

**IMPLEMENTASI METODE K-MEANS CLUSTERING DAN NAIVE
BAYES UNTUK PREDIKSI BALITA TERIDENTIFIKASI STUNTING
(STUDI KASUS BALAI PENYULUHAN KB KEC. DOLOK
SIGOMPULON)**

SKRIPSI

Untuk Memenuhi Persyaratan Memperoleh Gelar Sarjana (S1)

Program Studi Sistem Informasi Fakultas Sains dan Teknologi

Universitas Labuhanbatu



OLEH:

EGGI OK PERNANDA GAJA

2009102160

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS LABUHANBATU
RANTAUPRAPAT
2024**

LEMBAR PENGESAHAN/PERSETUJUAN TUGAS AKHIR

NAMA : EGGI OK PERNANDA GAJA
NPM : 2009102160
PROGRAM STUDI : S1 SISTEM INFORMASI
JUDUL SKRIPSI : IMPLEMENTASI METODE K-MEANS
CLUSTERING DAN NAIVE BAYES UNTUK
PREDIKSI BALITA TERIDENTIFIKASI STUNTING
(STUDI KASUS BALAI PENYULUHAN KB KEC.
DOLOK SIGOMPULON)

Disetujui pada Tanggal : 28 Agustus 2024

Pembimbing I



Syaiful Zuhri Harahap, S.Kom., M.Kom
NIDN. 0113129103

Pembimbing II



Rahma Muti'ah, S.Psi., M.Psi
NIDN. 0114068501

LAPORAN PENGESAHAN NASKAH TUGAS AKHIR

NAMA : EGGI OK PENANDA GAJA
NPM : 2009102160
PROGRAM STUDI : S1 SISTEM INFORMASI
JUDUL : IMPLEMENTASI METODE K-MEANS
CLUSTERING DAN NAIVE BAYES
UNTUK PREDIKSI BALITA
TERIDENTIFIKASI STUNTING (STUDI
KASUS BALAI PENYULUHAN KB KEC.
DOLOK SIGOMPULON)

Telah Diuji dan Dinyatakan Lulus pada Ujian Sidang Sarjana (S1)
Pada Tanggal 28 Agustus 2024

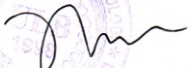
TIM PENGUJI

Penguji I (Ketua) Tanda Tangan
Nama : Syaiful Zuhri Harahap, S.Kom., M.Kom
NIDN : 0113129103

Penguji II (Anggota)
Nama : Rahma Muti'ah, S.Psi., M.Psi
NIDN : 0114068501

Penguji III (Anggota)
Nama : Irmayanti, S.Si., M.Pd
NIDN : 0124088404

Rantauprapat, 28 Agustus 2024


Kepala Program Studi Sistem Informasi
Budianto Bangun, S.Sos., M.Kom
NIDN. 0124047003


Dekan Fakultas Sains dan Teknologi
Dr. Iwan Purnama, S.Kom., M.Kom
NIDN. 0112029202

PERNYATAAN

Saya yang bertanda tangan dibawah ini:

NAMA : EGGI OK PERNANDA GAJA

NPM : 2009102160

JUDUL : IMPLEMENTASI METODE K-MEANS CLUSTERING DAN
NAIVE BAYES UNTUK PREDIKSI BALITA
TERIDENTIFIKASI STUNTING (STUDI KASUS BALAI
PENYULUHAN KB KEC. DOLOK SIGOMPULON)

Dengan ini menyatakan bahwa artikel ilmiah ini disusun sebagai syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Sains dan Teknologi di Universitas Labuhanbatu adalah hasil karya penulis sendiri. Semua kutipan maupun rujukan dalam penulisan artikel ilmiah ini telah penulis cantumkan sumbernya dengan benar sesuai dengan ketentuan yang berlaku. Jika dikemudian hasil ternyata ditemukan seluruh atau sebagian artikel ilmiah ini bukan hasil karya tulis atau plagiat, penulis bersedia menerima sanksi pencabutan gelar Akademik yang disandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Rantauprapat, 28 Agustus 2024

