

**KLASIFIKASI STATUS KESEHATAN BALITA
MENGUNAKAN ALGORITMA *VERY FAST* C4.5
TUGAS AKHIR**

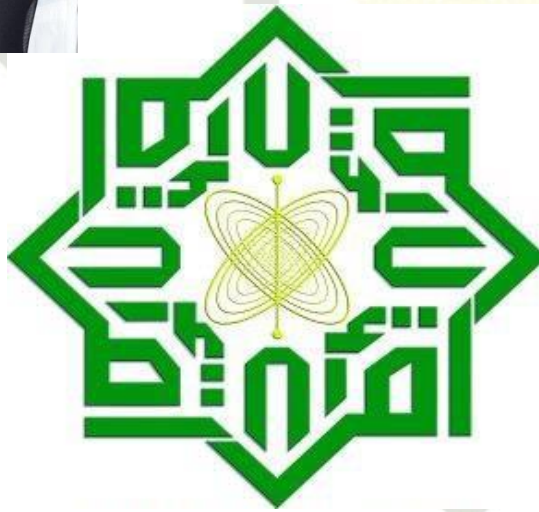
Diajukan Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika



Oleh:

PUSPITA SARI

11551202777



**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU**

2021

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN
KLASIFIKASI STATUS KESEHATAN BALITA
MENGGUNAKAN ALGORITMA VERY FAST C4.5
TUGAS AKHIR

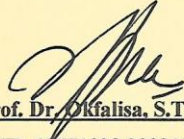
Oleh:

PUSPITA SARI

11551202777

Telah diperiksa dan disetujui sebagai laporan tugas akhir
di Pekanbaru, pada Desember 2021

Pembimbing,



Prof. Dr. Okfalisa, S.T., M.Sc.

NIP. 19771028 200312 2 004



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**LEMBAR PENGESAHAN
 KLASIFIKASI STATUS KESEHATAN BALITA
 MENGGUNAKAN ALGORITMA VERY FAST C4.5**

TUGAS AKHIR

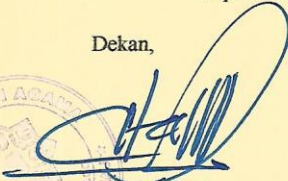
Oleh:

PUSPITA SARI
 11551202777

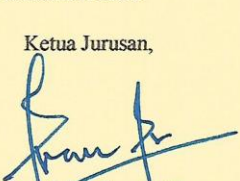
Telah dipertahankan di depan sidang dewan penguji
 sebagai salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika
 Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
 di Pekanbaru, pada Desember 2021

Dekan,

Ketua Jurusan,



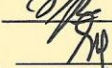
Dr. Hartono, M.Pd
 NIP. 19640301 199203 1 003

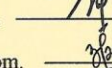


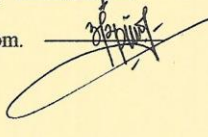
Iwan Iskandar, M.T
 NIP. 19821216 201503 1 003

DEWAN PENGUJI

KETUA : Reski Mai Candra, ST. M.Sc. 

SEKRETARIS : Prof. Dr. Okfalisa, S.T., M.Sc. 

PENGUJI I : Siska Kurnia Gusti, S.T., M.Sc. 

PENGUJI II : Eka Pandu Cynthia, S.T., M.Kom. 

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Lampiran Surat :
 Nomor : Nomor 25/2021
 Tanggal : 10 September 2021

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : PUSPITA SARI
 NIM : 11551202177
 Tempat/Tgl. Lahir : Dusun VII Sidorukun, 02-februari-1997
 Fakultas/Pascasarjana : SAINTEK
 Prodi : Teknik Informatika
 Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*:
KLASIFIKASI STATUS KESEHATAN BALITA MENGGUNAKAN ALGORITMA
VERY FAST C4-5

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis/Skripsi/(Karya Ilmiah lainnya)* saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 10 Desember 2021
 membuat pernyataan

 PUSPITA SARI
 NIM : 11551202177

*pilih salah satu sesuai jenis karya tulis

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis, diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Pekanbaru, 08 Desember 2021

Yang membuat pernyataan,

PUSPITA SARI

NIM. 11551202777

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Alhamdulillah Rabbil'alamiin...

Rasa syukur kuhaturkan kepada-Mu, Yaa Allah yang Maha Ber-Ilmu, hanya karena karunia-Mu lah hamba-Mu akhirnya dapat menyelesaikan Tugas Akhir ini

Allah Azza Wa Jalla

Sholawat serta salam untuk Rasulullah

Muhammad Shallallahu 'Alaihi Wasallam

Kupersembahkan karya sederhana ini, Tugas Akhir ini, untuk Ayah dan Ibu. Tentulah tidak akan tergantikan semua jasa, pengorbanan, tetes keringat, dan rasa letih itu, hanya dengan karya kecil dariku ini... Namun semoga dengan ini, aku dapat mengukir sebaris senyum bahagia di hati Ayah dan Ibu... Jika boleh kujabarkan cinta,, tentulah tidak pernah dapat seindah rasa syukurku menjadi anakmu... Terimakasih untuk semua rangkaian do'a, kasih sayang serta ilmu yang berharga...

Dan tidak lupa kupersembahkan untuk semua adik-adikku tersayang, terimakasih untuk semua do'a dan dukungan yang telah diberikan selama ini...

Juga, kupersembahkan untuk semua keluargaku dan kerabat... Semua kesulitan seolah lenyap saat mengingat bahwa aku memiliki dukungan dan do'a darimu semua. Aku tahu, engkau semua berjuang jauh lebih keras dariku, namun selalu memiliki energi yang hebat untuk menyemangatiku...

Alhamdulillah, Allah menganugerahiku keluarga, kerabat, dan kehidupan yang indah

KLASIFIKASI STATUS KESEHATAN BALITA MENGUNAKAN ALGORITMA *VERY FAST C4.5*

Puspita Sari
11551202777

Tanggal Sidang : 30 November 2021

Periode Wisuda :

Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau

ABSTRAK

Status kesehatan balita merupakan status yang sangat penting dimana status kesehatan balita dapat dinilai dari tiga indikator, yaitu Berat Badan menurut Umur (BB/U), Berat Badan menurut Tinggi Badan (BB/TB), dan Tinggi Badan menurut Umur (TB/U). Masa balita merupakan masa tumbuh kembang anak yang rentan terhadap masalah kesehatan. Sistem yang dirancang untuk penentuan status kesehatan balita dengan menggunakan model data *mining* yaitu dengan algoritma *Very Fast C4.5* menggunakan 2.790 data dari Puskesmas Kampar tahun 2020. Hasil penelitian ini berupa sistem klasifikasi status kesehatan balita (menurut BB/U) menggunakan algoritma *Very Fast C4.5* dengan tingkat akurasi 93,52% menggunakan ratio data uji dan data latih 90%:10%. Pengujian dilakukan menggunakan *K-fold cross validation* dan menghasilkan akurasi tertinggi pada *k-fold=8* yaitu sebesar 96,05%. Dengan demikian, dapat disimpulkan bahwa algoritma *Very Fast C4.5* berhasil digunakan untuk mengklasifikasikan status kesehatan balita.

Kata Kunci: Algoritma *Very Fast C4.5*, *Data Mining*, *K-Fold Cross Validation*, Klasifikasi, Status Kesehatan Balita.

THE HEALTH STATUS OF TODDLERS USING THE VERY FAST C4.5 ALGORITHM

Puspita Sari
11551202777

Date of final exam : 30th November 2021
Graduate period :

Department of Informatics Engineering
Faculty of Science and Technology
State Islamic University of Sultan Syarif Kasim Riau

ABSTRACT

The health status of toddlers is a very important status where the health status of toddlers can be assessed from three indicators, namely weight for age (W/U), weight for height (W/TB), and height for age (TB/U). Infancy is a period of growth and development of children who are vulnerable to health problems. The system designed to determine the health status of toddlers using a data mining model is the Very Fast C4.5 algorithm using 2,790 data from the Kampar Health Center in 2020. The results of this study are a classification system for the health status of toddlers (according to BB/U) using the Very Fast algorithm. C4.5 with an accuracy rate of 93.52% using the ratio of test data and training data 90%:10%. The test was carried out using K- fold cross validation and resulted in the highest accuracy at k- fold = 8, which was 96.05%. Thus, it can be concluded that the Very Fast C4.5 algorithm was successfully used to classify the health status of toddlers.

Keywords : *Very Fast C4.5 Algorithm , Data Mining , K- Fold Cross Validation , Classification, Toddler Health Status.*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR

Assalammu'alaikum wa rahmatullahi wa barakatuh.

