprocessing stage for the naïve bayes and KNN algorithms is done by changing the value of the numeric "X_train" attribute. The results of the preprocessing for the naïve bayes and KNN algorithms can be partially seen in [Table 3](#) below.

Table 3 Preprocessing Results

Drugs	Addicted	G001	...	G024	History	Disease
0.84	1.29	1	...	-1	0.25	GERD
0.18	0.06	1	...	1	1.75	Hepatitis
1.54	1.18	-1	...	-1	0.72	Acute Asthma
0.48	0.06	1	...	-1	0.96	Cyst
1.15	1.29	1	...	-1	0.35	Thypoid
...
1.54	1.29	1	...	-1	0.96	Cyst
0.48	0.06	-1	...	-1	0.81	Acute Asthma
1.15	1.29	1	...	1	0.81	Hepatitis

Based on [Table 3](#), the results of data processing describe the condition of the disease based on the type of drug used, the level of addiction, the symptoms experienced ranging from 1 to 24, as well as a history of previous illnesses. For example, record data set 1 indicates that the type of addiction was 0.84, the type of drug used was 'Cocaine', an addiction level was 1.29 categorized as 'Recreational' addiction, and symptoms that were felt according to code G001 were 'Nausea and Vomiting' and so on until symptom 24. If it was combined with history of previous illness it can be concluded the disease was 'GERD'. The same thing can be done in reading the next record in [Table 3](#) above. Information about the attributes used in reading [Table 3](#) can be seen with the attributes of the type of drug, the level of addiction, symptoms of the disease and history of the disease which can be explained in the following [Table 4-7](#).

Table 4. Description of Drug Type Attributes

Initial Attribute Value	Numerical Attribute Values
Crystal meth	1.54
Cocaine	0.84
Morphine	0.48
Gorilla Tobacco	1.15
Marijuana	0.18

[Table 4](#) Provides information regarding the attribute values for the drug types used and numerical values according to the types of drugs used. For example, narcotics of the methamphetamine type have a numerical attribute value of 1.54 and so on.

Table 5. Addiction Level Attribute Description

Initial Attribute Value	Numerical Attribute Values
Try Use	0.06
recreational	1.29
Active Dependence	1.18

Table 5 Provides a description of the level of addiction experienced by users such as a value of 0.06 indicating addiction at the 'Trying' stage and so on.

Table 6. Symptom Description

Initial Attribute Value	Numerical Attribute Values
Yes	1
Not	2

Table 6 Describes information on the symptoms experienced by addicts with a determination value of 1 'Yes' experience' and 2 'No' experience', while the symptoms experienced refer to **Table 2** above.

Table 7. Description of Disease History

Initial Attribute Value	Numerical Attribute Value
Asthma	0.72
Stomach Inflammation	0.25
Heart disease	1.75
Malaria	0.35
Pneumonia	1.66
Tuberculosis	0.44
Diabetes	0.81
Hypertension	0.96
There isn't any	0

Table 7 describes the previous history of the disease experienced by the addict as an additional variable in determining the conclusion of the illness.

B. Calculation results of the K-Nearest Neighbor method

The results of calculations with the KNN from the amount of existing data, 10 testing data were needed to be calculated with the results as in **Table 8**.

Table 8. Data Testing

Drug Type	Addiction Level	G001	...	G024	Disease History	Disease
0.48	1.18	1	...	-1	0	Thypoid
1.54	1.29	1	...	-1	1.75	GERD
1.15	0.06	-1	...	-1	0.96	Hepatitis
0.48	1.29	-1	...	-1	0	GERD
0.84	1.29	1	...	-1	0.81	Cyst
0.18	0.06	-1	...	1	1.75	Cyst
0.18	1.29	1	...	-1	0.81	Acute Asthma
1.15	0.06	1	...	1	1.75	Hepatitis
1.15	1.29	-1	...	-1	1.75	Cyst
1.54	0.06	1	...	-1	0.35	GERD

As seen in **Table 8**, it can be concluded that the results of testing from 10 data, with varying levels of disease experienced by addicts such as typhoid, GERD, and others according to the drugs consumed and their addiction level. From table 8 it can be calculated to determine the closest majority K value similarity or based on the distance between data sets with the Euclidean distance formula and the results were shown in **Table 9**.