Yang Membuat Pernyataan,



EGGI OK PERNANDA GAJA

NPM. 2009102160

DEDIKASI

1. Tuhan Yang Maha Esa Terima kasih yang sebesar-besarnya atas berkah, bimbingan, dan rahmat yang senantiasa mengiringi setiap langkah hidup saya hingga saya dapat menyelesaikan penelitian ini.
2. Pendiri Universitas Labuhanbatu Rasa hormat dan terima kasih yang tulus saya haturkan atas visi dan dedikasi dalam mendirikan lembaga pendidikan ini, yang telah memberikan saya kesempatan untuk menuntut ilmu dan berkembang.
3. Ketua Yayasan Universitas Labuhanbatu Terima kasih atas segala dukungan, kebijakan, dan perhatian yang diberikan kepada seluruh civitas akademika, sehingga proses pembelajaran dapat berjalan dengan baik dan kondusif.
4. Kepada Universitas Labuhanbatu Terima kasih atas lingkungan akademik yang mendukung dan fasilitas yang memungkinkan saya untuk terus belajar dan berkarya.
5. Rektor Universitas Labuhanbatu Ucapan terima kasih saya sampaikan atas arahan, kebijakan, dan kepemimpinan yang bijaksana dalam mengembangkan universitas ini.
6. Dosen yang Mengajar Terima kasih yang mendalam kepada para dosen yang telah membagikan ilmunya dengan penuh kesabaran dan dedikasi, serta terus mendorong saya untuk berpikir kritis dan kreatif.
7. Kepada Kabiro dan Staf di Universitas Labuhanbatu Terima kasih atas bantuan dan pelayanan yang diberikan selama saya menempuh pendidikan, yang telah memudahkan setiap proses administrasi dan kegiatan akademik.

8. Ayah dan Ibu Tercinta Terima kasih yang tak terhingga atas cinta, dukungan, doa, dan pengorbanan yang telah diberikan sepanjang hidup saya. Segala pencapaian ini saya persembahkan untuk kalian.
9. Abang dan Kakak Tercinta Terima kasih atas segala dukungan dan inspirasi yang diberikan, yang selalu menjadi penyemangat dalam setiap langkah saya.
10. Teman-Teman Terima kasih atas persahabatan, dukungan, dan kebersamaan yang tak ternilai selama menjalani masa-masa sulit dan penuh tantangan di bangku kuliah.
11. Almamater dan Kampus Terima kasih atas semua ilmu, pengalaman, dan kesempatan yang telah diberikan. Almamater ini akan selalu menjadi bagian dari perjalanan hidup dan karier saya.

ABSTRAK

Penelitian ini bertujuan untuk menganalisis stunting pada balita dengan menggunakan metode K-Means untuk clustering dan Naive Bayes untuk klasifikasi. Tujuan utama dari penelitian ini adalah untuk mengidentifikasi pola distribusi stunting pada balita dan mengklasifikasikan balita berdasarkan faktor-faktor risiko yang relevan. Dengan menggabungkan kedua metode ini, penelitian ini berupaya menyediakan informasi yang komprehensif tentang kondisi stunting, yang dapat digunakan oleh tenaga kesehatan dan pembuat kebijakan untuk mengimplementasikan intervensi yang lebih tepat sasaran. Selain itu, penelitian ini juga bertujuan untuk mengukur tingkat akurasi dan efektivitas dari kombinasi metode K-Means dan Naive Bayes dalam menangani data stunting pada balita. Metode yang digunakan dalam penelitian ini melibatkan beberapa tahapan. Pertama, data balita dikumpulkan dan dipersiapkan untuk analisis, termasuk pembersihan data dan normalisasi. Setelah itu, metode K-Means diterapkan untuk mengelompokkan data balita berdasarkan karakteristik yang telah ditentukan, seperti usia, berat badan, tinggi badan, dan status gizi. Proses clustering ini bertujuan untuk mengidentifikasi kelompok balita yang berisiko tinggi terhadap stunting. Setelah mendapatkan hasil clustering, metode Naive Bayes digunakan untuk mengklasifikasikan data berdasarkan faktor risiko stunting, termasuk pola makan, lingkungan sosial, dan faktor kesehatan lainnya. Proses klasifikasi ini bertujuan untuk memberikan prediksi mengenai status stunting pada balita baru yang tidak ada dalam data pelatihan. Hasil dari analisis ini menunjukkan bahwa kombinasi metode K-Means dan Naive Bayes menghasilkan akurasi yang cukup baik, mencapai 86%. Penggunaan perhitungan manual dalam proses analisis juga memberikan hasil yang akurat, memperkuat validitas temuan penelitian ini. Metode K-Means berhasil mengidentifikasi kelompok balita yang paling berisiko terkena stunting, sementara Naive Bayes secara efektif mengklasifikasikan balita ke dalam kategori stunting dan non-stunting berdasarkan variabel-variabel yang telah dianalisis. Dengan hasil yang menjanjikan ini, metode yang digunakan dapat menjadi alat yang berharga bagi praktisi kesehatan dalam mendeteksi dan menangani kasus stunting, serta membantu dalam merancang intervensi yang lebih tepat dan efisien untuk mengurangi angka stunting pada balita.