Alhamdulillah, tak henti-hentinya penulis ucapkan kehadiran Tuhan yang tiada Tuhan selain Dia, Allah *subhana wa ta'ala*, yang dengan rahmat dan hidayahNya penulis mampu menyelesaikan Tugas Akhir yang berjudul **"Klasifikasi Status Kesehatan Balita Menggunakan Algoritma Very Fast 4.5"** dengan baik. Tidak lupa dan tak akan pernah lupa bershalawat kepada Nabi dan RasulNya, Nabi Muhammad SAW yang hanya menginginkan keimanan dan keselamatan bagi umatnya dan sangat belas kasihan lagi penyayang kepada orang-orang mukmin.

Tugas Akhir ini disusun sebagai salah satu syarat untuk mendapatkan gelar kesarjanaan pada jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Banyak sekali pihak yang telah membantu penulis dalam penyusunan laporan ini, baik berupa bantuan materi ataupun berupa motivasi dan dukungan kepada penulis. Semua itu tentu terlalu banyak bagi penulis untuk membalasnya, namun pada kesempatan ini penulis hanya dapat mengucapkan terima kasih kepada

1. Bapak Prof Dr. Khairunnas M.Ag selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd. selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Bapak Iwan Iskandar, M.T. selaku Ketua Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Ibu Dr. Okfalisa, S.T., M.Sc. selaku Pembimbing. Terima kasih untuk waktu, motivasi, dan arahan yang telah diberikan sehingga Proposal TA ini dapat terselesaikan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

5. Ibu Siska Kurnia Gusti, S.T., M.Sc. selaku penguji I yang telah memberikan arahan, kritik, dan saran kepada penulis dalam peyusunan tugas akhir ini.
6. Ibu Eka Pandu Cynthia, S.T., M.Kom. selaku penguji II yang telah memberikan arahan, kritik, dan saran kepada penulis dalam peyusunan tugas akhir ini.
7. Ibu Fadhilah Syafria, S.T., M.Kom., CIBIA. selaku Koordinator Tugas Akhir Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Negeri Sultan Syarif Kasim Riau.
8. Seluruh Dosen Teknik Informatika yang telah memberikan ilmu dan bimbingan yang bermanfaat untuk kami.
9. Ayah Parsimin dan Ibu Saliah dan Sulis Tia Lani, Naysila Oktaviana yang telah memberikan do'a, motivasi, saran, dan semangat sehingga laporan ini dapat diselesaikan.
10. Gandhi Alwira Ponda, terimakasih atas bantuan serta dukungan selama ini sehingga laporan ini dapat diselesaikan.
11. Teman-teman kelas tif D'15 tidak bisa disebutkan satu per satu namanya, terima kasih atas do'a, bantuan, dan dukungannya selama ini sehingga laporan ini dapat diselesaikan.
12. Semua pihak yang ikut memberikan bantuan, dukungan, dan doa baik secara langsung ataupun tidak langsung.

Penulis menyadari bahwa dalam penulisan laporan ini masih banyak kesalahan dan kekurangan, oleh karena itu kritik dan saran yang sifatnya membangun sangat penulis harapkan untuk kesempurnaan laporan ini. Akhirnya penulis berharap semoga laporan ini dapat memberikan sesuatu yang bermanfaat bagi siapa saja yang membacanya. Amin.

Pekanbaru, November 2021

DAFTAR ISI

| | |
|--|-------|
| ABSTRAK | ii |
| ABSTRACT | iii |
| KATA PENGANTAR | iv |
| DAFTAR ISI | vi |
| DAFTAR GAMBAR | viii |
| DAFTAR TABEL | ix |
| DAFTAR PERSAMAAN | x |
| DAFTAR LAMPIRAN | xi |
| DAFTAR SIMBOL | xii |
| BAB I PENDAHULUAN | I-1 |
| 1.1 Latar Belakang | I-1 |
| 1.2 Rumusan Masalah | I-4 |
| 1.3 Batasan Masalah | I-4 |
| 1.4 Tujuan Penelitian | I-4 |
| 1.5 Sistematika Penulisan | I-4 |
| BAB II LANDASAN TEORI | II-1 |
| 2.1 <i>Knowledge Discovery in Database</i> | II-1 |
| 2.2 <i>Data Mining</i> | II-2 |
| 2.3 Algoritma <i>Very Fast C4.5</i> | II-3 |
| 2.4 Status Kesehatan | II-5 |
| 2.5 <i>K-Fold Cross Validation</i> | II-7 |
| 2.6 <i>Confusion Matrix</i> | II-7 |
| 2.7 <i>Matrix Laboratory</i> | II-8 |
| 2.8 Penelitian Terkait | II-9 |
| BAB III METODOLOGI PENELITIAN | III-1 |
| 3.1 Perencanaan | III-2 |
| 3.2 Pengumpulan Data | III-2 |
| 3.3 Analisa | III-2 |
| 3.5 Implementasi | III-3 |
| 3.7 Kesimpulan dan Saran | III-4 |
| BAB IV ANALISA DAN PERANCANGAN | IV-1 |
| 4.1 Analisis Kebutuhan Data | IV-1 |
| 4.1.1 <i>Data Selection</i> | IV-1 |
| 4.1.2 <i>Pre-Processing</i> | IV-1 |
| 4.1.3 <i>Transformation</i> | IV-3 |

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

| | |
|---|-------------|
| 4.1.4 Data Mining..... | IV-3 |
| 4.2 Perancangan Struktur Menu | IV-13 |
| 4.3 Perancangan <i>Interface</i> | IV-13 |
| 4.3.1 Rancangan Halaman Utama | IV-13 |
| 4.3.2 Rancangan Halaman Pengolaan Data..... | IV-14 |
| 4.3.3 Rancangan Halaman Pelatihan dan Pengujian | IV-15 |
| 4.3.4 Rancangan Halaman <i>Form</i> Klasifikasi..... | IV-15 |
| BAB V IMPLEMENTASI DAN PENGUJIAN | V-1 |
| 5.1 Pengertian dan Tujuan Implementasi | V-1 |
| 5.2 Batasan Implementasi..... | V-1 |
| 5.3 Implementasi | V-1 |
| 5.3.1 Interpretasi Pohon Keputusan..... | V-1 |
| 5.3.2 Halaman Utama | V-2 |
| 5.3.3 Halaman Pengolaan Data..... | V-3 |
| 5.3.4 Halaman Pelatihan dan Pengujian | V-3 |
| 5.3.5 Halaman <i>Form</i> Klasifikasi..... | V-4 |
| 5.4 Pengujian | V-5 |
| 5.4.1 Pengujian Akurasi Metode | V-5 |
| 5.4.3 Kesimpulan Pengujian..... | V-9 |
| BAB VI PENUTUP | VI-1 |
| 6.1 Kesimpulan..... | VI-1 |
| 6.2 Saran | VI-1 |
| DAFTAR PUSTAKA | xx |
| DAFTAR RIWAYAT HIDUP..... | xx |
| LAMPIRAN A | xx |
| LAMPIRAN B | 1 |
| LAMPIRAN C | 1 |
| LAMPIRAN D | 1 |
| LAMPIRAN E | 12 |

DAFTAR GAMBAR

| Gambar | Halaman |
|---|---------|
| 2.1 Tahapan Proses KDD | II-1 |
| 3.1 Tahapan Metodologi Penelitian | III-1 |
| 4.1 Pengecekan <i>Mising Values</i> | IV-1 |
| 4.2 Pengecekan Data Duplikat | IV-2 |
| 4.3 Pengecekan Pencarian Data Inkonsisten | IV-2 |
| 4.4 <i>Flowchat</i> algoritma <i>Very Fast C4.5</i> | IV-3 |
| 4.5 Pohon keputusan | IV-12 |
| 4.6 Rancangan struktur menu | IV-13 |
| 4.7 Rancangan Halaman Utama | IV-14 |
| 4.8 Rancangan Halaman Pengolaan Data | IV-14 |
| 4.9 Rancangan Halaman Pelatihan dan Pengujian | IV-15 |
| 4.10 Rancangan Halaman <i>Form</i> Klasifikasi | IV-16 |
| 5.1 Rule-rule pohon keputusan | V-2 |
| 5.2 Halaman Utama | V-3 |
| 5.3 Halaman Pengolaan Data | V-3 |
| 5.4 Halaman Pelatihan dan Pengujian | V-4 |
| 5.5 Halaman <i>Form</i> Klasifikasi | V-4 |
| A.1 Hasil Wawancara | xx |
| D.1 Cabang akar pohon keputusan | 12 |

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR TABEL

| Tabel | Halaman |
|---|---------|
| 2.1. Parameter Pemantauan Status Gizi Balita (PSG 2018)..... | II-6 |
| 2.2. Model <i>Confusion Matrix</i> | II-8 |
| 2.3. Penelitian Terkait | II-9 |
| 4.1. Parameter Data | IV-1 |
| 4.2 Data Penelitian | IV-1 |
| 4.3 Parameter <i>Data Selection</i> | IV-1 |
| 4.4 Hasil <i>Data Selection</i> | IV-1 |
| 4.5 Tabel pencarian nilai ambang batas | IV-4 |
| 4.6 Hasil perhitungan atribut..... | IV-10 |
| 4.7 Hasil perhitungan atribut..... | IV-11 |
| 5.1 Skenario Uji Akurasi dengan <i>10-fold cross validation</i> | V-5 |
| 5.2 Hasil Uji Akurasi dengan <i>10-fold cross validation</i> | V-6 |
| 5.3 Hasil <i>Confusion Matrix</i> 90%:10% | V-7 |
| 5.4 Hasil <i>Confusion Matrix</i> 80%:20% | V-8 |
| 5.5 Hasil <i>Confusion Matrix</i> 70%:30% | V-8 |
| 5.6 Hasil Pengujian Akurasi Menggunakan <i>Confusion Matrix</i> | V-9 |
| B.1 Parameter Status Gizi Balita (PSG 2018)..... | 1 |
| C.1 Data Penelitian..... | 1 |
| D.1 Tabel Hasil Perhitungan Manual | 6 |
| D.2 Hasil Perhitungan Manual | 12 |
| E.1 Tabel pencarian nilai ambang batas..... | 13 |
| E.2 Hasil perhitungan manual | 19 |
| E.3 Tabel pencarian nilai ambang batas..... | 20 |
| E.4 Hasil pencarian perhitungan manual | 26 |

© Hak Cipta Ilmiah UIN Suska Riau State Islamic University of Sultan Syarif Kasim Riau

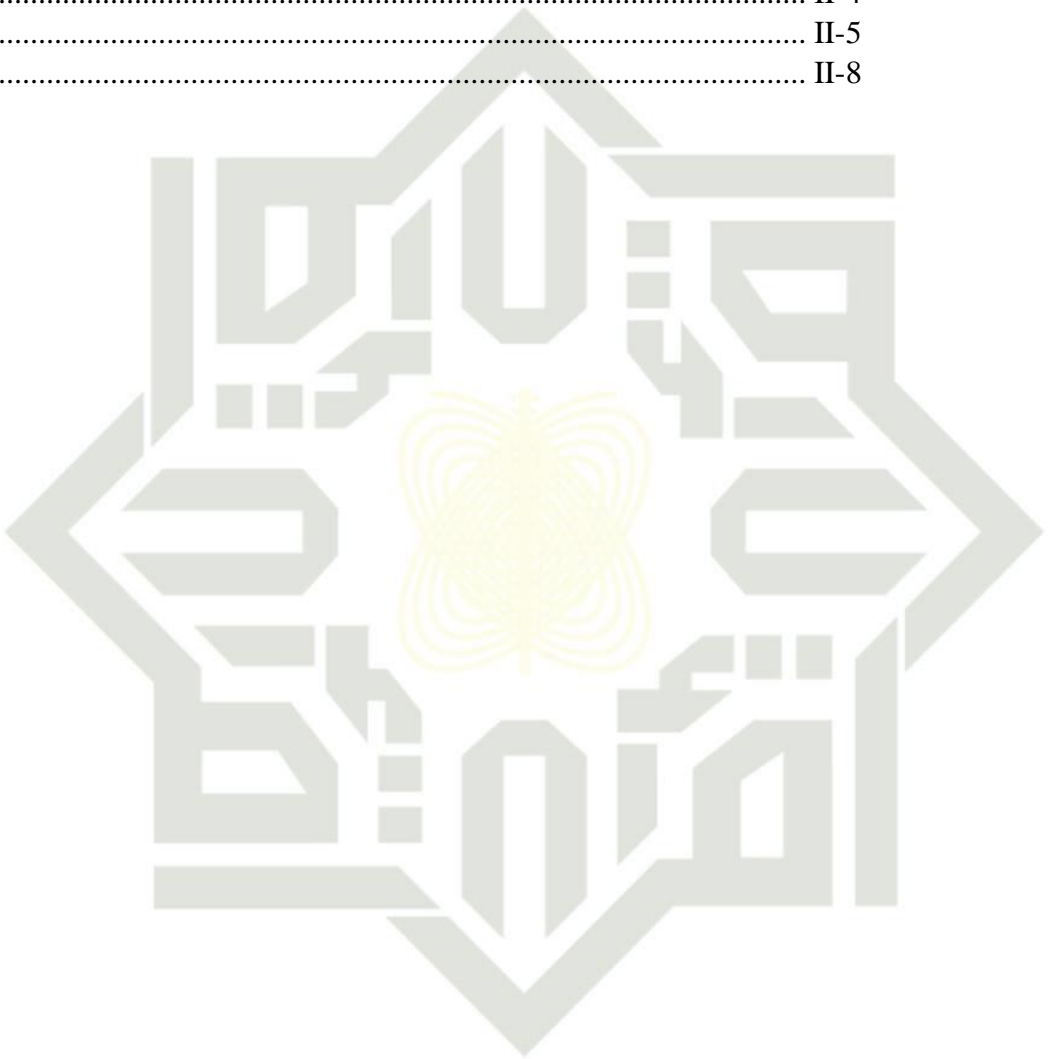
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR PERSAMAAN

| Persamaan | Halaman |
|-----------|---------|
| 1) | II-4 |
| 2) | II-4 |
| 3) | II-4 |
| 4) | II-4 |
| 5) | II-4 |
| 6) | II-5 |
| 7) | II-8 |



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR LAMPIRAN

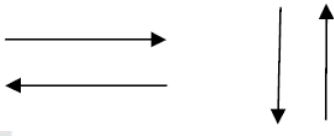

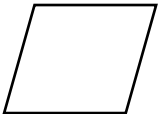


| Lampiran | Halaman |
|--|---------|
| A Hasil Wawancara..... | xx |
| B Parameter Pemantauan Status Gizi Balita 2018 | 1 |
| C Data Penelitian..... | 1 |
| D Hitungan Manual Lanjutan..... | 1 |
| E Hitungan Manual berdasarkan Pembagian Ratio | 12 |

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR SIMBOL

Simbol *Flowchart*

| | |
|---|--|
|  | <p><i>Flow Direction</i> Menghubungkan simbol yang satu dengan simbol yang lain.</p> |
|  | <p><i>Terminator</i> Simbol permulaan (<i>start</i>) atau akhir (<i>end</i>) dari suatu kegiatan.</p> |
|  | <p><i>Data</i> Simbol yang digunakan untuk mendeskripsikan data yang digunakan.</p> |
|  | <p><i>Process</i> Simbol yang berguna untuk melakukan proses data oleh <i>user</i> ataupun sistem.</p> |
|  | <p><i>Predefined Process</i> Simbol yang digunakan untuk proses pada subprogram pada sistem.</p> |

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB I PENDAHULUAN

I Latar Belakang

Melakukan peningkatan kesehatan merupakan keharusan bangsa Indonesia apabila ingin mencapai pembangunan manusia yang tinggi. Kesehatan merupakan hak asasi manusia yang harus dipenuhi dengan baik. Anak yang sehat menjadi investasi modal bagi manusia yang berkualitas di masa depan (Heri and Cicih 2011).

Menurut Hadi Siswanto tahun 2009 dalam buku yang berjudul Pendidikan Anak Usia Dini “Masalah kesehatan masyarakat dan masa depan bangsa ditentukan oleh kondisi kesehatan anak usia dini, pada saat lahir sampai kepada pola pengasuhan, asah, dan asih”. Masa balita merupakan masa tumbuh kembang anak yang rentan dengan masalah kesehatan. Terdapat dua jenis indikator yang dapat digunakan untuk menentukan status kesehatan masyarakat, yaitu angka kesakitan (morbiditas) dan angka kematian (mortalitas) (Putri 2012).

Masalah kesehatan disebabkan dua faktor yang saling mempengaruhi secara langsung atau tidak langsung. Sementara di tingkat rumah tangga, masalah kesehatan dipengaruhi oleh kemampuan rumah tangga menyediakan makanan yang cukup baik dari segi kualitas dan kuantitasnya, pola asuh anak, pengetahuan gizi, serta faktor sosio budaya lainnya (Zulhaida Lubis 2015).

Pemantauan pertumbuhan dan perkembangan anak harus dilakukan secara teratur dan berkesinambungan yang dilakukan oleh orang tua. Selain itu juga dapat dilakukan oleh masyarakat melalui kegiatan posyandu. Pemantauan tersebut meliputi pemantauan aspek fisik, psikologi, dan sosial. Parameter ukuran antropometrik yang dipakai dalam penilaian fisik adalah tinggi badan, berat badan, lingkar kepala, lipatan kulit, lingkar lengan atas, panjang lengan, proporsi tubuh, dan panjang tungkai. Oleh karena itu pentingnya pengetahuan orangtua, masyarakat tentang kesehatan balita perlu dimiliki (Chamidah 2009).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Kementrian Kesehatan RI telah melakukan berbagai upaya untuk meningkatkan kualitas kesehatan masyarakat. Menkes RI Nila Moeloek (2019) mengatakan pemerintah mendorong pembangunan kesehatan melalui Program Indonesia Sehat. Program tersebut dilakukan dengan pendekatan keluarga sehingga keluarga sehat dapat terwujud. Data pelaporan rutin dari setiap provinsi menunjukkan penurunan jumlah kematian ibu dan bayi. Angka Kematian Bayi (AKB) mengalami sedikit penurunan dari 34 per 1.000 kelahiran hidup (*Survey Demografi Kesehatan Indonesia* atau SDKI 2012) menjadi 24 per 1.000 kelahiran hidup (SDKI 2017).

Status kesehatan balita merupakan status yang sangat penting dimana status kesehatannya diukur dengan pemantauan gizi balita berdasarkan Umur (U), Berat Badan (BB), dan Tinggi Badan (TB). Status gizi balita dinilai dari tiga indikator, yaitu Berat Badan menurut Umur (BB/U), Berat Badan menurut Tinggi Badan (BB/TB), dan Tinggi Badan menurut Umur (TB/U) (Pemantauan Status Gizi Balita 2018). Hasil pemantauan kesehatan balita tahun 2020 dalam Laporan Akuntabilitas Kinerja Kementerian Kesehatan Republik Indonesia menyatakan bahwa hanya 23,33% balita di Indonesia yang mendapatkan pelayanan kesehatan ibu dan anak (Kementrian Kesehatan Republik Indonesia 2020).

Laporan Kerja Instansi Pemerintah Dinas Kesehatan Provinsi Riau tahun 2018, menyatakan hasil SDKI terakhir (tahun 2012) angka kematian bayi mencapai 24 per 1.000 kelahiran hidup dan angka kematian balita 28 per 1.000 kelahiran hidup sedangkan jumlah kematian balita tahun 2018 di Provinsi Riau mencapai 479 kasus. Jumlah kematian yang dilaporkan tersebut masih banyak yang belum tercatat khususnya dari Rumah Sakit Swasta, Bidan Praktek Perseorangan dan Rumah Bersalin (Kesehatan 2018).

Dari data diatas, menunjukkan pentingnya penyampaian peningkatan kesehatan bagi keluarga terutama anak balita yang masih dalam masa emas (*Golden age*) kepada masyarakat.

Penelitian sebelumnya pernah dilakukan untuk peningkatan kesehatan balita dengan mengklasifikasi kesehatan balita untuk mengetahui status kesehatan balita yang ada di Provinsi Riau pada tahun 2018 menggunakan gabungan metode

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

C4.5 dan *Pattern Swam Optimization* (PSO)(Akhyar 2018). Mengacu pada penelitian sebelumnya menyatakan bahwa klasifikasi status gizi balita menggunakan metode C4.5 dan PSO menghasilkan akurasi sebesar 94,49%(Akhyar 2018), sedangkan secara teori yang dibuktikan oleh Cherfi 2018 menyatakan bahwa algoritma *Very Fast C4.5* lebih akurat daripada metode klasifikasi lainnya(Cherfi, Nourira, and Ferchichi 2018).

Berdasarkan data yang ada, penulis ingin membuat sistem yang dapat mengklasifikasikan status kesehatan balita. Dengan sistem ini diharapkan dapat mengetahui tingkat kerja algoritma dalam mengklasifikasi status kesehatan balita. Dalam membangun sistem, penulis akan menerapkan algoritma *Very Fast C4.5* pada data balita di Puskesmas Kampar tahun 2020 untuk mengklasifikasikan status kesehatan balita.

Algoritma *Very Fast C4.5* merupakan algoritma baru yang dikembangkan untuk mempercepat dan meningkatkan kinerja algoritma C4.5. Dari hasil penelitian Anis Cherfi tahun 2018 yang melakukan perbandingan antara algoritma *Very Fast C4.5*, C4.5, *Very Fast Decision Tree* (VFDT), dan *Classification and Regression Tree* (CART) menyebutkan bahwa algoritma *Very Fast C4.5* memiliki akurasi tertinggi diantara algoritma lainnya. Penelitian tersebut menggunakan 49 dataset tanpa atribut untuk melatih algoritma *Very Fast C4.5*. Anis juga menyebutkan bahwa hasil percobaan menunjukkan algoritma *Very Fast C4.5* lebih cepat daripada algoritma VFDT dan lebih akurat dari C4.5 (Cherfi, Nourira, and Ferchichi 2018).

Penelitian terkait klasifikasi yang menggunakan algoritma lainnya pernah dilakukan seperti algoritma C4.5, *K-Nearest Neighbor* (KNN), *Modified K-Nearest Neighbor* (MKNN), *Support Vector Machine* (SVM), dan lainnya. Algoritma C4.5 dan KNN dinyatakan lebih akurat dari algoritma *naïve bayes* dalam mengklasifikasikan benih gandum (Fajri 2017). Penelitian klasifikasi *Short Message Service* (SMS) *Spam* juga pernah dilakukan dengan membandingkan algoritma SVM, C4.5, KNN, maupun *naïve bayes* dengan hasil bahwa algoritma SVM lebih akurat (Zuviyanto, Adji, and Setiawan 2018). Perbandingan algoritma lainnya yang dilakukan oleh Okfalisa, dkk yaitu perbandingan algoritma KNN

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dan *Linear Regression* menunjukkan bahwa algoritma KNN lebih efektif dan akurat daripada algoritma *Linear Regression* untuk klasifikasi penerima beasiswa (Okfalisa, Ratika Fitriani 2018). Algoritma MKNN dinyatakan lebih akurat dari algoritma KNN untuk klasifikasi dengan data 7.495 *records* (Okfalisa, Ikbal Gazalba, Mustakim 2019).

1.2 Rumusan Masalah

Berdasarkan permasalahan yang telah penulis paparkan di latar belakang tersebut, maka penulis akan membuat rumusan masalah yang akan dibahas, yaitu menerapkan algoritma *Very Fast C4.5* dalam mengklasifikasikan status kesehatan balita menurut Berat Badan menurut Umur (BB/U).

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini agar lebih jelas dan terarah adalah sebagai berikut:

1. Data balita yang digunakan adalah data kesehatan bayi dari Puskesmas Kampar tahun 2020 yang berjumlah 2.790 data dan menggunakan 13 atribut.
2. Target kelas yang digunakan untuk kalsifikasi adalah, sehat, kurang sehat, dan tidak sehat.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah diatas maka penulis mempunyai tujuan yang ingin dicapai melalui penelitian ini, yaitu:

1. Menerapkan algoritma *Very Fast C4.5* dalam mengklasifikasikan kesehatan balita.

1.5 Sistematika Penulisan

Agar lebih sistematis dan terarah dalam penyusunan laporan tugas akhir, penulis membagi dalam beberapa bab. Berikut ini adalah bab-bab yang disajikan:



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB I

PENDAHULUAN

Bab ini berisi deskripsi umum dari penelitian tugas akhir meliputi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, dan sistematika penulisan.

BAB II

LANDASAN TEORI

Bab ini berisi tentang penjelasan teori yang berhubungan dengan penelitian. Teori yang dijelaskan di bab ini antara lain mengenai *knowledge discovery in database*, *data mining*, algoritma *Very Fast C4.5*, status kesehatan balita, *confusion matrix*, dan penelitian terkait.

BAB III

METODOLOGI PENELITIAN

Bab ini berisi tentang langkah-langkah yang dilakukan dalam penelitian ini. Langkah-langkah tersebut antara lain adalah perencanaan, pengumpulan data, analisa, perancangan, implementasi, pengujian, serta kesimpulan dan saran.

BAB IV

ANALISA DAN PERANCANGAN

Bab ini membahas mengenai analisa data yang akan digunakan dan analisa aplikasi. Bab ini juga meliputi perancangan sistem yang akan dibangun.

BAB V

IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas mengenai implementasi metode yang digunakan terhadap data dan implementasi dari perancangan sistem yang telah dibuat sebelumnya. Bab ini juga meliputi pengujian dari hasil implementasi.

BAB VI

PENUTUP

Bab ini berisi tentang kesimpulan yang diperoleh sebagai hasil akhir penelitian yang telah dilakukan serta saran untuk penelitian selanjutnya.

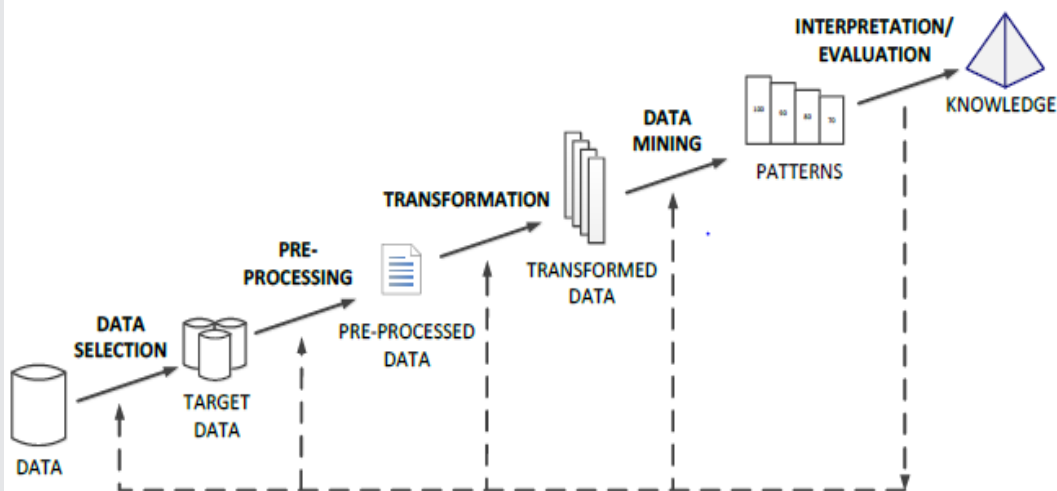
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II LANDASAN TEORI

2.1 Knowledge Discovery in Database

Knowledge discovery in database (KDD) merupakan metode untuk memperoleh pengetahuan dari *database* yang dapat digunakan untuk keperluan dalam pengambilan keputusan (Mardi 2017). Menurut pendapat Maimon dan Rokach *data mining* merupakan inti dari tahap KDD. Seringkali istilah KDD dan *data mining* diartikan sama. Sebenarnya kedua istilah tersebut memiliki perbedaan konsep, hanya saja keduanya saling berkaitan satu sama lain, dan salah satu tahapan di dalam proses KDD adalah *data mining* (Oded Maimon 2010). Berikut proses KDD secara garis besar:



Gambar 2.1 Tahapan Proses KDD

Penjelasan mengenai tahapan KDD adalah sebagai berikut:

1. Data selection

Pemilihan (seleksi) data dari sekumpulan data yang dilakukan sebelum tahap penggalian informasi dalam KDD. Data tersebut akan disimpan disuatu berkas yang kemudian akan digunakan untuk proses *data mining*.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. *Pre-processing/cleaning*

Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, mengisi nilai yang hilang, dan memperbaiki kesalahan pada data.

3. *Transformation*

Transformation merupakan proses perubahan format pada data yang dipilih, sehingga data tersebut sesuai untuk proses *data mining*.

4. *Data mining*

Data mining adalah proses pencarian pola atau informasi menarik dalam data yang terpilih dengan menggunakan teknik atau metode tertentu.

5. *Interpretation/evaluation*

Tahap ini merupakan bagian dari proses KDD dimana proses *data mining* sebelumnya perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

2.2 Data Mining

Menurut Gartnes Group, *data mining* adalah proses menemukan hubungan baru yang mempunyai arti, pola, dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan dengan menggunakan teknologi pengenalan pola seperti teknik statistik dan matematika. *Data mining* merupakan gabungan dari disiplin ilmu yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi data *database* yang besar (Parose 2005).

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dengan berbagai *database* besar. *Data mining* merupakan serangkaian proses untuk



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

menggal nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Bramer 2016).

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan (Mardi 2017), yaitu:

1. *Description* (deskripsi)
Pengamatan sederhana untuk menggambar pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.
2. *Estimation* (estimasi)
Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi yang lebih ke arah numerik daripada ke arah kategori.
3. *Prediction* (prediksi)
Prediksi juga hampir sama dengan estimasi dan klasifikasi, kecuali bahwa dalam prediksi nilai dai hasil akan ada dimasa mendatang.
4. *Classification* (klasifikasi)
Dalam klasifikasi, pengelompokan data yang terdapat target variabel kategori.
5. *Clustering* (pengklasteran)
Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan yang sama.
6. *Association* (asosiasi)
Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu.

Algoritma Very Fast C4.5

Algoritma *Very Fast C4.5* (VFC4.5) menggunakan tiga proses untuk mencapai kontruksi pohon keputusan, diantaranya: pemilihan atribut, binarisasi atribut, dan pemisahan dataset. Dalam proses lainnya *Very Fast C4.5* sama dengan algoritma C4.5 hanya saja algoritma *Very Fast C4.5* menggunakan *mean* dan



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

median sebagai nilai titik potong dalam proses binarisasi atribut. *Mean* atau rata-rata adalah nilai representatif dari seluruh kelompok data. Dalam hal ini mean digunakan sebagai nilai *threshold* yang akan meningkatkan klasifikasi. *Median* digunakan sebagai pembagi data menjadi dua bagian yang sama. Berikut merupakan langkah-langkah algoritma *Very Fast C4.5* untuk membangun *decision tree* (Cherfi, Nouira, and Ferchichi 2018):

1. Hitung nilai *Mean* dan *Median*:

$$Mean = \frac{1}{n} \sum_{i=1}^n x_i \tag{2.1}$$

$$Median = x_{(\frac{n+1}{2})} \tag{2.2}$$

Keterangan:

x_i = nilai rata-rata
 n = jumlah data

2. Hitung nilai *Entropy*:

$$Entropy(S) = \sum_{i=1}^n -p_i \times \log_2(p_i) \tag{2.3}$$

Keterangan:

P_i = Proporsi kelas ke- i dalam data yang di proses di node s

3. Hitung nilai *Gain*:

$$G(S, ccp_{ij}) = info(S) - \left(\frac{|S_1|}{|S|} info(S_1) + \frac{|S_2|}{|S|} info(S_2) \right) \tag{2.4}$$

Keterangan:

ccp_{1j} = *Mean*

ccp_{2j} = *Median*

S_1 = $[\min\{A_j\}, ccp_{1j}]$

S_2 = $[\max\{A_j\}, ccp_{2j}]$

4. Hitung nilai *Split*:

$$Split(S, cp_{oj}) = -\frac{|S_1|}{|S|} \log_2 \frac{|S_1|}{|S|} - \frac{|S_2|}{|S|} \log_2 \frac{|S_2|}{|S|} \tag{2.5}$$

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Keterangan:

$$cp_{oj} = \arg \max\{G\}_{i=1}^{n-1}$$

5. Hitung nilai *Gain Ratio*:

$$GainRatio(S, cp_{oj}) = \frac{Gain(S, cp_{oj})}{Split(S, cp_{oj})} \quad (2.6)$$

6. Pilih nilai A_o dan nilai titing potong cp_o yang paling tinggi untuk dijadikan nilai *threshold*.

Status Kesehatan

Status kesehatan adalah suatu keadaan kedudukan orang dalam tingkat sehat atau sakit. Kesehatan merupakan hak asasi manusia termasuk hak dasar anak yang harus dipenuhi dengan baik. Para ahli tumbuh kembang anak mengatakan periode 5 (lima) tahun pertama kehidupan anak sebagai “masa keemasan (*golden age*)” (Florentianus Tat 2018).

Terdapat dua indikator kesehatan balita yaitu status gizi balita dan status tumbuh kembang anak. Status gizi balita Status gizi balita dinilai berdasarkan tiga indikator, yaitu berat badan menurut umur (BB/U), berat badan menurut tinggi badan (BB/TB), dan tinggi badan menurut umur (TB/U) (Kementerian Kesehatan Republik Indonesia, 2017). Adapun cara deteksi tumbuh kembang anak menurut Depkes RI tahun 2009 (Sukesi, Rina, and Emilia 2011):

1. Mendeteksi tumbuh kembang pada anak diantaranya dengan pengukuran antropometri. Pengukuran antropometri ini dapat meliputi pengukuran berat badan, tinggi badan, lingkar kepala dan lingkar lengan atas.
 - a. Pengukuran berat badan yaitu pengukuran berat badan ini bagian dari antropometri yang digunakan untuk menilai hasil peningkatan atau penurunan semua jaringan yang ada pada tubuh.
 - b. Pengukuran tinggi badan yaitu pengukuran ini merupakan bagian dari pengukuran antropometrik yang digunakan untuk menilai status perbaikan gizi di samping faktor genetik.
 - c. Pengukuran lingkar kepala dan lingkar lengan.
2. Pertumbuhan dan perkembangan anak :



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- a. Anak pada usia 3-6 bulan mengangkat kepala dengan tegak pada posisi telungkup.
- b. Anak pada usia 9-12 bulan berjalan dengan berpegangan.
- c. Anak pada usia 12-18 bulan minum sendiri dari gelas tanpa tumpah.
- d. Anak pada usia 18-24 bulan mencorat-coret dengan alat tulis.
- e. Anak pada usia 2-3 tahun berdiri dengan satu kaki tanpa berpegangan, melepas pakaian sendiri.
- f. Anak pada usia 3-4 tahun mengenal dan menyebutkan paling sedikit 1 warna.
- g. Anak pada usia 4-5 tahun mencuci dan mengeringkan tangan tanpa bantuan.

Berikut tabel 2.1 penyajian parameter pemantauan status kesehatan balita menurut pengukuran status gizi balita dan perkembangan balita:

Tabel 2.1. Parameter Pemantauan Status Gizi Balita (PSG 2018)

| No | Nama Parameter | Deskripsi | Kode Pengisian |
|----|----------------|--|---|
| 1 | ID | Identitas individu | |
| 2 | B4R4 | Jenis kelamin | 1 = Laki-laki 2 = Perempuan |
| 3 | UMUR_BULAN | Umur (dalam satu bulan) | |
| 4 | BO1A | Apakah ketika baru lahir [NAMA] dilakukan Inisiasi Menyusui Dini (IMD) | 1 = Ya 2 = Tidak |
| 5 | B21 | Anak balita 6-11 bulan, sudah mendapat kapsul vitamin A warna biru | 1 = Ya 2 = Tidak pernah 3 = Belum pernah |
| 6 | B22 | Anak >12 bulan sudah mendapat Vitamin A | 1 = Ya, Vitamin A Biru 2 = Ya, Vitamin A Merah 3 = Ya, Vitamin A Biru dan Merah 4 = Tidak Pernah |
| 7 | B23 | Kepemilikan Kartu Menuju Sehat (KMS) | 1 = Ya, ada dokumen 2 = Ya, menurut pengakuan ibu 3 = Tidak |



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

| No | Nama Parameter | Deskripsi | Kode Pengisian |
|----|----------------|---|--|
| 8 | B26A | Berapa kali [NAMA] ditimbang dalam 6 bulan terakhir | |
| 9 | B27 | Apakah pernah dirawat karena kurang gizi | 1 = Ya 2 = Tidak |
| 10 | F01B | Nilai berat badan | |
| 11 | F02B | Nilai tinggi badan | |
| 12 | F02C | Cara ukur | 1 = Berdiri 2 = Terlentang |
| 13 | KAT_BBU | Status gizi berdasarkan indeks BB/U | 1 = Gizi buruk 2 = Gizi kurang 3 = Gizi baik |

K-Fold Cross Validation

Cross Validation merupakan teknik untuk memvalidasi keakuratan sebuah model yang dibangun berdasarkan data set tertentu. Data yang digunakan dalam proses pembentukan model disebut data latih atau *training* dan data yang digunakan untuk memvalidasi model tersebut sebagai data *testing* (Davison 2015).

Penggunaan *k-fold cross validation* untuk menghilangkan bias pada data. Pelatihan dan pengujian dilakukan sebanyak k kali. Pada percobaan pertama, subset S1 diperlakukan sebagai data pengujian dan subset lainnya diperlakukan sebagai data pelatihan, pada percobaan kedua subset S1,S3,...Sk menjadi data pelatihan dan S2 menjadi data pengujian, dan seterusnya. Tujuan digunakan *k-cross validation* agar hasil prediksi atau klasifikasi tidak hanya akurasi tinggi tetapi juga valid (Bramer 2016).

Confusion Matrix

Confusion matrix menurut Han dan Kamber dapat diartikan sebagai suatu alat yang memiliki fungsi untuk melakukan analisis apakah klasifikasi tersebut baik dalam mengenai nilai dari kelas yang berbeda (Han 2018). Tabel 2.1 berikut merupakan tabel penjelasan model *confusion matrix*:



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 2.2. Model *Confusion Matrix*

| Nilai Aktual | Nilai Hasil Prediksi | |
|--------------|----------------------|----|
| | TP | FN |
| | FP | TN |

Keterangan:

TP = nilai positif yang diklasifikasikan positif.

TN = nilai negatif yang diklasifikasikan negatif.

FP = nilai positif yang diklasifikasikan negatif.

FN = nilai negatif yang diklafisikasikan positif.

Untuk menghitung tingkat akurasi pada matriks digunakan persamaan berikut:

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.7}$$

2.7 Matrix Laboratory

Matrix Laboratory (MATLAB) merupakan perangkat lunak yang digunakan untuk program, analisis, serta komputasi teknis dan matematis berbasis matriks. MATLAB singkatan dari *Matrix Laboratory* karena mampu menyelesaikan masalah perhitungan dalam bentuk matriks. MATLAB versi pertama diliris pada tahun 1970 oleh Cleve Moler. Pada awalnya, MATLAB didesain untuk menyelesaikan masalah-masalah persamaan aljabar linear. Seiring berjalannya waktu, program ini terus mengalami perkembangan dari segi fungsi dan performa komputasi(Tjolleng 2019).

MATLAB merupakan merek *software* yang dikembangkan oleh Mathworks.Inc. merupakan *software* yang efisien untuk perhitungan numerik berbasis matriks. MATLAB yang merupakan bahasa pemrograman tingkat tinggi berbasis pada matriks sering digunakan untuk teknik komputasi numerik, untuk menyelesaikan masalah-masalah yang melibatkan operasi matematika elemen, matriks, optimasi, aproksimasi dan lain-lain. Sehingga MATLAB banyak digunakan pada(Cahyono 2013) :

- 1) Matematika dan Komputasi
- 2) Pengembangan dan Algoritma
- 3) Pemrograman modeling

- 4) Analisa Data, eksplorasi dan visualisasi
- 5) Analisa numeric dan statistik
- 6) Pengembangan aplikasi teknik

Penelitian Terkait

Penelitian yang pernah dilakukan sebelumnya tentang klasifikasi menggunakan algoritma VFC4.5, perbandingan metode algoritma C4.5 dengan *Effective* C4.5 akan dijelaskan secara rinci beserta keterangan masing-masing penelitian tersebut disajikan pada tabel 2.3 berikut:

Tabel 2.3. Penelitian Terkait

| No. | Peneliti | Metode | Tahun | Keterangan |
|-----|---|---|-------|---|
| 1. | Monalisa Mohanty, Santanu Sahoo, Pradyut Biswal, and Sukanta Sabut | <i>Efficient Classification of Ventricular Arrhythmias using Feature Selection and C4.5 Classifier</i> | 2018 | Klasifikasi menggunakan modifikasi algoritma C4.5 memiliki akurasi lebih baik dari algoritma SVM dengan hasil akurasi sebesar 97.02% dan 92.23% dengan menggunakan data yang sama yaitu 13 fitur berbasis waktu. |
| 2. | L.J. Muhammad, B. Z. Yahaya, Abdulkadir Ahmad, Muhammed Besirun Jibrin, I.A. Muhammed Besirun Jibrin, and Jamila Musa Amshi | <i>An Improved C4.5 Algorithm using Principle of Equivalent of Infinitesimal and Arithmetic Mean Best Selection Attribute for Large Dataset</i> | 2020 | Algoritma C4.5 baru yang diusulkan memiliki kompleksitas waktu yang lebih baik dari algoritma C4.5 biasa. Dan algoritma C4.5 baru yang ditingkatkan mempercepat pemrosesan atau fase pertumbuhan pohon keputusan ketika dataset yang lebih besar. |
| | Anis Cherfi, Kaouther Noura, and Ahmed Ferchichi | <i>Very Fast C4.5 Decision Tree Algorithm</i> | 2018 | Algoritma <i>Very Fast</i> C4.5 memberikan hasil akurasi yang lebih baik dibandingkan algoritma C4.5 dan algoritma VFDT (<i>Very Fast Decision Tree</i>). Dalam konstruksi data yang besar menunjukkan algoritma VFC4.5 lebih cepat dari algoritma lainnya. |

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

| No. | Peneliti | Metode | Tahun | Keterangan |
|-----|--|--|-------|--|
| | Ibomoiye Domor Mienye , Yanxia Sun, and Zenghui Wang | <i>Prediction Performance of Improved Decision Tree Based Algotihm</i> | 2019 | Algoritma C4.5 yang baru dalam pemrosesan data yang besar membutuhkan waktu yang lebih singkat dan lebih akurat dalam membangun pohon keputusan. |
| | Irfan Suhanri Damanik, Agus Perdana Windarto, Anjar Wanto, Poningsih, Sundari Retno Andani, Widodo Saputra | <i>Decision Tree Optimization in C4.5 Algorithm using Genetic Algorithm</i> | 2021 | Algoritma ini akan sangat efektif jika kasus yang dihadapi sangat kompleks dengan kondisi percabangan yang lebih banyak dengan ketepatan. |
| | Yawen Zhong | <i>Research on Construction Engeenering Project Management Optimization Based on C4.5 Improved Algorithm</i> | 2019 | Algoritma C4.5 baru yang disederhanakan untuk mempertahankan presisi tinggi yang memastikan akurasi klasifikasi sekaligus mempercepat kecepatan klasifikasi. |
| 7. | A.P. Subapriya, M. Kalimuthu, and P. Sengottuvelan | <i>Efficient Decision Tree Construction in Unrealized Dataset using C4.5 Algorithm</i> | 2016 | Pohon keputusan dari algoritma C4.5 baru dapat digunakan untuk menghasilkan fungsi keputusan yang cukup kecil dan sangat akurat. |
| | Hartana Wijaya | Optimasi Penerapan Algoritma Genetika Menggunakan Metode C4.5 untuk Memprediksi Penyakit Kanker Payudara | 2019 | Dari Penelitian yang dilakukan telah terbukti bahwa algoritma C4.5 yang di optimalkan menggunakan algoritma Genetika mampu meningkatkan akurasi prediksi pada kanker payudara. |



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

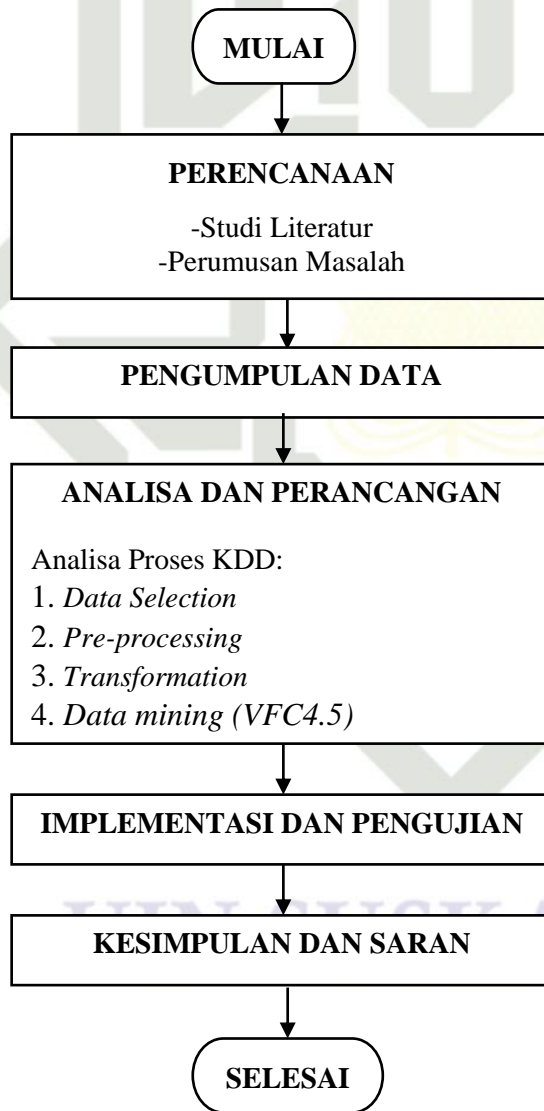
| No. | Peneliti | Metode | Tahun | Keterangan |
|-----|---|---|-------|--|
| 10. | Ahmed Abba Haruna, B. Z. Yahaya, N. D. Oye, L. J. Muhammad, E. J. Garba, and Low Tan Jung | <i>An Improved C4.5 Data Mining Driven Algorithm for the Diagnosis of Coronary Artery Disease</i> | 2020 | Algoritma C4.5 yang ditingkatkan telah menunjukkan sedikit lebih baik kinerja untuk diagnosis dataset CAD dengan akurasi keseluruhan 97.23%. |
| 10. | I.I Sinam and Abdulwahab Lawan | <i>AN Improved C4.5 Model Classification Algorithm Based on Taylor's Series</i> | 2019 | Penelitian ini mengusulkan model yang dimodifikasi yang disebut EC4.5. Eksperimen menunjukkan bahwa EC4.5 mengungguli dengan akurasi 99.40% sedangkan C4.5 hanya 51.27%. |

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB III METODOLOGI PENELITIAN

Pada metodologi penelitian ini tentunya memiliki tahapan-tahapan yang rinci dan terorganisasi, guna menjelaskan seluruh tahapan yang akan dilalui pada penelitian ini, sehingga tidak ada tahapan yang tidak dijalankan dan tahapan yang lebih dulu dilakukan sebelum tahapan sebelumnya selesai dilakukan. Berikut *flowchart* dan tahapan-tahapan pada metodologi penelitian ini :



Gambar 3.1 Tahapan Metodologi Penelitian

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.1 Perencanaan

Pada tahap perencanaan, penulis mencari jurnal dari berbagai sumber (studi literatur) hingga menemukan permasalahan yang dapat diteliti lebih lanjut. Hasil dari tahap pencarian permasalahan tersebut adalah penerapan algoritma *Very Fast C4.5* dalam mengklasifikasikan status kesehatan balita berdasarkan BB/U, pengaruh 13 parameter (atribut) terhadap data balita pada status kesehatan balita, dan membangun sistem untuk mengklasifikasikan kesehatan balita.

3.2 Pengumpulan Data

Pada tahap ini, dilakukan pengumpulan data balita untuk dianalisa. Data balita yang diperoleh dari hasil Pemantauan Status Gizi balita (PSG) tahun 2020 di Puskesmas Kampar. Data diperoleh melalui tenaga kesehatan di Puskesmas Kampar yaitu Bidan Susmita yang didapatkan data yang berjumlah 2.790 data. Terdapat 213 parameter pemantauan gizi balita tahun 2018, namun pada penelitian ini hanya 13 parameter yang digunakan sesuai dengan yang telah didiskusikan oleh tenaga kesehatan di Puskesmas Kampar (selengkapnya pada Lampiran A).

3.3 Analisa

Tahap analisa dilakukan setelah perumusan masalah dan pengumpulan data, dengan tahapan KDD sebagai berikut:

1. *Data Selection*

Pada tahap ini pemilihan parameter/atribut dari data untuk digunakan pada tahap selanjutnya. Adapun atribut yang akan diseleksi adalah ID, B4R4, UMUR_BULAN, B01A, B21, B22, B23, B26A, B27, F01B, F02B, F02C, KAT_BBU.

2. *Pre-Processing*

Tahapan ini akan dilakukan pembersihan data (*cleaning*) yang telah diseleksi untuk menghilangkan *noise* atau membersihkan data yang *missing value*, data tidak konsisten, data duplikat, dan *outlier* data balita dari Puskesmas Kampar tahun 2020. Proses pengecekan *outlier* tersebut menggunakan Ms. Excel.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. *Transformation*

Tahapan ini akan dilakukan pada data yang telah melalui proses seleksi data dan *pre-processing*, selanjutnya ditransformasi dan disimpan dalam format yang sesuai dengan *tools* yang akan digunakan. Pada tahap ini mengubah data menjadi format yang sesuai dengan yang dibutuhkan *data mining*.

4. *Data Mining*

Pada tahap *data mining* akan diterapkan teknik *data mining* yaitu menerapkan algoritma *Very Fast C4.5* menggunakan Matlab. Hasil keluaran (*output*) yang diperoleh dari penerapan algoritma *Very Fast C4.5* yaitu klasifikasi kesehatan balita yang menjadi tiga target yaitu sehat, kurang sehat, dan tidak sehat.

3.4 Perancangan

Pada tahap ini mendeskripsikan perancangan sistem yang akan dibuat berdasarkan analisa yang akan dilakukan. Perancangan sistem pada penelitian ini adalah sebagai berikut:

1. Perancangan struktur menu

Merancang menu-menu pada sistem sesuai dengan fungsi masing-masing.

2. Perancangan antarmuka (*interface*)

Merancang tampilan antarmuka (*interface*) sistem yang akan membuat interaksi antar pengguna (*user*) dengan sistem. Tampilan yang dibuat dapat memberikan gambaran umum implementasi dari sistem yang dibuat.

Implementasi

Tahap implementasi dan pengujian merupakan tahapan yang dilakukan setelah tahapan analisa dan perancangan selesai dilakukan. Pada tahap ini mendeskripsikan dan menjelaskan secara rinci bagaimana penerapan sistem yang telah selesai dianalisa dan dirancang sebelumnya. Adapun kebutuhan alat dan pendukung untuk proses implementasi yaitu perangkat keras (*hardware*) dan perangkat lunak (*software*) sebagai berikut :

1. Perangkat keras (*hardware*) yang dibutuhkan adalah:



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- a. *Processor*: Intel Core i5
 - b. *Memory*: RAM 4GB
 - c. *Harddisk*: 500GB
2. Perangkat lunak (*software*) yang dibutuhkan adalah:
 - a. *Operating System*: Microsoft Windows 10 Pro 64-bit
 - b. *Tools Implementation*: Matlab R2014a

3.6 Pengujian

Pada tahap pengujian, penulis akan mengevaluasi (*evaluation*) hasil klasifikasi algoritma *Very Fast C4.5* untuk mendapatkan nilai akurasi. Pengujian hasil klasifikasi ini akan menggunakan *confusion matrix* dan *K-fold cross validation*.

3.7 Kesimpulan dan Saran

Tahapan kesimpulan dan saran adalah tahapan terakhir dari penelitian. Penulis akan menarik kesimpulan hasil akhir penelitian yang telah dilakukan untuk mengetahui apakah hasil penelitian telah mencapai tujuan yang diinginkan dan memenuhi kriteria sesuai dengan kebutuhan serta dapat dioperasikan dengan baik sehingga dapat bermanfaat. Penulis juga akan memberikan saran untuk mengembangkan penelitian ini lebih lanjut.



BAB VI PENUTUP

6.1 Kesimpulan

Penerapan algoritma *Very Fast C4.5* untuk klasifikasi status kesehatan balita telah berhasil diuji dengan menggunakan *confusion matrix* dan memperoleh akurasi sebesar 93.52% dengan ratio data latih dan data uji 90%:10%. Pengujian kinerja algoritma dengan menggunakan *k-fold cross validation* berhasil dengan menggunakan menghasilkan akurasi tertinggi pada *k-fold=8* yaitu sebesar 96,05% dan menggunakan data 2.790 data.

6.2 Saran

Saran untuk engembangan penelitian ini di masa depan diantaranya adalah sebagai berikut:

1. Penelitian selanjutnya diharapkan untuk mengkaji lebih banyak sumber maupun referensi yang terkait dengan status kesehatan balita untuk pengembangan penelitian yang lebih baik
2. Penelitian selanjutnya diharapkan dapat mengembangkan sistem untuk mengetahui kesehatan balita berdasarkan faktor-faktor lainnya seperti faktor lingkungan.
3. Penelitian selanjutnya diharapkan dapat melakukan pengembangan aplikasi *mobile* dengan sistem status kesehatan balita.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR PUSTAKA

- Akhyar, Amany. 2018. “Klasifikasi Status Gizi Balita Menggunakan Algoritma C4.5 Dan Particle Swarm Optimization.”
- Bramer, Max. 2016. *Principles of Data Mining*. <https://doi.org/10.1007/978-1-84628-766-4>.
- Cahyono, Budi. 2013. “Penggunaan Software Matrix Laboratory (MATLAB) Dalam Pembelajaran Aljabar Linear” 1:45–62.
- Chamidah, Atien Nur. 2009. “Deteksi Dini Gangguan Pertumbuhan Dan Perkembangan Anak.”
- Cherfi, Anis, Kaouther Nouira, and Ahmed Ferchichi. 2018. “Very Fast C4 . 5 Decision Tree Algorithm.” *Applied Artificial Intelligence* 0 (0). Taylor & Francis:1–19. <https://doi.org/10.1080/08839514.2018.1447479>.
- Davison, Anthony. 2015. “Bootstrap Methods and Their Application,” no. September. <https://doi.org/10.2307/1271471>.
- Fajri, Ika Nur. 2017. “Analisi Perfoma Algoritma Klasifikasi Pada Pengelompokan Benih Gandum” 18.
- Florentianus Tat, Aben B.Y.H. Rimana. 2018. “Hubungan Tingkat Pengetahuan Dan Perilaku Merawat Bayi Dengan Status Kesehatan Bayi” 2 (April).
- Han, Jiawei. 2018. “Data Mining : Concepts and Techniques.”
- Heri, Lilis, and Mis Cicih. 2011. “Pengaruh Perilaku Ibu Terhadap Status Kesehatan Anak Baduta Di Provinsi Jawa Tengah” 13 (1).
- Indonesia, Kementerian Kesehatan Republik. 2018. “Bahan Ajar Gizi: Penilaian Status Gizi.” <http://bpsdmdk.kemkes.go.id/pusdiksmdk/wp-content/uploads/2017/11/PENILAIAN-STATUS-GIZI-FINAL-SC.pdf?opwvc=1>.
- Kesehatan, Dinas. 2018. “Laporan Kinerja Instansi Pemerintah Provinsi Riau Tahun 2018.”
- Larose, Daniel T. 2005. *Discovering Knowledge in Data*.
- Mardi, Yuli. 2017. “Klasifikasi Menggunakan Algoritma C4.5.”
- Oded Maimon, Lior Rokach. 2010. *Data Mining and Knowledge Discovery Handbook*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Okfalisa, Ikbal Gazalba, Mustakim, Nurul Gayatri Indah Reza. 2019. "Comparative Analysis of K-Nearest Neighbor and Modified K-Nearest Neighbor Algorithm for Data Classification," 1–5.

Okfalisa, Ratika Fitriani, Yelfi Vitriani. 2018. "The Comparison of Linear Regression Method and K-Nearest Neighbors in Scholarship Recipient." IEEE, 194–99.

Putri, Desti Sagita dan Dadang Sukandar. 2012. "Keadaan Rumah, Kebiasaan Makan, Status Gizi, Dan Status Kesehatan Balita Di Kecamatan Tamansari, Kabupaten Bogor" 7 (November):163–68.

Republik, Indonesia Kementerian Kesehatan. 2020. "Laporan Akuntabilitas Kinerja Tahun 2020." <https://e-renggar.kemkes.go.id/file2018/e-performance/1-131313-1tahunan-211.pdf>.

Sukei, Niken, Desi Rina, and Kurniawati Emilia. 2011. "Manajemen Penatalaksanaan Deteksi Dini Tumbuh Kembang Anak Di Posyandu Kelurahan Manyaran."

Tjolleng, Amir. 2019. "Panduan Praktis Belajar MATLAB," no. August.

Zulhaida Lubis, Isyatun Mardiyah Syahri. 2015. "Pengetahuan Dan Tindakan Kader Posyandu Dalam Pemantauan Pertumbuhan Anak Balita" 11 (1):65–73.

Zuviyanto, Edi, Teguh Bharata Adji, and Noor Akhmad Setiawan. 2018. "Perbandingan Algoritme- Algoritme Pembelajaran Mesin Pada Klasifikasi SMS Spam," 20–26.



DAFTAR RIWAYAT HIDUP

DATA PRIBADI



| | |
|----------------------|---------------------------------------|
| Nama | Puspita Sari |
| Tempat/Tanggal Lahir | Dusun VII Sidorukun, 02 Februari 1997 |
| Jenis Kelamin | Perempuan |
| Anak Ke- | 1(Satu) dari 3 Bersaudara |
| Tinggi Badan | 163 cm |
| Berat Badan | 57 kg |
| Kebangsaan | Indonesia |

KONTAK

| | |
|-----------------|--|
| Alamat | Jl. By Pass Dusun II Sidorukun |
| Nomor Handphone | 081270337508 |
| Email | puuspitaa@gmail.com |

RIWAYAT PENDIDIKAN

| | |
|---|-----------------|
| SD NEGERI 116244 SIDORUKUN | Tahun 2003-2009 |
| SMP NEGERI 1 BILAH HULU | Tahun 2009-2012 |
| SMA NEGERI 1 BILAH HULU | Tahun 2012-2015 |
| UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU | Tahun 2015-2021 |

UIN SUSKA RIAU

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.