Table 9. Euclidean distance result

Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
0.677495387	1.18156676	1.64693655	1.834557167	1.4066272	1.731964203	0.43104524	1.614218077	1.0694391	1.33574
1.094257739	1.34331679	1.52715422	1.698293261	1.5270887	2.021113554	0.67749539	1.771242502	0.8258935	1.22793
1.923174459	0.88977525	1.09535382	0.849352695	0.8713208	2.141144554	1.11919614	1.345696846	0.5374942	0.47223
1.718924082	0.44922155	1.83338485	1.697910481	0.4114608	2.255393536	0.94175368	1.67002994	0.9393615	0.54452
1.479932431	1.70120545	0.66272166	1.04517941	1.8645643	1.638566447	1.21268298	1.315028517	1.1157509	1.59706
1.706399719	0.27037012	1.53590364	1.3713497	0.2014944	2.021014597	0.91869473	1.336674979	0.868159	0.50359
1.263566381	1.59176003	1.56626945	1.762498227	1.7652762	2.167694628	0.93648278	1.934760967	0.9404786	1.39893
0.992219734	1.22711043	1.16357209	1.380543371	1.4259383	1.683627037	0.55506756	1.401784577	0.7598684	1.19537
1.158015544	1.3095419	1.66679333	1.824883558	1.4974645	0.671490879	1.23798223	0.820548597	1.8714967	1.91348

Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
2.067147793	1.10498869	0.93193347	0.624579859	1.0901835	1.801887899	1.43405021	0.872811549	1.1983322	1.13723
0.517010638	1.05763888	1.42369238	1.637101096	1.3042622	1.288953064	0.37802116	1.206067991	1.1540364	1.38351
0.933380951	1.31244047	1.20307938	1.44934468	1.5182226	1.65562677	0.60712437	1.434747365	0.8673523	1.30407
1.741895519	0.53037722	1.77476759	1.634441801	0.4987986	2.298543017	0.93455872	1.693664666	0.8190849	0.41833
1.293290377	1.40730238	0.74040529	1.04890419	1.5836666	0.881873007	1.1319894	0.516623654	1.3944533	1.6533
...
1.443502684	1.64972725	0.52316345	0.962756459	1.8177184	1.251958466	1.257417989	0.940638081	1.3234047	1.70678

From **Table 9** the selection of the smallest distance value that was included in the ranking of a number of K values is shown. In **Table 9** the value of K is 10, so you have to find the smallest distance value of 10. Furthermore, the results of the highest number of classes from a predetermined K value can be obtained as predicted in **Table 10** below.

Table 10. The highest number of classes from the value of K

N th Test Data	Original Class	Acute Asthma	Cyst	GERD	Hepatitis	Typoid
1	Thypoid	False	False	False	False	True
2	Hepatitis	True	False	False	False	False
3	Asma Akut	False	False	False	True	False
4	Kista	False	False	False	True	False
5	Asma Akut	True	False	False	False	False
6	Hepatitis	False	False	False	False	False
7	Kista	True	False	False	False	False
8	Asma Akut	False	False	False	False	False
9	GERD	False	False	True	False	False
10	GERD	False	False	True	False	False

Table 10 explains that the test results indicate a 'True' value, so addicts are diagnosed with a disease according to the predicted results as presented in **Table 9** test results.

C. Calculation results with Naive Bayes

The calculation process with the naïve bayes method was started by retrieving training data from the data that had been obtained. The variables that were used in the classification of drug addicts were disease, drug types, level of drug addiction, illness records, and the results of the calculation showed that the average probability of the disease class was 20% while for the probability of drug type attributes as shown in **Table 11**.

Table 11. Attribute Probability of Drug Types

Drug Type	Acute Asthma	GERD	Typhoid	Cyst	Hepatitis
Crystal meth	10%	30%	30%	30%	10%
Cocaine	50%	20%	20%	20%	20%
Morphine	0%	10%	0%	10%	10%
Gorilla Tobacco	30%	30%	40%	10%	0%
Marijuana	10%	10%	10%	30%	60%
Total number	100%	100%	100%	100%	100%

Table 11. Attribute Probability of Drug Types

Addiction Disease	Acute Asthma	GERD	Typhoid	Cyst	Hepatitis
Recreational	40%	30%	20%	30%	40%
Active dependency	0%	30%	50%	30%	40%
Try it on	60%	40%	30%	40%	20%
Total	100%	100%	100%	100%	100%

The resulting model was evaluated with the aim of assessing the performance of the naïve bayes algorithm and the KNN algorithm. Evaluation was conducted by calculating the accuracy, precision, and recall of each algorithm used using the Confusion Matrix Technique. If the evaluation results showed good performance, then the two algorithms were feasible in making predictions on new data whose labels were unknown.

D. Testing the Naïve Bayes

At this stage the test was carried out on the performance of the naïve bayes classification model made when making predictions on data testing, using the confusion matrix and the confusion matrix plot. The test results can be seen in **Figure 2** of the confusion chart.

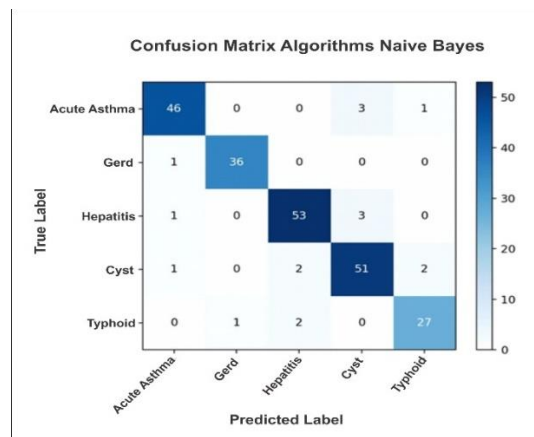


Figure 2. Naive Bayes Confusion Matrix Graph

In the confusion matrix f above, it can be seen that the prediction for the Acute Asthma class has a true value of 46, the prediction for the GERD has a value of 36, the prediction for the Hepatitis has a value of 53, the prediction for the class of Cyst is 51, and the prediction for the Typhoid class with a value of 27, wrong prediction in the Acute Asthma class by 4, GERD class by 1, Hepatitis class by 4, Cyst class by 2, and Typhoid class by 3. The results of the calculation of the Naive Bayes confusion matrix can be seen in [Table 12](#).

Table 12. Confusion Matrix Naïve Bayes

		Acute Asthma	Gerd	Typhoid	Cyst	Hepatitis
Predictions	Acute Asthma	46	0	0	3	1
	GERD	1	36	0	0	0
	Hepatitis	1	0	53	3	0
	Cyst	1	0	2	51	0
	Typhoid	0	1	2	0	27

Based on [Table 12](#), prediction accuracy can be calculated using the naïve bayes algorithm which produces a prediction accuracy of 93% as in the following [Equation 3](#) calculation.

$$\text{Prediction Accuracy} = \frac{46+36+53+51+27}{230} = 0,93\% \quad (3)$$

E. Comparison Results of the Naïve Bayes and K-Nearest Neighbor Methods

This test was carried out by dividing training data and data testing into 4 variations of data, and testing was carried out on naïve bayes and KNN modeling on variations in the percentage of training data and testing data. The results of testing with the confusion matrix can be seen in the comparison of accuracy, precision and recall of the two algorithms as shown in [Table 13](#).

Table 13. Results of Comparison of Accuracy, Precision and Recall

Data Training	Data Testing	Accuracy (%)		Precision (%)		Recall (%)	
		Naïve Bayes	K-NN	Naïve Bayes	K-NN	Naïve Bayes	K-NN
335	201	93	93	93	93	93	93
335	180	96	90	95	90	95	90
335	230	93	91	93	91	93	91
335	181	96	96	96	96	96	96
Average		94.5	92.5	94.25	92.5	94.25	92.5

Based on [Table 13](#), the calculations after being validated showed that the naïve bayes method had a higher accuracy than the KNN, with a value of 94.5%, while KNN after being validated had an accuracy value of 92.5%.