Kata kunci: Metode K-Means; Metode Naive Bayes; Stunting Balita

ABSTRACT

This study aims to analyze stunting in toddlers using the K-Means method for clustering and Naive Bayes for classification. The main objectives of this study are to identify the distribution patterns of stunting in toddlers and classify toddlers based on relevant risk factors. By combining these two methods, this study seeks to provide comprehensive information on stunting conditions, which can be used by health workers and policy makers to implement more targeted interventions. In addition, this study also aims to measure the level of accuracy and effectiveness of the combination of the K-Means and Naive Bayes methods in handling stunting data in toddlers. The method used in this study involves several stages. First, toddler data is collected and prepared for analysis, including data cleaning and normalization. After that, the K-Means method is applied to group toddler data based on predetermined characteristics, such as age, weight, height, and nutritional status. This clustering process aims to identify groups of toddlers at high risk of stunting. After obtaining the clustering results, the Naive Bayes method is used to classify data based on stunting risk factors, including diet, social environment, and other health factors. This classification process aims to provide predictions regarding the stunting status of new toddlers that are not in the training data. The results of this analysis show that the combination of the K-Means and Naive Bayes methods produces quite good accuracy, reaching 86%. The use of manual calculations in the analysis process also provides accurate results, strengthening the validity of the findings of this study. The K-Means method successfully identified the group of toddlers most at risk of stunting, while Naive Bayes effectively classified toddlers into stunting and non-stunting categories based on the variables that have been analyzed. With these promising results, the method used can be a valuable tool for health practitioners in detecting and treating stunting cases, as well as assisting in designing more appropriate and efficient interventions to reduce stunting rates in toddlers.

Keywords: K-Means Method; Naive Bayes Method; Toddler Stunting

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan YME berkat Rahmat, Hidayah, dan Karunia-Nya kepada kita sehingga saya dapat menyelesaikan proposal skripsi dengan judul “ Implementasi Metode K-Means Clustering Dan Naive Bayes Untuk Prediksi Balita Teridentifikasi Stunting Studi Kasus Balai Penyuluhan Kb Kec.Dolok Sigompulon”. Laporan proposal skripsi ini disusun sebagai salah satu syarat untuk mengerjakan tugas akhir pada program Studi Sistem Informasi Fakultas Sains Dan Teknologi universitas labuhanbatu.

Saya sebagai Penulis menyadari dalam penyusunan proposal skripsi ini tidak akan dapat terselesaikan tanpa bantuan dari berbagai pihak. Karena itu pada kesempatan ini saya ingin mengucapkan terima kasih kepada:

1. Bapak Pendiri Universitas Labuhanbatu, Alm Bapak Dr. H. Amarullah Nasution, S.E., M.B.A.
2. Bapak Ketua Yayasan Universitas Labuhanbatu, Bapak Halomoan Nasution, S.H., M.H.
3. Bapak Rektor Universitas Labuhanbatu, Bapak Assoc, Prof. Ade Parlaungan Nasution, Ph.D
4. Bapak Dr. Iwan Purnama, S.Kom., M.Kom selaku Dekan Fakultas Sains Dan Teknologi Universitas Labuhanbatu.
5. Bapak Budianto Bangun, S.Kom., M.Kom selaku Kepala Program Studi Sistem Informasi.
- 6.

6. Bapak Syaiful Zuhri Harahap, S.Kom., M.Kom selaku Dosen Pembimbing 1 (Satu).
7. Ibu Rahma Mutia, S.Psi., M.Psi selaku Dosen Pembimbing 2 (Dua)
8. Ibu Irmayanti, S.Si., M.Pd selaku Dosen Penguji saya ucapkan terimakasih kepada orang tua saya yang selalu mensupport kuliah saya sampai menyanggah gelar Sarjana Komputer (S.Kom) dan kepada teman-teman seperjuangan dengan saya, terimakasih telah berjuang Bersama dalam penelitian dan pengerjaan skripsi ini. Dan teman seperjuangan kelas sistem informasi. saya menyadari skripsi ini tidak luput dari berbagai kekurangan. Penulisan mengharapkan saran dan kritik untuk perbaikannya sehingga skripsi ini dapat memberikan manfaat bagi banyak orang-orang dalam bidang komputer.

Rantauprapat, 28 Agustus 2024

Penulis



Eggi Ok Pernanda Gaja

NIM. 2009102160

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN/PERSETUJUAN TUGAS AKHIR.....	ii
LAPORAN PENGESAHAN NASKAH TUGAS AKHIR	iii
PERNYATAAN	iv
DEDIKASI	v
ABSTRAK	vii
ABSTRACT	viii
KATA PENGANTAR	ix
DAFTAR ISI.....	xi
BAB I	14
PENDAHULUAN	14
1.1. Latar Belakang Masalah.....	14
1.2. Perumusan Masalah	14
1.3. Ruang Lingkup Masalah	16
1.4. Tujuan dan Manfaat Penelitian.....	17
1.4.1. Tujuan Penelitian	17
1.4.2. Manfaat Penelitian	17
1.5. Sistematika Penulisan	18
BAB II	19
LANDASAN TEORI.....	21
2.1. Data Mining	21
2.2. <i>Knowledge Discovery in Database (KDD)</i>	22

2.3. Metode Naïve Bayes	23
2.3.1. Uji Performa	24
2.4. Metode K-Means	25
2.5. Alat Bantu Pemrograman/Tools Pendukung	27
2.5.1. <i>Orange</i>	27
2.6. Metodologi Penelitian	28
2.6.1. Penelitian Terdahulu	28
2.7. Kerangka Penelitian	31
BAB III	35
ANALISIS DAN PERANCANGAN	35
3.1. Akrsitektur Sistem	35
3.2. Desain Aktifitas Sistem	36
3.3. Pengolahan Data Metode K-Means	36
3.4. Naive Bayes	48
3.4.1. Selection	18
3.4.2. Preprocessing	49
3.4.3. Transformation	51
3.4.4. Data Mining	51
3.4.5. Interpretation/Evaluation	53
3.4.6. Uji Performa dengan Confusion Matrix	54
BAB IV	56
HASIL DAN PEMBAHASAN	56
4.1. Metode K-Means	56
4.1.1. Seleksi Data	56
4.1.2. Perancangan Model Clustering	57

4.1.3. Hasil Cluster	58
4.2. Metode Naive Bayes	60
4.2.1. Pemilihan Data.....	60
4.2.2. Preprocessing Data.....	60
4.2.3. Pembagian Data	61
4.2.4. Pembentukan dan Pengujian Model Metode Naive.....	64
4.2.5. Hasil Klasifikasi Model Metode Naive Bayes	65
4.2.6. Evaluasi Model Metode Naive Bayes	67
BAB V	69
KESIMPULAN DAN SARAN.....	69
5.1. Kesimpulan	69
5.2. Saran	69
DAFTAR PUSTAKA.....	70