PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA NAÏVE BAYES, C4.5, DAN K-NEAREST NEIGHBOR UNTUK KLASIFIKASI KEMISKINAN DI DKI JAKARTA

Denis Nila Cahyani

41820010111

Aisyah Tri Deanita

41820010122

Muhammad Anand Rizki Andinta

41820010069

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2023

PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA NAÏVE BAYES, C4.5, DAN K-NEAREST NEIGHBOR UNTUK KLASIFIKASI KEMISKINAN DI DKI JAKARTA

Laporan Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Denis Nila Cahyani
41820010111
Aisyah Tri Deanita
41820010122
Muhammad Anand Rizki Andinta
41820010069

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2023

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini:

Nama Mahasiswa (1)	:	Denis Nila Cahyani
NIM	:	41820010111
Nama Mahasiswa (2)	:	Aisyah Tri Deanita
NIM	:	41820010122
Nama Mahasiswa (3)	:	Muhammad Anand Rizki Andinta
NIM	:	41820010069
Judul Tugas Akhir	:	Penerapan Data Mining Menggunakan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kemiskinan Di DKI Jakarta

Menyatakan bahwa Tugas Akhir ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Tugas Akhir saya terdapat terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 19 Januari 2024

Denis Nila Cahyani

LEMBAR PENGESAHAN

Nama Mahasiswa (1)	:	Denis Nila Cahyani
NIM		41820010111
Nama Mahasiswa (2)	:	Aisyah Tri Deanita
NIM		41820010122
Nama Mahasiswa (3)	:	Muhammad Anand Rizki Andinta
NIM		41820010069
Judul Tugas Akhir	:	Penerapan Data Mining Menggunakan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kemiskinan di DKI Jakarta

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 20 Desember 2023

Menyetujui

Pembimbing	: Dr. Bambang Jokonowo, S.Si., M.T.I.	( )
NIDN	: 0320037002	
Ketua Penguji	: Ratna Mutu Manikam, S.Kom, MT	( )
NIDN	: 0308017101	
Penguji 1	: Rinto Priambodo, ST, MTI	( )
NIDN	: 0327057905	
Penguji 2	: Abdi Wahab, S.Kom, MT	( )
NIDN	: 0305068502	

Mengetahui,

Dr. Bambang Jokonowo, S.Si., M.T.I.
Dekan Fakultas Ilmu Komputer

Dr. Ruci Meiyanti, M.Kom
Ka.Prodi Sistem Informasi

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini yang berjudul "Penerapan Data Mining Menggunakan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kemiskinan di DKI Jakarta". Penulis menyadari bahwa tanpa bantuan dan bimbingan tugas akhir ini mungkin tidak dapat diselesaikan. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Dr. Bambang Jokonowo, S.Si., M.T.I., selaku Dekan Fakultas Ilmu Komputer dan Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan Tugas Akhir ini.
2. Dr. Ruci Meiyanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi atas dukungan dan bimbingannya selama saya menjalani perkuliahan.
3. Kepada pihak Keluarga khususnya kedua Orang Tua kami yang tanpa henti memberikan dukungan, semangat, dan doa yang sangat luar biasa kepada penulis.
4. Kepada Rekan Tim yang telah terlibat dalam penyusunan Tugas Akhir ini hingga dapat terselesaikan.
5. Kepada Seluruh pihak yang membantu selama proses pengerjaan Tugas Akhir yang tidak dapat disebutkan satu persatu.

Akhir kata, penulis berharap Tugas Akhir ini bisa bermanfaat untuk semua pihak. Selain itu, kritik dan saran yang membangun sangat penulis harapkan dari para pembaca sekalian agar Tugas Akhir ini bisa lebih baik lagi.