Conclusion

Narcotics are drugs that are prohibited from being widely distributed, because they can lead to addiction for their users and can cause various deadly diseases. This research was conducted to classify various diseases of Narcotic addicts in the form of application programs using two methods, namely naïve bayes and KNN. This research has shown that there was a very close relationship between drug users and various deadly diseases. The prediction results showed that the naïve bayes method provided a prediction accuracy of 94.5% while the KNN showed a prediction accuracy of 92.5%. This shows that the naïve bayes method provides better predictive performance than the KNN with the data set used in this study.

References

- [1] B. Priambodo, Y. Jumaryadi, and U. Salamah, "Penggunaan Aplikasi Deteksi Pecandu Narkoba Di Meruya Utara," *J. Pasopati Pengabd. Masy. dan Inov. Pengemb. Teknol.*, vol. 4, no. 3, 2022, doi: [10.14710/pasopati.2022.14468](https://doi.org/10.14710/pasopati.2022.14468).
- [2] M. Mustaqim, H. Dafitri, and D. Dharmawati, "Edukasi Digital Pengenalan Bahaya Narkoba Bagi Anak Usia Dini Berbasis 3D Dan Augmented Reality," *Djtechno J. Teknol. Inf.*, vol. 2, no. 2, pp. 170–176, 2021, doi: [10.46576/djtechno.v2i2.1623](https://doi.org/10.46576/djtechno.v2i2.1623).
- [3] B. C. Laksono and N. W. K. Projo, "Pemodelan Analisis Rantai Markov untuk Mengestimasi Potensi Kasus Narkoba di Indonesia," *Semin. Nas. Off. Stat.*, vol. 2021, no. 1, pp. 715–722, 2021, doi: [10.34123/semnasoffstat.v2021i1.1016](https://doi.org/10.34123/semnasoffstat.v2021i1.1016).
- [4] A. Sinjar and T. Sahuri, "Bahaya Narkoba Terhadap Masa Depan Generasi Muda," vol. 2, no. 2, pp. 154–160, 2021.
- [5] D. Priyanto, M. Zarlis, H. Mawengkang, and S. Efendi, "Analysis of earthquake hazards prediction with multivariate adaptive regression splines," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 3, pp. 2885–2893, 2022, doi: [10.11591/ijece.v12i3.pp2885-2893](https://doi.org/10.11591/ijece.v12i3.pp2885-2893).
- [6] H. Hairani, G. S. Nugraha, M. N. Abdillah, and M. Inuddin, "Komparasi Akurasi Metode Correlated Naive Bayes Classifier dan Naive Bayes Classifier untuk Diagnosis Penyakit Diabetes," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 3, no. 1, pp. 6–11, 2018, doi: [10.30743/infotekjar.v3i1.558](https://doi.org/10.30743/infotekjar.v3i1.558).
- [7] Y. F. Safri, R. Arifudin, and M. A. Muslim, "K-Nearest Neighbor and Naive Bayes Classifier Algorithm in Determining The Classification of Healthy Card Indonesia Giving to The Poor," *Sci. J. Informatics*, vol. 5, no. 1, p. 18, 2018, doi: [10.15294/sji.v5i1.12057](https://doi.org/10.15294/sji.v5i1.12057).
- [8] R. M. Abarca, "Sistem Pakar Diagnosis Penyakit Mata," *Nuevos Sist. Comun. e Inf.*, no. sistem pakar penyakit mata, pp. 2013–2015, 2021.
- [9] A. H. Bisri Merluarini, Diah Safitri, "Perbandingan Analisis Klasifikasi menggunakan Metode K-Nearest Neighbor (K-Nn) dan Multivariate Adaptive Regression Spline (Mars) pada Data Akreditasi Sekolah Dasar Negeri di Kota Semarang," vol. 3, pp. 313–322, 2014.
- [10] P. I. Lestari, D. E. Ratnawati, and L. Muflikhah, "Implementasi Algoritme K-Means Clustering Dan Naive Bayes Classifier Untuk Klasifikasi Diagnosa Penyakit Pada Kucing," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 3, no. 1, pp. 968–973, 2019.
- [11] A. Rajkumar, "Comparison of Fuzzy Diagnosis with K-Nearest Neighbor and Naive Bayes-2010.pdf," *Glob. J. Comput. Sci. Technol.*, vol. 10, no. 10, pp. 38–43, 2010.
- [12] A. Desiani, "Perbandingan Implementasi Algoritma Naive Bayes dan K-Nearest Neighbor Pada Klasifikasi Penyakit Hati," *Simkom*, vol. 7, no. 2, pp. 104–110, 2022, doi: [10.51717/simkom.v7i2.96](https://doi.org/10.51717/simkom.v7i2.96).
- [13] A. B. Wibisono and A. Fahrurrozi, "Perbandingan Algoritma Klasifikasi Dalam Pengklasifikasian Data Penyakit Jantung Koroner," *J. Ilm. Teknol. dan Rekayasa*, vol. 24, no. 3, pp. 161–170, 2019, doi: [10.35760/tr.2019.v24i3.2393](https://doi.org/10.35760/tr.2019.v24i3.2393).
- [14] S. Sulastri, K. Hadiono, and M. T. Anwar, "Analisis Perbandingan Klasifikasi Prediksi Penyakit Hepatitis Dengan Menggunakan Algoritma K-Nearest Neighbor, Naive Bayes Dan Neural Network," *Dinamik*, vol. 24, no. 2, pp. 82–91, 2020, doi: [10.35315/dinamik.v24i2.7867](https://doi.org/10.35315/dinamik.v24i2.7867).
- [15] A. Suragala, P. Venkateswarlu, and M. China Raju, "A Comparative Study of Performance Metrics of Data Mining Algorithms on Medical Data," *Lect. Notes Electr. Eng.*, vol. 698, no. February 2021, pp. 1549–1556, 2021, doi: [10.1007/978-981-15-7961-5_139](https://doi.org/10.1007/978-981-15-7961-5_139).
- [16] V. Biksham, V. Srujana, I. Meghana, B. Harshath, and G. Tarun, "Heart Disease Prediction Using Machine Learning," *Ymer*, vol. 21, no. 4, pp. 489–494, 2022, doi: [10.37896/YMER21.04/48](https://doi.org/10.37896/YMER21.04/48).
- [17] E. Wijaya, "Implementation Analysis of GLCM and Naive Bayes Methods in Conducting Extractions on Dental Image," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 407, no. 1, 2018, doi: [10.1088/1757-899X/407/1/012146](https://doi.org/10.1088/1757-899X/407/1/012146).
- [18] J. N. Peksi, B. Yuwono, and Y. M. Florestiyanto, "Classification of Anemia with Digital Images of Nails and Palms using the Naive Bayes Method," *J. Inform. dan Teknol. Inf.*, vol. 18, no. 1, pp. 118–130, 2021, doi: [10.31515/telematika.v18i1.4587](https://doi.org/10.31515/telematika.v18i1.4587).
- [19] P. Chalekar, S. Shroff, S. Pise, and S. Panicker, "Use of K-Nearest Neighbor in Thyroid Disease Classification," *Int. J. Curr. Eng. Sci. Res.*, vol. 1, no. 2, pp. 2394–0697, 2014.

-
- [20] J. M. Martínez-Otzeta and B. Sierra, "Analysis of the iterated probabilistic weighted k nearest neighbor method, a new distance-based algorithm," *ICEIS 2004 - Proc. Sixth Int. Conf. Enterp. Inf. Syst.*, no. January, pp. 233–240, 2004, [doi: 10.5220/0002605402330240](https://doi.org/10.5220/0002605402330240).
- [21] A. K. Saputro, K. A. Wibisono, and F. P. Pratiwi, "Identification of Disease Types on Tea - Plant Varieties Based Image Processing with K-Nearest Neighbor Method," *J. Phys. Conf. Ser.*, vol. 1569, no. 3, 2020, [doi: 10.1088/1742-6596/1569/3/032078](https://doi.org/10.1088/1742-6596/1569/3/032078).