Jakarta, 8 Desember 2023

Penulis

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan dibawah ini:

Nama Mahasiswa (1)	:	Denis Nila Cahyani
NIM	:	41820010111
Nama Mahasiswa (2)	:	Aisyah Tri Deanita
NIM	:	41820010122
Nama Mahasiswa (3)	:	Muhammad Anand Rizki Andinta
NIM	:	41820010069
Judul Tugas Akhir	:	Penerapan Data Mining Menggunakan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kemiskinan Di DKI Jakarta

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 19 Januari 2024

Yang menyatakan,

Denis Nila Cahyani

ABSTRAK

Nama Mahasiswa (1)	: Denis Nila Cahyani
NIM	: 41820010111
Nama Mahasiswa (2)	: Aisyah Tri Deanita
NIM	: 41820010122
Nama Mahasiswa (3)	: Muhammad Anand Rizki Andinta
NIM	: 41820010069
Pembimbing TA	: Dr. Bambang Jokonowo, S.Si., M.T.I.
Judul Tugas Akhir	: Penerapan Data Mining Menggunakan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kemiskinan di DKI Jakarta

Kemiskinan masih menjadi salah satu permasalahan fundamental yang sulit dihadapi, termasuk di Provinsi DKI Jakarta sebagai Ibu kota negara Indonesia, yang juga tidak luput dari permasalahan kemiskinan. Penelitian ini menggunakan *data mining* untuk melakukan klasifikasi terhadap tingkat kemiskinan di Provinsi DKI Jakarta dengan data yang diperoleh dari Badan Pusat Statistik (BPS). Tujuan penelitian ini adalah untuk secara efektif dan akurat mengklasifikasikan data kemiskinan dan membandingkan hasil dari algoritma Naïve Bayes, C4.5N dan K-NN guna mencari hasil akurasi terbaik. Melalui penerapan algoritma tersebut, didapatkan hasil bahwa performa algoritma dapat dipengaruhi oleh pembagian data, dan peningkatan rasio data uji cenderung meningkatkan akurasi serta konsistensi hasil klasifikasi. Pada rasio pembagian data 70:30, Naïve Bayes mencapai akurasi 81%, C4.5 sebesar 76%, dan K-NN sebesar 71%. Pada rasio 80:20, Naïve Bayes menunjukkan akurasi 93%, C4.5 sebesar 79%, dan K-NN sebesar 86%. Sementara pada rasio 90:10, Naïve Bayes mencapai akurasi 100%, C4.5 sebesar 71%, dan K-NN sebesar 86% Dengan mempertimbangkan variasi performa ketiga algoritma, dapat disimpulkan bahwa Naïve Bayes lebih unggul sebagai algoritma yang stabil dan dapat diandalkan dalam berbagai skenario pembagian dataset, menunjukkan akurasi tinggi dan kemampuan baik dalam mengidentifikasi kasus positif.

Kata kunci:

kemiskinan, klasifikasi, Naïve Bayes, C4.5, K-Nearest Neighbor

ABSTRACT

Name (1)	: Denis Nila Cahyani
Student Number	: 41820010111
Name (2)	: Aisyah Tri Deanita
Student Number	: 41820010122
Name (3)	: Muhammad Anand Rizki Andinta
Student Number	: 41820010069
Counsellor	: Dr. Bambang Jokonowo, S.Si., M.T.I.
Title	: Penerapan Data Mining Menggunakan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kemiskinan di DKI Jakarta

Poverty remains one of the fundamental problems that is difficult to overcome, including in DKI Jakarta Province as the capital of Indonesia, which is also not immune to poverty. This research uses data mining to classify the poverty level in DKI Jakarta Province with data obtained from the Central Bureau of Statistics (BPS). The purpose of this study is to effectively and accurately classify poverty data and compare the results of the Naive Bayes, KNN, and C4.5 algorithms to find the best accuracy results. Through the application of such algorithms, the results were obtained that the performance of the algorithm could be influenced by the division of data, and increased ratio of test data tended to improve the accuracy as well as consistency of classification results. At the data division ratio of 70:30, Naïve Bayes achieved accuracy of 81%, C4.5 of 76%, and K-NN of 71%. At the ratio of 80:20, Naïve Bayes showed accurateness of 93%, C4.5 of 79%, and K-NN of 86%. Whereas at the ratios of 90:10, Naïve Bayes achieves accurateness of 100%, C4.5 of 71%, and K-NN of 86%. By considering the performance variations of the third algorithm, it can be concluded that Naïve Bayes is superior as a stable and reliable algorithm in a variety of data set splitting scenarios, showing high accuracy and good ability in identifying positive cases.

Keywords:

poverty, classification, Naïve Bayes, C4.5, K-Nearest Neighbor

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN